



# Spatial patterns of linear and nonparametric long-term trends in Baltic sea-level variability

R. V. Donner<sup>1</sup>, R. Ehrcke<sup>1</sup>, S. M. Barbosa<sup>2</sup>, J. Wagner<sup>1</sup>, J. F. Donges<sup>1,3</sup>, and J. Kurths<sup>1,3,4</sup>

<sup>1</sup>Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany

<sup>2</sup>Instituto Dom Luiz, University of Lisbon, Campo Grande, Edifício C8, 1749-016 Lisboa, Portugal

<sup>3</sup>Department of Physics, Humboldt University Berlin, Newtonstr. 15, 12489 Berlin, Germany

<sup>4</sup>Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB243UE, UK

Correspondence to: R. V. Donner (reik.donner@pik-potsdam.de)

Received: 7 October 2011 – Revised: 6 February 2012 – Accepted: 6 February 2012 – Published: 13 February 2012

**Abstract.** The study of long-term trends in tide gauge data is important for understanding the present and future risk of changes in sea-level variability for coastal zones, particularly with respect to the ongoing debate on climate change impacts. Traditionally, most corresponding analyses have exclusively focused on trends in mean sea-level. However, such studies are not able to provide sufficient information about changes in the full probability distribution (especially in the more extreme quantiles). As an alternative, in this paper we apply quantile regression (QR) for studying changes in arbitrary quantiles of sea-level variability. For this purpose, we chose two different QR approaches and discuss the advantages and disadvantages of different settings. In particular, traditional linear QR poses very restrictive assumptions that are often not met in reality. For monthly data from 47 tide gauges from along the Baltic Sea coast, the spatial patterns of quantile trends obtained in linear and nonparametric (spline-based) frameworks display marked differences, which need to be understood in order to fully assess the impact of future changes in sea-level variability on coastal areas. In general, QR demonstrates that the general variability of Baltic sea-level has increased over the last decades. Linear quantile trends estimated for sliding windows in time reveal a wide-spread acceleration of trends in the median, but only localised changes in the rates of changes in the lower and upper quantiles.

Not only the uncertainty on the future rate of global sea-level rise is very large (e.g. Milne et al., 2009), but also sea-level change is subject to considerable spatial heterogeneity. As demonstrated by satellite altimetry observations (e.g. Cazenave and Nerem, 2004), sea-level exhibits significant regional variability, reflecting the variety of regional and local geophysical phenomena that determine changes in the mass and volume of the Earth's oceans.

Regional sea-level variability, rather than global sea-level rise, is the main concern for risk assessment. The threat to coastal populations results from the combination of global sea-level rise with regional changes in sea-level associated with changes in atmospheric patterns and winds, tidal ranges, waves, coastal morphology, or land elevation (IPCC, 2007).

For a largely confined water body such as the Baltic sea, regional sea-level variability is particularly important for coastal management and socio-economic activities related to future changes. The Baltic is a shallow and semi-enclosed sea, with long-term sea-level variations being mainly driven by the exchange of water through the Danish Straits due to persistent winds over the North Sea and the Baltic entrance (Samuelsson and Stigebrandt, 1996). Hence, sea-level in the Baltic area is strongly influenced by atmospheric circulation patterns (Ekman, 2009). Specifically, the analysis of tide gauge records shows a correlation with the North Atlantic Oscillation (NAO) index (e.g. Jevrejeva et al., 2005). Other meteorological factors such as temperature and precipitation have considerably influence as well (Hünicke and Zorita, 2006). Due to this multiplicity of different processes driving sea-level variability, the accurate estimation of long-term sea-level trends in the Baltic area is a challenging problem, particularly in a climate change context.

## 1 Introduction

Global sea-level rise is one of the most direct expected consequences of climate change and associated with substantial socio-economic risks (IPCC, 2007; Church et al., 2010).

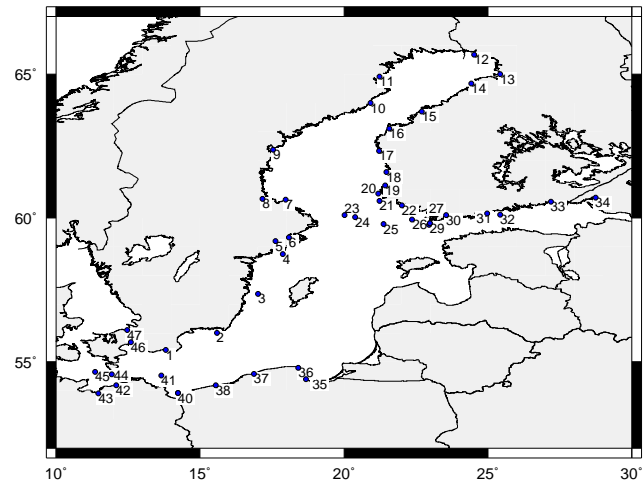
Studies of long-term sea-level variability based on the analysis of tide gauge records traditionally assume trends to be linear, and obtainable from the ordinary least squares fit of a first-order polynomial (linear function) to the tide gauge observations. However, among other issues such as auto-correlation (e.g. Barbosa et al., 2008), the assumption of a linear monotonic trend can be highly questionable, particularly in the case of strong inter-annual and decadal variability (e.g. Holgate, 2007). Furthermore, slopes derived from linear trend models only provide information on the rate of change of the mean tide gauge observations, whereas long-term variability in other parts of the data distribution is equally or even more relevant, particularly in terms of risk assessment and coastal protection. In this work quantile regression is applied to derive information on long-term variability for the entire probability distribution of sea-level.

This paper is organised as follows: in Sects. 2 and 3, we describe the data and methods considered in this work in some detail. The results of our analysis are presented and thoroughly discussed in Sect. 4. Finally, the main findings of this study are summarised and put into an oceanographic and climate change context (Sect. 5).

## 2 Baltic tide gauge data

The longest available observational records of sea-level variability are from coastal tide gauge stations. Since tide gauges measure relative sea-level (RSL, i.e. the height of the sea surface relative to a reference point on land), these records include both the rise and drop of the sea surface as well as the vertical movements (uplift or subsidence) of the adjacent land. The interest in studying the postglacial rebound of Fennoscandia motivated the precocious set-up of a dense network of tide gauges in the Baltic area, and as a result a significant number of long and continuous tide gauge records are presently available for studies of long-term sea-level variability.

Monthly tide gauge records of relative sea-level are provided by the Permanent Service for Mean Sea Level (PSMSL, see <http://www.psmsl.org>) (Woodworth and Player, 2003). In this work, we have selected 47 records from the Baltic Sea area with sufficiently continuous data, which are shown in Fig. 1 (for detailed information on the properties of these records, see Table 1). Among these tide gauges, 30 time series provide a sufficiently complete temporal coverage of the second half of the 20th century (1951–1999). Since the data have been obtained from its repository, the basic quality-control procedures defined by the PSMSL have already been applied. For most records only a few non-consecutive observations have been missing, and the total amount of missing values is less than 2%, as can be seen from Table 1. For a few records (specifically RAA, VAA, LYP, DEG and RUS) one complete year is missing. For one single record (SAS) more than one consecutive year is



**Fig. 1.** Location of the tide gauges considered in this study (cf. Table 1).

missing (1988–1992). However, gaps in the monthly time series data need not to be filled in, since quantile regression (see Sect. 3) as the method of choice in this work does not require equidistant observations and is thus able to cope with missing data.

As it can be seen from Table 1, a large part of the Baltic region experiences a notable land uplift due to postglacial rebound, reaching  $\sim 90 \text{ mm dec}^{-1}$  in the northern area of the Gulf of Bothnia (e.g. Ekman, 1996). In this work, the uplift effect is considered in terms of a glacial isostatic adjustment (GIA) model based on the VM4 earth model (Peltier, 1998, 2004). Corresponding average trends in relative sea-level have also been obtained from the PSMSL. In turn, the tide gauge records under study have not been corrected for the inverse barometric effect, since we are interested in investigating the observed sea-level variability in the Baltic sea as such, irrespective of the causative forcings, atmospheric or other.

Sea-level in the Baltic exhibits in general an annual cycle peaking in the winter months (e.g. Hünicke and Zorita, 2008). For the purpose of studying long-term sea-level variability, the mean annual cycle is estimated by averaging all values for each calendar month contained in the respective time series, and then subtracted from each sea-level record. For this purpose, we have used the STL ((S)easonal-(T)rend decomposition procedure based on (L)ocally weighted regression) algorithm (Cleveland et al., 1990) for a seasonal time series decomposition by means of locally weighted regression (LOESS) (Cleveland, 1979; Cleveland and Devlin, 1988) in its implementation in the R package *stats*. For simplicity, we considered a fixed seasonal cycle instead of a potentially changing one, which might provide an even more appropriate description of the corresponding annual variability component. Note that the proper removal of

**Table 1.** Basic properties of the tide gauge records used in this study. The station ID refers to the coastline (first three digits) and station codes from the PSMSL.  $T$ ,  $N$  and  $NA$  give the total length of the time series (in years), the number of data points, and the number of missing data points, respectively, LON and LAT denote the station coordinates. GIA-VM4 lists the average trends in sea-level due to glacial isostatic adjustment obtained from the VM4 earth model (see [http://www.psmsl.org/train\\_and\\_info/geo\\_signals/gia/peltier/](http://www.psmsl.org/train_and_info/geo_signals/gia/peltier/)).

