AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

REVIEW

10.1002/2015JC010716

**Key Points:**

- A review is given of 30 trend models applied in the field of sea level research
- Varying trend patterns can be found for the same data depending on the method chosen
- Contradictory trend inferences can be avoided by applying good modeling practices

**Correspondence to:**
H. Visser,
hans.visser@pbl.nl

**Citation:**
Visser, H., S. Dangendorf, and A. C. Petersen (2015), A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate", *J. Geophys. Res. Oceans*, *120*, 3873–3895, doi:10.1002/2015JC010716.

Received 12 JAN 2015
Accepted 8 MAY 2015
Accepted article online 13 MAY 2015
Published online 5 JUN 2015



# A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate"

**Hans Visser[1], Sönke Dangendorf[2], and Arthur C. Petersen[3,4]**

[1]PBL Netherlands Environmental Assessment Agency, Bilthoven, Netherlands, [2]Research Institute for Water and Environment, University of Siegen, Siegen, Germany, [3]Department of Engineering, Science, Technology and Public Policy, University College London, London, UK, [4]Institute for Environmental Studies, VU University Amsterdam, Amsterdam, Netherlands

**Abstract** Global sea levels have been rising through the past century and are projected to rise at an accelerated rate throughout the 21st century. This has motivated a number of authors to search for already existing accelerations in observations, which would be, if present, vital for coastal protection planning purposes. No scientific consensus has been reached yet as to how a possible acceleration could be separated from intrinsic climate variability in sea level records. This has led to an intensive debate on its existence and, if absent, also on the general validity of current future projections. Here we shed light on the controversial discussion from a methodological point of view. To do so, we provide a comprehensive review of trend methods used in the community so far. This resulted in an overview of 30 methods, each having its individual mathematical formulation, flexibilities, and characteristics. We illustrate that varying trend approaches may lead to contradictory acceleration-deceleration inferences. As for statistics-oriented trend methods, we argue that checks on model assumptions and model selection techniques yield a way out. However, since these selection methods all have implicit assumptions, we show that good modeling practices are of importance too. We conclude at this point that (i) several differently characterized methods should be applied and discussed simultaneously, (ii) uncertainties should be taken into account to prevent biased or wrong conclusions, and (iii) removing internally generated climate variability by incorporating atmospheric or oceanographic information helps to uncover externally forced climate change signals.

## 1. Introduction

It is expected that sea level rise will have large impacts along coastal zones of the world [*Nicholls and Cazenave*, 2010; *Church et al.*, 2013]. Therefore, the estimation of long-term trends and the establishment of incremental curves (i.e., the first derivatives) thereof, showing sea level changes in mm/yr, are of eminent importance. Accelerations or decelerations follow directly from these incremental patterns. Knowledge about accelerations is crucial because their detection may have important consequences for coastal protection planning and climate change related measures.

This has led to an intensive debate on the existence of significant accelerations in regional and global sea level in recent years. Some researchers report accelerations where others find linear trends or even decelerations, sometimes based on the same data. Also the direct relation between historic and future trends is debated. Some authors state that a lack of acceleration of (global) sea levels contradicts increasing sea levels in the coming decades [e.g., *Houston and Dean*, 2011a]. Others show that making such a link between historic and future trends is far too simple [e.g., *Dangendorf et al.*, 2014a; *Haigh et al.*, 2014], as the signal requires decades to become distinguishable from the superimposed noise [*Lyu et al.*, 2014].

While the search for the existence of accelerations is a topic with a long lasting history [e.g., *Douglas*, 1991; *Church et al.*, 2006; *Jevrejeva et al.*, 2008], the most recent debate was stimulated by the publications of *Houston and Dean* [2011a, 2011b, 2011c]. The authors reported that the relative sea level patterns along the U.S. coast and a selection of global tide gauges were slightly decelerating. *Rahmstorf and Vermeer* [2011] found the arguments of Houston and Dean not convincing and showed that accelerations are present if the data are analyzed in a different way, using different starting years for the trends. Furthermore, *Donoghue and Parkinson* [2011] concluded that the Houston and Dean publication has no relevance for future sea level rise, since climate models project that the current century will be quite dissimilar from the past. *Baart et al*. [2012b] made a summary of this debate and provided a list of six guidelines for best modeling practices in this field of research. In a more recent article, *Houston and Dean* [2013] accuse "previous authors" of neglecting decadal variations due to natural cycles, such as a 60 year cycle, resulting in invalid conclusions. *Houston and Dean* [2013] argue that tide gauge series shorter than 75 years should not be analyzed at all to avoid erroneous conclusions on accelerations.

Next to the 2011a study of Houston and Dean, *Watson* [2011] reported decelerating rather than accelerating tide gauge series along the Australian coastline. His publication led to a fierce debate in the media and blogosphere [*Rintoul*, 2011]. The debate on (a lack of) accelerations around Australia was continued by *Boretti* [2012] who also claimed that tide gauge data around Australia do not show any sign of acceleration at present time and that the same holds for global data. He directly coupled these findings to future sea level patterns which would not accelerate either. However, *Hunter and Brown* [2012] responded to this study by showing that accelerations are in line with those presented in *IPCC* [2007] if *uncertainty bands* in the acceleration estimates are taken into account. Again, this discussion was continued in the media. The "Australia debate" had a continuation, with Boretti publishing under a different name (Parker). This led to renewed discussions on the same theme in a different journal (*Parker et al*. [2013] and the comments thereon by *Hunter* [2013]). Also see *Hunter et al*. [2013a], *Watson and Parker* [2013], and *Hunter et al*. [2013b].

In this article, we want to shed light on these controversies and show how researchers could avoid misunderstandings. Our approach is of a mathematical/statistical (methodological) rather than an oceanographic nature since we believe that much of the misunderstandings/controversies are due to mathematical or statistical characteristics of the models applied.

First, we provide a number of trend definitions and, subsequently, a review of trend methods which have been applied in this field of research. We show that 30 methods are in use, all with different characteristics. Subsequently, we argue that researchers should be open in their arguments for choosing one or more trend methods in their analysis, since trend patterns may depend on these choices. Just by choosing a different model one can find decelerations where another trend model shows a linear or even accelerating pattern. The same holds for choosing different flexibilities within one particular trend model. Also, the addition of external factors or information (such as knowledge about climate internal processes driving large parts of the observed ups and downs in sea level along the coast) to trend estimation may seriously influence trend estimates and the uncertainties attached to these estimates. Finally, many trend models appear to be sensitive to the sample period chosen. Based on these findings, we review varying techniques present in the literature which might point to the suitability or nonsuitability of trend models. Additionally, we formulate a number of recommendations or "good modeling practices."

The content of this article is as follows. We start with preliminary information and definitions in section 2. A review of trend methods is given in section 3. In section 4, we give four examples of how conflicting trend patterns may arise from the same data set, in this case, the tide gauge series of San Francisco (1855–2014). Methods to overcome such conflicting results are reviewed in section 5 (model selection techniques) and section 6 (good modeling practices). Conclusions are given in section 7.

## 2. Preliminaries

### 2.1. Why Estimating Trends?

Why are trends in sea level records estimated? Generally, three goals/applications of trend estimation can be discerned. The first is the one most frequently encountered in sea level research, and climate research in general. Trends are presented to highlight nonstationary behavior in a series of sea levels, temperatures, precipitation totals, wind speeds, etc. In doing so, trend estimation may play a role in the *detection* of climate change [*IPCC*, 2013, chap. 10]. The trend, a low-frequency signal, is interpreted as a climate signal, while the residuals can be seen as natural variability. Thus, trends are seen as a smooth signal where the smoothness is chosen to filter out shorter-term (decadal) natural fluctuations.

Second, trends may be viewed in terms of *prediction*. The estimated trend is conceived as that part of the series which, when extrapolated, provides the clearest indication of the future long-term movements in the series. This application of trends is very popular in fields such as business planning, financial market analyses, and econometrics [e.g., *Harvey*, 1989, 2001; *Makridakis and Hibon*, 2000; *Armstrong*, 2001; *Soyer and Hogarth*, 2012]. Forecasting by extrapolation is sometimes applied in short-term weather prediction, and it has shown up recently in the "sea level debate," described in the introduction. The prediction performance of various models can also be used as tool for model selection, as we will point out in section 5.2.

Third, sea level data and trends therein are used for *calibration* of recent attempts to model historic thermal expansion as well as glacier and ice sheet melting. See *Church et al*. [2013, Table 13.1 and Figure 13.7] and literature cited therein for examples. The outputs, based on global climate models such as CMIP5 AOGCMs, are then used to project global and regional sea level rise in the coming century [e.g., *Church et al*., 2013, Figures 13.23 and 13.24]. A recent example for the validation of global surface temperature modeled with 114 CMIP5 AOGCMs has been given by *Marotzke and Forster* [2015], who compare 15 and 62 year linear trends from models and observations.

### 2.2. Definitions

In the following, we introduce a number of definitions used throughout the review. We review analyses of sea level data on global, regional, or local scales. The analyzed records are mostly presented as series $y_t$ of annual or monthly means (expressed in mm). Since we are interested in trend patterns and derivatives thereof, the particular choice of a fixed zero point is not of importance here.

We denote a trend in data $y_t$ by $\mu_t$, which is again expressed in mm. The first derivative or *trend increment* is then denoted by $\Delta\mu_t \equiv [\mu_t - \mu_{t-1}]$, expressed in mm/yr. Accelerations in trends, or accelerations in short, are defined as the second derivative of the trend:

$\Delta^2\mu_t \equiv [\mu_t - \mu_{t-1}] - [\mu_{t-1} - \mu_{t-2}] = [\mu_t - 2\mu_{t-1} + \mu_{t-2}]$, expressed in mm/yr$^2$.

There are many ways that trends can be estimated in sea level records, and it might be helpful to group trend methods into distinct classes. The following categorization follows the chapters in the book of *Chandler and Scott* [2011], which we see as the main reference for this review of methods. We discern five classes:

1. Exploratory data analysis. Trends in this group are simple low-pass filters (or smoothed time series), such as moving averages, and the wide class of linear filters [*Chandler and Scott*, 2011, chap. 2]. Also, visual inspection and expert judgment of patterns is a form of exploratory analysis.

2. Parametric trend estimation. Parametric models give a precise description of the trend shape and, possibly, the relation to external variables. Examples are the linear trend and the second-order polynomial (parabolic) fit, which can be seen as part of the group of multiple regression models. Most of the models in this group are of a statistical nature. Herewith we mean that the model reads as $y_t = f(\mu_t, \varepsilon_t)$, where $\varepsilon_t$ represents an error term (a white or red noise process or a long-term correlated process). In the following, only linear relations are dealt with, thus $y_t = \mu_t + \varepsilon_t$ or $y_t = \mu_t + \alpha_1 X_{1,t} + \alpha_1 X_{2,t} + \varepsilon_t$, where $X_{1,t}$ and $X_{2,t}$ are external variables such as climate indices. Also cycles can be added here. For the well-known linear trend model, it holds that accelerations are zero for all times t, or in mathematical terms:

$$\mu_t - 2\mu_{t-1} + \mu_{t-2} = 0 \tag{1a}$$

For the parabolic fit, we have a constant acceleration, or

$$\mu_t - 2\mu_{t-1} + \mu_{t-2} = c \tag{1b}$$

In some cases, other deterministic curves are fitted to the data, such as an exponential curve with three unknown parameters [see *Chandler and Scott*, 2011, chap. 3]. For the modeling of extreme sea levels, the group of nonstationary extreme value trend models are of importance [*Coles*, 2001; *Mudelsee*, 2014].

3. Nonparametric trend estimation. Trend models in this group do not have a predefined functional relation as models in class 2. Thus, the shape of trends is more "data driven." Examples are spline smoothing, Lowess smoothing, and wavelets [see *Chandler and Scott*, 2011, chap. 4].

4. Stochastic trend models. Stochastic trends are trends for which the behavior is assumed to originate from underlying noise processes. Examples are ARIMA models and Structural Time series Models (STMs). An example of the latter group is the Integrated Random Walk (IRW model) which reads as

$$\mu_t - 2\mu_{t-1} + \mu_{t-2} = \eta_t \tag{1c}$$

with $\eta_t$ being a white noise process [see *Chandler and Scott*, 2011, chap. 5; *Mudelsee*, 2014, chap. 2].

5. Miscellaneous models. A number of models are based on parametric trend estimation (class 2), but applied on shifting windows, thus revealing possibly nonstationary behavior. Other models are not related to classes 1–4. Examples are Artificial Neural Networks (ANN) or Singular Spectrum Analysis (SSA).

### 2.3. Are Some Models Better Than Others?

A natural question that arises in the light of so much choice in trend models is: is there a best trend method? Strictly speaking, the answer should be "no." Since we do not know what the "best" or "true" trend is, we also do not know which trend model comes closest to this "true" trend. Furthermore, the definition of "trend" is not unique. In section 2.1, we have formulated two definitions: (i) trends are seen as a smooth signal where the smoothness is chosen to filter out shorter-term (decadal) natural fluctuations, and (ii) trends may be viewed in terms of *prediction* where the estimated trend is conceived as that part of the series which, when extrapolated, provides the clearest indication of the future long-term movements in the series [*Harvey*, 2001]. And both definitions are pretty vague: what is "smoothness" or what are "long-term movements"?

