https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

Technical note: Extending sea level time series for extremes analysis with machine learning and neighbouring station data

Kévin Dubois[1,2], Morten Andreas Dahl Larsen[3,4], Martin Drews[3], Erik Nilsson[1,2], Anna Rutgersson[1,2]

[1]Department of Earth Sciences, Uppsala University, Uppsala, 752 36, Sweden
[2]Centre of Natural Hazards and Disaster Science (CNDS), Uppsala University, Uppsala, 752 36, Sweden
[3]Department of Technology, Management and Economics, Technical University of Denmark, Lyngby, 2800, Denmark
[4]Danish Meteorological Office, Copenhagen, 2100, Denmark

Correspondence to: Kévin Dubois (kevin.dubois@geo.uu.se)

Abstract.

Extreme sea levels may cause damage and disruption of activities in coastal areas. Thus, predicting extreme sea levels is essential for coastal management. Statistical inference of robust return level estimates critically depends on the length and quality of the observed time series. Here we compare two different methods for extending a very short (~10 years) time series of tide gauge measurements using a longer time series from a neighbouring tide gauge: Linear Regression and Quantile Regression Forest machine learning. Both methods are applied to stations located in the Kattegat basin between Denmark and Sweden. Reasonable results are obtained using both techniques with the machine learning method providing a better reconstruction of the observed extremes. Generating a set of stochastic time series reflecting uncertainty estimates from the machine learning model and subsequently estimating the corresponding return levels using extreme value theory, the spread of the return levels is found to agree with results derived from more physically-based methods.

1 Introduction

Extreme sea levels (ESL) can have disastrous consequences in the coastal zone in terms of flooding vulnerable assets, loss of lives, and disturbances (Brown et al., 2018; Vousdoukas et al., 2020; Wahl et al., 2017). Coastal floods generally result from a combination of ESLs, wind, waves, tides and local conditions, including bathymetry and terrain features. Climate change also affects ESL events due to sea level rise and changes in storm frequency and/or intensity particularly (Rutgersson et al., 2021). Reliable estimates of current and future ESLs are urgently needed to mitigate the impacts of disaster risks and to inform adaptation to climate change. Long time series of observed sea levels are essential for improving confidence in statistically inferred return levels (RLs) (Menéndez et al., 2010; Woodworth et al., 2011) and are often considered essential for coastal planning. International initiatives such as the Global Sea Level Observing System (GLOSS) (Caldwell et al., 2012; Merrifield et al., 2012) and other works (Woodworth et al., 2010) have highlighted this necessity and called for recovering historical records in what is called "data archaeology". Still, the temporal paucity of sea level time series (Holgate et al., 2013) remains a limitation for adequately estimating RLs and ESLs in many places.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

This technical note evaluates a machine learning (ML) method for extending the sea level time series obtained by a tide gauge of interest using a longer time series at a neighbouring tide gauge in the context of analysing sea level extremes. This is particularly relevant when the initial time series is very short e.g., in the order of ~10 to 20 years, which is principally too short to allow reliable statistical inference of ESLs e.g., return levels corresponding to a 100-year event. The ML methodology is compared to a linear regression (LR) model, which for short time series could also be expected to perform adequately. Our study area lies within the Kattegat basin, located on the west coast of Sweden around the city of Halmstad (Fig. 1). Here the highest recorded Swedish sea level of 235 cm was observed in November 2015 according to the Swedish Meteorological and Hydrological Institute (SMHI). In this area, tides vary with an amplitude of around 20 cm during spring tides, and current ESLs are mainly due to storm surge effects. But also, other factors could play a role, such as the preconditioning of the Baltic Sea (Andrée et al., 2022). Hieronymus et al. (2020) have shown that the Swedish West Coast is expected to be one of Sweden's most exposed areas due to rising sea levels.

Different methods have been proposed for extending sea level records, such as Bernier et al. (2007), who use short observation time series associated with a 40-year hindcast surge model. Reconstructions by Cid et al. (2018) are based on tide gauge data and atmospheric conditions. Hieronymus et al. (2019) show a good performance of neural networks in predicting sea levels at tide gauges located along the Swedish coast based on different atmospheric variables and tide gauge records. Granata et al. (2021) find similar results when forecasting tides in the Venice region using different ML methods: Random Forest, Regression Tree and Multilayer Perceptron. Recently, Bellinghausen et al. (2023) have demonstrated the utility of using a Random Forest Classifier to satisfactorily predict the occurrence of ESLs at a few stations around the Baltic Sea within three days based on surface wind and pressure fields, precipitation and the prefilling state of the Baltic Sea.

In the following, we systematically evaluate the performance of Quantile Regression Forests (QRF) (Meinshausen, 2006) as means of extending a very short time series of only ten years, reconstructing past sea level variations based on a more extended time series from a neighbouring station. This approach is compared to the linear regression approach. Both methods have previously been found to reduce biases efficiently and are relatively computationally inexpensive with low complexity, when applied to a small number of input variables as is the case in this study. To evaluate the sensitivity of the reconstructed sea level with respect to the geographic distance to neighbouring stations, we apply the method to data from different stations. Finally, we consider the method's potential and limitations with respect to the reconstruction of sea level extremes, when the time series of interest is very short, and inherently provide a poor sampling of even moderately extreme events.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