|    | ID     | Station           | Observation<br>Period | $T$<br>(yr) | $N$  | $NA$ | LON<br>(°E) | LAT<br>(°N) | GIA-VM4<br>(mm dec <sup>-1</sup> ) |
|----|--------|-------------------|-----------------------|-------------|------|------|-------------|-------------|------------------------------------|
| 1  | 050071 | Ystad             | Jan 1887–Dec 1981     | 95          | 1140 | 0    | 13.817      | 55.417      | 2.4                                |
| 2  | 050081 | Kungholmsfort     | Jan 1887–Dec 2006     | 120         | 1440 | 1    | 15.583      | 56.01       | 1.6                                |
| 3  | 050091 | Olands Norra Udde | Jan 1887–Dec 2006     | 120         | 1440 | 0    | 17.01       | 57.367      | −2.9                               |
| 4  | 050121 | Landsort          | Jan 1887–Dec 2006     | 120         | 1440 | 0    | 17.867      | 58.75       | −16.1                              |
| 5  | 050131 | Nedre Sodertalje  | Jan 1869–Dec 1970     | 102         | 1224 | 0    | 17.617      | 59.2        | −23.1                              |
| 6  | 050141 | Stockholm         | Jan 1889–Dec 2006     | 118         | 1416 | 0    | 18.083      | 59.317      | −23.5                              |
| 7  | 050161 | Bjorn             | Jan 1892–Dec 1976     | 85          | 1020 | 0    | 17.967      | 60.633      | −44.2                              |
| 8  | 050171 | Nedre Gavle       | Jan 1896–Dec 1986     | 91          | 1092 | 1    | 17.167      | 60.667      | −45.9                              |
| 9  | 050183 | Spikarna          | Jan 1969–Dec 2006     | 38          | 456  | 1    | 17.533      | 62.367      | −58.8                              |
| 10 | 050191 | Ratan             | Jan 1892–Dec 2006     | 115         | 1380 | 3    | 20.917      | 64          | −70.7                              |
| 11 | 050210 | Furuogrund        | Jan 1916–Dec 2006     | 91          | 1092 | 5    | 21.233      | 64.9167     | −73.9                              |
| 12 | 060001 | Kemi              | Jan 1920–Dec 2004     | 85          | 1020 | 40   | 24.5167     | 65.667      | −81.8                              |
| 13 | 060011 | Oulu              | Jan 1889–Dec 2004     | 116         | 1392 | 68   | 25.4167     | 65.0033     | −79.1                              |
| 14 | 060021 | Raahe             | Jan 1923–Dec 2004     | 82          | 984  | 75   | 24.4167     | 64.667      | −79.0                              |
| 15 | 060041 | Pietarsaari       | Jan 1915–Dec 2004     | 90          | 1080 | 21   | 22.7        | 63.7        | −69.9                              |
| 16 | 060051 | Vaasa             | Jan 1922–Dec 2004     | 83          | 996  | 75   | 21.567      | 63.1        | −64.3                              |
| 17 | 060071 | Kaskinen          | Jan 1927–Dec 2004     | 78          | 936  | 23   | 21.2167     | 62.333      | −56.1                              |
| 18 | 060101 | Mantyluoto        | Jan 1911–Dec 2004     | 94          | 1128 | 19   | 21.467      | 61.6        | −44.7                              |
| 19 | 060121 | Rauma             | Jan 1933–Dec 2004     | 72          | 864  | 6    | 21.4333     | 61.1333     | −37.6                              |
| 20 | 060221 | Lyokki            | Jan 1858–Dec 1936     | 79          | 948  | 3    | 21.1833     | 60.85       | −34.6                              |
| 21 | 060231 | Lypyrsti          | Jan 1858–Dec 1936     | 79          | 948  | 17   | 21.2333     | 60.6        | −30.2                              |
| 22 | 060241 | Turku             | Jan 1922–Dec 2004     | 83          | 996  | 21   | 22.01       | 60.433      | −22.4                              |
| 23 | 060271 | Lemstrom          | Jan 1889–Dec 1936     | 48          | 576  | 0    | 20.0167     | 60.1        | −28.6                              |
| 24 | 060281 | Degerby           | Jan 1924–Dec 2004     | 81          | 972  | 49   | 20.3833     | 60.0333     | −25.7                              |
| 25 | 060291 | Uto               | Jan 1866–Dec 1936     | 71          | 852  | 4    | 21.367      | 59.783      | −16.8                              |
| 26 | 060311 | Jungfrusund       | Jan 1858–Dec 1934     | 77          | 924  | 0    | 22.367      | 59.95       | −14.1                              |
| 27 | 060316 | Russaro           | Jan 1866–Dec 1936     | 71          | 852  | 21   | 22.95       | 59.767      | −9.0                               |
| 28 | 060331 | Hanko             | Jan 1888–Dec 1935     | 48          | 576  | 37   | 22.9833     | 59.8167     | −9.4                               |
| 29 | 060331 | Hanko             | Jan 1943–Dec 1997     | 55          | 660  | 15   | 22.9833     | 59.8167     | −9.4                               |
| 30 | 060344 | Skuro             | Jan 1900–Dec 1936     | 37          | 444  | 2    | 23.55       | 60.1        | −10.3                              |
| 31 | 060351 | Helsinki          | Jan 1879–Dec 2004     | 126         | 1512 | 2    | 24.967      | 60.15       | −5.3                               |
| 32 | 060354 | Soderskar         | Jan 1866–Dec 1936     | 71          | 852  | 0    | 25.4167     | 60.1167     | −3.7                               |
| 33 | 060361 | Hamina            | Jan 1929–Dec 2001     | 73          | 876  | 14   | 27.183      | 60.567      | −4.6                               |
| 34 | 080002 | Vyborg            | Jan 1889–Dec 1938     | 50          | 600  | 0    | 28.733      | 60.7        | −4.0                               |
| 35 | 110022 | Gdansk            | Jan 1951–Dec 1999     | 49          | 588  | 0    | 18.683      | 54.4        | 0.0                                |
| 36 | 110047 | Wladyslawowo      | Jan 1951–Dec 1999     | 49          | 588  | 0    | 18.4167     | 54.8        | 0.8                                |
| 37 | 110057 | Ustka             | Jan 1951–Dec 1999     | 49          | 588  | 0    | 16.867      | 54.583      | 0.5                                |
| 38 | 110072 | Kolobrzeg         | Jan 1951–Dec 1999     | 49          | 588  | 0    | 15.55       | 54.183      | −0.1                               |
| 39 | 110092 | Swinoujscie       | Jan 1824–Dec 1941     | 118         | 1416 | 0    | 14.233      | 53.9167     | −0.1                               |
| 40 | 110092 | Swinoujscie       | Jan 1951–Dec 1999     | 49          | 588  | 0    | 14.233      | 53.9167     | −0.1                               |
| 41 | 120004 | Sassnitz          | Jan 1946–Dec 2004     | 59          | 708  | 61   | 13.65       | 54.52       | 1.6                                |
| 42 | 120012 | Warnemuinde       | Jan 1856–Dec 2005     | 150         | 1800 | 2    | 12.0833     | 54.1833     | 2.3                                |
| 43 | 120022 | Wismar            | Jan 1849–Dec 2005     | 157         | 1884 | 2    | 11.467      | 53.9        | 2.3                                |
| 44 | 130001 | Gedser            | Jan 1898–Dec 2002     | 105         | 1260 | 13   | 11.93       | 54.57       | 3.3                                |
| 45 | 130011 | Rødbyhavn         | Jan 1955–Dec 2002     | 48          | 576  | 28   | 11.35       | 54.65       | 4.1                                |
| 46 | 130021 | København         | Jan 1889–Dec 2002     | 114         | 1368 | 28   | 12.6        | 55.68       | 3.1                                |
| 47 | 130031 | Hornbæk           | Jan 1898–Dec 2002     | 105         | 1260 | 27   | 12.47       | 56.1        | 2.2                                |

all seasonal effects is a non-trivial and still not completely solved problem of geoscientific time series analysis (Donner et al., 2008). In general, the results of quantile regression depend to some degree on whether or not seasonality effects have been removed from the data prior to analysis (see Sect. 4.1). In order to further address the related problems, it would be necessary to systematically compare the performance of different methods for removing the annual cycle from the data in order to verify the robustness of the obtained results of quantile regression. However, a corresponding detailed study is clearly beyond the scope of this work. Even more, for practical purposes such as planning and managing of adaptation measures to counter future sea-level rise in coastal areas, we may argue that the net effect of long-term trends plus seasonality is most relevant.

Barbosa (2008) already analysed a subset of the records considered in this work by means of linear QR. It has been demonstrated that the slopes of the linear quantile functions depend strongly on the chosen quantile, and that the higher quantiles of the sea-level distribution rise clearly faster (or fall slower, respectively) than the mean. In this work, we extend these previous results in different ways. (i) We consider a larger set of tide gauges for an improved coverage of the spatial structure of sea-level trends in the Baltic Sea. (ii) We carefully examine the effect of removing the annual cycle from the data. (iii) We investigate how strong the linearity assumption influences the estimated trends in different quantiles by comparing the results of linear and nonparametric QR. (iv) We explicitly consider the effect of GIA on the observed sea-level variability by correcting the obtained results for the mean rise/fall rates due to vertical land movements.

### 3 Quantile regression (QR) analysis

Quantile regression (Koenker, 2005; Yu et al., 2003) is a well-defined statistical framework that allows evaluating dependences of the quantiles of a given variable of interest on a set of independent covariates or predictors. In this sense, it generalises classical regression analysis which characterises corresponding relationships for the mean. Given a random variable  $Y$  with a continuous cumulative distribution function  $F_Y(y)$ , the  $\alpha$ -quantile  $q_{Y,\alpha}$  is defined as the value of  $Y$  for which  $P(Y \leq q_{Y,\alpha}) = F_Y(q_{Y,\alpha}) = \alpha$  ( $0 < \alpha < 1$ ) (here,  $P(A)$  is the probability of the condition  $A$  to apply). Thus, the quantile  $q_{Y,\alpha}$  can be determined by evaluating the inverse function associated with the cumulative distribution at the value  $\alpha$ , i.e.  $q_{Y,\alpha} = F_Y^{-1}(\alpha)$ .

In many practical situations, one is interested in the conditional distribution of  $Y$  given the values of one or more covariates  $X$  (for our purposes,  $X$  will be the time coordinate). Then, the corresponding conditional quantile function  $q_{Y|X,\alpha}(x)$  has to satisfy  $P(Y \leq q_{Y|X,\alpha}(x) | X = x) = \alpha$ . While classical regression analysis considers the conditional mean, QR is based on the conditional quantile functions and

a minimisation of the sum of asymmetrically weighted absolute residuals (see Sect. 3.1). In the following, we will omit the subscripts indicating the variable of interest in order to simplify the notation.

Although it has been originally introduced and widely applied in econometrics, in the last years, an increasing number of applications of QR to geoscientific problems has been reported. In a time series analysis context, variations in the distribution of temperature and precipitation records have been studied by various authors (Koenker and Schorfheide, 1994; Draghicescu, 2002; Zhou and Wu, 2009; Timofeev and Sterin, 2010; Cannon, 2011; Barbosa et al., 2011). Besides time as a unique predictor, problems interrelating different geoscientific variables with each other have been extensively discussed, including the effect of meteorological variables on ozone concentration (Baur et al., 2004), the modelling of tropical cyclone intensity based on an additive QR model with different climatic covariates (Elsner et al., 2008; Jagger and Elsner, 2009), or the soil-moisture impact on hot extremes in southeastern Europe (Hirschi et al., 2011). Kyselý et al. (2010) used QR for obtaining threshold values for time-dependent extreme value analysis of climate simulations. In the context of sea-level research, Barbosa (2008) studied linear QR models for selected tide gauge records from the Baltic Sea. Park et al. (2010) investigated the interrelationships between local extreme sea-level in Florida and the Atlantic Multidecadal Oscillation (AMO). In addition to many other applications as well as intensive methodological work mainly done in the econometrics community, these examples demonstrate the wide applicability of QR. To our knowledge, there are no other conceptually different methods for estimating conditional quantiles available so far that perform equally well – or even better – for the purpose of estimating quantile trends.

#### 3.1 Basic idea: linear QR

In linear QR, the unknown quantile function  $q_\alpha(x)$  is expressed in terms of a linear model function  $f_\alpha(x) = \beta_\alpha x + \gamma_\alpha$ . In order to properly estimate the values of the slope  $\beta_\alpha$  and intercept  $\gamma_\alpha$ , one has to modify the classical (least-squares) regression approach as

$$\hat{q}_\alpha(x) = \min_f \sum_{i=1}^n \rho_\alpha(y_i - f_\alpha(x_i)) \quad (1)$$

with the so-called check function

$$\begin{aligned} \rho_\alpha(u) &= \alpha u I_{[0,\infty)}(u) - (1-\alpha) u I_{(-\infty,0)}(u) \\ &= u \left( \alpha - I_{(-\infty,0)}(u) \right) \\ &= \left[ \frac{1}{2} + \left( \alpha - \frac{1}{2} \right) \text{sgn}(u) \right] |u| \end{aligned} \quad (2)$$

(where  $I_A(\cdot)$  is the indicator function of the set  $A$ ). The solution of the linear estimation problem (1) can be obtained by

standard linear optimisation algorithms (Koenker, 2005). In this work, we use the R package `quantreg` (function `rq`) for performing the corresponding analyses.

We emphasise that the above setup generalises the symmetric loss functions for the mean ( $\rho(u) = u^2$ ) and median ( $\rho_{0.5}(u) = 0.5|u|$ ) from classical regression. In this spirit, the results of QR for intermediate quantiles are typically more robust against outliers than those of standard least-squares regression for trends in the mean. However, this robustness necessarily decreases towards more extreme quantiles. These general statements do not exclusively apply to linear QR, but also to its nonlinear or nonparametric counterparts (see below).

Another typical problem of practical importance when studying sea-level variability are possible shifts in the data, e.g. due to imperfect calibration or changes of the measurement devices. In such cases, it is likely that the whole probability distribution of observed values is shifted by a fixed value, so that a constant shift involving the entire time series is no problem to the analysis. In turn, having structural break points in the time series due to some intermittent shift of the observations will clearly influence the outcome of linear QR depending on the magnitude of the shift and the total number of available data. We note that this problem is partially solved when using nonparametric QR methods (see below) that interpolate the observed probability distribution locally.