As a consequence, the practitioner's choice of a specific trend model might be steered by the characteristics of the method, as reflected by the five classes in section 2.2. We will describe these characteristics in more detail in section 3. However, within the field of statistical models (classes 2 and 4; class 5 in part), a number of model selection techniques have been developed. To use a quote from *Burnham and Anderson* [2002], a much-cited book on model selection: *while a model can never be "truth", a model might be ranked from very useful, to useful, to somewhat useful to, finally, essentially useless*. Here Burnham and Anderson go one step further than statistician G.E.P. Box in his quote: *all models are wrong, but some are useful*.

We will discuss model selection techniques in section 5 where we will address results from two "schools" of statistical inference, the classical or frequentist approach, and the Bayesian approach where probability is interpreted as degrees of belief. As for Bayesian model selection, the following ratio is often evaluated:

$$p(y_t|M_1)/p(y_t|M_2) = [p(M_1|y_t)/p(M_1)]/[p(M_2|y_t)/p(M_2)] \tag{2}$$

where $M_1$ and $M_2$ are two models to be compared, and $y_t$ the data observed. The term $p(M_1|y_t)$ stands for the so-called posterior belief we have in model $M_1$, given the data $y_t$. The term $p(M_1)$ stands for the prior belief in model $M_1$. This ratio is denoted as the Bayes factor [*Gill*, 2008, section 7.3]. Clearly, a ratio above 1.0 will support model $M_1$. We will introduce other selection indices, such as Akaike's Information Criterion (AIC) in section 5.2.

However, selection criteria such as given in equation (2) do not give a definite list of model performance (such as suggested by the quote of Burnham and Anderson). First, many models are not statistical in nature, such as models in classes 1, 3, and 5. Second, the belief in model $M_1$ or model $M_2$ will depend on expert judgment and is thus necessarily subjective. Third, indices such as AIC have underlying assumptions. Therefore, we will formulate a number of good modeling practices in section 6 which are *complementary* to techniques given in section 5.

## 3. A Review of Trend Methods

### 3.1. Trends in the Field of Sea Level Research

We have scanned the sea level literature to see which mathematical/statistical methods have been used to estimate trends. It appears that a wide set of methods are in use, varying from simple methods, such as moving averages or linear trends, to complex approaches such as extended multiple regression methods based on generalized additive models, with corrections for correlated noise (GAMs) [*White et al*., 2014]. We identified 30 trend methods, all given in Table 1.

The fact that so many different trend methods are found in the literature is not unique. This situation also applies to estimating trends in global or hemispheric temperature data. *Mills* [2010] gives a discussion in this field. He concludes with an old saying: "There are many ways to skin a cat."

**Table 1.** Summary of Trend Models[a]

| Number | Trend Model | Short Description | Literature Examples |
|---|---|---|---|
| 1-1 | Moving averages | Arithmic mean of data within a moving window. Width of the window steers the trend flexibility | *Rhein et al.* [2013, Figure 3.12], *Watson* [2011, Figure 5], *Houston and Dean* [2013], and *Hay et al.* [2015, Figure 3] |
| 1-2 | Median filter | Median of data within a moving window. The width of the window steers the trend flexibility | *Douglas* [1991] |
| 1-3 | Low-pass FIR filter | Filter method using an impulse response function to weigh data lying within a window | *White et al.* [2014, Figure 4] |
| 1-4 | Low-pass cosine-bell filter | Filter method using cosine-shaped weighing factors | *Woodworth et al.* [2009, Figure 2] |
| 1-5 | Mann Kendall test | Test for monotonic trends. This is a classical trend test, which may be extended to the Theil-Slope estimator | *Wahl et al.* [2011] and *Chandler and Scott* [2011, section 2.4] |
| 1-6 | Expert judgment | Trends are identified by visual inspection of the data, combined with expert judgment | *IPCC WG I Reports* [1990, 1995, 2001] |
| 2-1 | Linear OLS | Linear trend based on ordinary least squares (OLS) optimization | *Church and White* [2011], *Wahl et al.* [2013, Figure 2], *Mudelsee* [2014, section 4], and *Marotzke and Forster* [2015] |
| 2-2 | Linear with nonwhite noise | As model 2-1 but with a correction for confidence limits, based on the assumption of AR(1) residuals (so-called $\rho$ correction). Another approach is that of *Prais and Winsten*. Here both intercept and slope are corrected, along with confidence limits | *IPCC* [2013, chapter 2/SM-11], *White et al.* [2014], *Prais and Winsten* [1954], and *Brown et al.* [2011] |
| 2-3 | Piecewise linear OLS | As model 2-1, but now for fixed segments of the time series at hand (consecutive segments) | *Bromirski et al.* [2011] and *Olivieri and Spada* [2013, Figure 3] |
| 2-4 | Linear with cycle | Linear trend and cycle based on nonlinear optimization | *Chambers et al.* [2012] and *Baart et al.* [2012a] |
| 2-5 | Quadratic OLS | Second-order polynomial trend (parabole fit) | *Houston and Dean* [2011a], *Woodworth et al.* [2011], and *Boon* [2012] |
| 2-6 | Quadratic with cycle | Second-order trend and cycle, based on nonlinear optimization | *Houston and Dean* [2011d] |
| 2-7 | Multiple regression | Regression model where the influence of explanatory variables is modeled in an additive way, next to noise | *Zhang and Church* [2012], *Albrecht and Weisse* [2012], *Dangendorf et al.* [2013a, 2014b] |
| 2-8 | Exponential growth curve | Fitting an exponential curve to the data with three shape parameters. Based on nonlinear optimization | *Parker et al.* [2013] |
| 2-9 | Nonstationary extreme value trend models | Trends where residuals are drawn from extreme value GEV distributions. Applications in annual maximum sea levels | *Coles* [2001, section 6.3.1] and *Mudelsee* [2014, section 6] |
| 3-1 | Extended regression with splines (GAMs) | Generalized additive models, or GAMs, are models in which measurements depend linearly on unknown smooth functions | *Chandler and Scott* [2011, section 4.2.1] and *White et al.* [2014] |
| 3-2 | Lowess smoothing | Nonparametric regression technique using kernel weight functions to define the "neighborhood" of each covariate value of interest | *Barbosa et al.* [2004] and *Chandler and Scott* [2011, section 4.3.1] |
| 3-3 | Quantile regression | In quantile regression (linear) trends are estimated for predefined quantiles, such as medians, 0.10 and 0.90 quantiles | *Donner et al.* [2012] and *Chandler and Scott* [2011, section 4.3.5] |
| 3-4 | Wavelets | A wavelet consists of brief wave-like oscillations with amplitudes beginning and ending with zero | *Jevrejeva et al.* [2006] and *Chandler and Scott* [2011, section 4.3.2] |
| 4-1 | ARIMA models | Autoregressive integrated moving average models, or ARIMA(p,d,q) in short. The parameters p, d, and q are 0, 1, or 2 in practice | *Chandler and Scott* [2011, section 5.2.1], *Imani et al.* [2014], and *Mudelsee* [2014, section 2.3] |
| 4-2 | STMs | Structural time series models, composed of additive components (flexible trend, explanatory variables, cycles, and noise) | This article, *Chandler and Scott* [2011, section 5.5], and *Visser and Petersen* [2012] |
| 4-3 | Long-memory models (ARFIMA) | Slow-decaying persistence models, based on fractional ARIMA models (cf. model 4-1). Now, the difference operator d is fractional | *Chandler and Scott* [2011, section 5.3], *Franzke* [2012], and *Mudelsee* [2014, section 2.4] |
| 5-1 | Linear with shifting windows | Linear OLS trends (model 2-1), estimated on shifting windows. Visual display by showing trend slopes | *Holgate* [2007], *Church and White* [2011, Figure 8], and *Haigh et al.* [2014] |
| 5-2 | Quadratic with a variety of windows | Second-order polynomial trends (model 2-5), estimated on shifting windows and varying window lengths. Presentation by color graphs | *Jevrejeva et al.* [2008, 2014] and *Haigh et al.* [2014, Figures 3 and 4] |
| 5-3 | MSDA diagrams | Second-order polynomial trends, estimated on shifting windows and varying window lengths (scales). Presentation by color graphs | *Scafetta* [2014] |
| 5-4 | SLRD with $\rho$ correction | Sea level rate differences, computed as the differences between two linear trend slopes, computed over two | *Sallenger et al.* [2012] and *Calafat and Chambers* [2013] |

**Table 1.** (continued)

| Number | Trend Model | Short Description | Literature Examples |
|---|---|---|---|
| | | consecutive windows. The analysis is repeated for shifted windows | |
| 5-5 | SLRD with cycle | Idem, with cycle added, based on nonlinear optimization | *Houston and Dean* [2013] |
| 5-6 | EEMD or EMD | Empirical mode decomposition, decomposes data into a set of intrinsic mode functions (IMFs). The residual is the trend function | *Breaker and Ruzmaikin* [2013], *Chen et al.* [2014], and *Ezer et al.* [2013] |
| 5-7 | Singular spectrum analysis | Data are decomposed with data-adaptive orthogonal filters separating the data into a nonlinear low-frequency trend, quasiperiodic signals and noise | *Jevrejeva et al.* [2006] and *Wahl et al.* [2011, 2013] |
| 5-8 | MARS regression | Multivariate adaptive regression splines (MARS) is a nonparametric regression technique, able to deal with nonlinearities and interactions between variables | *Feng et al.* [2004] |
| 5-9 | Artificial neural networks (ANNs) | Artificial neural networks or ANNs mimic nervous systems by formulating a complex set of regression equations | *Imani et al.* [2014] |
| 5-10 | Single and multitaper spectral analysis | Spectral smoothing using a general data weighing technique (tapering) | *Baker and McGowan* [2014] and *Mudelsee* [2014, section 5.2] |

[a]All models except 4-2 and 4-3 have been applied in sea level research (cf. section 3.2). Trend models are grouped into five classes, defined in section 2.2, and described in more detail in Appendix A.

Since trend characteristics are important for the choice of a specific trend model (cf. section 2.3), we have shortly summarized these characteristics in the third column of Table 1. More details are given in Appendix A where characteristics are treated for each of the five classes defined in section 2.2. For example, if one is interested in trend estimates for the full sample period, one should not choose moving averages since the method does not yield information for the final years of the data (due to the window chosen). Consequently, short-term predictions are not generated either. Furthermore, the method is not statistical in nature and uncertainties are not provided. On the other hand, the method is mathematically simple and by choosing specific window lengths one can discard short-term variability or cyclic behavior in the data at hand. As a result, the intended readers/users will understand the background of the method more easily, a communicational advantage.

As an illustration of various trend methods we choose the tide gauge series of San Francisco, since it is long and well documented [*Smith*, 2002]. This series extends over the period 1855–2014 and can be downloaded from the PSMSL database (site code 10) [*Holgate et al.*, 2013]. We found the following 12 trend analyses for this single tide gauge series (be it over slightly different sample periods):

1. Moving averages: *Douglas* [1991, Figure 2], *Smith* [2002, Figure 6], and *Rhein et al.* [2013, Figure 3.12b].
2. Median filter: *Douglas* [1991, Figure 1].
3. Smoothed curve, method unclear: *Breaker and Ruzmaikin* [2013, Figure 1].
4. Low-pass cosine-bell filter: *Woodworth et al.* [2009, Figure 2b].
5. Linear trend: *Smith* [2002, Figure 6].
6. Linear trend plus 55 year oscillation: *Chambers et al.* [2012, Table 2].
7. Piecewise linear trend: *Bromirski et al.* [2011, Figure 1] and *Woodworth et al.* [2011, Figure 1b].
8. Second-order polynomial fit: *Houston and Dean* [2011a, Table 1] and *Breaker and Ruzmaikin* [2013, Figure 4].
9. The Sea Level Rate Difference technique (SLRD): *Sallenger et al.* [2012, Figure 2].
10. SLRD technique combined with sine functions: *Houston and Dean* [2013, Figure 5].
11. Emperical Mode Decomposition (EMD): *Breaker and Ruzmaikin* [2013, Figure 5].
12. Comparing detrended sea levels with air pressure (detrending method unclear): *Miller and Douglas* [2007, Figures 4 and 5].

### 3.2. Trends in the Broader Field of Climate Research

So far, we have discussed the wide range of trend methods used in sea level research. However, that is not all there is. We found the following trend models in the broader field of climate change research (without being complete): low-pass filters with various binomial weights (with or without end point estimates by padding), kernel smoothers, Restricted Maximum Likelihood AR(1)-based linear trends, Bayesian trend models, GARCH models, smooth transition models, exponential smoothing, Seidel-Lanzante trends incorporating abrupt changes, Holmes double-detrending methods, and Students *t* test on subperiods in time and long-memory trend models. Many of these methods are described in *Chandler and Scott* [2011] and *Mudelsee* [2014].

A number of these methods are interesting to apply in sea level research. Here we draw attention to two of these trend models, namely Structural Time series Models (STMs) and long-memory trend models (models 4-2 and 4-3 in Table 1, respectively). STMs are ideally suited to analyze sea level data since they can be seen as a natural extension of the linear Ordinary Least Squares (OLS) trend line, yielding more flexibility, and of the multiple regression model while retaining all relevant uncertainty information [Harvey, 1984; *Visser and Molenaar*, 1995; *Visser*, 2004; *Visser et al*., 2014, supporting information].

STMs are estimated by the Kalman filter and are fully based on the principle of one-step-ahead predictions, as is the basis of the second definition of a trend, named in sections 2.1 and 2.3. The Kalman filter equations were formulated around 1960 and were directly applied in many research fields such as aerospace [*Grewal and Andrews*, 2010]. The filter is often applied for data assimilation and is not unknown in oceanographic research, e.g., *Carmillet et al*. [2001], *Wenzel and Schröter* [2014], or *Hay et al*. [2015]. The filter is attractive since it yields the so-called minimum mean square estimator (MMSE) for relevant elements in the model, such as the trend estimate $\mu_t$.