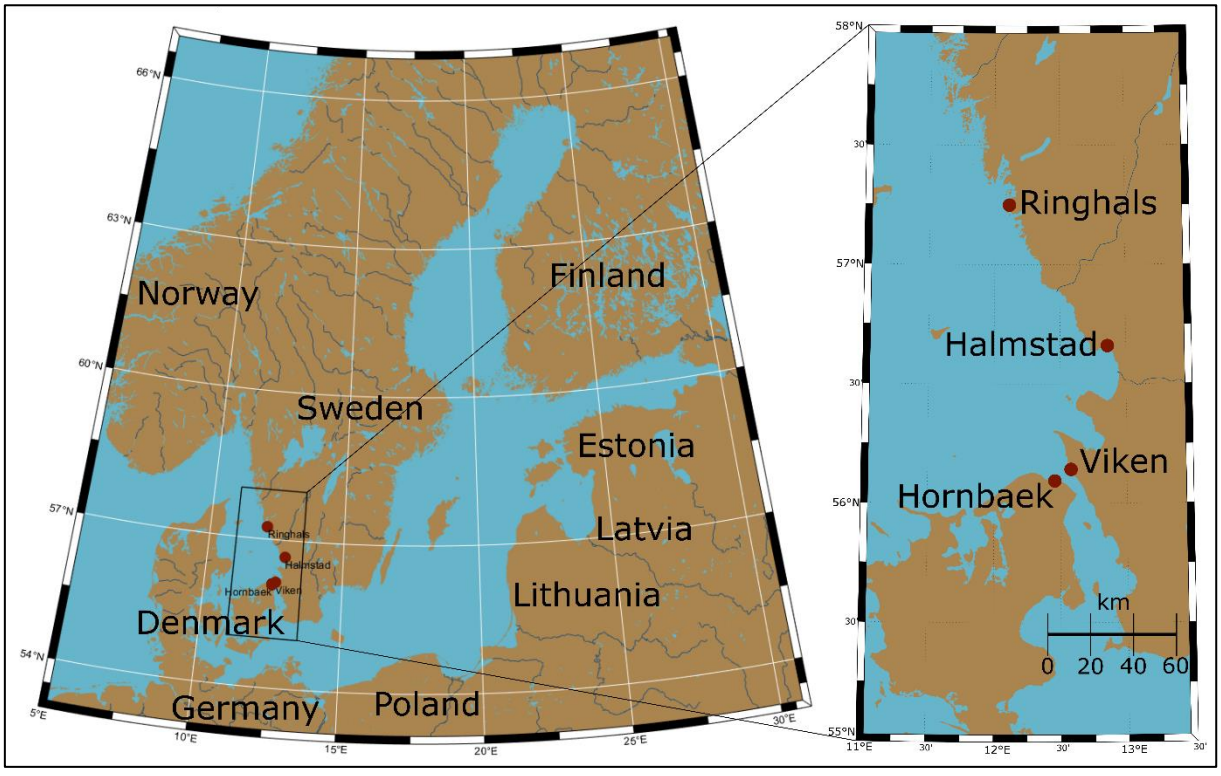

Figure 1: Map of Northern Europe with indication of the study area and the tide gauge stations (red dots).

2 Data and Methods

2.1 Sea level data

The datasets used in the analysis are hourly sea level observations from different stations available from SMHI (SMHI, 2022) and the Danish Meteorological Institute (DMI). Three stations are located on the west coast of Sweden: Ringhals (station number: 2105 "RINGHALS"), Halmstad (station number: 35115 "HALMSTAD SJÖV") and Viken (station number: 2228 "VIKEN"), and one station is located on the east coast of Denmark: Hornbaek (Hansen, 2007) (Fig. 1). The distance between Hornbaek and Viken is around 9 km, around 130 km between Hornbaek and Ringhals and 127 km between Viken and Ringhals.

Each hourly time series is first transformed into a time series of daily maxima from which the annual (daily) maximum is determined for each year in the series. When determining the annual daily maximum, we enforce a minimum temporal separation of 2 days to assert the independence of events at each station. The data sets are of varying lengths (Fig. 2) ranging from twelve years (Halmstad station) to 129 years (Hornbaek station). We assume that the time series are stationary at the decadal scale. Conversely, long-term linear trends (i.e., sea level rise) were estimated for all time series and found to range between 0.34 and 1.47 cm per decade.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