### 3.2 Nonparametric QR using splines

In contrast to the linearity assumption made in traditional QR, processes in nature, and resulting statistical interrelationships are typically nonlinear and/or non-stationary. Such behaviour implies that for a changing value of a certain covariate  $x$  (e.g. time), changes in the distribution of an observable  $y$  can often not be appropriately described by linear functions. Therefore, various extensions of linear QR have been developed. On the one hand, it is possible to explicitly prescribe nonlinear parametric models to the desired quantile functions, for which the appropriate model parameters can be estimated by directly generalising the least-squares based approach from linear QR (Koenker, 2005). However, the latter approach requires a priori knowledge on the functional form of the trends under study, which is often not available. In the latter cases, it can therefore be desirable to follow some nonparametric statistical approach, which involves the appropriate choice of suitable strategies for estimation and possible smoothing. Among other methods, the approximation of the conditional quantiles by means of spline functions (Koenker et al., 1994; Koenker and Schorfheide, 1994; Thompson et al., 2010) has been intensively studied in the statistical literature and offers a method with a particularly solid theoretical foundation.

In its basic setting, spline QR is a simple generalisation of traditional (linear or nonlinear) parametric quantile regression methods, where the globally defined model quantile

function  $f_\alpha(x)$  in Eq. (1) is replaced by a spline function with predefined boundary conditions, but without extensive additional constraints. This strategy corresponds to a piecewise polynomial regression with multiple (unknown) break-points. The disadvantage of this conceptually still rather simple strategy is, however, that the estimated quantile functions may display strong fluctuations, which is not desired when studying trends that mainly reflect smooth long-term changes.

As a possible solution to the aforementioned regularisation problem, the desired smoothness of the quantile functions to be estimated can be included as an additional constraint in the estimation problem. In this case, the minimisation problem for the least-squares “fidelity” (or risk/loss) term is extended by an additional penalty term which characterises the smoothness of the desired quantile function (see below), which leads to a so-called *quantile smoothing spline* (QSS). Specifically, Koenker et al. (1994) proposed solving the following problem with a general  $L_p$  penalty term:

$$\hat{q}_\alpha(x|\lambda) = \min_{f \in S} \left[ \sum_{i=1}^n \rho_\alpha(y_i - f(x_i)) + \lambda \left( \int_0^1 dx |f''(x)|^p \right)^{\frac{1}{p}} \right], \quad (3)$$

where  $S$  denotes the set of admissible spline functions. At  $\lambda = 0$ , the estimate  $\hat{q}_\alpha(x)$  interpolates the  $\alpha$ -quantile at the selected design points of the spline function, whereas for  $\lambda \rightarrow \infty$ , the linear QR solution is asymptotically approached. Koenker et al. (1994) demonstrated that different choices of  $p$  imply different types of spline functions, with linear splines (for  $L_1$  penalty) and quadratic splines (for  $L_\infty$  penalty) as the limiting cases. Consequently,  $L_1$ -smoothing splines correspond to a piecewise linear change point model.

In this study, we use two different implementations of the respective algorithms in R: for  $L_1$  smoothing splines, the corresponding implementation in the `quantreg` package (`rqss` function) has been utilised. In addition, the COBS algorithm (He and Ng, 1999) based on constrained B-splines allows studying both  $L_1$  and  $L_\infty$  smoothing splines. In the latter case, we have used the `cobs` function from the R package `cobs99`<sup>1</sup>.

### 3.3 Parameter selection

An appropriate regularisation of the desired solution (balancing the fidelity and smoothness) is an omnipresent problem in nonparametric QR. For properly choosing the corresponding smoothing parameter, a variety of criteria has been suggested.

<sup>1</sup>The more recent implementation (Koenker and Ng, 2005; Ng and Maechler, 2007) of the same function in the R package `cobs`, which makes use of different numerical routines, has been found to run less stably for the data studied in this work.

On the one hand, the number  $p(\lambda)$  of “active knots” of the spline (i.e. the number of interpolated data points) determines the solution of the optimisation problem in Eq. (3) to a large extent. This implies that the obtained solution changes only at discrete values of  $\lambda$  due to changes in the number or positions of the active knots. Specifically, under rather general conditions there is exactly one choice of the number  $p$  and positions of active knots (which is taken for a finite interval of  $\lambda$ ) that corresponds to some optimum parsimony of the resulting QSS. In this sense, determining an appropriate value of  $\lambda$  is a model selection problem with  $p(\lambda)$  determining the “model order”. For determining a reasonable choice of  $\lambda$ , it is therefore possible to consider standard penalised-likelihood criteria such as Akaike’s information criterion

$$\text{AIC} = \log \left( \frac{1}{n} \sum_{i=1}^n \rho_{\alpha}(y_i - \hat{q}_{\alpha}(x_i|\lambda)) \right) + \frac{p(\lambda)}{n} \quad (4)$$

or the Schwarz (Bayesian) information criterion (Koenker et al., 1994)

$$\text{SIC} = \log \left( \frac{1}{n} \sum_{i=1}^n \rho_{\alpha}(y_i - \hat{q}_{\alpha}(x_i|\lambda)) \right) + \frac{p(\lambda) \log n}{2n}. \quad (5)$$

These criteria should take their global minimum (or maximum when equivalently expressed in terms of the likelihood associated with the fidelity term) for the optimal model. However, since they involve certain assumptions regarding the distribution of the data by construction (which may not be fulfilled in real-world problems), it is possible that both AIC and SIC may lead to sub-optimal choices of  $\lambda$  in practice. Specifically, it has been argued that SIC is not feasible for selecting  $\lambda$  for extreme quantiles (Koenker et al., 1994).

On the other hand, the optimum value of  $\lambda$  can be determined by means of cross-validation techniques, adopting classical approaches from nonparametric regression and density estimation (Härdle, 1990). For example, the leave-one-out estimators  $\hat{q}_{\alpha}^{(-i)}(x|\lambda)$  estimating the desired quantile function  $q_{\alpha}(x|\lambda)$  using all available data but  $(x_i, y_i)$  can be used for choosing the optimum  $\lambda$  as the value minimising the cross-validation score

$$\text{CV}(\lambda) = \frac{1}{n} \sum_{i=1}^n \rho_{\alpha}(y_i - \hat{q}_{\alpha}^{(-i)}(x_i|\lambda)). \quad (6)$$

Based on ideas from Nychka et al. (1995), Yuan (2006) derived the generalised approximate cross-validation score

$$\text{GACV}(\lambda) = \frac{1}{n - \text{tr}(H)} \sum_{i=1}^n \rho_{\alpha}(y_i - \hat{q}_{\alpha}(x_i|\lambda)) \quad (7)$$

(where  $\text{tr}(H)$  is the trace of the matrix  $H_{ij} = \partial \hat{q}_{\alpha}(x_i|\lambda) / \partial y_j$ ), which leads to a significant reduction of computational demands compared to leave-one-out estimators. We note that both  $p(\lambda)$  and  $\text{tr}(H)$  are measures for the degrees of freedom  $df$  of the regularisation problem. A detailed inspection

reveals that GACV can be transformed into a form closely related to AIC and SIC, with the penalty term being replaced by  $\log(1 - df/n)$  (Li et al., 2007). For large  $n$  ( $df/n \ll 1$ ), the latter term can be approximated by  $df/n$ , which leads to AIC in the asymptotic limit.

It should be underlined that in general the optimum choice of  $\lambda$  depends on the considered quantile  $\alpha$ . In this work, we will restrict our attention to the penalised-likelihood criteria, particularly AIC, as the most widely used approach.

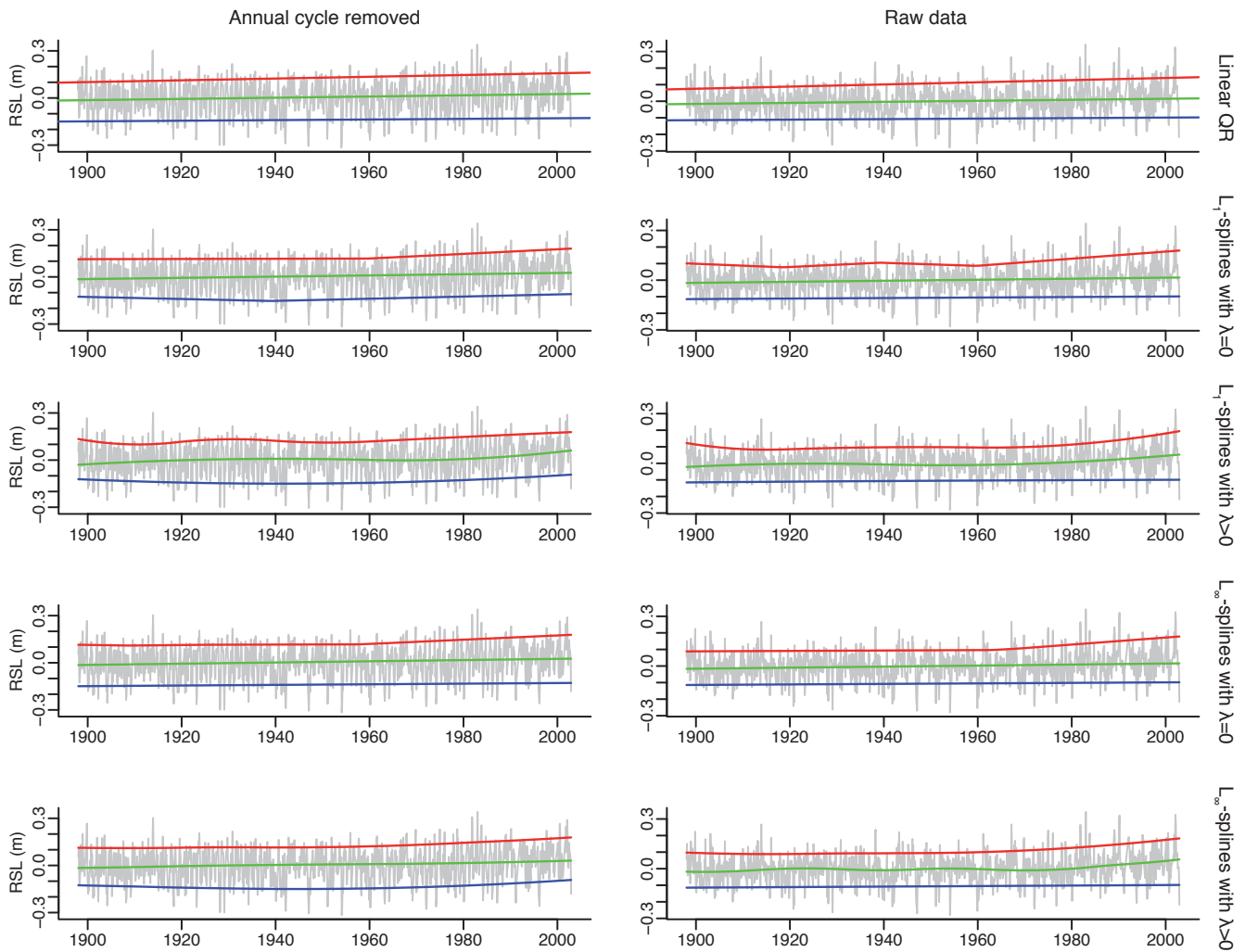
### 3.4 Quantile crossing

An inherent problem of linear QR is that the individual quantile functions cross each other (possibly outside the studied interval of the covariate  $x$ ) given that the slopes  $\beta_{\alpha}$  are not the same for all quantiles  $\alpha$ . This fact is an unavoidable methodological disadvantage of the linear model, which particularly motivates the use of nonparametric methods. It should be noted that even in the latter case, the fact that the different quantile functions are estimated independently often leads to quantile crossings. There are, however, different approaches for circumventing this problem, including the explicit consideration of non-crossing constraints in the optimisation problem (Takeuchi et al., 2006) and the double-kernel approach (Yu and Jones, 1998) where the indicator function  $I_A(\cdot)$  is replaced by a continuous distribution function associated with a kernel function acting on the dependent variable  $y$ . Since in this work, we will be mainly interested in the mean slope of each quantile and its variance, we will not consider such approaches explicitly. Instead, in the remainder of this paper we will focus our attention on both linear QR and quantile smoothing splines without considering any specific non-crossing constraint.