We give an example of STMs in section 4.4. We will compare two STMs. The first STM reads as:

$$y_t = \mu_t + \varepsilon_t \quad \text{and} \quad \mu_t - 2\mu_{t-1} + \mu_{t-2} = \eta_t \tag{3a}$$

where $y_t$ denotes a measurement at time t; $\eta_t$ and $\varepsilon_t$ are independent, normally distributed, white noise processes with zero mean. This model is denoted as the Integrated Random Walk (IRW) trend model. The second STM is a regression-like extension and reads as:

$$y_t = \mu_t + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \varepsilon_t \quad \text{and} \quad \mu_t - 2\mu_{t-1} + \mu_{t-2} = \eta_t \tag{3b}$$

One of the important advantages of models (3a) and (3b) is their treatment of uncertainties. Uncertainties are available not only for trend estimates $\mu_t$ but also for any trend increment $[\mu_t - \mu_s]$.

The second group of trend models are called long-memory models, meaning that noise processes show slow-decaying autocorrelated behavior (following a power law), possibly around a (deterministic) trend. Such a noise behavior could arise in complex systems such as climate due to natural processes which can be seen as a superposition of smaller-scale processes, many of which have quite different dynamics [*Bos et al*., 2013; *Chandler and Scott*, 2011, section 5.3; *Mudelsee*, 2014, section 2.5; *Dangendorf et al*., 2014a; *Becker et al*., 2014].

Long-memory models are increasingly applied in climate research such as the analysis of hemispheric temperature series. One promising approach is the use of ARIMA(p,d,q) models where the differencing parameter d may be fractional. These models are denoted as fractional autoregressive integrated moving average models or ARFIMA(p,$\delta$,q) in short. ARFIMA models have been applied in three ways: (i) as fractional noise only, thus as ARFIMA(0,$\delta$,0) models [*Franzke*, 2012], (ii) as fractional noise around a linear trend [*Baillie and Chung*, 2002], and (iii) as full ARFIMA models [*Yuan et al*., 2014]. Methods related to ARFIMA models are also introduced in the latter reference.

In the following section, we show that the choice of a particular trend model, in combination with certain parameter choices, is not without consequences. It could be that one model shows an accelerating trend, another model a linear pattern, a third method a decelerating pattern, and a fourth method a trend pattern with alternating periods of accelerations and decelerations.

## 4. Examples With Conflicting Outcomes

It is well known that calculations from varying models may lead to different output results. These differing results may be due to uncertainties in parameters within the model as well as due to varying model structures [*Visser et al*., 2000; *Refsgaard et al*., 2006]. In some cases, results may be contradictory. Since such situations play a role in the "acceleration-deceleration debate," we have chosen here to give four situations

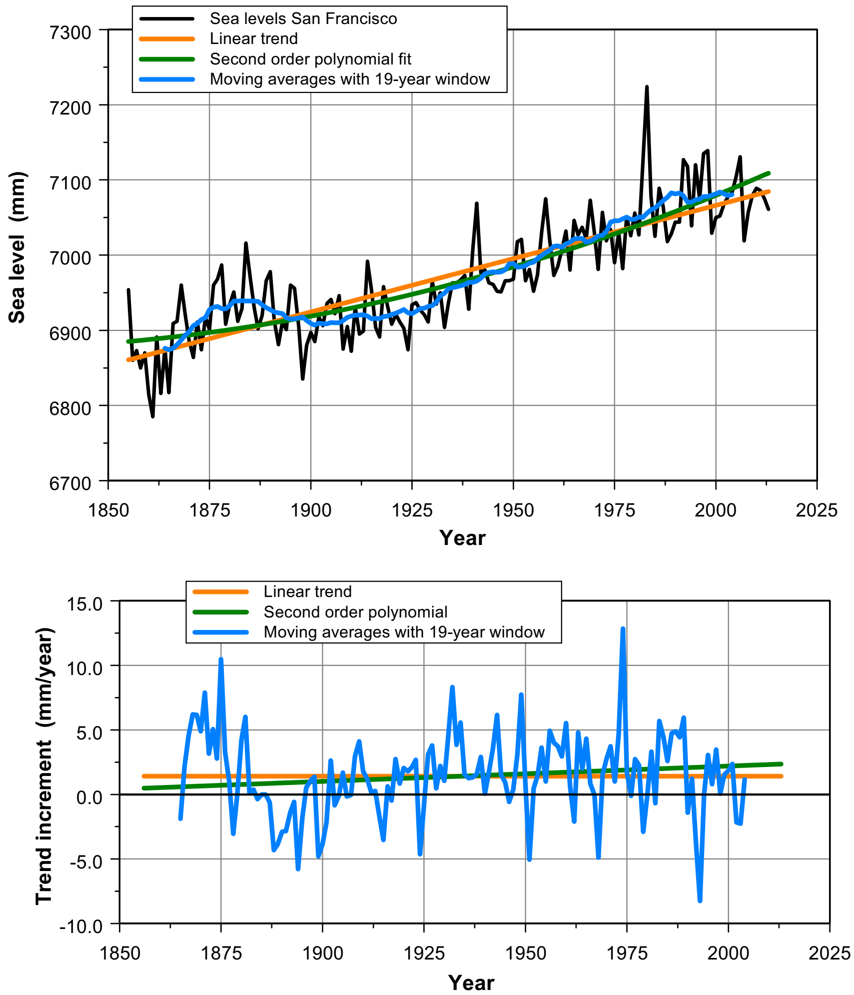



**Figure 1.** Three trend models $\mu_{1,t}$, $\mu_{2,t}$, and $\mu_{3,t}$, all estimated on the San Francisco tide gauge data: linear, second-order polynomial, and moving averages with a 19 year window. Sample period is 1855–2013. (bottom) The three corresponding trend increment series $[\mu_{1,t} - \mu_{1,t-1}]$, $[\mu_{2,t} - \mu_{2,t-1}]$, and $[\mu_{3,t} - \mu_{3,t-1}]$.

where such contradictory results arise. All illustrations here are based on the San Francisco tide gauge series, 1855–2013.

### 4.1. Varying Models, Different Results

Figure 1 shows the estimates of three trend models for the San Francisco tide gauge data: (i) moving averages (MA) with a window of 19 years, (ii) a linear trend, and (iii) a second-order polynomial fit. The top plot shows that the results for the linear and second-order polynomial are quite similar although the second-order polynomial shows a small acceleration. The MA model shows a different pattern: alternating periods with both accelerations and decelerations.

This is highlighted more clearly in the bottom plot where the corresponding trend increments $[\mu_t - \mu_{t-1}]$ are shown. It demonstrates the important role of internal variability in the record as suggested by many authors before [e.g., *Chelton and Davis*, 1982; *Meyers et al.*, 1998; *Miller and Douglas*, 2007; *Bromirski et al.*, 2011; *Thompson et al.*, 2014]. See section 6.3 for a more elaborate discussion.

The curves shown do not have uncertainty bands. This is logical for the MA model since this model is not statistical in nature. For the linear and the second-order polynomial model, uncertainties are not given since residuals are not white noise. We return to this point in section 5.1.

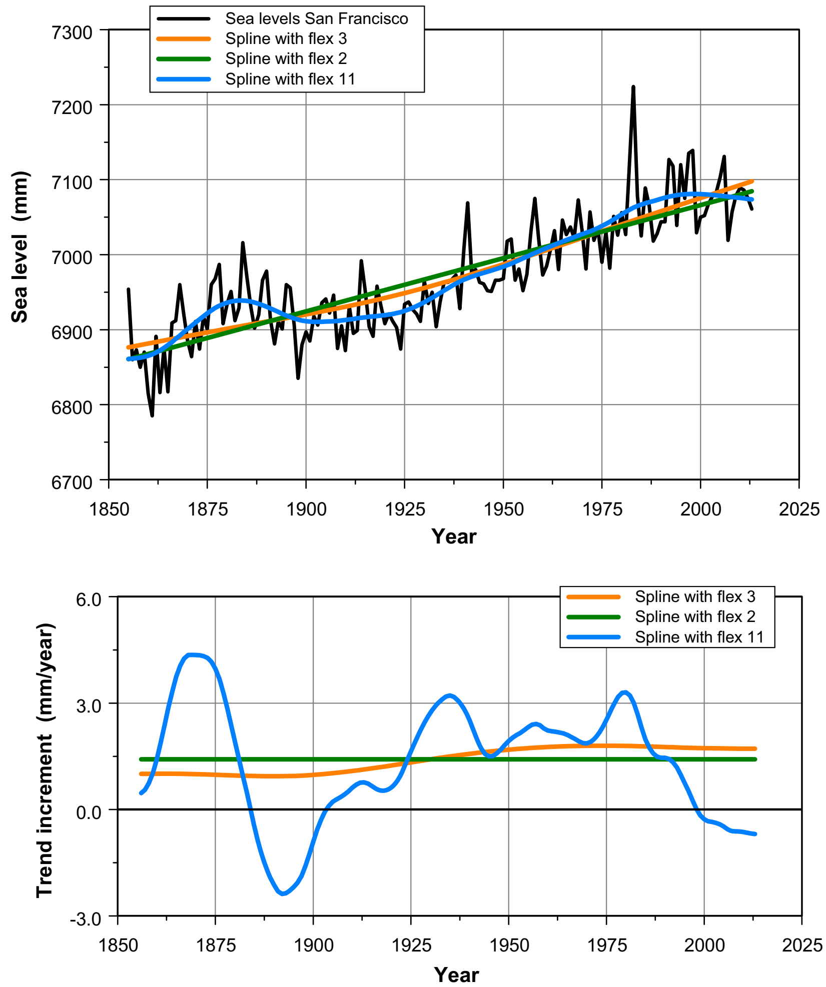



**Figure 2.** The spline trend model, estimated for three flexibilities. Data are for the San Francisco tide gauge series, as in Figure 1. (bottom) The three corresponding trend increment series.

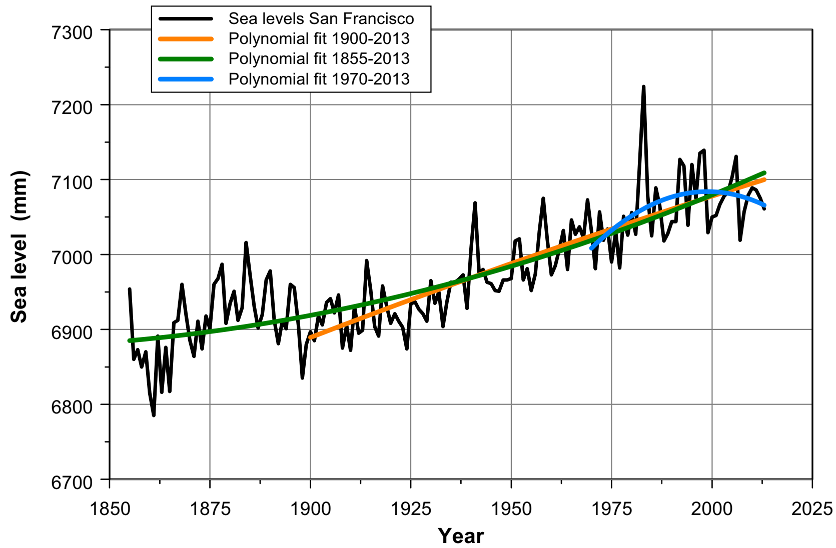
### 4.2. Varying Parameters, Different Results

Figure 2 shows estimates using a spline trend model with varying flexibility, again for the San Francisco tide gauge data. The trend with lowest flexibility ("2") yields a trend close to a straight line. The top plot shows that the results for flexibilities "2" and "3" are quite similar, although the trend based on the latter flexibility shows a small acceleration. The spline for flexibility "11" shows alternating periods of accelerations and decelerations. This is highlighted more clearly in the bottom plot where the corresponding trend increments are shown. Note that uncertainties are lacking in Figure 2 since the spline trend model is not statistical in nature.

### 4.3. Varying Sampling Periods, Different Results

We choose a second-order polynomial trend fit and apply it to the San Francisco data. Three sample periods are chosen: 1855–2013, 1900–2013, and 1970–2013. The results are given in Figure 3 (this example is similar to that given by *Breaker and Ruzmaikin* [2013, Figure 4]). The results appear to be quite different. The polynomial fit over the full sample period shows an accelerating pattern, the 1900–2013 fit shows an almost linear trend, and the 1970–2013 fit shows a strong decelerating pattern. These observations also follow from the bottom plot in Figure 3 where the corresponding trend increments are given.