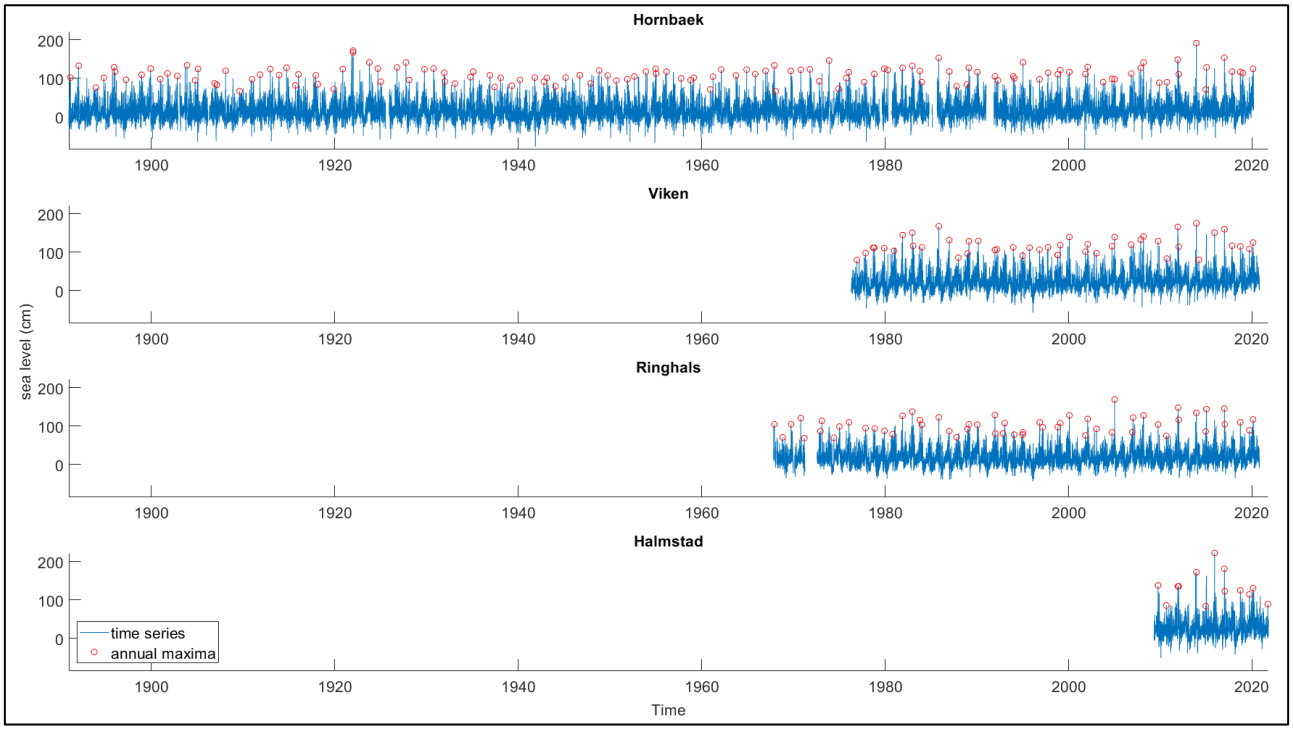

Figure 2: Sea level time series from the four tide gauge stations, i.e., daily maximum values (blue) and yearly maximum values (red circles).

2.2 Methods

The proposed approach for extending short sea level time series uses one neighbouring station as predictor (station x) of past sea level data at the station of interest (station y). For each x/y station pair (Table 1), we use ten years of temporally coinciding time series to set up each prediction model. We will denote this the setup period. Within the setup period, models are fitted or trained using both time series for eight years of overlapping data (training dataset), whereas the last two years serve as validation dataset to provide an unbiased evaluation of the model fit. The coinciding period outside the 10-year setup period constitutes the testing period. Two overall predictor approaches are employed – one using simple LR and one based on QRF machine learning.

2.2.1 Linear Regression

Based on each x/y predictor-reconstruction station pair, a linear equation is found using the least squares method as means of determining the best fit coefficients. Based on the resulting equation, the sea level at station y is predicted from the sea level at station x.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

2.2.2 Quantile Regression Forest

A probabilistic ML model is trained using the sea level at one station as the predictor and the sea level at another station as the predictand. The QRF method yields a mean and a standard deviation for each predicted value (Breiman, 2001; Meinshausen, 2006). The QRF model is implemented using the MATLAB function TreeBagger where the regression method is based on a number of trees and minimum leaf size hyper-parameters. These parameters are here set to 500 and 1, respectively.

The LR and the QRF models are trained using the same setup period for each station pair (Table 1).

2.2.3 Model Testing

To evaluate the proposed methodology, different analyses with different combinations of stations are used to test the spatial sensitivity (Table 1). Six analyses using different combinations of station data obtained at Hornbaek, Viken and Ringhals are carried out using the recent 10 full years (2010-2020) as the common setup period for model training and validation (cf. section 2.2.2). Two additional analyses are carried out, where we predict Hornbaek sea levels from Viken data using the two previous time periods (2000-2010) and (1990-2000) for training and validation to evaluate the temporal sensitivity. Finally, we compare reconstructed sea levels at Halmstad using the station data from Hornbaek, Viken and Ringhals, respectively (cf. section 3.2) for the period (2010-2020). In the latter case, we also estimate RLs based on the reconstructed time series and compare them to previous results reported for Halmstad.

Table 1: Experimental setup and summary of analyses.

Predictor	Predictand	Setup period			Coinciding period	Study
station x	station y	1	2	3		
Hornbaek	Viken	2010-2020			1977-2020	Spatial correlation analysis
Ringhals	Viken	2010-2020			1977-2020	
Hornbaek	Ringhals	2010-2020			1968-2020	
Viken	Ringhals	2010-2020			1977-2020	
Ringhals	Hornbaek	2010-2020			1968-2020	
Viken	Hornbaek	2010-2020			1977-2020	
Viken	Hornbaek		2000-2010		1977-2020	Temporal correlation analysis
Viken	Hornbaek			1990-2000	1977-2020	
Hornbaek	Halmstad	2010-2020			2010-2020	case study
Viken	Halmstad	2010-2020			2010-2020	

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

Ringhals	Halmstad	2010-2020			2010-2020	

To assess the performance of each model, different goodness-of-fit metrics (GOFs) are chosen i.e., the root mean square error (RMSE) and the Pearson correlation coefficient (*r*). Also, the general bias (*bias*) and the 95th percentile bias (*perc95-bias*) between the observations and both model reconstructions (LR and QRF) within the validation period are calculated.

To evaluate the model's performance towards the extremes, annual maxima and values above the 95th, 97th and 99th percentiles from observations are compared with the corresponding predicted values.

