**CHAPTER 3 - Supplementary material**

**3.4 Consistent data set of coastal sea level: The synergy between tidal gauge data and numerical modelling**

Authors: Sebastian Grayek, Emil Stanev, Nam Pham, Antonio Bonaduce and Joanna Staneva

**3.4.6 Appendix A; Analyses of proxy data used for the training of the reconstruction method.**

The following section presents analyses of the correlation patterns in the leading EOFs of the proxy data used for the training of the reconstruction. The cumulative explained variance across the leading 30 EOFs is shown in Fig. 3.4.7A. The explained variance of the first four and six EOFs is more than 90% and 95%, respectively. From EOF 7, we gain less than 1% of the explained variance when we add higher modes. Fig. 3.4.7B-E show the four leading EOFs and their corresponding principal components (PCs) explaining 64.08%, 14.89%, 6.26%, and 5.58% of the total variance in the model simulations. EOF 1 shows positive values throughout the region and is related to the filling and emptying of the basin due to external conditions. The spatial patterns in EOF 2 through EOF 4 have positive and negative values and are related to the relocation of water within the region. To capture the annual variability of the PCs, we show the 14-day low-pass filtered PCs in Fig. 3.4.7F. Higher frequencies are shown in Fig. 3.4.7G for a shorter period in June 2019 by the original PCs, without low-pass filtering. All four PCs clearly show semimonthly to trimonthly variations. PC 3 and PC 4 contain higher frequencies with distinct amplitudes and a period of about 12 hours, most likely related to the M2 tides. The relatively low explained variance of PC 3 and PC 4 and the small amplitudes of these frequencies indicate that tides play only a minor role in the Baltic Sea.



*Figure 3.4.7 The graph in picture A displays the accumulated explained variance across the leading 30 EOFs, which are used for the background covariance matrix* $P$*. The numbers show the individual variance explained by the first seven EOFs. Panel B, C, D and E show the leading four EOFs. Picture F displays the respective PCs low-pass filtered with a time window of 14 days. Picture G presents the original PCs, without low-pass filtering, for the first 14 days in June 2019. Products used: ref. 3.4.1 (ANL).*

**3.4.7 Appendix B; Extended validation and statistics of reconstructed data.**

In Fig. 3.4.8A and B, we validate the yearly statistics of the reconstruction and compare its performance with the CMEMS product. For all tide gauges, we present the scatter, linear regression and QQ plot of the reconstructed (left, REC) and CMEMS data (right, ANL) against the observations (TGD-P28 and TGD-P46). Since the amount of data for the entire reconstruction period is too large, we show the plots for the year 2020 as an example. Both data sets agree quite well with the observations, with the reconstruction performing slightly better. The statistical values for the model simulations from other years in the period (not shown here) fluctuate slightly, while the statistics for the reconstruction remain quite stable. Fig. 3.4.8C shows the statistics of the reconstructed extreme values (>99th percentile) for the entire reconstruction period from 1993 to 2020. Similarly to Fig. 3.4. 8A and B, we show scatter, linear regression, and QQ plots of the reconstructed data versus the observations for all tide gauges, but only for pairs of data (reconstructed and observed values) that are above the 99th percentile in both data sets. Not surprisingly, the statistics for the reconstruction of the extreme values are not as good as for the entire data set. The dispersion of the reconstruction data around the observation is considerably larger than for the full 2020 period data (see Fig. 3.4.8A), resulting in lower correlation and smaller values for the coefficient of determination (R^2) in the linear regression with the observations. However, there is fairly good agreement between the observed and reconstructed percentiles (black crosses in the plot) up to ~1.4 m. The percentiles between ~1.4m and ~1.8m are slightly overestimated in the reconstruction. Above ~1.8m, the deviation becomes very large.



*Figure 3.4.8 Scatter, linear regression and QQ plot of the reconstructed (A) and CMEMS data (B) for the year 2020 and the reconstructed extreme values (>99th percentile, C) for the whole data set (1993-2020) against the tide gauge observations; Products used: ref. 3.4.1 (ANL), ref. 3.4.3 (TGD-P26 and TGD-P46) and ref. 3.4.6 (REC).*