Uncertainty inferences on accelerations are as follows. For the 1855–2013 period, we find an acceleration "2c" of $0.0118 \pm 0.0068$ mm/yr$^2$. The significance of this estimate is not given here since residuals are not white noise and uncertainties should be corrected, similar to model 2-2 in Table 1. For the 1900–2013



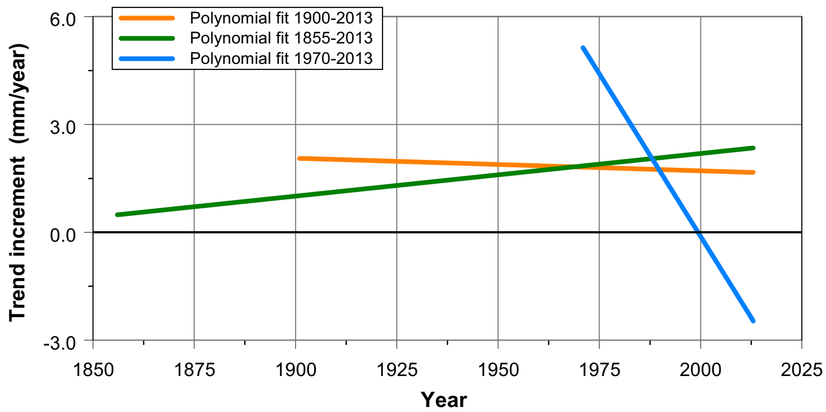




**Figure 3.** Second-order polynomial trend fit, estimated for three sample periods: 1855–2013, 1900–2013, and 1970–2013. Data are for the San Francisco tide gauge series are as in Figures 1 and 2. Due to the definition of second-order polynomials, the trend increment curves are always a straight line as shown in the bottom plot.

period, we find a nonsignificant deceleration of $-0.0036 \pm 0.0144$ mm/yr$^2$; for the 1970–2013 period we find a significant deceleration of $-0.182 \pm 0.180$ mm/yr$^2$ (all uncertainties are $2\sigma$).

The fact that the sample size chosen influences one's findings on accelerations, has been demonstrated in many articles [e.g., *Wahl et al.*, 2011; *Houston and Dean*, 2011a, 2011b, 2011c, 2011d; *Rahmstorf and Vermeer*, 2011; *Church and White*, 2011; *Haigh et al.*, 2014; *Jevrejeva et al.*, 2014] and, again, simply illustrates the role of internal variability [*Calafat and Chambers*, 2013; *Woodworth et al.*, 2011].

### 4.4. Adding External Information, Different Results

The examples discussed so far demonstrate the importance of climate internal processes on the estimation of acceleration and deceleration patterns, especially on shorter time scales of a few decades. This raises the question whether conventional trend models can be extended for such influences. In recent years, a number of authors, therefore, enhanced existing trend models with additional information provided by indices, which are commonly used to describe climate internal processes [e.g., *Zhang and Church*, 2012; *Dangendorf et al.*, 2013a; *Calafat and Chambers*, 2013; *Dangendorf et al.*, 2014b; *Haigh et al.*, 2014]. Here we give such an example by applying the stochastic Integrated Random Walk (IRW) model, given in equation (3a) [*Visser*, 2004], to the raw sea level data from the San Francisco tide gauge. Estimation results are given in Figure 4 (left). Trend increments are shown in the bottom left plot and show an alternating pattern.

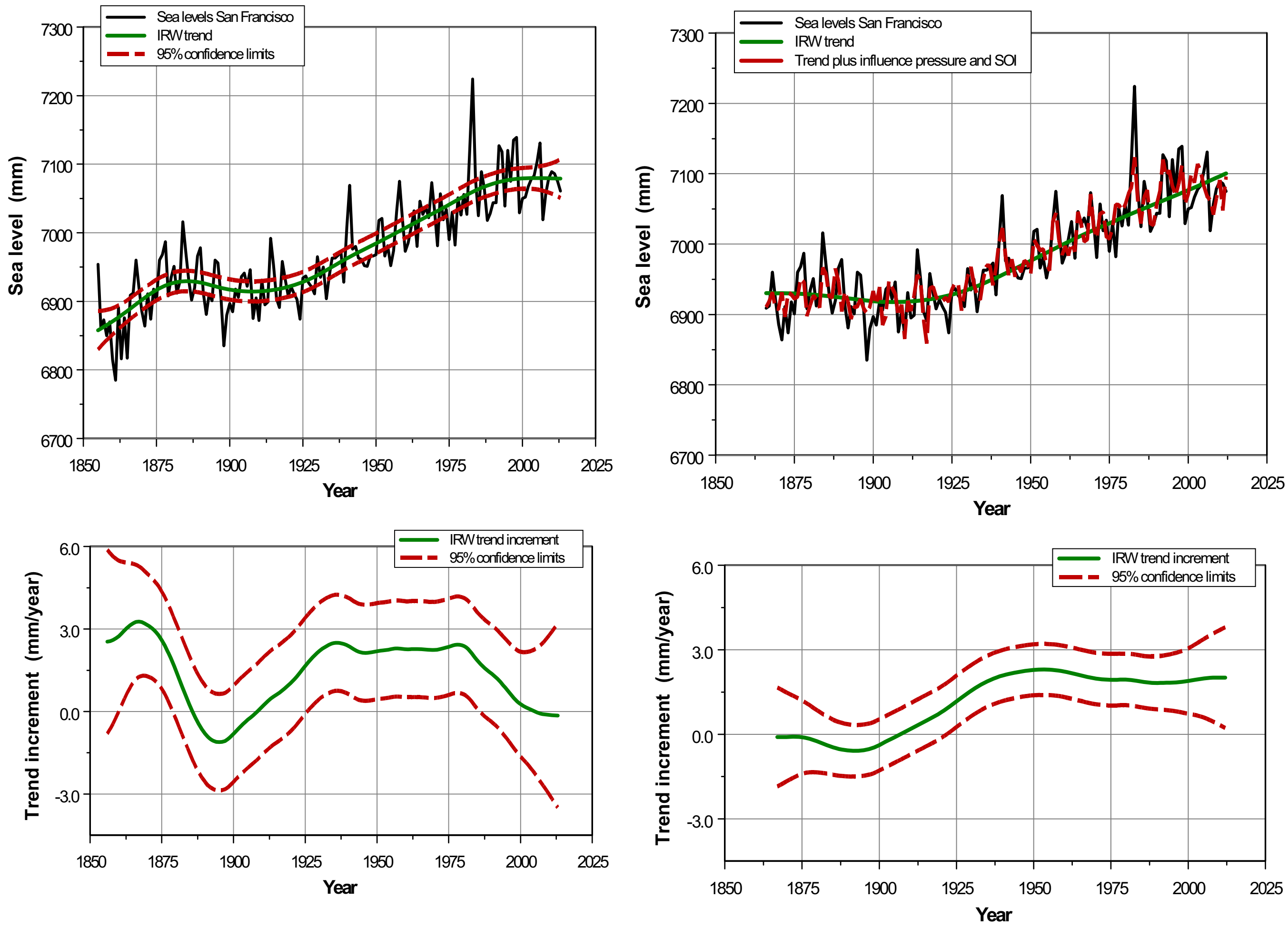



**Figure 4.** Two models estimated by Structural Time series Models (STMs). In the left plot, the IRW trend model is applied to the San Francisco data (1855–2013). Dashed lines represent 95% confidence limits. In the right plot, IRW trend model is extended by adding two explanatory variables: local pressure and the SOI index (1866–2012). The IRW trend is given by the green line; the trend plus influence of local pressure and SOI is given by the red line. The explanatory variables yield a variance reduction of 42% in the sea level variability around the trend. Both bottom plots show the corresponding trend increment series, along with 95% confidence limits.

We have also extended the IRW trend by adding two explanatory variables, as in equation (3b). These variables are local sea level pressure (SLP) and the Southern Oscillation Index (SOI). Both series are downloaded from the KNMI Climate Explorer website, where a wide range of climate indices can be found. Gridded pressures since 1850 are described by *Allan and Ansell* [2006], and the SOI since 1866 by *Allan et al*. [1991]. The two series do not capture all known features of climate internal variations as measured at the San Francisco tide gauge [*Thompson et al*., 2014] but they describe two major contributors and should therefore be understood as generic. The SLP affects sea level locally via the inverse barometer effect, i.e., the hydrostatic adjustment of the ocean to local pressure fluctuations. The SOI, in turn, mainly approximates variations in the alternating equatorial trade winds [*Merrifield et al*., 2012] and influences sea level around San Francisco remotely via wind-driven wave propagation [*Thompson et al*., 2014]. Results of the extension with climate variables are shown in Figure 4 (right). The explanatory variables of local SLP and the SOI yield a variance reduction of 42% in the sea level variability around the trend.

Comparing the left and right plots of Figure 4 shows considerable differences between trends and trend increments. Especially we would like to emphasize the vanishing trend increment since 1980 in the left plot, which is known to be mainly forced by the steady increase in equatorial trade winds since the early 1990s [*Merrifield et al*., 2012]. Due to the inclusion of the SOI, this latter pattern has disappeared in the trend in the right plot. Here trend increments stay on a constant level from 1940 onward, consistent with recent investigations by *Thompson et al*. [2014].

As for uncertainties shown in Figure 4, we note that adding additional variables helps in generating white noise residuals ("innovations" or "one-step-ahead prediction errors" in Kalman filter terms). The innovation series for the IRW trend model (Figure 4, left) shows a small but significant AR(1) value of 0.22; for the IRW trend model plus influence of SOI and SLP innovations are white noise. Thus, the addition of external information improves the whiteness of residuals, an assumption underlying these models.

## 5. Model Selection

In section 3, we have shown that 30 trend models have been applied in the field of sea-level research thus far, and that many (promising) methods are available outside this research field. We have also shown that these methods have varying characteristics and that practitioners may choose for a specific model based on these specific characteristics. This section is a elaboration of the questions raised in section 2.3: (i) How can we test the suitability or unsuitability of a certain trend model for the application at hand? and (ii) Are some trend methods better than others?

### 5.1. Suitability of Trend Models

A number of models given in Table 1 have underlying assumptions. That means that a trend model for which these assumptions are not met is less suitable for the application at hand. This holds for the statistical models 2-1 up to 2-9, 4-1–4-3, and 5-1 up to 5-5 given in Table 1. These models have the underlying assumption of noise processes being white noise (at least if one wants to present model estimates along with uncertainty information). Thus, one has to check that noise processes have (i) zero mean, (ii) a variance which is constant over time, and (iii) errors which are uncorrelated. Additionally, normality of errors is often assumed for hypothesis testing.

A range of tests for checking model assumptions are given by *Montgomery and Peck* [1982, chap. 3], *Harvey* [1989, chap. 5], and *Chandler and Scott* [2011, section 3.3]. We note that some statistical models estimate hundreds to thousands of statistical trends to construct *one* graph (models 5-2 and 5-3, and to a lesser extent models 5-1, 5-4, and 5-5 in Table 1). For these models, the checking of model assumptions is a huge task, since the residuals of each individual trend model should be analyzed. The importance of checking model assumptions in relation to the "acceleration-deceleration debate" has been underlined by *Baart et al*. [2012b, best practices #5].

An example of violations of assumptions is given in Figure 1. Residuals of the linear and second-order polynomial are not white noise. The autocorrelation functions (ACFs) show a damping pattern and a high AR(1) parameter, respectively. Therefore, these two models are not suited for this application, at least if one wants to present trends along with their uncertainty estimates. We note that methods are available for correcting the width of confidence limits (the so-called $\rho$ correction) [*Zieba*, 2010; *Bos et al*., 2013; *Tamazian et al*., 2015], or correcting both model parameters and confidence limits [*Prais and Winsten*, 1954; *Brown et al*., 2011].

### 5.2. Techniques for Model Selection

#### 5.2.1. Fitting Performance and Model Complexity

There is an extensive literature available on model selection within the context of statistical (trend) models. Statistical models are those listed in Table 1 in groups 2 and 4 (models 5-1 up to 5-5 are less appropriate here). Methods for model selection are based on the fitting performance of various models, often in relation to the complexity of models. Generally, more parsimonious models are favored over more complex models, provided that their fitting performance is not much worse. Here complex models are those having more unknown parameters. Clearly, more complex models are more difficult to interpret.

A number of complexity indices have been introduced in the literature. We name the adjusted R squared, Mallows's Cp, Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Deviance Information Criterion (DIC). Fitting performance can be tested by the *F* test or the Likelihood Ratio (LR) test. A description of these methods is beyond the scope of this article. We refer to *Harvey* [1989, chap. 5], *von Storch and Zwiers* [1999], *Burnham and Anderson* [2002], *Claeskens and Hjort* [2008], and *Chandler and Scott* [2011, section 3.2.3]. As for the Bayesian-oriented methods, such as BIC and DIC, we refer to *Spiegelhalter et al*. [2002] and *Gill* [2008, sections 7.4 and 7.5].

**Table 2.** A Series of Six Nested Models, Estimated for the San Francisco Tide Gauge Data, 1855–2013[a]

| Trend Model | Adjusted R Squared | Akaike's Information Criterion (AIC) | Bayesian Information Criterion (BIC) | *F* Test for Consecutive Model Pairs |
|---|---|---|---|---|
| 1 $y_t = a + \varepsilon_t$ | −0.006 | 1386 | 1389 | |
| 2 $y_t = a + bt + \varepsilon_t$ | 0.70 | 1194 | 1200 | **373** |
| 3 $y_t = a + bt + ct^2 + \varepsilon_t$ | 0.72 | 1185 | 1194 | **11.7** |
| 4 $y_t = a + bt + ct^2 + dt^3 + \varepsilon_t$ | 0.72 | 1186 | 1198 | 0.68 |
| 5 $y_t = a + bt + ct^2 + dt^3 + et^4 + \varepsilon_t$ | **0.75** | **1166** | **1181** | **22.8** |
| 6 $y_t = a + bt + ct^2 + dt^3 + et^4 + fX_t + \varepsilon_t$ | 0.75 | 1168 | 1186 | 0.034 |

[a]The adjusted R squared, Akaike's information criterion (AIC), and the Bayesian information criterion (BIC) are measures for goodness of fit where there is a penalty for more complex models (i.e., more parameters). As for AIC and BIC, the lower the index value, the better the model. The *F* test values generate a test for significant improvement if a parameter is added (values lower than 3.9 are nonsignificant). It is based on the fitting performance of models. Optimal values are given in bold.