2.2.4 QRF method with random sampling to evaluate return levels (RLs)

The QRF estimates the standard deviation associated with the predicted daily maximum at each time point. Based on the QRF daily means and standard deviations, we calculate the corresponding annual maxima from the reproduced time series and their associated standard deviations. We assume that a Gaussian distribution describes the probability for each predicted annual maximum. In this light, 10 000 sets of 30-year maxima are randomly sampled with repeatability from the Gaussian distributions. RLs are subsequently calculated using a generalised extreme value (GEV) distribution fitted to the annual maxima (Coles, 2001). This yields an ensemble of randomly drawn RL curves. The 95th percentile of the ensemble spread is calculated.

3 Results and discussion

3.1 Model validation

To validate the models, GOFs are calculated (Table 2). For the time series of daily maxima, roughly similar statistics are found for all datasets whether using the QRF or LR. In general, we find slightly (but not significantly) better *r* and RMSE values associated with the LR and slightly better *perc95-bias* for the QRF (Table 2). For the annual maxima, the 95th, 97th, and 99th percentiles sets, marginally higher *r* and lower RMSE values are found for the LR in nearly all cases, with a maximum difference of 6 cm for the RMSE and 0.10 for the *r* value. Overall, RMSE values are between 10 and 40 cm, and *r* values are between 0.6 and 0.9 in most cases. For extremes, represented by the high percentile datasets, bias values range from -30 to -5 cm i.e., a slight underestimation of the observed extreme values for both the LR and QRF. As shown in Table 2, biases are similar with a maximum difference between models of ~5 cm in almost all cases. Figure 3 depicts the correlation between observations and models in predicting Viken sea levels from the Hornbaek data trained on the 2010-2018 period. A similar picture is observed in nearly all cases. As shown in Fig. 3, the QRF model returns significantly higher sea levels and shows higher variability towards the most extreme range than the LR, except when predicting Hornbaek sea levels from Viken data, where the model is trained on observations from the 1990-2000 and 2000-2010 setup periods (not shown). In those two cases, the QRF does not correctly reproduce the extreme range, as they are out-of-sample while the predicted values are bounded,

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

since the ML can only reproduce in-sample events. Compared to an LR, it is clear that the inherently non-linear QRF is better able to account for the few moderate extremes that occur during the 8-year training period, whereas they are likely to be suppressed in a linearized model.

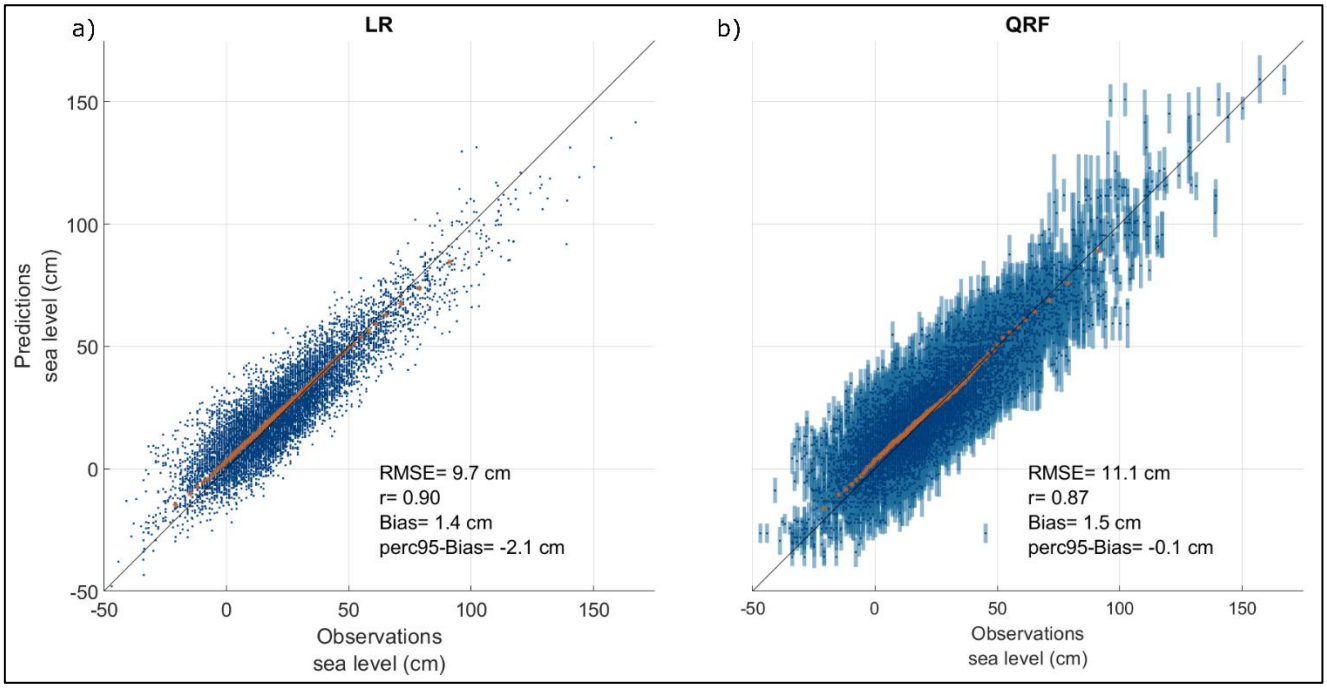

Figure 3: Scatter plot between observations and LR (a) and QRF (b) for predicting Viken sea level (station y) from the Hornbaek tide gauge (station x). Blue points show the daily maxima and each corresponding blue bar shows the standard deviation (b). Coloured stars correspond to the 1st to 99th percentile values of the dataset.