## 4 Results

### 4.1 Example: quantile trends for København

In order to get an impression of the differences between linear and nonparametric quantile trends, we first focus on results obtained for individual stations. As an example, Fig. 2 shows the estimated trend functions for some quantiles  $\alpha$  for the København tide gauge obtained using linear QR and different variants of spline QR. The associated mean slopes for a larger range of quantiles are displayed in Fig. 3. As expected, the qualitative behaviour of the obtained quantile functions clearly depends on the particular method chosen. In general, the estimated slopes show a clear tendency towards a broadening of the underlying probability distribution, i.e. lower quantiles rise slower than higher ones (note the positive GIA slope at the København tide gauge, cf. Table 1). This result is highlighted by the fact that the linear quantile slopes  $\beta_{\alpha}$  (as well as the mean slopes  $\hat{\beta}_{\alpha}$  of the corresponding spline estimates) scatter around the slope of a linear regression for the mean (Fig. 3) with a tendency towards smaller slopes for

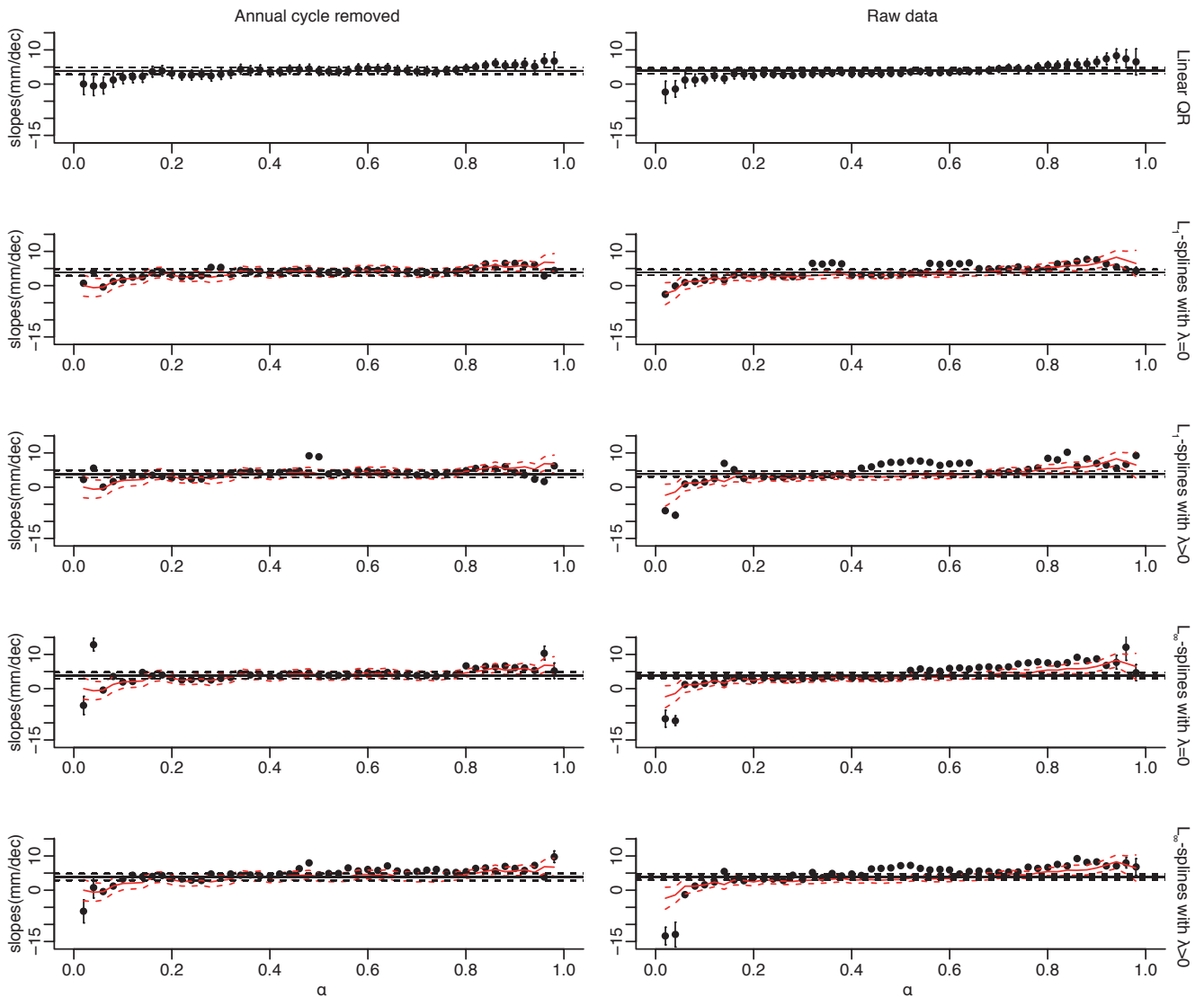


**Fig. 2.** Estimated trends in the 10 % (blue), 50 % (green) and 90 % (red) sea-level quantiles from the København tide gauge, obtained using linear and spline QR (from top to bottom: linear QR,  $L_1$ -splines (linear) with  $\lambda = 0$  and  $\lambda > 0$ ,  $L_\infty$ -splines (quadratic) with  $\lambda = 0$  and  $\lambda > 0$ ) obtained with (left) and without (right) removing the annual cycle from the original record. All spline models have been obtained with the `cobs99` package in R using an automatic parameter selection for  $\lambda$  based on AIC.

lower, and larger slopes for higher quantiles. As in a previous study (Barbosa, 2008), this appears to be a generic feature of monthly tide gauge records from the Baltic Sea.

Careful inspection of the average quantile slopes obtained with different variants of spline QR reveals that the trends inferred by nonparametric QR are qualitatively consistent with those shown by linear QR. There are, however, some distinct exceptions: On the one hand, there is a clear tendency for the extreme high and low quantiles estimated in a non-parametric way not to fall into the confidence bounds of the linear model. We relate this to the fact that for properly estimating extreme quantiles (e.g. below 5 % and above 95 %), a very high number of data is required, which is not available in the case of monthly records. On the other hand, there are distinct outliers for the spline-based methods where the

estimated mean slopes differ strongly from those obtained for neighbouring quantiles (Fig. 3). This results from the automatic selection of the number of active knots and the resulting smoothing parameter  $\lambda$ , which is carried out independently for each  $\alpha$  here. Specifically, in typical optimisation problems, there are multiple very similar optima (in terms of the associated AIC values) which seem to be taken here for different quantiles. In fact, both the extreme quantiles and these “outliers” have the strongest variability, which is expressed by the highest standard errors for the mean monthly increments ( $SE(\hat{\beta}) = \sigma_\beta / \sqrt{n}$  with  $\hat{\beta}$  and  $\sigma_\beta$  being the mean value and standard deviations of monthly increments of the estimated quantile functions). If we use a fixed value of  $\lambda$  for all quantiles  $\alpha$ , this effect vanishes (see Sect. 4.3).



**Fig. 3.** Average slopes  $\bar{\beta}_\alpha$  of the estimated quantile functions for the København tide gauge, obtained using linear and spline QR (from top to bottom: linear QR,  $L_1$ -splines (linear) with  $\lambda = 0$  and  $\lambda > 0$ ,  $L_\infty$ -splines (quadratic) with  $\lambda = 0$  and  $\lambda > 0$ ) obtained with (left) and without (right) removing the annual cycle from the original record. All spline models have been estimated with the `cobs99` package in R using an automatic parameter selection for  $\lambda$  based on AIC. Error bars indicate confidence intervals corresponding to  $\pm 1$  standard errors of the respective linear slope estimates  $\beta_\alpha$  (for linear QR models) and of the mean monthly increments  $\bar{\beta}_\alpha$  (for the nonparametric quantile estimates), respectively. For the nonparametric average quantile trends, the confidence intervals for linear QR estimates are additionally shown as red lines for comparison. Horizontal lines indicate the linear trend obtained for mean sea-level (solid line, estimated using ordinary least-squares regression) and the corresponding  $\pm 1$  standard error confidence intervals (dashed lines).

The obtained results underline that local sea-level trends are hardly uniform, but characterised by temporary increases as well as decreases in slope. This holds for trends in mean and median as well as for those in arbitrary quantiles. The temporal changes in slope could originate from long-term variations in other covariates, particularly meteorological parameters such as air temperature, pressure, or solar irradiation. We will briefly come back to such effects in the further course of this paper when studying temporal variations

in long-term sea-level trends in some more detail. However, a detailed discussion of the possible co-evolution of meteorological observables and RSL (Hünicke and Zorita, 2006) is beyond the scope of the present work and will remain a subject of future research. Besides these limitations of the present study, we emphasise that a detailed investigation of the robustness of the estimated nonparametric trend functions and the identification of possible periods with accelerating or decelerating sea-level trends is necessary in order to



derive insights into the complex interplay between triggering factors and sea-level response at a local scale.

In addition to the differences between different methods, Figs. 2 and 3 also allow evaluating the impact of deseasoning on the results of QR. While details in the quantile trends indeed change qualitatively when removing the annual component from the monthly tide gauge records, it can be seen that the general trend pattern persists. Specifically, the quantitative differences between the mean trends in lower and upper quantiles are a robust feature that is not altered by the corresponding preprocessing. Hence, as long as one is interested in the average long-term trends only, deseasoning of monthly tide gauge records is not necessary. In turn, if one is interested in temporal changes in the trends (in particular, the acceleration or deceleration of sea-level rise at a given site), the annual component plays a considerable role as it significantly interferes with the trend on annual to decadal time-scales. This calls for a careful treatment of the annual cycle depending on the specific research question under study.

#### 4.2 Spatial patterns of linear quantile trends

Previous research on Baltic sea-level variability from tide gauge data has mainly focussed on the consideration of individual tide gauges (Barbosa, 2008). Given the complete amount of records provided by PSMSL, in this work we are able to study the underlying spatial patterns of long-term sea-level trends. A first insight is gained by an inspection of linear quantile trends obtained for all 47 available tide gauges. Note that although only data from the period 1898–2002 will be considered in the following, the actual time intervals covered by the individual records are substantially different (see Table 1).