In many cases, model selection techniques are applied to a set of so-called *nested* models. As an illustration, we give an example of nested models in combination with a number of indices named above for station San Francisco. We choose six trend models in the following nested manor [*Montgomery and Peck*, 1982, chap. 5]:

(1) $y_t = a + \varepsilon_t$, (2) $y_t = a + bt + \varepsilon_t$, (3) $y_t = a + bt + ct^2 + \varepsilon_t$, (4) $y_t = a + bt + ct^2 + dt^3 + \varepsilon_t$, (5) $y_t = a + bt + ct^2 + dt^3 + et^4 + \varepsilon_t$, and (6) $y_t = a + bt + ct^2 + dt^3 + et^4 + fX_t + \varepsilon_t$.

Here $y_t$ is the San Francisco tide gauge series (1855–2013), "t" is time in years, and $\varepsilon_t$ is a white noise process. The parameters "a," "b," "c," "d," "e," and "f" are estimated by multiple regression models (model 2–7 in Table 1). Finally, the variable $X_t$ is an arbitrarily chosen explanatory variable. Clearly, the models "grow" as for complexity, starting was a simple constant as trend, and ending with a fourth-order polynomial, along with the external variable $X_t$. We note that accelerations $[\mu_t - 2\mu_{t-1} + \mu_{t-2}]$ are quite different for models (1)–(5), that is 0.0, 0.0, 2c, 2c + 6dt, 2c + 6dt + 12et$^2$, respectively.

The estimation results for these nested models have been summarized in Table 2. The table shows that the complexity measures (adjusted R squared, AIC, and BIC) yield consistent results: the fourth-order polynomial (model (5)) is optimal for all three measures. As for AIC and BIC values, it holds that: the lower the index, the better the model with respect to fitting performance and complexity.

The final column yields the *F* test values for paired models: the value 373 is the F value for comparing model 1 and 2, the value 11.7 the F value for comparing model 2 and 3, etc. F values lower than 3.9 are nonsignificant (values for models (4) and (6)). The F values show that the fourth-order polynomial is optimal (model (5)), and that the addition of an external variable yields no improvements (F value of 0.034). It is interesting to note that third and fourth-order polynomials are not listed in Table 1.

Another example of nested models is given in Figure 4: the IRW trend model as formulated in equation (3a) and the extended model given by equation (3b). For the trend-only model, we find an AIC value of 1295 and for the extended model an AIC value of 1240, which shows a clear preference for the extended model (the lower value), this despite the fact that the extended model is more complex (three unknown parameters versus one unknown parameter for the trend-only model). BIC values are similar: 1298 for the one-parameter model and 1249 for the three-parameter model.

Model selection based on fitting performance and complexity measures has not been applied yet in the field of sea level research but we conclude that it certainly needs further exploration. Given certain statistical paradigms, such as the classical (frequentist) or the Bayesian school of inference, one has instruments to rank trend models as for their suitability for the data at hand. As for Bayesian techniques, we point to Bayesian Hierarchical Modeling as a promising tool [e.g., *Gill*, 2008, chap. 10].

### 5.2.2. Comparing Prediction Performance of Models

There is a danger in relying too much on the fitting performance of models. For example, any series of measurements $y_1$, ., $y_N$ can be modeled by the sum of N/2 sine functions with doubling frequencies (by Fourier series expansion). The fitting performance of the model is perfect but the prediction for year N + 1 will be poor: $\hat{y}_{N+1}$ will always be equal to $y_1$ due to the periodicity of sine functions. Therefore, it may be advantageous to select models on the basis of their *prediction performance* rather that their fitting performance (in fact, the second definition of "trend," as we have described in sections 2.1 and 2.3).

The importance of prediction criteria for model selection has been underlined by *Harvey* [1989, section 5.5], *Makridakis and Hibon* [2000], *Armstrong* [2001], and *Chandler and Scott* [2011, section 1.4.5]. We also refer to the recommendations in *Baart et al*. [2012b, best practices #5 and #6]. As for applying and comparing the prediction performance of models, the only characteristic we need is the possibility of extrapolating the trend beyond the sample period chosen. Trend models which yield predictions are models in the groups 2 and 4 and model 5–9 in Table 1.

Model selection based on prediction performance has not been applied yet in the field of sea level research. Since extrapolation of historic trends into the future plays an important role in the "acceleration-deceleration debate," the evaluation of prediction performance of trend models needs further exploration. We will return to this point in section 6.2.

### 5.3. Model Selection and Monte Carlo Simulation

The model selection methods named in section 5.2 have one drawback: they are limited to a subset of trend models named in Table 1. Thus, we cannot compare the performance of the linear trend model with wavelets, or the exponential growth curve with the EEMD model, etc. A way out here is the application of Monte Carlo (MC) simulation.

The rationale of MC is simple: generate a certain trend pattern (linear, quadratic, exponential, sigmoid, or alike) and add artificial noise, drawn by a random generator. The characteristics of the noise can be varied in numerous ways: white noise with normal distributions or extreme value distributions, correlated noise created by ARMA(1,0,1) or ARFIMA(1,0,1) models, etc. Subsequently, a synthetic sea level series is composed by adding a certain trend and additional noise. A first step in generating synthetic data in sea level research has been given by *Watson* [2015].

MC simulation is often applied to check to the performance of *one* specific (trend) model. Several software implementations can be compared, model parameters can be varied, characteristics of trend patterns and/or noise structures can be adapted, etc. A number of such applications have been given in the broader field of climate research: *Zhang and Zwiers* [2004], *Trömel and Schönwiese* [2008], and *Mudelsee* [2014, sections 4.1.6, 7.3, and 8.3].

In the same way, MC simulation could be used for *model selection*. Since we know the "true trend," varying trend methods can be compared as for their ability to reconstruct this trend. *Franzke* [2012] applied this technique to select trend models using synthetic temperature data. He generated five types of (correlated) noise in combination with six trend models and two trend shapes (his graphs 3 and 4).

However, the MC approach has also a serious drawback. Since we do not know the true trend shape(s) and the true noise structure(s), there is a danger of selecting simulation examples which favor the ideas we already have. For example, a second-order polynomial will accurately reproduce simulated trends which are based on a priori chosen linear, quadratic or exponential trend shapes (latter two shapes chosen by *Franzke*). Variation in the structure of the noise will not alter this good result. However, if we chose a sigmoid curve at the basis of our simulations, the second-order polynomial will be outperformed by many other trend methods.

We conclude that, on the one hand, MC simulation certainly has potential as a tool for model selection. It should therefore be explored in the field of sea level research [*Watson*, 2015]. On the other hand, however, the approach has drawbacks too. Since we do not know the true trends and noise structure in tide gauge series or regionally averaged sea level data, there is a danger of circular reasoning: we find what we hope to find.

## 6. Good Modeling Practices

As argued in section 2.3, techniques such as given in the preceding section are not conclusive in giving definite rankings of trend models. For example, we cannot conclude that a method such as wavelets (model 3–4 in Table 1) is better than an OLS linear trend (model 2-1 in Table 1) for the simple reason that the wavelet model is not statistical in nature. And if two models both are statistical in nature, there are still underlying

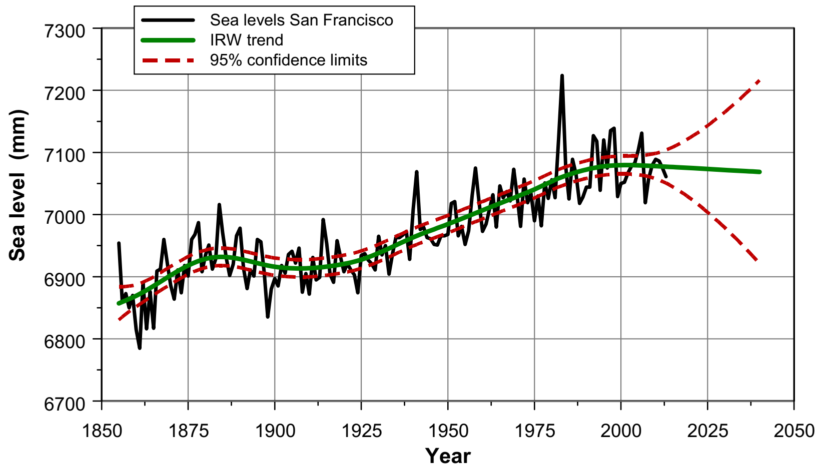



**Figure 5.** Trend extrapolation for station San Francisco. Period is 2014–2040. Predictions are based on the IRW trend model shown in the left plot of Figure 4. Red dashed lines represent 95% confidence limits for the trend.

assumptions attached to model selection. Given this inconclusive situation, we draw attention to three elements of "good modeling practices."

### 6.1. Model and Parameter Uncertainty

Contradictory findings such as illustrated in Figures 1 and 2 are not exclusive for the modeling of sea level data. Every modeler will experience that other choices for parameters, or changes in the structure of his or her model might lead to differing estimation results. A good way to proceed in these cases is to present the sensitivity of model output for such variations. See *Refsgaard et al*. [2006] for an extensive treatment of this topic.

If such variations are similar, or at least lead to the same conclusion, a multimodel approach strengthens the credibility of the output results, and inferences will be more robust. If, on the other hand, results differ considerably, leading to contradictory conclusions, one can discuss these results and give explanations. Suppose that coastal managers are confronted with well-presented but contradictory sea level trends (both accelerations and decelerations). For those cases, they could choose for the precautionary principle to be prepared for a worst-case situation.

A number of examples of varying trend models estimated on the same data can be found in the sea level literature, similar to the example in Figure 1. *Boon* [2012, Table 2] analyzes tide gauge data along the coasts of U.S. and Canada using linear and quadratic trend fits over the period 1969–2011. *Chambers et al*. [2012, Table 1] present three global data sets and give two trend estimates: *with* and *without* a 55 year cyclic component. *Jevrejeva et al*. [2008] present analyses using a second-order polynomial fit, trend accelerations based on second-order fits with varying windows and window lengths, and trends based on MC-SSA. *Church and White* [2011] present linear, quadratic, and moving window linear trends (16 year window) for global mean sea level data. Their models highlight various aspects of the same data. Similar to *Church and White*, sea level conclusions in *IPCC* [2013] are based on three trend models: linear, second-order polynomials over more than 100 years, and linear trends with 18 year shifting windows [*Rhein et al*., 2013, Figure 3.14].

An example of model estimates in relation to varying flexibilities for the same data are given by *Haigh et al*. [2014]. Their Figure 5 shows linear trends with varying window lengths of 10, 20, 30, 40, and 50 years (horizontal rows). The estimated trends appear to be quite different. They also present quadratic trend fits with shifting windows in their Figures 3 and 4. *Jevrejeva et al*. [2014] show quadratic trend fits with shifting windows and varying window lengths for regional and global sea levels in their Figure 15.

It is interesting to note that trend methods vary across subsequent IPCC WG I reports. The First Assessment Report (FAR), Second Assessment Report (SAR), and Third Assessment Report (TAR), published in 1990, 1995, and 2001, respectively, use linear trends and a visual presentation to conclude that (i) no accelerations are seen in the twentieth century, but (ii) that rates of sea level rise have increased between the mid-nineteenth and the mid-twentieth centuries. Expert judgment plays an important role in these assessments. The fourth assessment report (AR4 in 2007) draws similar conclusions on a wider set of studies and adds the method of linear trends with 10 year shifting windows. Moreover, sea level rates, based on satellite data and a linear trend, show higher rates over the period 1993–2003.

In relation to the "acceleration-deceleration debate," *Baart et al*. [2012b] also recommend the "multimodel approach" as propagated by *Refsgaard et al*. [2006]. *Baart et al*. suggest to present a broad range of methods, rather than to rely on a single method only. In this context, we note that the contested study of *Boretti* [2012] uses second-order polynomials only. The same holds for *Parker et al*. [2013] who propose to fit an exponential growth curve over the period 1990–2100.

We conclude that the multimodel approach is highly recommendable. It adds to the robustness of acceleration-deceleration inferences.

### 6.2. Uncertainties and Prediction

Statistical models are of importance since an estimate of a rising or declining trend is not necessarily statistically significant. Significance depends on the signal/noise ratio and the autocorrelation of the data at hand. Therefore, the availability of uncertainty estimates for trends, trend increments and trend accelerations has an added value (models in classes 2, 3, 4, and some in class 5). The importance of uncertainties is also highlighted in the IPCC definition of "detection of change" [*IPCC*, 2013, Annex III, Glossary]. Detection of change is "the process of demonstrating that climate or a system affected by climate has changed in some *statistical sense*, without providing a reason for that change."

The use of statistical models is therefore to be recommended since it yields uncertainty information for trend estimates $\mu_t$ and trend extrapolations. We note that some models also generate uncertainties for trend increments $[\mu_t - \mu_s]$, with times "t" and "s" within the sample period. Models which give the latter information are Ordinary Least Squares (OLS) linear trends and Structural Time series Models (STMs) [*Visser and Petersen*, 2012].