Table 2 analyses the sensitivity to the distance between the two tide gauges for the 2010-2020 period. When the distance between two stations grows, the model accuracy clearly decreases. For example, RMSE values of around 10 to 20 cm and *r* values of around 0.7 to 0.9 are found for the set reproducing Viken sea level from the Hornbaek tide gauge (9 km apart) compared to RMSE values of around 10 to 40 cm and *r* coefficients around 0.3 to 0.8 when reproducing Ringhals sea levels from Hornbaek data (130 km apart) (Table 2). Higher differences are also observed for the extremes (annual maxima or 95th, 97th and 99th percentiles). Similar results are found when comparing the sea level time series for Hornbaek and Ringhals, based on Viken data, and when comparing predictions for Viken and Hornbaek sea levels based on Ringhals data. Comparing results from mirrored sets, for example, when predicting Viken sea levels from Ringhals data or predicting Ringhals sea levels from Viken data, we do not always find the same performance as measured by the GOFs. This can probably be explained on physical grounds however, this is beyond the current technical note. In general, the QRF method seems to be more accurate than the LR when predicting local sea levels from stations located further away e.g., between Ringhals and Viken / Hornbaek as

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

compared to Viken and Hornbaek. This is also true for the extremes, where it is essential to capture the non-linear behaviour and variability associated with the complex natural interactions between drivers of ESL events. These are likely to become more prominent when observations are obtained at sites further away. Conversely, an LR is inherently constrained by a linearity assumption.

Table 2: Pearson correlation coefficients (*r*) and biases between different datasets evaluated in the validation period(s). Noticeable improvements (>5 cm) in terms of model bias with respect to annual maxima using the ML model are highlighted in bold. A negative bias (blue background) corresponds to an underestimation of the predicted values, and a positive one (red background) an overestimation. Error metrics calculated over the testing period for the case study of Halmstad city are displayed with a grey background. Because of the short length of the testing period, we do not calculate the bias on the annual maxima.

Predictor station x	Predictand station y	Setup period	*r* on daily maxima		*Bias* on daily maxima (cm)		*Bias* on annual maxima (cm)	
			LR	*QRF*	*LR*	*QRF*	*LR*	*QRF*
Hornbaek	Viken	2010-2020	0.90	0.87	1.4	1.5	**-14.0**	**-5.0**
Ringhals	Viken	2010-2020	0.91	0.88	1.9	1.2	**-25.4**	**-18.3**
Hornbaek	Ringhals	2010-2020	0.83	0.77	-0.9	-0.9	-17.0	-18.2
Viken	Ringhals	2010-2020	0.91	0.89	-1.7	-1.7	-8.1	-11.7
Ringhals	Hornbaek	2010-2020	0.83	0.78	1.1	0.4	**-30.5**	**-25.4**
Viken	Hornbaek	2010-2020	0.90	0.87	-1.3	-1.1	-7.7	-7.5
Viken	Hornbaek	2000-2010	0.91	0.89	0.6	1.4	-12.4	-17.8
Viken	Hornbaek	1990-2000	0.91	0.90	-1.6	-1.8	-12.1	-15.0
Hornbaek	Halmstad	2010-2020	0.93	0.88	1.0	1.0		
Viken	Halmstad	2010-2020	0.97	0.95	1.0	0.9		
Ringhals	Halmstad	2010-2020	0.95	0.93	0.4	0.3		

3.2 Halmstad

The highest sea level recorded in Sweden occurred in Halmstad, indicating that Halmstad is highly susceptible to ESLs. However, the length of the local sea level time series is very short. Subsequently, the three stations: Hornbaek, Viken and Ringhals, are used for reconstructing the Halmstad sea level time series (Table 2). As shown above, using a QRF or LR method, we can in principle reconstruct Halmstad sea levels back until 1891 for the period before observations became available in 2009 with reasonable confidence, using the station Hornbaek as a predictor, since this has the longest observed time series. It is worth noting that since we use observed tide gauge data, long-term trends, that is, climate change induced sea level rise are implicitly considered, although site-specific changes in the relative sea levels due to, e.g., human activities may introduce biases.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

Because of the short length of Halmstad's time series, the setup period is almost identical to the full time series which in practice makes it difficult to assess the model behaviour on extremes. When analysing the error metrics over the testing period, the model based on Viken station presents the best results. However, all sites present similar results for QRF predictions from Hornbaek and Ringhals, respectively (Table 2). When comparing the QRF method to LR, slightly better RMSE and *r* values are found for the LR, but when looking at higher percentile levels, the QRF results in higher corresponding values than for LR in all sets (not shown).

In previous studies, Halmstad's RLs were calculated for current and future climate scenarios based on reconstructed sea levels from local wind speed observations and Viken tide gauge data (Andersson, 2021). For Halmstad, RLs based on extended time series using the three neighbouring stations permit a reduction of the 95th percentile confidence interval (CI) compared with observations. Here, observed values are added to the extended time series to get the longest time series possible before a GEV fit is applied. Even so, RLs are still lower than the ones displayed by Andersson (2001) when based on the ML mean outputs (fig. 4-a; Table 3).

Table 3: Halmstad's RLs from reconstructed time series using the outputs from the QRF and the QRF method with random sampling applied from the station Hornbaek (blue background) compared with assessment by Andersson (2001) in grey background.

