We would like to thank the two reviewers for taking the time to review our manuscript and for their valuable comments and suggestions to improve our manuscript. We have written our responses below their points in blue.

RC1:
The modelling of storm tide events is highly important for understanding extreme sea levels and coastal flooding, especially from a climatic perspective (past and future). Moreover, because of its complex dynamics, long-term modelling and analysis are challenging. The paper presents the analysis of different reanalysis forcing for storm tide modelling in the North Sea. The paper is well written and well structured, and results are clearly presented, with figures and tables being all relevant.

I compliment the authors for their huge amount of modelling work. I suggest the author use such a database of model results for performing a deeper analysis and discussion on the uncertainty in reproducing extreme events using ensembles (20CRv2c and 20CRv3).

We have extended the discussion of uncertainties in the results section and expanded the appendix with the statistics of water level extremes for both ensembles. The main differences between 20CRv2c and 20CRv3 are the increased temporal and spatial resolution and changed data assimilation scheme for the 20CRv3 ensemble. This results in the more realistic representation of extreme water levels forced by 20CRv3 for the low-pressure systems, which are moving over the central North Sea. More realistic in this case means that the ensemble values are more centred around the observed one rather than overestimate the observations as in the case of 20CRv2. The model results for southern tracks show a higher variability in extreme water levels compared to northern tracks and in general purer accuracy. Still, the 20CRv3 ensemble demonstrates significant improvements with respect to previous version.

From line 382 we will complement the paragraph with:

Another point worth mentioning is that the ensemble mean values of maxima water levels underestimate the observed values almost for all events and both ensembles, with the exception of 1962 and 21 Jan 1976 storms simulated with 20CRv2 (Table A5). Whereas the selected ensemble members can reproduce or come close to the observed extreme water levels, especially for the Scandinavia type storms, the ensemble mean underestimation reaches up to 1.6m. Specifically for Jutland type events, the use of ensemble mean values for the representation of water level extremes is not recommended.

This table is added to the Appendix.

Cuxhaven	**20CRv2c**				**20CRv3**			
	Max - Obs.	Mean - Obs.	Max - Min	STD	Max - Obs.	Mean - Obs.	Max - Min	STD
16 Feb 1962	0.71	0.52	0.38	0.09	0.54	-0.26	1.41	0.28
21 Jan 1976	0.16	0.02	0.31	0.07	0.08	-0.48	1.03	0.20
06 Dec 2013	-0.24	-0.35	0.21	0.05	0.28	-0.17	0.95	0.20
03 Jan 1976	-1.46	-1.63	0.36	0.09	-0.58	-1.23	1.27	0.30
03 Dec 1999	-0.88	-1.07	0.38	0.09	-0.32	-1.08	1.29	0.30
23 Feb 1967	-0.64	-0.70	0.15	0.04	-0.12	-0.79	1.06	0.18

Tab A5: Statistics of modelled water levels by 20CRv2c and by 20CRv3 and observation (Obs.) for the location Cuxhaven