Uncertainties play an important role in the acceleration-deceleration discussion between *Boretti* [2012] and *Hunter and Brown* [2012]. Hunter and Brown argue that the absence of uncertainty estimates is the main reason for the lack of acceleration reported by Boretti. Furthermore, a central topic in the debate is a lack of acceleration in relation to future projections. Some authors claim that the lack of accelerations (or the presence of small decelerations) in tide gauge data questions future projections such as given by the IPCC [e.g., *Houston and Dean*, 2011a, 2011b, 2011c, 2011d; *Boretti*, 2012; *Parker et al*., 2013]. However, such conclusions do not take uncertainties into account and neglect the fact that the anthropogenic signal in projections requires decades to become distinguishable from climate noise [*Dangendorf et al*., 2014b; *Haigh et al*., 2014; *Lyu et al*., 2014; *Richter and Marzeion*, 2014].

As an example, we have extrapolated the Integrated Random Walk (IRW) trend as shown in Figure 4 (left), up to the year 2040, along with uncertainties. The result is shown in Figure 5. It shows that predictions are stable over the prediction horizon 2014–2040. However, the 95% confidence limits show that both accelerations and deceleration could arise based on the historic information 1855–2013. Clearly, uncertainties grow considerately as function of the prediction horizon. Thus, results as shown in Figure 5 do not help coastal managers much, at least not for predictions over decades.

The way to proceed here is the application of emission scenarios and GCMs, incorporating physical mechanisms. Such projections should not be seen as *absolute* predictions of future sea levels but more as *conditional* predictions, or what-if "projections," without making statements on their probabilities. Examples of scenario-based projections are given by *Church et al*. [2013, Figures 13.23 and 13.24], *Dangendorf et al*. [2013a] and *Nicholls et al.* [2014]. An example for San Francisco is provided in Figure 6 and is based on the 21 CMIP5 models used in the AR5 reports [*Church et al*., 2013; *Slangen et al*., 2014].

We conclude that extrapolations of historic trends should be handled with care, and at least presented along with their corresponding uncertainties. Long-term predictions cannot be based on such extrapolations. Scenario-based projections are the only physical meaningful way to proceed.

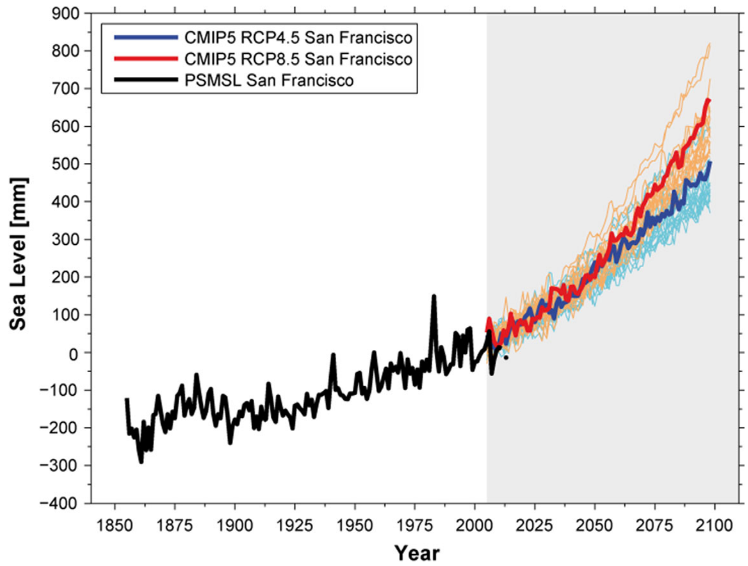



**Figure 6.** Projections (red and blue) and observations (black) of annual MSL for the station San Francisco, as provided by 21 CMIP5 models (thin colored lines) and presented in the AR5 [*Church et al.*, 2013]. These projections are based on two scenarios (RCP4.5 (blue) and RCP8.5 (red)) and a range of processes, such as dynamic sea level changes, glacier melting, melting of the large ice sheets in Greenland and Antarctica, hydrological changes and glacial isostatic adjustment. For each scenario, one particular model projection has been highlighted using a thick colored line.

### 6.3. The Importance of Additional Information

We have illustrated in Figure 4 that the addition of external information such as climate indices can alter trend patterns considerably. This phenomenon has been reported in a number of recent articles. *Woodworth et al*. [2009, sections 6 and 7] give an overview of estimating trends along with information on sea level pressure (SLP) and climate indices such as NAO, AO, PDO, AMO, and SOI. *Albrecht and Weisse* [2012], *Dangendorf et al*. [2013a, Figure 6], and *Dangendorf et al*. [2014b] show trend results with and without atmospheric pressure correction for the German Bight.

*Calafat and Chambers* [2013] apply a regression model with atmospheric pressure, wind, and climate indices (MEI, NAO, and PDO) as explanatory variables. *Zhang and Church* [2012] apply a linear trend model next to a trend based on multiple regression models (with a linear trend). They use gridded sea levels derived from satellites. They apply the PDO and MEI indices as explanatory variables. Their Figure 4e shows that the trends with and without the addition of external variables differ considerably. *Cazenave et al*. [2014] analyze satellite-derived sea level data and show trends with and without a mass and thermosteric correction. See their Figures 1 and 2. *White et al*. [2014, Figure 10 and Table 4] show large differences between trend models with and without removal of an El Niño-Southern Oscillation (ENSO) signal.

The studies mentioned here show the importance of modeling long-term natural variability. However, the inclusion of external information in regression-based models also has other important advantages. First, uncertainty bands for trends and their increments become smaller due to the shrinkage of noise levels [*Dangendorf et al*., 2013b, Figure 9]. Second, residuals may become white noise which might not be the case without the addition of explanatory variables. Thus, corrections such as applied in model 2-2 in Table 1, are not needed.

Third, the role of cyclic patterns can be dealt with *without* an explicit (uncertain) modeling of cycles (as in *Chambers et al*. [2012] and *Houston and Dean* [2013]). The reason is that a number of climate indices also contain these cycles. Figure 7 illustrates that point for three prominent indices: SOI, PDO, and AMO. It shows that the PDO and AMO contain a (damping) cycle of around 60 years. Thus, if these indices are used in a regression context, the role of this cycle is implicitly taken into account (at least if the right lead-lag relationships are taken into account).

Part of the "acceleration-deceleration debate" deals with (weak) decelerations around the Australian mainland. *White et al*. [2014] conclude that a large part of the interannual and decadal variability in sea level

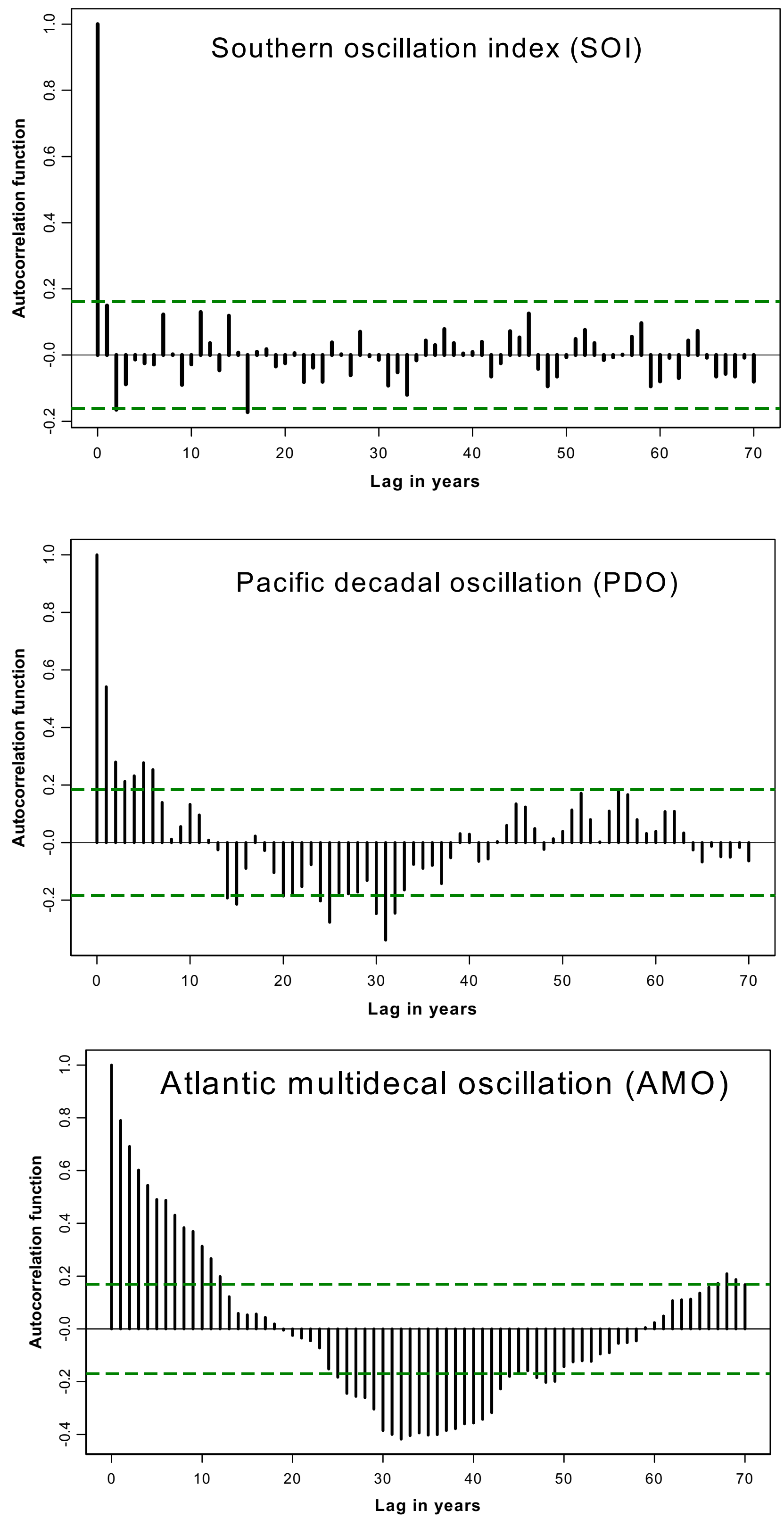



**Figure 7.** Autocorrelation functions with lags up to 70 years for three climate indices: (top) SOI, (middle) PDO, and (bottom) AMO.

around the whole coast of Australia is coherent. This basin-scale coherence (also visible in an inverse manner at San Francisco tide gauge [*Hogarth*, 2014; *Thompson et al*., 2014]) has its physical origin in the equatorial trade winds leading to water mass propagation through the Indonesian waveguides and along the continental slope of the Australian mainland. *White et al*. [2014] report alternating periods with strong accelerations and decelerations from 1890 onward, with clear accelerations since 1980 (remember the opposite development at San Francisco in section 4). Their approach and results deviate from that presented by *Watson* [2011] and *Boretti* [2012] who do not incorporate SOI information into their models.

We conclude that the addition of external information may seriously alter the pattern of trend estimates in sea level data in both their magnitude and their noise level. Therefore, it strengthens the robustness of acceleration-deceleration inferences if such analyses are included in sea level studies.

## 7. Summary and Conclusions

We have given a review of trend methods as applied in the field of sea level research. The reason for this review is the ongoing debate amongst researchers as for (alternating) accelerations and decelerations in such trends, sometimes estimated for the *same* data. This debate has implications for sea level rise and coastal management in the coming decades, and, therefore, needs further exploration.

Our review reveals a wide range of trend methods, 30 in total. We have categorized these methods into five classes, all with different characteristics: (i) explorative methods, (ii) parametric trend methods, (iii) nonparametric trend methods, (iv) stochastic trend methods, and (v) "miscellaneous methods." The number of models found in these classes are 6, 9, 4, 1, and 10, respectively. Methods have been summarized in Table 1. We have also drawn attention to two trend methods with attractive properties, which have not been applied in sea level research thus far (and not present in this set of 30 methods): structural time series models and long-memory trend models.

To make a choice between all these trend methods, it is important to know the specific characteristics. A number of these characteristics are as follows:

1. Some methods generate trends which are available over a part of the sample period due to a window chosen, while others give trends over the full period.
2. Some methods give uncertainty estimates for trends and their increments, others do not.
3. Some methods allow for the addition of cyclic or additional information, such as the SOI or the PDO, others do not.
4. Some methods allow for extrapolation to future years, along with uncertainties, others do not.

Details on characteristics have been summarized in Appendix A.

One of our findings is that contradictory trend estimates can be found, based on the same data. This phenomenon has been illustrated in Figures 1–4. To avoid such inferences, we have explored methods for model selection in sections 2.3 and 5: are some models better than others? We have shown that ways of ranking trend methods are available for statistics-oriented models. First, it should be checked if assumptions underlying the model(s) chosen are fulfilled. Second, a ranking of methods can be found using methods based on (i) fitting performance and complexity, or (ii) the prediction performance of methods. Furthermore, these rankings can be based on two "schools of inference": the classical or frequentist approach (AIC, *F* tests) or the Bayesian approach (BIC or DIC, Bayesian factors). Also the approach of generating synthetic sea level data and testing trend models on those data has been discussed.

However, since all these methods have implicit assumptions, we have formulated three modeling guidelines which are *complementary* to model selection. First, it is advisable to explore trend patterns from more than one method, and compare/discuss differences if present. Second, the choice for statistical models that generate uncertainty information around trend estimates and increments is essential since it prevents one from drawing conclusions on accelerations or decelerations which occur only by chance. The addition of uncertainties is especially important (i) if extrapolations are made for future sea levels, as illustrated in Figure 5, or (ii) model projections are compared to observations [*Hunter and Brown*, 2012].

Finally, it is shown that the addition of external information has considerable consequences for trend estimates.

## Appendix A: Characteristics of Methods

Here we will shortly discuss the characteristics for each of the five trend-model classes defined in section 2.2. A more elaborated discussion can be found in *Chandler and Scott* [2011], which is the major reference for this section.