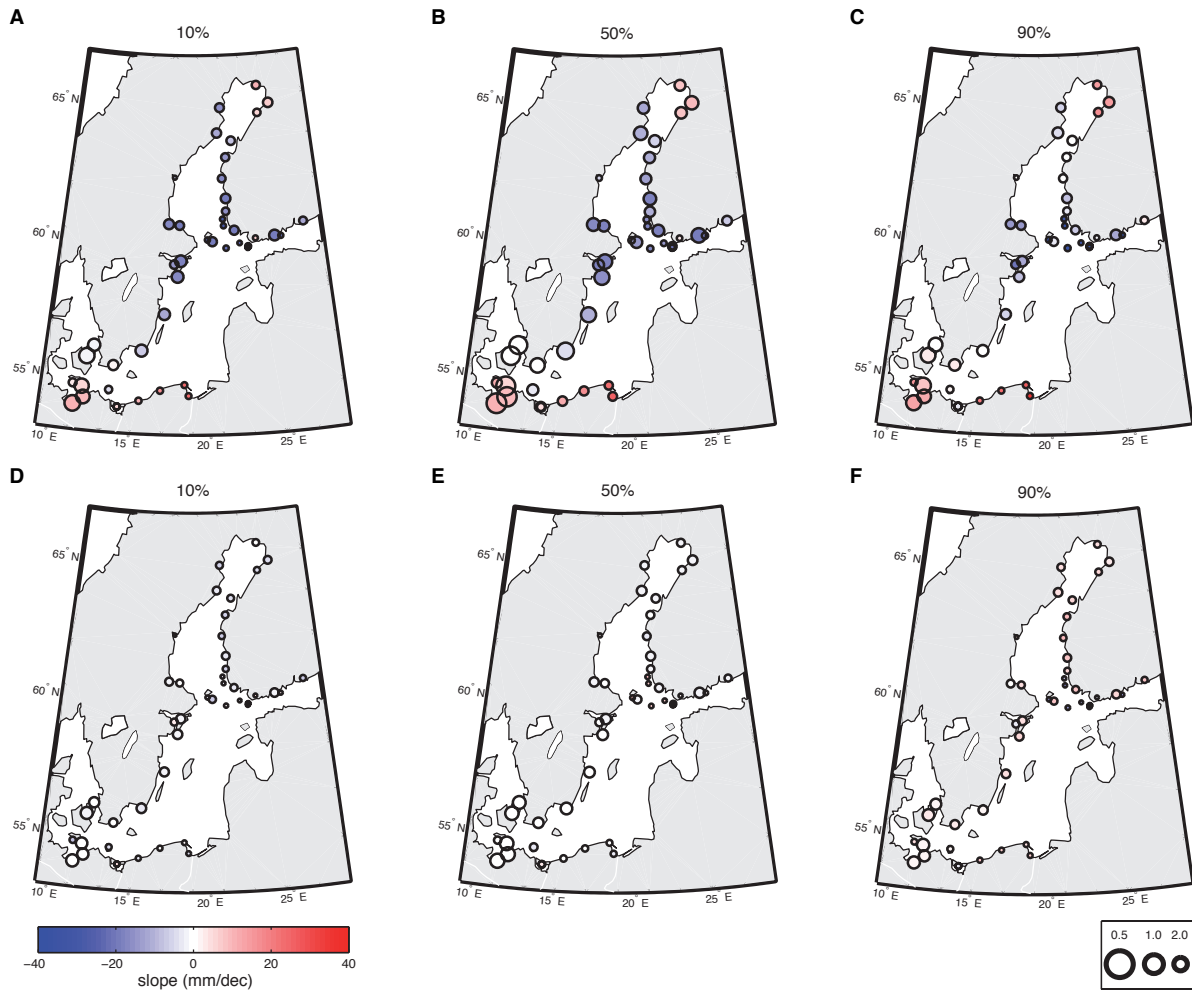
In general, the trends of RSL in the Baltic area derived from QR include both changes due to vertical land movements as well as changes in the height of the sea surface itself. In order to separate both effects, Fig. 4a–c shows the results of linear QR corrected for the influence of land movements by subtracting the trend from the GIA model. A spatially-consistent pattern is found for the low and high quantiles ( $\alpha = 0.1$  and  $0.9$ , respectively) as well as the median, with positive slopes in the southern area, negative slopes in the Gulf of Finland and Bothnian Sea, and positive slopes in the northernmost stations in the Bothnian Bay. Especially large positive trends are found along the Polish coast, whereas the results obtained for the south-western part of the Baltic Sea (Germany, Denmark, southern Sweden) show a slower rate of increase. On the one hand, these findings could just result from the different time coverage of the individual tide gauge records. In particular, the available data from Poland cover only the time period starting in 1951 (i.e. the most recent decades), whereas many of the other records contain measurements from considerably earlier times. This would have a significant effect especially if the trends in RSL quantiles are not constant in time. We will explicitly

study the influence of a homogenous reference period on the obtained spatial trend patterns below, whereas the particular question of possibly changing trends will be further addresses in Sect. 4.5. On the other hand, the spatial pattern could also result from uncertain estimates of the postglacial rebound rates in the considered GIA model. However, the information available to us does not allow further evaluating this possibility in this work.

The major conceptual advantage of QR in comparison with conventional methods of trend analysis for the mean is its ability to provide information on changes of the entire probability distribution of RSL. In order to highlight the differences, Fig. 4d–f shows the residual quantile trends relative to the corresponding trend in mean sea-level. For the 10 % quantile most slopes are consistently, but only very weakly negative, indicating that the trend in lower RSL quantiles due to global sea-level rise is somewhat less positive than the trend in the mean. For the 90 % quantile the slopes are positive in most of the Baltic with the exception of some tide gauges in the Archipelago Sea and the Gulf of Finland (where the associated uncertainties of the trend estimates are, however, rather large), indicating that upper RSL quantiles rise generally faster than the mean. Both observations together indicate that the total variability of RSL is increasing in the entire Baltic Sea, which can be attributed to a general intensification of atmospheric dynamics forcing short-term sea-level variability.

From the analysis presented so far, it has not yet been possible to draw systematic conclusions due to the different periods covered by the individual tide gauge records. In order to solve this problem, we next apply linear QR to all records completely covering the second half of the 20th century (1951–1999) without significant gaps, which are available for 30 tide gauges. In this case, the observed spatial pattern becomes more coherent. For the relative sea-level trends corrected for GIA effects (Fig. 5a–c), the upper quantiles show consistent positive trends in the entire Baltic Sea, whereas median and lower quantiles show negative relative trends in vast parts of the study area with the exception of the southwestern Baltic Sea (Poland, Germany, Denmark, southern Sweden) and northernmost Gulf of Bothnia, where also the high quantiles show the strongest positive trends. When considering the difference between the quantile slopes and the mean sea-level trend, we find consistently positive relative trends in the higher quantiles and negative ones in intermediate (median) and lower quantiles (Fig. 5d–f) with only few local exceptions. These results are in excellent agreement with those obtained by Barbosa (2008) for individual stations. Note, however, that these relative quantile trends have considerably lower magnitudes and higher standard errors than those in comparison with the mean GIA slope.

Since long-term sea-level variability in the Baltic is mainly determined by westerly winds, a plausible explanation for the observed trends, particularly in the upper quantiles, is the change in regional wind patterns and specifically the



**Fig. 4.** Slopes  $\beta_{0.1}$ ,  $\beta_{0.5}$  and  $\beta_{0.9}$  of the linear trends in the (A, D) 10 %, (B, E) 50 % and (C, F) 90 % quantiles of the 47 tide gauge records in the time period 1898–2002 (cf. Barbosa, 2008), corrected for the overall mean trends obtained from the GIA model (A–C) and trends in mean sea-level (D–F, including contributions from global sea-level changes, but not explicitly corrected for glacial isostatic adjustment processes). Colours indicate the slope values, sizes of the associated circles their standard errors in  $\text{mm dec}^{-1}$  (large circles: low uncertainty, small circles: high uncertainty). Note that the statistical confidence of trends corrected for linear trends in the mean is smaller, since the standard errors of estimated total quantile trends and trends in the mean add up, whereas we have implicitly assumed the absence of uncertainty in the GIA model as a simplification.

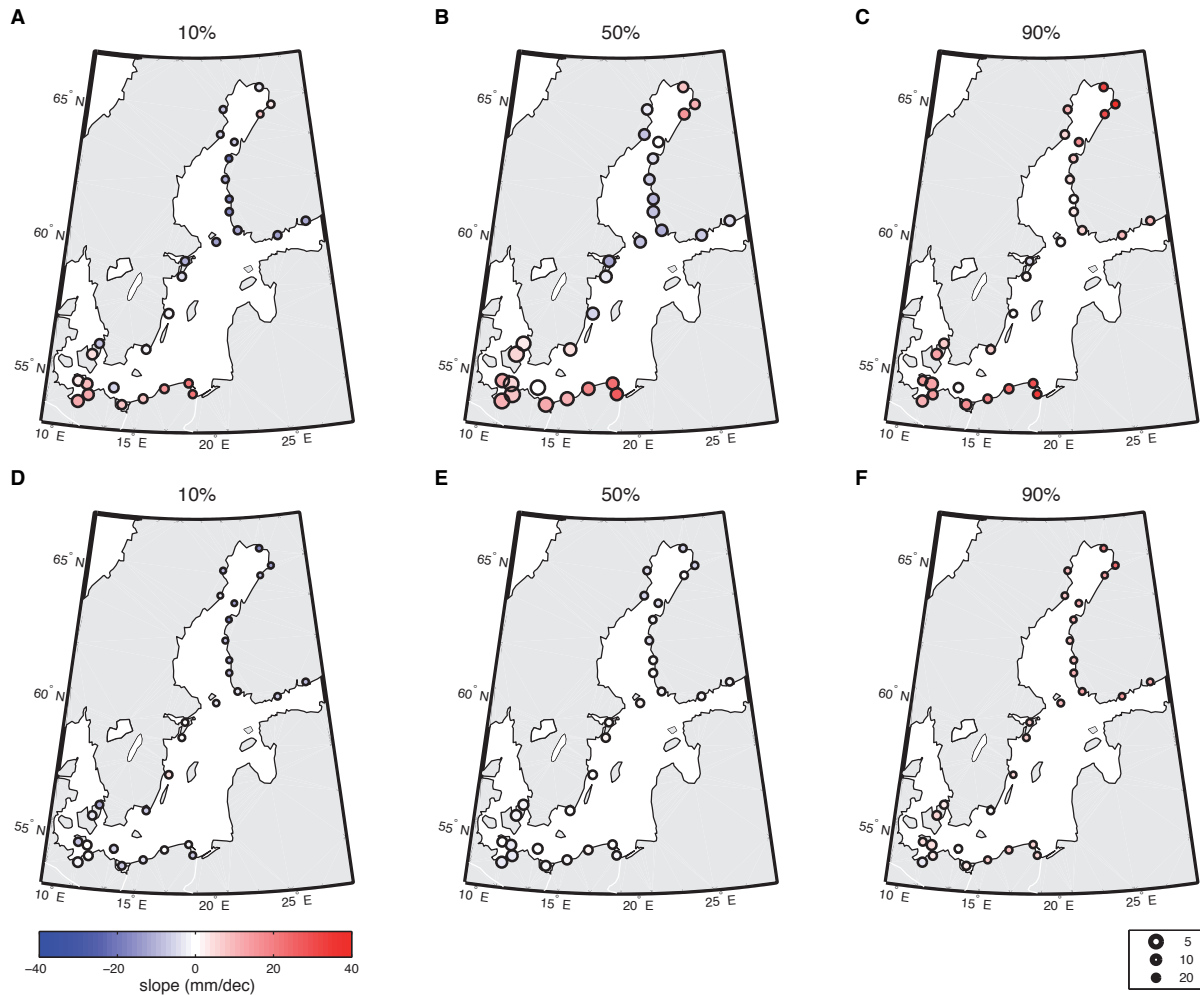
intensification of zonal wind and cyclones. In general, a detailed interpretation of the results at this point would be difficult and speculative, since sea-level is influenced by a multiplicity of different variables (wind, temperature,...) that are mutually interdependent and change over time in a complicated way. In turn, much more detailed future studies – specifically involving information on possible triggering factors as covariates – would be necessary to develop and test corresponding hypotheses.

As an intermediate summary, we conclude that (i) the quantile trends obtained from linear QR are distinctively different from trends in the mean, and that (ii) the GIA processes cannot explain the observed changes in Baltic sea-level quantiles.

### 4.3 Parameter selection for nonparametric QR

In order to systematically compare the results of linear QR presented above with those of a particular nonparametric QR method, a reasonable choice of the regularisation parameter  $\lambda$  has to be determined. In the following, we will illustrate this choice for some exemplary tide gauges. Subsequently, in Sect. 4.4 the resulting spatial trend patterns for selected quantiles will be compared with the outcomes of linear QR.