## 3.4.8 Auxiliary Data

|  |  |  |
| --- | --- | --- |
|  | TIDE GAUGES FOR RECONSTRUCTION | TIDE GAUGES FOR VALIDATIONS |
|  | STATIONS | LAT | LON |  |  | STATIONS | LAT | LON |
| Aa | Aarhus | 56.15 | 10.217 |  | 1 | Skagsudde | 63.191 | 19.013 |
| De | Degerby | 60.032 | 20.385 |  | 2 | Spikarna | 62.363 | 17.531 |
| Fo | Forsmark | 60.409 | 18.211 |  | 3 | Stockholm | 59.324 | 18.082 |
| Fu | Furuogrund | 64.916 | 21.231 |  | 4 | Landsort1 | 58.75 | 17.867 |
| Ham | Hamina | 60.563 | 27.179 |  | 5 | Marviken | 58.554 | 16.837 |
| Han | Hanko | 59.823 | 22.977 |  | 6 | OlandsNorraUdde | 57.366 | 17.097 |
| Hei | Heiligenhafen | 54.373 | 11.006 |  | 7 | Oskarshamn | 57.275 | 16.478 |
| Hel | Helsinki | 60.154 | 24.956 |  | 8 | Tejn | 55.25 | 14.833 |
| Hes | Hesnaes | 54.817 | 12.133 |  | 9 | Ronne | 55.1 | 14.683 |
| Kal | KalixStoron | 65.697 | 23.096 |  | 10 | Rodvig | 55.254 | 12.373 |
| Kas | Kaskinen | 62.344 | 21.215 |  | 11 | Klagshamn | 55.522 | 12.894 |
| Kem | Kemi | 65.673 | 24.515 |  | 12 | Drogden | 55.536 | 12.712 |
| Kol | Kolka | 57.733 | 22.583 |  | 13 | NordreRose | 55.636 | 12.687 |
| Kro | Kronstadt | 59.967 | 29.75 |  | 14 | Kobenhavn | 55.7 | 12.6 |
| Ku | Kungsholmsfort | 56.105 | 15.589 |  | 15 | Barseback | 55.756 | 12.903 |
| La | LandsortNorra | 58.769 | 17.859 |  | 16 | Hornbaek | 56.1 | 12.467 |
| Raa | Raahe | 64.666 | 24.407 |  | 17 | Viken | 56.142 | 12.579 |
| Rat | Ratan | 63.986 | 20.895 |  | 18 | GoteborgTorshamnen | 57.685 | 11.791 |
| Rau | Rauma | 61.133 | 21.426 |  | 19 | Stenungsund | 58.093 | 11.833 |
| Ri | Ringhals | 57.25 | 12.113 |  | 20 | Kungsvik | 58.997 | 11.127 |
| Sa | Sassnitz | 54.511 | 13.643 |  | 21 | Frederikshavn | 57.433 | 10.567 |
| Sim | Simrishamn | 55.557 | 14.358 |  | 22 | Grena | 56.411 | 10.931 |
| Skag | Skagen | 57.719 | 10.586 |  | 23 | SjaellandsOdde | 55.975 | 11.372 |
| Skan | Skanor | 55.417 | 12.829 |  | 24 | Ballen | 55.817 | 10.644 |
| Sli | Slipshavn | 55.283 | 10.833 |  | 25 | Juelsminde | 55.717 | 10.017 |
| So | Sonderborg | 54.917 | 9.783 |  | 26 | Fredericia | 55.567 | 9.75 |
| Ud | Udbyhoej | 56.6 | 10.3 |  | 27 | Korsor | 55.333 | 11.133 |
| Vi | Visby | 57.639 | 18.284 |  | 28 | Fynshav | 55 | 9.983 |
|  |  |  |  |  | 29 | Spodsbjerg | 54.933 | 10.833 |
|  |  |  |  |  | 30 | Bagenkop | 54.753 | 10.678 |
|  |  |  |  |  | 31 | Rodby | 54.65 | 11.35 |
|  |  |  |  |  | 32 | Gedser | 54.567 | 11.933 |
|  |  |  |  |  | 33 | KielHoltenau | 54.372 | 10.157 |
|  |  |  |  |  | 34 | Travemuende | 53.958 | 10.872 |
|  |  |  |  |  | 35 | Warnemuende | 54.17 | 12.103 |
|  |  |  |  |  | 36 | Koserow | 54.06 | 14.001 |
|  |  |  |  |  | 37 | Daugavgriva | 57.05 | 24.017 |
|  |  |  |  |  | 38 | Lehtma | 59.069 | 22.697 |
|  |  |  |  |  | 39 | Paldiski | 59.335 | 24.08 |
|  |  |  |  |  | 40 | Tallinn | 59.444 | 24.764 |
|  |  |  |  |  | 41 | StPetersburg | 59.933 | 30.267 |
|  |  |  |  |  | 42 | Turku | 60.428 | 22.101 |
|  |  |  |  |  | 43 | Pori | 61.594 | 21.463 |
|  |  |  |  |  | 44 | Vaasa | 63.082 | 21.571 |
|  |  |  |  |  | 45 | Pietarsaari | 63.709 | 22.69 |
|  |  |  |  |  | 46 | Oulu | 65.04 | 25.418 |