The **first group** (models 1-1 up to 1-6 in Table 1) falls in the characterization of exploratory tools. Examples are linear filters such as moving averages with a fixed window size. These filters are easy to understand and software is widely available. Furthermore, short-term natural variations can be smoothed by taking window lengths of a few decades or even longer, depending on the tide gauge location and the underlying physics. However, linear filters have three drawbacks: (i) they are not based on statistical formulation and, thus, provide no uncertainty bands for trend estimates, (ii) they depend on the choice of a specific window size, which results in missing trend estimates at the end of the sample period, and (iii) they cannot generate short-term predictions. The latter drawback might play a role if coastal management is of importance and no physical projections are available. The Mann-Kendal trend test (and variants thereof) is important if one is only interested in the significance of a trend. However, the method is a test and no actual trend is given.

Models from the **second group** (models 2-1 up to 2-9 in Table 1) are most often encountered in sea level research. They fall in the characterization of parametric models for deterministic trends. The term parametric means that the shape of the trend is fixed or “deterministic.” A general form for these models is

$$y_t = a_1 X_{1,t} + a_2 X_{2,t} + a_3 X_{3,t} \ldots + \varepsilon_t \tag{A1}$$

where the terms $X_{i,t}$ can be polynomial functions of time or explanatory variables such as climate indices. Models of the form (A1) have a number of attractive properties. First, their formulation is transparent and intuitive; trends are estimated over the full sample period, and software is easily accessible. Second, trend estimates are gained along with uncertainties therein due to the noise term $\varepsilon_t$ in equation (A1). This property closely follows the definition of climate change signal detection in the AR5 [*IPCC*, 2013, Glossary]. Third, additional information can be added to the model, which is of importance in terms of distinguishing natural variability from external causes as discussed in section 6.3. Finally, deterministic trends can be extrapolated (with the respective data-driven uncertainties) into future. Parametric trends have also limitations. For uncertainty estimates to hold, residuals should be tested for being white noise, which is generally questionable in raw climate data (seasonality, natural persistence). Methods for testing this aspect are given by *Chandler and Scott* [2011, section 3.1.1]. We note that a number of authors quantify uncertainties but do not mention any test on the residuals [cf., *Baart et al*., 2012b, best practice #5]. Ways of correcting uncertainties in case of correlated noise, are (i) the application of a $\rho$ correction [e.g., *Chandler and Scott*, 2011, section 3.3.3] or (ii) the approach of *Prais and Winsten* [1954]. Note that in case of extreme values, rather than monthly or annual averages, more sophisticated nonstationary extreme value statistics exist. Important references are *Coles* [2001] and *Mudelsee* [2014].

The **third group** (models 3-1 up to 3-4, Table 1) falls in the class of nonparametric models. These models do not assume fixed trend shapes and have the more general form of

$$y_t = f_1(X_{1,t}) + f_2(X_{2,t}) + f_3(X_{3,t}) \ldots + \varepsilon_t \tag{A2}$$

The functions $f_1$, $f_2$, $f_3$ are not predefined but follow directly from the data. These trend models have the advantage of being more flexible and data driven. However, the estimation and interpretation of these models is more complex, and less open-source software is available. Uncertainties can be treated in a similar way as discussed for the second group.

The **fourth group** (models 4-1 up to 4-3, Table 1) falls in the class of stochastic models. Here, trend models have a stochastic formulation, as in equation (1c). The idea is that the data at hand are a realization of stochastic processes and trends therein should be treated accordingly. This group of models is rarely applied in sea level research but often encountered in the wider field of climate research or econometry.

The **fifth group** (models 5-1 up to 5-10, Table 1) do not fall in one of the preceding four classes and we denote them here as "miscellaneous." We note that a number of methods are based on parametric trends but put in a different context. This holds for models 5-1 up to 5-5. Models 5–6 up to 5-10 are based on other mathematical principles, such as neural networks. It is beyond the scope of this article to define these methods further.

**Acknowledgments**
The historic tide gauge data for San Francisco as used in this study are freely accessible through the Permanent Service for Mean Sea Level (PSMSL). Station San Francisco has been given code "10" and information can be found at http://www.psmsl.org/data/obtaining/stations/10.php. Sea level projections for station San Francisco are freely available from the Coupled Model Intercomparison Project Phase 5 (CMIP5). See their website: http://www.psmsl.org/data/obtaining/stations/10.php. We thank two anonymous reviewers for helpful comments on the manuscript, especially those concerning model selection techniques.

## References

Albrecht, F., and R. Weisse (2012), Pressure effects on past regional sea level trends and variability in the German Bight, *Ocean Dyn.*, *62*, 1169–1186.

Allan, R. J., and T. Ansell (2006), A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004, *J. Clim.*, *19*, 5816–5842.

Allan, R. J., N. Nicholls, P. D. Jones, and I. J. Butterworth (1991), A further extension of the Tahiti-Darwin SOI, early ENSO events and Darwin pressure, *J. Clim.*, *4*, 743–749.

Armstrong, J. S. (2001), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Springer, Berlin.

Baart, F., P. H. A. J. M. van Gelder, J. de Ronde, and M. van Koningsveld (2012a), The effect of the 18.6-year lunar cycle on regional sea-level rise estimates, *J. Coastal Res.*, *28*(2), 511–516.

Baart, F., M. Koningsveld, and M. J. F. Stive (2012b), Trends in sea-level trend analysis, *J. Coastal Res.*, *28*(2), 311–315.

Baillie, R. T., and S. K. Chung (2002), Modeling and forecasting from trend-stationary long memory models with applications to climatology, *Int. J. Forecasting*, *18*, 215–226.

Baker, R. G. V., and S. A. McGowan (2014), Periodicities in mean sea-level fluctuations and climate change proxies: Lessons from the modelling for coastal management, *Ocean Coastal Manage.*, *98*, 187–201.

Barbosa, S. M., M. J. Fernandes, and M. E. Silva (2004), Nonlinear sea level trends from European tide gauge records, *Ann. Geophys.*, *22*, 1465–1472.

Becker, M., M. Karpytchev, and S. Lennartz-Sassinek (2014), Long-term sea level trends: Natural or anthropogenic?, *Geophys. Res. Lett.*, *41*, 5571–5580, doi:10.1002/2014GL061027.

Boon, J. D. (2012), Evidence of sea level acceleration at U.S. and Canadian tide stations, Atlantic coast, North America, *J. Coastal Res.*, *28*(6), 1437–1445.

Boretti, A. (2012), Is there any support in the long term tide gauge data to the claims that parts of Sydney will be swamped by rising sea levels?, *Coastal Eng.*, *64*, 161–167.

Bos, M. S., S. D. P. Williams, I. B. Araújo, and L. Bastos (2013), The effect of temporal correlated noise on the sea level rate and acceleration uncertainty, *Geophys. J. Int.*, *196*(3), 1423–1430, doi:10.1093/gji/ggt481.

Breaker, L. C., and A. Ruzmaikin (2013), Estimating rates of acceleration based on the 157-year record of sea level from San Francisco, California, U.S.A., *J. Coastal Res.*, *29*(1), 43–51.

Bromirski, P. D., A. J. Miller, R. E. Flick, and G. Auad (2011), Dynamical suppression of sea level rise along the Pacific coast of North America: Indications for imminent acceleration, *J. Geophys. Res.*, *116*, C07005, doi:10.1029/2010JC006759.

Brown, B. E., R. P. Dunne, N. Phongsuwan, and P. J. Somerfield (2011), Increased sea level promotes coral cover on shallow reef flats in the Andaman Sea, eastern Indian Ocean, *Coral Reefs*, *30*, 867–878, doi:10.1007/s00338-011-0804-9.

Burnham, K. P., and D. R. Anderson (2002), *Model Selection and Multimodel Inference*, Springer, Berlin.

Calafat, F. M., and D. P. Chambers (2013), Quantifying recent acceleration in sea level unrelated to internal climate variability, *Geophys. Res. Lett.*, *40*, 3661–3666, doi:10.1002/grl.50731.

Carmillet, V., J.-M. Brankart, P. Brasseur, H. Drange, G. Evensen, and J. Verron (2001), A singular evoluvative extended Kalman filter to assimilate ocean color data in a coupled physical-biochemical model of the North Atlantic ocean, *Ocean Modell.*, *3*(3–4), 167–192.

Cazenave, A., H. B. Dieng, B. Meyssignac, K. von Schuckmann, B. Decharme, and E. Berthier (2014), The rate of sea-level rise, *Nat. Clim. Change*, *4*, 358–361, doi:10.1038/nclimate2159.

Chambers, D. P., M. A. Merrifield, and R. S. Nerem (2012), Is there a 60-year oscillation in global mean sea level?, *Geophys. Res. Lett.*, *39*, L18607, doi:10.1029/2012GL052885.

Chandler, R. E., and E. M. Scott (2011), *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*, 368 p., John Wiley, Chichester, U. K.

Chelton, D. B., and R. E. Davis (1982), Monthly mean sea level variability along the west coast of North America, *J. Phys. Oceanogr.*, *12*(8), 757–784.

Chen, X., Y. Feng, and N.E. Huang (2014), Global sea level trend during 1993–2012, *Global Planet. Change*, *112*, 26–32.

Church, J. A., and N. J. White (2006), A 20th century acceleration in global sea-level rise, *Geophys. Res. Lett.*, *33*, doi:10.1029/2005GL024826.

Church, J. A., and N. J. White (2011), Sea-level rise from the late 19th to the early 21st century, *Surv. Geophys.*, *32*(4–5), 585–602, doi:10.1007/s10712-011-9119-1.

Church, J. A., et al. (2013), Sea level change, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifths Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., pp. 1137–1216, Cambridge Univ. Press, Cambridge, U. K.

Claeskens, G., and N. L. Hjort (2008), *Model Selection and Model Averaging*, Cambridge Univ. Press, Cambridge, U. K.

Coles, S. (2001), An Introduction to Statistical Modeling of Extreme Values*, Springer Ser. Stat.*, Springer, London, U. K.

Dangendorf, S., T. Wahl, E. Nilson, B. Klein, and J. Jensen (2013a), A new atmospheric proxy for sea level variability in the southeastern North sea: Observations and future ensemble projections, *Clim. Dyn.*, *43*, 447–467, doi:10.1007/s00382-013-1932-4.

Dangendorf, S., C. Mudersbach, T. Wahl, and J. Jensen (2013b), Characteristics of intra-, inter-annual and decadal sea level variability and the role of meteorological forcing: The long record of Cuxhaven, *Ocean Dyn.*, *63*, 209–224, doi:10.1007/s10236-0598-0.

Dangendorf, S., D. Rybski, C. Mudersbach, A. Müller, E. Kaufmann, E. Zorita, and J. Jensen (2014a), Evidence for long-term memory in sea level, *Geophys. Res. Lett.*, *41*, 5530–5537, doi:10.1002/2014GL060538.

Dangendorf, S., F. M. Calafat, A. Arns, T. Wahl, I. D. Haigh, and J. Jensen (2014b), Mean sea level variability in the North sea: Processes and implications, *J. Geophys. Res. Oceans*, *119*, 6820–6841, doi:10.1002/2014JC009901.

Donner, R. V., R. Ehrcke, S. M. Barbosa, J. Wagner, J. F. Donges, and J. Kurths (2012), Spatial patterns of linear and nonparametric long-term trends in Baltic sea-level variability, *Nonlinear Processes Geophys.*, *19*, 95–111, doi:10.5194/npg-19-95-2012.

Donoghue, J. F., and R. W. Parkinson (2011), Discussion of: Houston, J.R. and Dean, R.G., 2011. Sea-level acceleration based on U.S. tide gauges and extensions of previous global-gauge analyses, Journal of Coastal Research, 27(3), 409–417, *J. Coastal Res.*, *27*(5), 994–996.

Douglas, B. C. (1991), Global sea level rise, *J. Geophys. Res.*, *96*, 6981–6992.

Ezer, T., L. P. Atkinson, W. B. Corlett, and J. L. Blanco (2013), Gulf Stream's induced sea level rise and variability along the U.S. mid-Atlantic coast, *J. Geophys. Res. Oceans*, *118*, 685–697, doi:10.1002/jgrc.20091.

Feng, M., Y. Li, and G. Meyers (2004), Multidecadal variations of Fremantle sea level: Footprint of climate variability in the tropical Pacific, *Geophys. Res. Lett.*, *31*, L16302, doi:10.1029/2004GL019947.

Franzke, C. (2012), Nonlinear trends, long-range dependence, and climate noise properties of surface temperature, *J. Clim.*, *25*, 4172–4183.

Gill, J. (2008), *Bayesian Methods: A Social and Behavioral Sciences Approach*, Chapman and Hall, N. Y.

Grewal, M. S., and A. P. Andrews (2010), Applications of Kalman filtering in Aerospace 1960 to the present, *IEEE Control Syst. Mag.*, *30*, 69–78.

Haigh, I. D., T. Wahl, E. J. Rohling, R. M. Price, C. B. Pattiaratchi, F. M. Calafat, and S. Dangendorf (2014), Timescales for detecting a significant acceleration in sea level rise, *Nat. Commun.*, *5*, 3635, doi:10.1038/ncomms4635.

Harvey, A. C. (1984), A unified view of statistical forecasting procedures, *J. Forecasting*, *3*, 245–275.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge Univ. Press, Cambridge, U. K.

Harvey, A. C. (2001), Trend analysis, in *Encyclopedia of Environmetrics*, vol. 4, edited by D. R. Brillinger, pp. 2243–2257, John Wiley, Chichester, U. K.