As discussed in Sect. 3.3, there are two widely accepted possibilities for determining proper values for  $\lambda$ , i.e. penalised-likelihood and cross-validation criteria. In case of nonparametric quantile regression, we request the solution of the underlying estimation problem (i) to be optimal and



**Fig. 5.** As in Fig. 4 for the 30 available tide gauge records for the time period 1951–1999.

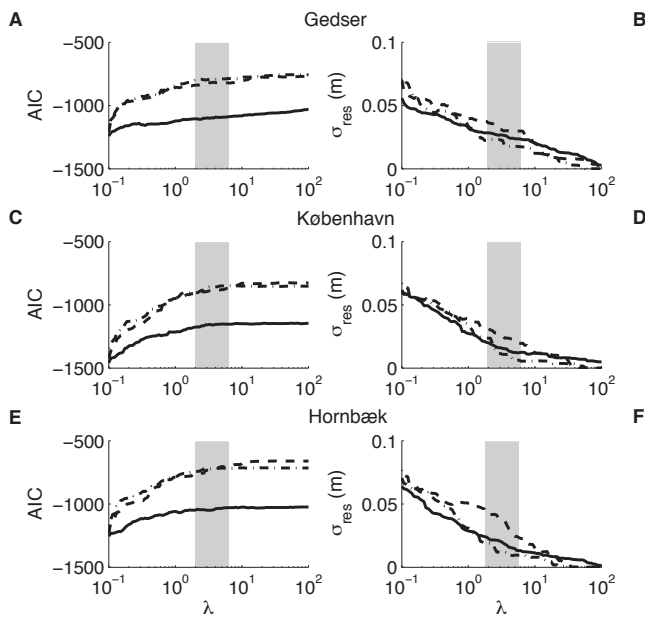
sparse in the sense of a high fidelity and a low number of parameters and (ii) not to differ “too much” from the linear QR model. The second requirement allows for moderate long-term variations in the instantaneous slope of the estimated quantile trends, but does not permit strong short-term fluctuations. While the first requirement is quantified by means of AIC or SIC, the second constraint is captured by the variance  $\sigma_{\text{res}}^2$  of the residual nonparametric trend model with respect to the linear one.

Although  $\lambda$  should be in principle selected independently for each quantile  $\alpha$ , an automatic parametric selection based on the optimisation of some individual criterion can lead to inconsistencies between the mean slopes obtained for different quantiles (cf. Sect. 4.1). As a consequence, we search for a reasonable trade-off between optimising AIC and keeping the deviations from the linear trend model in an acceptable range simultaneously for low, intermediate, and high quantiles. As an example, Fig. 6 shows the values of AIC and  $\sigma_{\text{res}}$  in dependence on  $\lambda$  for the 10%, 50%, and 90%

quantiles estimated with  $L_1$ -smoothing splines (using the R package `quantreg`) for three Danish tide gauges (Gedser, København, and Hornbæk). The grey bars highlight values of  $\lambda$  for which both requirements of high AIC and low residual variance are still fulfilled for all three considered quantiles (these criteria obviously remain valid for higher values of  $\lambda$  as well, however, we are seeking for a solution that allows for a maximum degree of variability with a still reasonable smoothness). From these three examples, we conclude that a value of  $\lambda = 4$  is a reasonable choice for the following analyses.

#### 4.4 Spatial patterns of nonparametric quantile trends

In order to understand the potential influence of nonlinearities in the quantile trends, the 30 selected stations previously analysed by means of linear QR have been subjected to an additional nonparametric QR using  $L_1$ -smoothing splines with a fixed  $\lambda = 4$  for all considered quantiles (see above). In Fig. 7, the average slopes of the nonparametric quantile



**Fig. 6.** Dependence of the AIC criterion (left panels) and the residual standard deviations  $\sigma_{res}$  with respect to the corresponding linear quantile models (right panels) on the smoothing parameter  $\lambda$  for the 10% (dashed), 50% (solid), and 90% (dash-dotted) quantile functions obtained for three example records from Denmark (time period 1951–1999) using  $L_1$ -smoothing (piecewise linear) splines. We propose that a reasonable choice for  $\lambda$  is obtained where AIC and  $\sigma_{res}$  start to significantly decrease and increase, respectively (indicated by grey bars).

trends taken over the entire period of observation are shown. From a conceptual perspective, using average quantile trends in situations when the actual changes of the probability distribution functions depend on both the considered quantile and time obviously leads to a loss of information. Even more, the possibility to interpret average trends in a meaningful way clearly depends on the temporary variations of the instantaneous trends. For this purpose, information on the latter aspect is encoded in Fig. 7 in terms of different symbol sizes.

Comparing the average slopes of the nonparametric trends with the linear quantile trends (Fig. 5), distinct differences are observed. Specifically, for the sea surface heights corrected for GIA, we find consistent positive trends in the southern Baltic Sea and the Gulf of Finland for all quantiles. In the central part of the Baltic Sea as well as most tide gauges in the Gulf of Bothnia, the relative trends with respect to the GIA model are positive for lower quantiles, but become negative for higher quantiles in the northern Gulf of Bothnia. Differences between linear and spline model are mainly found in the lower quantiles, where we observe negative trends along the Swedish and Finnish coastlines for the linear, but positive ones for the spline estimates. We have to emphasise, however, that especially for quantiles deviating strongly from the median, the results obtained with the

spline model show a considerable degree of sensitivity with respect to changes at both ends of the considered time series, e.g. the observed average quantile trends can change strongly if the records are extended by another year or so. This is due to the generally strong effect of boundary conditions on splines (Rice and Rosenblatt, 1983), which does not apply to linear spline models and is more severe for the more extreme quantiles, since they are statistically less well determined.

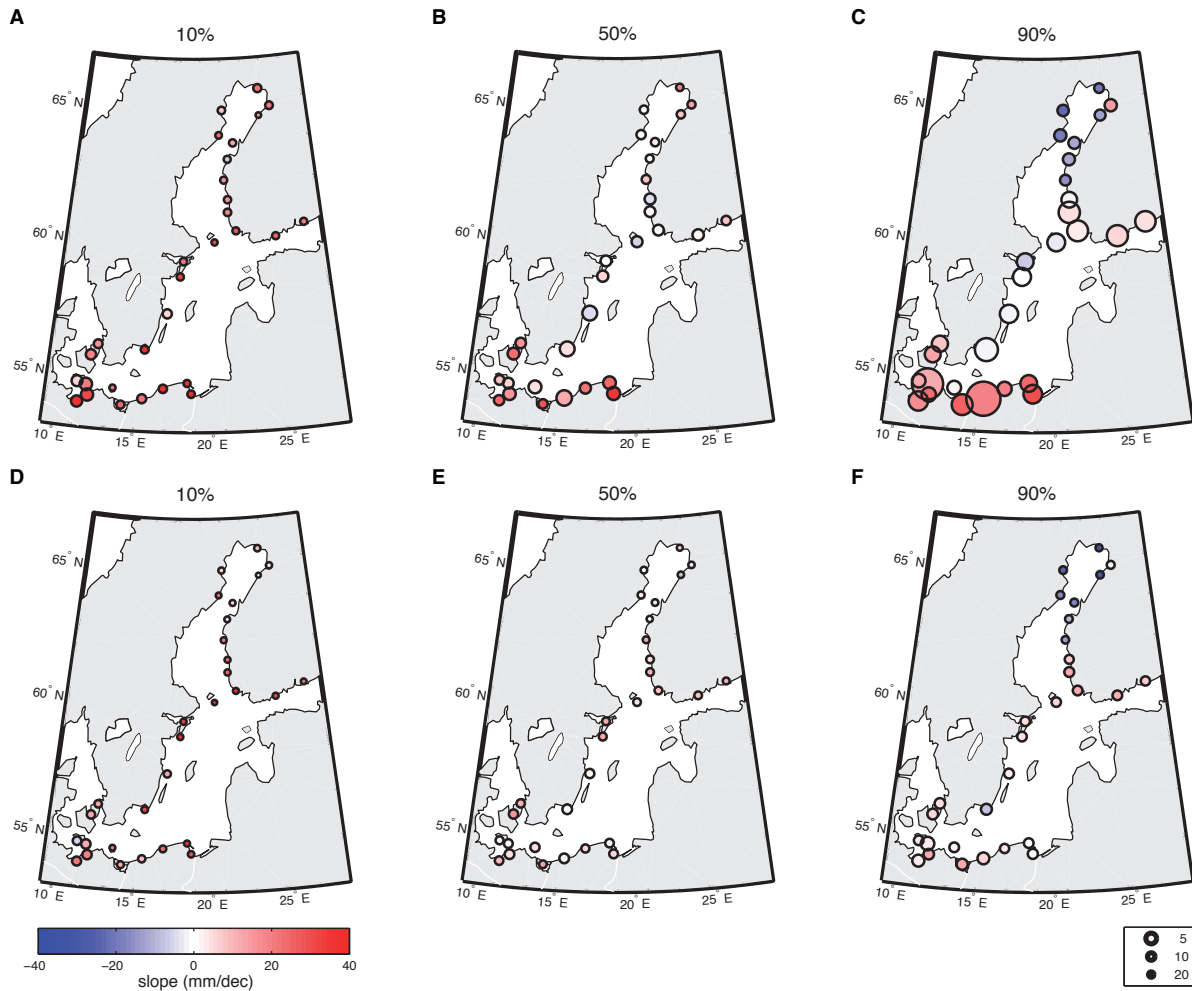
Taking the trends in mean sea-level instead of the mean GIA rates as a basis, the overall spatial pattern does not change qualitatively. Specifically, when comparing average trends from the spline model with the linear quantile trends, all results become more pronounced underlining the importance of the nonlinear characteristics for studying low-frequency sea-level variability in the Baltic Sea. In general, we have to note that the absolute differences between linear and *mean* nonparametric trends are not statistically significant when considering the standard errors of the estimates.

With respect to the results discussed above, we emphasise that we have only studied the behaviour of  $L_1$ -smoothing splines. From the present analysis, it cannot be ruled out that  $L_\infty$ -smoothing splines could display a somewhat different spatial pattern for selected quantiles. A detailed comparison of different approaches is beyond the scope of the present study, but should be performed in future work. In general, we note that estimating nonparametric quantile models is computationally more demanding than linear QR, which is reflected in the corresponding CPU times required for performing our analyses. With respect to the different algorithms used for spline QR, the algorithm implemented in the `cobs99` package is somewhat more efficient than that used in the `quantreg` package when run on the same hard- and software environment.

#### 4.5 Acceleration of local sea-level trends

The possible acceleration of global sea-level rise has recently attracted considerable interest (Woodworth, 1990; Douglas, 1992; Church and White, 2006; Jevrejeva et al., 2008; Woodworth et al., 2009; Merrifield et al., 2009; Houston and Dean, 2011; Rahmstorf and Vermeer, 2011). In order to complement our previous analysis, in the following, we study possible evidence for changes in sea-level quantile trends on the local scale. For this purpose, we estimate linear models for the quantile trends for running windows of width 10 years (Holgate, 2007) with a mutual offset of 1 yr. To assure comparability of our results, we again restrict our attention to the 30 tide gauge records that cover the time period 1951–1999 completely, yielding in total 40 linear trend values for each record. As examples, the temporal variability of local linear quantile trends for the three Danish tide gauges from Fig. 6 is shown in Fig. 8.