Station x	**5-year RL (m)**	**10-year RL (m)**	**50-year RL (m)**	**100-year RL (m)**	**200-year RL (m)**	
Hornbaek	1.6	1.8	2.1	2.2	2.3	
Ringhals	1.6	1.8	2.3	2.5	2.7	
Viken	1.6	1.7	2.0	2.1	2.2	QRF
Viken	1.8	2.0	2.4	2.6	2.8	Andersson
uncertainties	1.5 - 2.0	1.7 - 2.3	2.0 - 2.8	2.1 - 3.0	2.3 - 3.3	(2021)
Hornbaek	*1.7*	*1.9*	*2.4*	*2.6*	*2.8*	*QRF method with random sampling*
95th percentile ensemble spread	*1.6 - 1.8*	*1.8 - 2.0*	*2.2 - 2.7*	*2.3 - 3.0*	*2.5 - 3.3*	

Conversely, we apply a QRF-based random sampling to predict RLs probabilistically as described in section 2.2.4 (fig. 4-b), at Hornbaek, Viken and Ringhals (which permits extended time series of around 120 years, 35 years and 45 years, respectively). As would be expected due to the long time series, estimates based on Hornbaek data deliver the best performance and yield what seems like a reasonable 95th percentile ensemble spread (Fig. 4). The inferred RLs are slightly higher than the RLs derived directly from observations, which are associated with a very large 95th percentile CI due to the short length of the time series.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

The predictions using Viken data present the lowest RLs with a 95th percentile ensemble spread (upper values) corresponding almost to the median RLs from observations probably underestimating the extremes. On the other hand, predictions from Ringhals result in the highest RLs but like Viken are also associated with a rather large ensemble spread. Because of the lengths of the respective time series, there is low confidence in return periods of rare occurrences such as a 200-year event (a little less pronounced for Hornbaek-based predictions). This challenge of rare occurrences is evident when looking at the 95th percentile CIs for each RL curve resulting from the QRF method with random sampling. For Halmstad, RLs based on inputs from the Hornbaek station following the QRF method with random sampling are close to the ones displayed by Andersson (2001), highlighting the importance of considering the full uncertainty range when predicting high sea levels from a small sample of such events (Table 3).

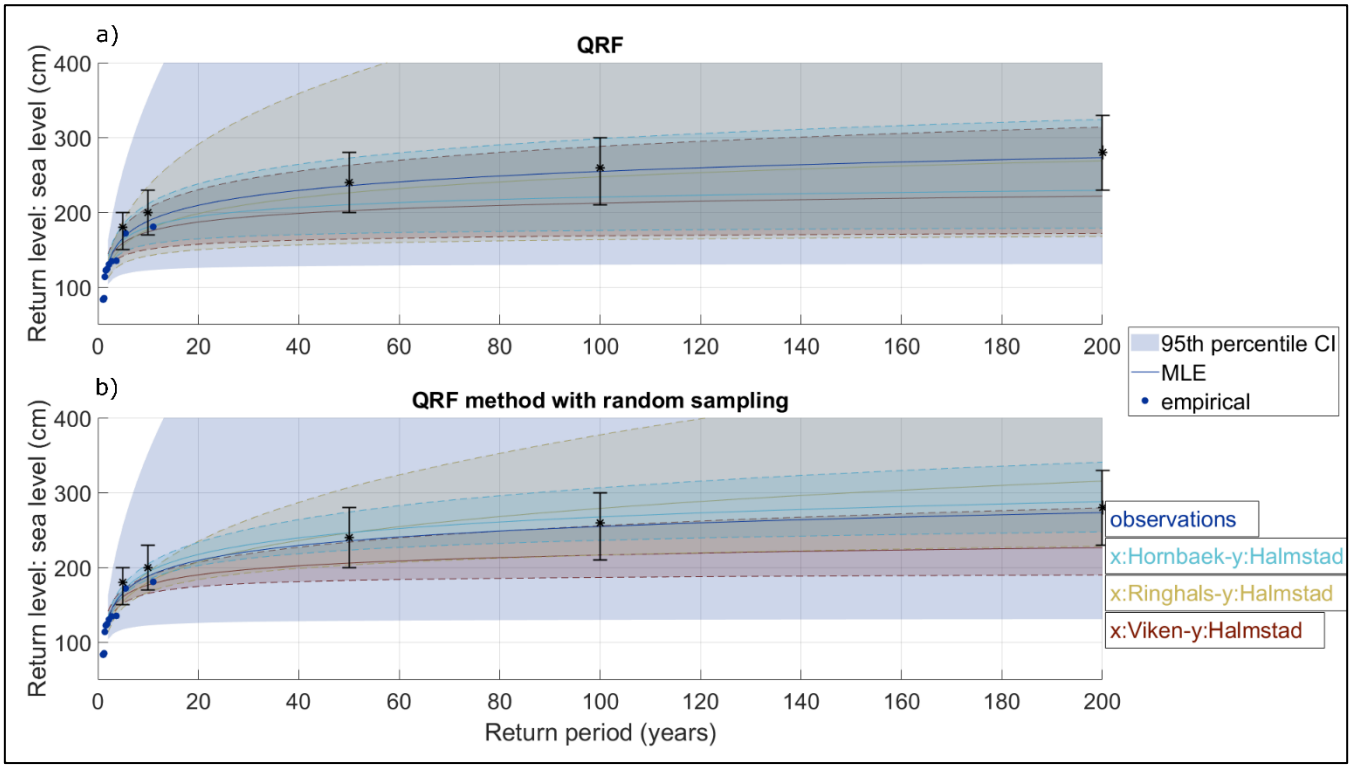

Figure 4: RLs from each reconstructed time series to predict Halmstad based on ML model mean outputs (a) and following the QRF method with random sampling (b). It displays GEV distribution fits of each dataset associated with 95th percentile confidence intervals (CIs) (a) and 95th percentile ensemble spread (b) in background colours, and the dots are the empirical data from observations. Black error bars show RLs and 95th percentile CI calculated from Andersson (2021).