Hay, C. C., E. Marrow, R. E. Kopp, and J. X. Mitrivica (2015), Probabilistic reanalysis of twentieth-century sea-level rise, *Nature*, *517*, 481–484.

Hogarth, P. (2014), Preliminary analysis of acceleration of sea level rise through the twentieth century using extended tide gauge data sets (August 2014), *J. Geophys. Res. Oceans*, *119*, 7645–7659, doi:10.1002/2014JC009976.

Holgate, S. J. (2007), On the decadal rates of sea level change during the twentieth century, *Geophys. Res. Lett.*, *34*, L01602, doi:10.1029/2006GL028492.

Holgate, S. J., A. Matthews, P. L. Woodworth, L. J. Rickards, M. E. Tamisiea, E. Bradshaw, P. R. Foden, K. M. Gordon, S. Jevrejeva, and J. Pugh (2013), New data systems and products at the permanent service for mean sea level, *J. Coastal Res.*, *29*(3), 493–504.

Houston, J. R., and R. G. Dean (2011a), Sea-level accelerations based on U.S. tide gauges and extensions of previous global-gauge analyses, *J. Coastal Res.*, *27*(3), 409–417.

Houston, J. R., and R. G. Dean (2011b), Reply to: Rahmstorf, S. and Vermeer, M., 2011. Discussion of: Houston, J.R. and Dean, R.G., 2011. Sea-level acceleration based on U.S. tide gauges and extensions of previous global-gauge analyses, Journal of Coastal Research, 27(3), 409–417, *J. Coastal Res.*, *27*(4), 788–790.

Houston, J. R., and R. G. Dean (2011c), Reply to: Donoghue, J.F. and Parkinson, R.W., 2011. Discussion of: Houston, J.R. and Dean, R.G., 2011. Sea-level acceleration based on U.S. tide gauges and extensions of previous global-gauge analyses, Journal of Coastal Research, 27(3), 409–417, *J. Coastal Res.*, *27*(5), 997–998.

Houston, J. R., and R. G. Dean (2011d), Accounting for the nodal tide to improve estimates of sea level acceleration, *J. Coastal Res.*, *27*(5), 801–807.

Houston, J. R., and R. G. Dean (2013), Effects of sea-level decadal variability on acceleration and trend difference, *J. Coastal Res.*, *29*(5), 1062–1072.

Hunter, J. R. (2013), Comment on ''Sea-level trend analysis for coastal management'' by A. Parker, M. Saad Saleem and M. Lawson, *Ocean Coastal Manage.*, *87*, 114–115.

Hunter, J. R., and M. J. I. Brown (2012), Discussion of Boretti, A., 'Is there any support in the long term tide gauge data to the claims that parts of Sydney will be swamped by rising sea levels?', Coastal Engineering, 64, 161–167, June 2012, *Coastal Eng.*, *75*, 1–3.

Hunter, J. R., J. A. Church, N. J. White, and X. Zhang (2013a), Towards a global regionally varying allowance for sea-level rise, *Ocean Eng.*, *71*, 17–27.

Hunter, J. R., J. A. Church, N. J. White, and X. Zhang (2013b), Reply to ''Discussion of 'Towards a global regionally varying allowance for sea-level rise' by J.R Hunter, J.A. Church, N.J. White, X. Zhang [Ocean Engineering 71(1) 17–27 (1 October 2013)],'' *Ocean Eng.*, *72*, 473–473.

Imani, M., R.-J. You, and C.-Y. Kuo (2014), Caspian sea level using satellite altimetry by artificial neural networks, *Int. J. Environ. Sci. Technol.*, *11*, 1035–1042.

IPCC (1990), *Climate Change: the IPCC Scientific Assessment. Contribution of Working Group I to the First Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., Cambridge Univ. Press, Cambridge, U. K.

IPCC (1995), *Climate Change 1995: the Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., Cambridge Univ. Press, Cambridge, U. K.

IPCC (2001), *Climate Change 2001: the Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., Cambridge Univ. Press, Cambridge, U. K.

IPCC (2007), *Climate Change 2007: the Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K.

IPCC (2013), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifths Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K.

Jevrejeva, S., A. Grinsted, J. C. Moore, and S. Holgate (2006), Nonlinear trends and multiyear cycles in sea level records, *J. Geophys. Res.*, *111*, C09012, doi:10.1029/2005JC003229.

Jevrejeva, S., J. C. Moore, A. Grinsted, and P. L. Woodworth (2008), Recent global sea level started over 200 years ago?, *Geophys. Res. Lett.*, *35*, L08715, doi:10.1029/2008GL033611.

Jevrejeva, S., J. C. Moore, A. Grinsted, A. P. Matthews, and G. Spada (2014), Trends and acceleration in global and regional sea levels since 1807, *Global Planet. Change*, *113*, 11–22.

Lyu, K., X. Zhang, J. A. Church, A. B. A. Slangen, and J. Hu (2014), Time of emergence for regional sea-level change, *Nat. Clim. Change*, *4*, 1006–1010, doi:10.1038/nclimate2397.

Makridakis, S., and M. Hibon (2000), The M3-competition: Results, conclusions and implications, *Int. J. Forecasting*, *16*, 451–476.

Marotzke, J., and P. M. Forster (2015), Forcing, feedback and internal variability in global temperature trends, *Nature*, *517*, 565–570.

Merrifield, M. A., P. R. Thompson, and M. Lander (2012), Multidecadal sea level anomalies and trends in the western tropical Pacific, *Geophys. Res. Lett.*, *39*, L13602, doi:10.1029/2012GL052032.

Meyers, G., A. Melsom, G. T. Mitchum, and J. J. O'Brian (1998), Detection of the fast Kelvin wave teleconnection due to El Nino southern oscillation, *J. Geophys. Res.*, *103*, 27,655–27,663.

Miller, L., and B. C. Douglas (2007), Gyre-scale atmospheric pressure variations and their relation to 19th and 20th century sea level rise, *Geophys. Res. Lett.*, *34*, L16602, doi:10.1029/2007GL030862.

Mills, T. C. (2010), 'Skinning a cat': Alternative models of representing temperature trends, *Clim. Change*, *101*, 415–426.

Montgomery, D. C., and E. A. Peck (1982), *Introduction to Linear Regression Analysis*, 504 pp., John Wiley, N. Y.

Mudelsee, M. (2014), *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, Springer, N. Y.

Nicholls, R. J., and A. Cazenave (2010), Sea-level rise and its impact on coastal zones, *Science*, *328*, 1517–1520.

Nicholls, R. J., S. E. Hanson, J. A. Lowe, R. A. Warrick, X. Lu, and A. J. Long (2014), Sea-level scenarios for evaluating coastal impacts, *WIREs Clim. Change*, *5*, 129–150.

Olivieri, M., and G. Spada (2013), Intermittent sea-level acceleration, *Global Planet. Change*, *109*, 64–72.

Parker, A., M. S. Saleem, and M. Lawson (2013), Sea-level trend analysis for coastal management, *Ocean Coastal Manage.*, *73*, 63–81.

Prais, S. J., and C. B. Winsten (1954), Trend Estimators and Serial Correlation, Cowles Commission Discussion Paper, Statistics no. 383.

Rahmstorf, S., and M. Vermeer (2011), Discussion of: Houston, J.R. and Dean, R.G., 2011. Sea-level acceleration based on U.S. tide gauges and extensions of previous global-gauge analyses, Journal of Coastal Research, 27(3), 409–417, *J. Coastal Res.*, *27*(4), 784–787.

Refsgaard, J. C., J. P. van der Sluijs J. Brown, and P. van der Keur (2006), A framework for dealing with uncertainty due to model structure error, *Adv. Water Resour.*, *29*, 1586–1597.

Rhein, M., et al. (2013), Observations: Ocean, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifths Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K.

Richter, K., and B. Marzeion (2014), Earliest local emergence of forced dynamic and steric sea-level trends in climate models, *Environ. Res. Lett.*, *9*, 115007.

Rintoul, S. (2011), Sea-level rises are slowing, tidal gauge records show. The Australian. [Available at http://www.theaustralian.com.au/national-affairs/sea-level-rises-are-slowing-tidal-gauge-records-show/story-fn59niix-1226099350056.]

Sallenger, A. H., Jr., K. S. Doran, and P. A. Howd (2012), Hotspot of accelerated sea-level rise on the Atlantic coast of North America, *Nat. Clim. Change*, *2*, 884–888, doi:10.1038/nclimate1597.

Scafetta, N. (2014), Multi-scale dynamical analysis (MSDA) of sea level records versus PDO, AMO, and NAO indexes, *Clim. Dyn.*, *43*, 175–192, doi:10.1007/s00382-013-1771-3.

Slangen, A. B. A., M. Carson, C. A. Katsman, R. S. W. van de Wal, A. Köhl, L. L. A. Vermeersen, and D. Stammer (2014), Projecting twenty-first century regional sea-level changes, *Clim. Change*, *124*, 317–332.

Smith, R. A. (2002), Historical golden gate tidal series, *NOAA Tech. Rep. NOS CO-OPS 035*.

Soyer, E., and R. M. Hogarth (2012), The illusion of predictability: How regression statistics mislead experts, *Int. J. Forecasting*, *28*, 695–711.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002), Bayesian measures of model complexity and fit, *J. Stat. Soc., Ser. B*, *64*(4), 583–639.

Tamazian, A., J. Ludescher, and A. Bunde (2015), Significance of trends in long-term correlated records, *Phys. Rev. E*, *91*, 032806.

Thompson, P. R., M. A. Merrifield, J. R. Wells, and C. M. Chang (2014), Wind-driven coastal sea level variability in the Northwest Pacific, *J. Clim.*, *27*, 4733–4751.

Trömel, S., and C. D. Schönwiese (2008), Robust trend estimation of observed German precipitation, *Theor. Appl. Climatol.*, *93*, 107–115.

Visser, H. (2004), Estimation and detection of flexible trends, *Atmos. Environ.*, *38*, 4135–4145.

Visser, H. and J. Molenaar (1995), Trend estimation and regression analysis in climatological time series: An application of structural time series models and the Kalman filter, *J. Clim.*, *8*(5), 969–979.

Visser, H. and A. C. Petersen (2012), Inferences on weather extremes and weather-related disasters: A review of statistical methods, *Clim. Past*, *8*, 265–286.

Visser, H., R. J. M. Folkert, J. Hoekstra and J. J. de Wolf (2000), Identifying key sources of uncertainty in climate change projections, *Clim. Change*, *45*, 421–457.

Visser, H., A. C. Petersen, and W. Ligtvoet (2014), On the relation between weather-related disaster impacts, vulnerability and climate change, *Clim. Change*, *125*, 461–477.

von Storch, H., and F. W. Zwiers (1999), *Statistical Analysis in Climate Research*, Cambridge Univ. Press, Cambridge, U. K.

Wahl, T., J. Jensen, T. Frank, and I. D. Haigh (2011), Improved estimates of mean sea level changes in the German Bight over the last 166 years, *Ocean Dyn.*, *61*, 701–715, doi:10.1007/s10236-011-0383-x.

Wahl, T., I. D. Haigh, P. L. Woodworth, F. Albrecht, D. Dillingh, J. Jensen, R. J. Nicholls, R. Weisse, and G. Wöppelmann (2013), Observed mean sea level changes around the North sea coastline from 1800 to present, *Earth Sci. Rev.*, *124*, 51–67.

Watson, P. J. (2011), Is there evidence yet of acceleration in mean sea level rise around mainland Australia?, *J. Coastal Res.*, *27*(2), 368–377.

Watson, P. J. (2015), Development of a unique synthetic data set to improve sea-level research and understanding, *J. Coastal Res.*, *31*, 758–770.

Watson, T., and A. Parker (2013), Discussion of "Towards a global regionally varying allowance for sea-level rise" by J.R Hunter, J.A. Church, N.J. White, X. Zhang [Ocean Engineering 71(1) 17–27 (1 October 2013)], *Ocean Eng.*, *72*, 470–472.

Wenzel, M., and J. Schröter (2014), Global and regional sea level change during the 20th century, *J. Geophys. Res. Oceans*, *119*, 7493–7508, doi:10.1002/2014JC009900.

White, N. J., et al. (2014), Australian sea levels—Trends, regional variability and influencing factors, *Earth Sci. Rev.*, *136*, 155–174.

Woodworth, P. L., N. J. White, S. Jevrejeva, S. J. Holgate, J. A. Church, and W. R. Gehrels (2009), Review: Evidence for the acceleration of sea level on multi-decade and century timescales, *Int. J. Climatol.*, *29*, 777–789.

Woodworth, P. L., M. Menéndez, and W. R. Gehrels (2011), Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels, *Surv. Geophys.*, *32*(4/5), 603–618.

Yuan, N., F. Zuntao, and S. Liu (2014), Extracting climate memory using fractional integrated statistical model: A new perspective on climate prediction, *Nat. Sci. Rep.*, *4,6577*, doi:10.1038/srep06577.

Zhang, X., and J. A. Church (2012), Sea level trends, interannual and decadal variability in the Pacific Ocean, *Geophys. Res. Lett.*, *39*, L21701, doi:10.1029/2012GL053240.

Zhang, X., and F. W. Zwiers (2004), Monte Carlo experiments on the detection of trends in extreme values, *J. Clim.*, *17*, 1945–1952.

Zieba, A. (2010), Effective number of observations and unbiased estimators of variance for autocorrelated data—An overview, *Metrol. Meas. Syst.*, *17*(1), 3–16.