The observed temporal variability of the obtained trends shows significant changes that are different even for these spatially close locations. At the Gedser tide gauge (Fig. 8a)



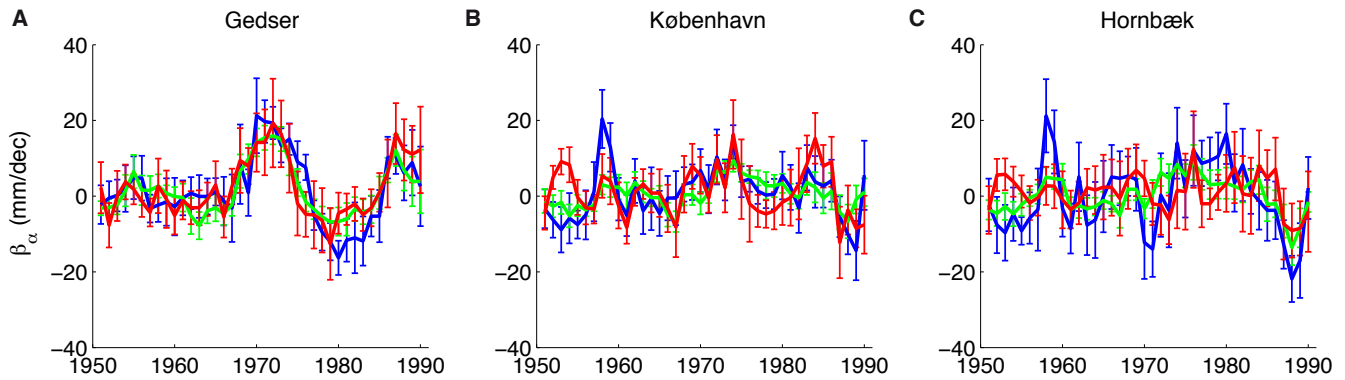
**Fig. 7.** As in Fig. 5 for the average slopes  $\bar{\beta}_{0.1}$ ,  $\bar{\beta}_{0.5}$  and  $\bar{\beta}_{0.9}$  of the nonparametric quantile trends ( $L_1$ -smoothing splines with  $\lambda = 4$ ). The size of the symbols represents the standard errors of the estimated mean trends, which are directly proportional to the standard deviations of instantaneous quantile trends.

there is a marked multi-decadal variability that is common to all three considered quantiles. København and Hornbæk also display decadal-scale changes in trends, though less pronounced, and a similar qualitative behaviour, particularly for the lower quantiles. A detailed analysis of these features and possible interpretations in terms of atmospheric or other forcing factors acting on decadal to multi-decadal scales (Hünicke and Zorita, 2006) will be subject of future research.

For statistically assessing whether or not quantile trends in sea-level variability have changed systematically over the last 50 yr, we have performed two different statistical tests. On the one hand, a (two-sided) t-test has been considered, which evaluates whether the time series of individual trends obtained for the 40 subsequent time windows display significant *linear* trends themselves. On the other hand, we have used a Mann-Kendall test for identifying *arbitrary monotonous* changes of these temporary trend values, which

are reflected by the presence of ordered structures within the associated time series of rank numbers significantly deviating from the expectations for randomly shuffled (disordered) ranks. Since the distribution of both test statistics is crucially influenced by the autodependence between subsequent temporary trend values<sup>2</sup> (i.e. due to the use of overlapping time windows), we apply bootstrapping by considering an AR[1] model fitted to the respective trend data and calculating the distributions of both test statistics for 1000 realisations of this model. We emphasise that AR[1] models are frequently used for estimating the significance of certain statistical properties (e.g. power or wavelet spectra) of geophysical records (Gilman et al., 1963; Torrence and Compo, 1998).

<sup>2</sup>It has been shown both analytically and numerically that deviations from the independence assumption can strongly change the distribution of the Mann-Kendall test statistics (Hamed and Rao, 1998; Hamed, 2009).



**Fig. 8.** Linear trends estimates and associated  $\pm 1$  standard error confidence intervals obtained for 10-yr running windows (step size 1 yr, the time axis denotes the start of each window) for the 10 % (blue), 50 % (green), and 90 % (red) sea-level quantiles observed at the tide gauges at (A) Gedser, (B) København, and (C) Hornbæk.

The results of our evaluations can be found in Fig. 9. For comparison, we also consider the significance of the supposed overall trends in the temporary linear quantile trends under the milder assumption of mutually independent data, which is, however, obviously violated in our case. Since the number of available time windows is rather low (40), we consider potential indications for accelerating sea-level changes on a 90 % significance level only. (The consideration of longer records might further increase a possible significance of the obtained results.) In this case, we observe significantly accelerating trends for the 10 % quantiles exclusively in central Sweden, and for the median in vast parts of the Baltic Sea. All these results are, however, only significant when neglecting the auto-dependence structure of the data. Comparing the empirical values of the test statistics with those obtained for AR[1] surrogates adjusted to the data, all accelerating trends in the 10 % and 50 % quantiles become insignificant with respect to both  $t$  and Mann-Kendall test.

For the 90 % quantiles, the spatial pattern is, however, distinctively different. Here, the only significant variations are found in the southwestern part of the Baltic Sea, where the  $t$ -test (but not the Mann-Kendall test) reveals three tide gauges with significant *deceleration* of trends, of which two (Warnemünde and Wismar) are even positive when considering the data as dependent, whereas there are no gauges with significant accelerations of the associated quantile trends. We emphasise, however, that the non-detection of significant accelerations or decelerations does not allow to conclude that long-term changes in trends are absent, since the detection of changes in slopes of a few  $\text{mm dec}^{-1}$  is very challenging in face of the still rather short length of the available records and the presence of substantial interannual and inter-decadal variability.

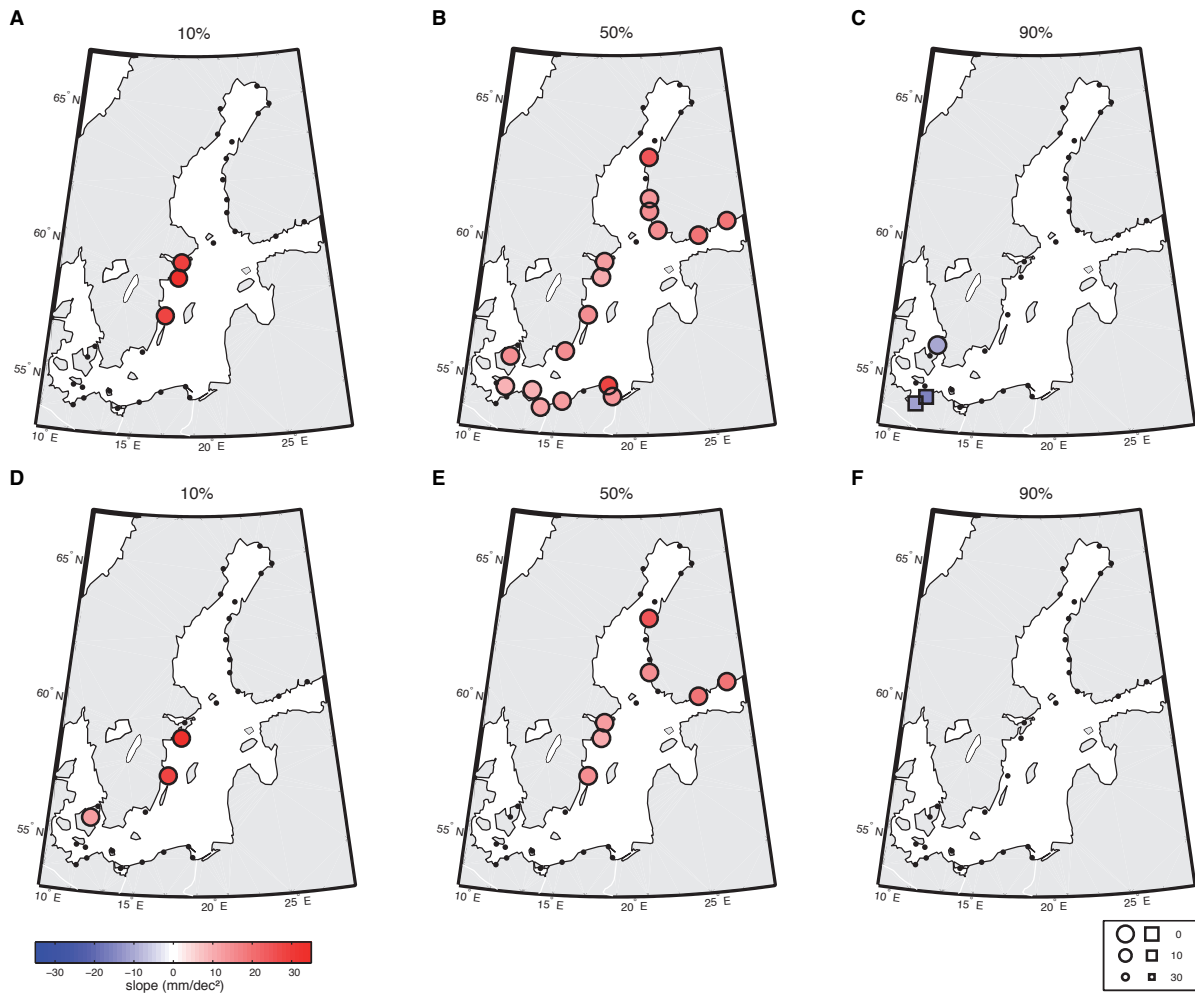
Regarding the reliability of the aforementioned results, one has to keep in mind that for the considered problem, it is not clear whether the intrinsic autodependences are sufficiently well described by the very simple stochastic model of

an AR[1] process. Furthermore, we note that there are more sophisticated bootstrapping approaches that could be alternatively used to obtain more reliable results (such as block or sieve bootstrap). In general, a similar analysis as done above for linear temporary trends could also be performed directly on the results of nonparametric QR. However, in this case, the problem of mutual dependence between subsequent temporary trends becomes even more obvious and possibly enhanced by the applied smoothing constraints. Because of this, we have restricted our attention here to the more easily interpretable piecewise linear quantile trends.

## 5 Conclusions

Sea-level variability is affected by many different meteorological and oceanographic processes over a wide range of both temporal and spatial scales. Therefore, even in a relatively small and confined sea such as the Baltic, long-term sea-level variability exhibits substantial spatial heterogeneity, reflecting the diverse local and regional processes having a relevant influence. The large dataset of tide gauge records considered in this study allows to characterise the spatial patterns of sea-level variability in the Baltic. By considering quantile instead of ordinary linear regression, a more complete description of trends is achieved, since the entire probability distribution of sea-level, rather than just the mean, is assessed. Furthermore, since the assumption of a monotonic linear trend is often unrealistic, nonparametric quantile regression is considered for a more flexible description of long-term sea-level variability in the Baltic area.

Sea-level in the Baltic exhibits an annual cycle with a maximum in winter. The estimation of trends from tide gauge records is traditionally performed after subtraction of the seasonal component. However, removing seasonality can impact the derived slopes. The results obtained in the present study allow to conclude that differences between lower and



**Fig. 9.** Spatial distribution of potentially significant accelerations (red) and decelerations (blue) of different local RSL quantiles obtained from the running 10-yr linear trend estimates for the time period 1951–1999 (30 tide gauges). Symbols represent tide gauges where the trends pass a t-test (two-sided 90 % significance level, upper panels A–C) and a Mann-Kendall test (one-sided 90 % significance level, lower panels D–F) for dependent (squares) and independent (circles) data, respectively. The colours of the symbols represent the linear trends in the piecewise linear quantile trends, their size the associated standard errors (in  $\text{mm dec}^{-2}$ ). Black dots indicate tide gauges with insignificant trend changes.

upper quantile trends are robust and that average long-term quantile trends are not affected qualitatively by the corresponding preprocessing of the analysed tide gauge records.

After correcting for postglacial rebound (i.e. land uplift/subsidence) processes by means of a glacial isostatic adjustment model, linear quantile regression reveals consistent spatial patterns for lower and median quantiles, with positive slopes in the southern and northernmost parts of the Baltic and negative trends in the central area. The former areas are the most dynamic regions, since sea-level is substantially affected by water exchange at the Baltic entrance and storms normally resulting in strongly deviating sea-levels at the “ends” of the Baltic (Ekman, 2007). The aforementioned general pattern relates exclusively to processes acting on the height of the sea surface. At most tide gauges, lower

sea-level quantiles change at less positive rates than mean sea-level when considering a linear trend model. In contrast, the upper quantiles increase faster than mean sea-level, indicating an overall increase in the dynamical range of sea-level in the second half of the 20th century. When relaxing the linearity assumption, the average trends typically exceed the linear trend of the mean even for the lower quantiles, pointing to the particular relevance of low-frequency modes for long-term sea-level variability.