4 Limitations

It is evident that our statistical reconstructions are limited by several factors in particular local ocean dynamics and the length of the time series used. Both are especially important for extreme analysis. We implicitly assume that a time window of only

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

10 years is sufficient for describing the relationship between two stations under normal ocean conditions, and while this study seems to support this hypothesis, it is by no means assured that this will be the case for any two neighbouring stations. Especially when the relationship is found to be highly non-linear. For non-normal situations like ESLs, it is evident that our setup period is principally much too short to learn the (inherently non-linear) dynamics related to rare sea level extremes, and that our modelling essentially yields an extrapolation of the normal ocean dynamics relating two sites, which may introduce significant biases in the subsequent RL estimates. Even, so it is trivial to assume that non-linear and non-parametric methods like the QRF outperform other methods in terms of capturing extreme trends within a very short time window.

As indicated earlier, the QRF is limited in range by the input values. Hence, in principle, this method is not suitable for extrapolating to higher values than what is seen in the training period, as highlighted when predicting Hornbaek sea levels from the Viken tide gauge based on the 1990-2000 and 2000-2010 setup periods. This limitation is a known issue when applying ML-based prediction models (Tyralis et al., 2019; Hengl et al., 2018); it can be mitigated to some extent by using many extended time series for model training as new data becomes available. In this study, we did not find out-of-sample issues to have a strong influence as the ML model reproduced extremes rather well. Adding additional sources, e.g. observed wind information, may also improve predictions (Johansson et al., 2018) or reanalysis (Hieronymus et al., 2019). However, these approaches were outside the scope of this technical note, focusing on exploring the limitations and advantages of only using neighbouring observations of sea level. If more complex methods can achieve additional accuracy, this is of course of great value, but it may also confuse the interpretations at times, which is not preferable. In preliminary tests, additional improvements from adding reanalysis and hindcast data did not appear to add enough value to warrant the decreased interpretability, but this is certainly an active research area.

Finally, this study focused on a limited area of the Swedish West Coast. The methodology is generally applicable but contingent on local conditions, and hence further research is needed to investigate if similar performance can be found when applying the proposed method to other areas with different ocean dynamics.

5 Summary and conclusions

This study demonstrates that sea level time series of daily maxima can be relatively successfully reconstructed from a neighbouring station using the LR or QRF approaches using even very short overlapping intervals (10 years). As expected due to the short length of the overlap, ESLs are somewhat underestimated. The QRF model is better able to capture the inherent non-linearities and hence proves to be more accurate during those conditions. The corresponding absolute bias values are generally lower than those found from the LR. The best reconstructions are generally achieved for stations spatially closer to each other, though this can be partially offset using the QRF, which is found to yield better results than the LR for stations spatially located further away from each other. We tested the QRF method with random sampling in the case of Halmstad. When applied to reconstructed time series from a 10-years dataset, the method confirmed the results from a previous more physics-based study, reproducing RLs with a reasonable uncertainty range given by the 95th percentile ensemble spread.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

The method is easily applicable to other sites and can also be applied across regions as long as two neighbouring stations' sea-level time series are available. Overall, using the QRF method with random sampling to represent the uncertainty of extremes could be an advantage compared to many single-output ML predictions.

Code availability. The code used to generate the figures and tables can be acquired by contacting the first author (kevin.dubois@geo.uu.se).

Data availability. The Hornbaek data used are available upon request. The data from the different Swedish stations are available at https://www.smhi.se/data/oceanografi/ladda-ner-oceanografiska-observationer#param=sealevelrh2000,stations=core (last access: 14 October 2021).

Author contributions. KD developed the code and conducted the analysis. KD prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgements. The work forms part of the project: Extreme events in the coastal zone – a multidisciplinary approach for better preparedness.

Financial support. This research was funded by the Swedish Research Council FORMAS (Grant No. 2018-01784) and the Centre of Natural Hazards and Disaster Science (CNDS).

References

Andersson, M. (2021). Climate Adaptation by Managed Realignment. Future mean and extreme sea levels (Issue April). https://www.sgi.se/globalassets/smhi-report-camel-sea-levels-v1_1.pdf

Andrée, E., Su, J., Andreas, M., Larsen, D., Drews, M., Stendel, M., & Madsen, K. S. (2022). The role of preconditioning for extreme storm surges in the western Baltic Sea. Natural Hazards and Earth System Sciences, May, 1–25. https://doi.org/10.5194/nhess-2022-149

Bellinghausen, K., Hünicke, B., & Zorita, E. (2023). Short-term prediction of extreme sea-level at the Baltic Sea coast by Random Forests. Natural Hazards and Earth System Sciences, March. https://doi.org/10.5194/nhess-2023-21

Bernier, N. B., Thompson, K. R., Ou, J., & Ritchie, H. (2007). Mapping the return periods of extreme sea levels: Allowing for short sea level records, seasonality, and climate change. Global and Planetary Change, 57(1–2), 139–150. https://doi.org/10.1016/j.gloplacha.2006.11.027

Breiman, L. (2001). Random Forests. Machine Learning, 12343 LNCS, 503–515. https://doi.org/10.1007/978-3-030-62008-0_35

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

Brown, S., Nicholls, R. J., Goodwin, P., Haigh, I. D., Lincke, D., Vafeidis, A. T., & Hinkel, J. (2018). Quantifying Land and People Exposed to Sea-Level Rise with No Mitigation and 1.5°C and 2.0°C Rise in Global Temperatures to Year 2300. Earth's Future, 6(3), 583–600. https://doi.org/10.1002/2017EF000738

Caldwell, P. (2012). Tide Gauge Data Rescue. Proceedings of The Memory of the World in the Digital Age: Digitilization and Preservation, 134–149. https://doi.org/10.1029/2008GL036185.1

Cid, A., Wahl, T., Chambers, D. P., & Muis, S. (2018). Storm Surge Reconstruction and Return Water Level Estimation in Southeast Asia for the 20th Century. Journal of Geophysical Research: Oceans, 123(1), 437–451. https://doi.org/10.1002/2017JC013143

Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Technometrics, 44(4), 397–397. https://doi.org/10.1198/tech.2002.s73

Crameri, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. Nature Communications, 11(1), 1–10. https://doi.org/10.1038/s41467-020-19160-7

Granata, F., & Di Nunno, F. (2021). Artificial Intelligence models for prediction of the tide level in Venice. Stochastic Environmental Research and Risk Assessment, 35(12), 2537–2548. https://doi.org/10.1007/s00477-021-02018-9

Hansen, L. (2007). Technical Report 07-09 Hourly values of sea level observations from two stations in Denmark . Hornbæk 1890-2005 and Gedser 1891-2005 Colophon.

Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., & Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 2018(8). https://doi.org/10.7717/peerj.5518

Hieronymus, M., & Kalén, O. (2020). Sea-level rise projections for Sweden based on the new IPCC special report: The ocean and cryosphere in a changing climate. Ambio, 49(10), 1587–1600. https://doi.org/10.1007/s13280-019-01313-8

Hieronymus, M., Hieronymus, J., & Hieronymus, F. (2019). On the application of machine learning techniques to regression problems in sea level studies. Journal of Atmospheric and Oceanic Technology, 36(9), 1889–1902. https://doi.org/10.1175/JTECH-D-19-0033.1

Holgate, S. J., Matthews, A., Woodworth, P. L., Rickards, L. J., Tamisiea, M. E., Bradshaw, E., Foden, P. R., Gordon, K. M., Jevrejeva, S., & Pugh, J. (2013). New data systems and products at the permanent service for mean sea level. Journal of Coastal Research, 29(3), 493–504. https://doi.org/10.2112/JCOASTRES-D-12-00175.1

Johansson, L. (2018). Extremvattenstånd i Halmstad. https://www.msb.se/siteassets/dokument/amnesomraden/skydd-mot-olyckor-och-farliga-amnen/naturolyckor-och-klimat/oversvamning/oversvamningskartering-kust/halmstad.pdf

Meinshausen, N. (2006). Quantile Regression for Left-Truncated Semicompeting Risks Data. Journal of Machine Learning Research, 67(3), 701–710. https://doi.org/10.1111/j.1541-0420.2010.01521.x

Menéndez, M., & Woodworth, P. L. (2010). Changes in extreme high water levels based on a quasi-global tide-gauge data set. Journal of Geophysical Research: Oceans, 115(10), 1–15. https://doi.org/10.1029/2009JC005997

Merrifield, Mark; Holgate, Simon; Mitchum, Gary; Pérez, Begoña; Rickards, Lesley; Schöne, Tilo; Woodworth, Philip; Wöppelmann, G. (2012). Global Sea-level Observing System (GLOSS) Implementation plan – 2012.

https://doi.org/10.5194/egusphere-2023-1159
Preprint. Discussion started: 8 June 2023

Rutgersson, A., Kjellström, E., Haapala, J., Stendel, M., Danilovich, I., Drews, M., Jylhä, K., Kujala, P., Guo Larsén, X., Halsnæs, K., Lehtonen, I., Luomaranta, A., Nilsson, E., Olsson, T., Särkkä, J., Tuomi, L., & Wasmund, N. (2021). Natural Hazards and Extreme Events in the Baltic Sea region. Earth System Dynamics Discussions, 1–80. https://doi.org/10.5194/esd-2021-13

SMHI. (2022). Retrieved from Ladda ner oceanografiska observationer: https://www.smhi.se/data/oceanografi/ladda-ner-oceanografiska-observationer#param=sealevelrh2000,stations=core

Tyralis, H., Papacharalampous, G., & Langousis, A. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water (Switzerland), 11(5). https://doi.org/10.3390/w11050910

Vousdoukas, M. I., Mentaschi, L., Hinkel, J., Ward, P. J., Mongelli, I., Ciscar, J. C., & Feyen, L. (2020). Economic motivation for raising coastal flood defenses in Europe. Nature Communications, 11(1), 1–11. https://doi.org/10.1038/s41467-020-15665-3

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slangen, A. B. A. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. Nature Communications, 8(May), 1–12. https://doi.org/10.1038/ncomms16075

Woodworth, P. L., Menéndez, M., & Gehrels, W. R. (2011). Evidence for Century-Timescale Acceleration in Mean Sea Levels and for Recent Changes in Extreme Sea Levels. Surveys in Geophysics, 32(4–5), 603–618. https://doi.org/10.1007/s10712-011-9112-8

Woodworth, P. L., Pouvreau, N., & Wöppelmann, G. (2010). The gyre-scale circulation of the North Atlantic and sea level at Brest. Ocean Science, 6(1), 185–190. https://doi.org/10.5194/os-6-185-2010