The comparison of linear and nonparametric quantile trends reveals a somewhat different behavior: while relative to GIA processes, the linear sea-level trends are positive in the southern and northernmost part of the Baltic Sea and negative in the central part for all considered quantiles, the corresponding patterns for the average nonparametric quantile

trends differ for lower and upper quantiles. Strong linear trends for upper quantiles in the Bothnian Bay can be interpreted as resulting from changes in regional wind patterns and associated changes in storminess and wave set-up (Barbosa, 2008). A possible explanation for the apparently contradicting results obtained with nonparametric quantile regression is that the more flexible (but less robust) nonparametric approach is able to better capture inter-annual variability and the punctual character of storm events. Understanding the different results given by linear and nonparametric quantile regression requires further investigation, but the nonlinear (spline-based) framework already presents a promising technique for the study of sea-level variability based on the complete sea-level data distribution.

Regarding possible temporary changes in sea-level trends, both nonparametric and piecewise linear quantile regression reveal that there is indeed some decadal-scale acceleration and deceleration of local sea-level trends. A detailed analysis of the associated long-term variability will be subject of future research. In general, we emphasise that quantile regression is a promising tool for detecting and properly describing general long-term changes in the distribution of relative sea-level. However, the capabilities of this approach are intrinsically limited when considering changes in the extreme quantiles (in our case of monthly data, quantiles below 10 % and above 90 %), since the corresponding statistical confidence is not sufficient for drawing any far reaching conclusions. If one is interested in changes of distinct sea-level extremes, alternative methods such as time-dependent extreme value analysis are therefore more appropriate, but will probably require the consideration of at least daily instead of monthly data. In this respect, an important advantage of nonparametric quantile regression methods in comparison with methods from extreme value statistics is that the corresponding data requirements are typically more moderate.

*Acknowledgements.* This work has been financially supported by the DAAD/FCT project “Temporal changes in sea-level variability: from local to regional scales” (project ID 50750929), the Leibniz association (project ECONS), the FCT (contract under Programme Ciencia2008), the German National Academic Foundation (JFD), and the Federal Ministry for Education and Research (BMBF) via the Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability (PROGRESS).

Edited by: P. Miranda

Reviewed by: two anonymous referees

## References

- Barbosa, S. M.: Quantile trends in Baltic sea level, *Geophys. Res. Lett.*, 35, L22704, doi:10.1029/2008GL035182, 2008.
- Barbosa, S. M., Silva, M. E., and Fernandes, M. J.: Time series analysis of sea-level records: characterising long-term variability, Springer, Vol. 112 of Lecture Notes in Earth Sciences, 157–173, 2008.
- Barbosa, S. M., Scotto, M. G., and Alonso, A. M.: Summarising changes in air temperature over Central Europe by quantile regression and clustering, *Nat. Hazards Earth Syst. Sci.*, 11, 3227–3233, doi:10.5194/nhess-11-3227-2011, 2011.
- Baur, D., Saisana, M., and Schulze, N.: Modelling the effects of meteorological variables on ozone concentration – a quantile regression approach, *Atmos. Environ.*, 38, 4689–4699, 2004.
- Cannon, A. J.: Quantile regression neural networks: Implementation in R and application to precipitation downscaling, *Comput. Geosci.*, 37, 1277–1284, 2011.
- Cazenave, A. and Nerem, R. S.: Present-day sea level change: observations and causes, *Rev. Geophys.*, 42, RG3001, doi:10.1029/2003RG000139, 2004.
- Church, J. A. and White, N. J.: A 20th century acceleration in global sea-level rise, *Geophys. Res. Lett.*, 33, L01602, doi:10.1029/2005GL024826, 2006.
- Church, J. A., Woodworth, P. L., Aarup, T., and Wilson, W. S. (Eds.): *Understanding Sea-level Rise and Variability*, Wiley-Blackwell, Oxford, 2010.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, *J. Official Stati.*, 6, 3–33, 1990.
- Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, *J. Am. Stat. Assoc.*, 74, 829–836, 1979.
- Cleveland, W. S. and Devlin, S. J.: Locally weighted regression: An approach to regression analysis by local fitting, *J. Am. Stat. Assoc.*, 83, 596–610, 1988.
- Donner, R. V., Sakamoto, T., and Tanizuka, N.: Complexity of spatio-temporal correlations in Japanese air temperature records, Springer, Vol. 112 of Lecture Notes in Earth Sciences, 124–154, 2008.
- Douglas, B. C.: Global sea-level acceleration, *J. Geophys. Res.–Oceans*, 97, 12699–12706, 1992.
- Draghicescu, D. F.: Nonparametric quantile estimation for dependent data, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2002.
- Ekman, M.: A consistent map of the postglacial uplift of Fennoscandia, *Terra Nova*, 8, 158–165, 1996.
- Ekman, M.: A secular change in storm activity over the Baltic Sea detected through analysis of sea level data, *Small Publications in Historical Geophysics*, 16, 13 pp., 2007.
- Ekman, M.: The changing level of the Baltic Sea during 300 years: A clue to understanding the Earth, Summer Institute for Historical Geophysics, Godby, 2009.
- Elsner, J. B., Kossin, J. P., and Jagger, T. H.: The increasing intensity of the strongest tropical cyclones, *Nature*, 455, 92–95, 2008.
- Gilman, D., Fuglister, F., and Mitchell, J.: On the power spectrum of red noise, *J. Atmos. Sci.*, 20, 182–184, 1963.
- Hamed, K.: Exact distribution of the Mann-Kendall trend test statistic for persistent data, *J. Hydrol.*, 365, 86–94, 2009.
- Hamed, K. H. and Rao, A. R.: A modified Mann-Kendall trend test for autocorrelated data, *J. Hydrol.*, 204, 182–196, 1998.
- Härdle, W.: *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, 1990.
- He, X. and Ng, P.: COBS: Qualitatively constrained smoothing via linear programming, *Comput. Stat.*, 14, 315–337, 1999.
- Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orłowsky, B., and Stepanek, P.: Observational evidence for soil-moisture im-



- pact on hot extremes in southeastern Europe, *Nat. Geosci.*, 4, 17–21, 2011.
- Holgate, S. J.: On the decadal rates of sea level change during the twentieth century, *Geophys. Res. Lett.*, 34, L0160, doi:10.1029/2006GL028492, 2007.
- Houston, J. R. and Dean, R. G.: Sea-level acceleration based on US tide gauges and extensions of previous global-gauge analyses, *J. Coast. Res.*, 27, 409–417, 2011.
- Hünicke, B. and Zorita, E.: Influence of temperature and precipitation on decadal Baltic Sea level variations in the 20th century, *Tellus*, 58A, 141–153, 2006.
- Hünicke, B. and Zorita, E.: Trends in the amplitude of Baltic Sea level annual cycle, *Tellus*, 60A, 154–164, 2008.
- IPCC: *Climate Change 2007: The Physical Science Basis*, Cambridge University Press, Cambridge, 2007.
- Jagger, T. H. and Elsner, J. B.: Modeling tropical cyclone intensity with quantile regression, *Int. J. Climatol.*, 29, 1351–1361, 2009.
- Jevrejeva, S., Moore, J. C., Woodworth, P. L., and Grinsted, A.: Influence of large-scale atmospheric circulation on European sea level: results based on the wavelet transform method, *Tellus*, 57A, 183–193, 2005.
- Jevrejeva, S., Moore, J. C., Grinsted, A., and Woodworth, P. L.: Recent global sea level acceleration started over 200 years ago?, *Geophys. Res. Lett.*, 35, L08715, doi:10.1029/2008GL033611, 2008.
- Koenker, R.: *Quantile Regression*, Cambridge University Press, Cambridge, vol. 38 of *Econometric Society Monographs*, 2005.
- Koenker, R. and Ng, P.: Inequality constrained quantile regression, *Sankhya*, *Indian J. Statistics*, 67, 418–440, 2005.
- Koenker, R. and Schorfheide, F.: Quantile spline models for global temperature change, *Climatic Change*, 28, 395–404, 1994.
- Koenker, R., Ng, P., and Portnoy, S.: Quantile smoothing splines, *Biometrika*, 81, 673–680, 1994.
- Kyselý, J., Pícek, J., and Beranová, R.: Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold, *Global Planet. Change*, 72, 55–68, 2010.
- Li, Y., Liu, Y., and Zhu, J.: Quantile regression in reproducing kernel Hilbert spaces, *J. Am. Stat. Assoc.*, 102, 255–268, 2007.
- Merrifield, M. A., Merrifield, S. T., and Mitchum, G. T.: An anomalous recent acceleration of global sea level rise, *J. Climate*, 22, 5772–5781, 2009.
- Milne, G. A., Gehrels, W. R., Hughes, C. W., and Tamisiea, M. E.: Identifying the causes of sea-level change, *Nat. Geosci.*, 2, 471–476, 2009.
- Ng, P. and Maechler, M.: A fast and efficient implementation of qualitatively constrained quantile smoothing splines, *Stat. Model.*, 7, 315–328, 2007.
- Nychka, D., Gray, G., Haaland, P., Martin, D., and O’Connell, M.: A nonparametric regression approach to syringe grading for quality improvement, *J. Am. Stat. Assoc.*, 90, 1171–1178, 1995.
- Park, J., Obeysekera, J., Irizarry-Ortiz, M., Barnes, J., and Park-Said, W.: Climate links and variability of extreme sea-level events at Key West, Pensacola, and Mayport, Florida, *Journal of Waterway, Port, Coast. Ocean Eng.*, 136, 350–356, 2010.
- Peltier, W. R.: Postglacial variations in the level of the sea: Implications for climate dynamics and solid-earth geophysics, *Rev. Geophys.*, 36, 603–689, 1998.
- Peltier, W. R.: Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G(VM2) model and GRACE, *Ann. Rev. Earth Planet. Sci.*, 33, 111–149, 2004.
- Rahmstorf, S. and Vermeer, M.: Discussion of: Houston, J. R. and Dean, R. G., 2011, Sea-level acceleration based on US tide gauges and extensions of previous global-gauge analyses, *Journal of Coastal Research*, 27(3), 409–417, *J. Coast. Res.*, 27, 784–787, 2011.
- Rice, J. and Rosenblatt, M.: Smoothing splines: regression, derivatives and deconvolution, *Ann. Stat.*, 11, 141–156, 1983.
- Samuelsson, M. and Stigebrandt, A.: Main characteristics of the long-term sea level variability in the Baltic sea, *Tellus*, 48A, 672–683, 1996.
- Takeuchi, I., Le, Q. V., Sears, T. D., and Smola, A. J.: Nonparametric quantile estimation, *J. Machine Learning Res.*, 7, 1231–1264, 2006.
- Thompson, P., Cai, Y., Moyeed, R., Reeve, D., and Stander, J.: Bayesian nonparametric quantile regression using splines, *Comput. Stat. Data An.*, 54, 1138–1150, 2010.
- Timofeev, A. A. and Sterin, A. M.: Using the quantile regression method to analyze changes in climate characteristics, *Russ. Meteorol. Hydrol.*, 35, 310–319, 2010.
- Torrence, C. and Compo, G.: A practical guide to wavelet analysis, *B. Am. Meteorol. Soc.*, 79, 61–78, 1998.
- Woodworth, P. L.: A search for accelerations in records of European mean sea-level, *Int. J. Climatol.*, 10, 129–143, 1990.
- Woodworth, P. L. and Player, R.: The Permanent Service for Mean Sea Level: an update to the 21st century, *J. Coast. Res.*, 19, 287–295, 2003.
- Woodworth, P. L., White, N. J., Jevrejeva, S., Holgate, S. J., Church, J. A., and Gehrels, W. R.: Evidence for the accelerations of sea level on multi-decade and century timescales, *Int. J. Climatol.*, 29, 777–789, 2009.
- Yu, K. and Jones, M. C.: Local linear quantile regression, *J. Am. Stat. Assoc.*, 93, 228–237, 1998.
- Yu, K., Lu, Z., and Stander, J.: Quantile regression: applications and current research areas, *The Statistician*, 52, 331–350, 2003.
- Yuan, M.: GACV for quantile smoothing splines, *Comput. Stat. Data An.*, 50, 813–829, 2006.
- Zhou, Z. and Wu, W. B.: Local linear quantile estimation for non-stationary time series, *Ann. Stat.*, 37, 2696–2729, 2009.