Answers to Referee 2

Masini et al have analyzed the data associated with two storm events in the Gotland region that impacted water desalination plants on the island. They have effectively utilized available oceanographic and atmospheric data to present an observational analysis of the hazards from those events to the desalination plants, and proposed a methodology to estimate future hazard forecasts.

While the observational and analysis part of the framework is sound and well documented, the forecast and broader applicability beyond the 2 historical events and identified locations is questionable based on the results presented in the manuscript. I have further described this concern and other minor issues below. I hope that the authors and the editor find these comments helpful for further improving this manuscript for future readers.

We thank the referee for the constructive feedback that led to a considerable improvement of our manuscript. In the following, we answer point by point to the comments raised. Regarding the implemented changes, **we refer to the lines in the submitted version of the manuscript**, i.e., the preprint available at `https://egusphere.copernicus.org/preprints/2026/egusphere-2026-1344/`).

- *With the linear regression based forecast model as a primary contribution of the manuscript, the fit is generally quite poor (max R2=0.62 at KVA and 0.46 at HER; and negative R2 for 2024 cold spells at KEV). Additionally due to lack of many events in the prediction period (F2024), the validation of the model is difficult. In fact the model failed to predict any of the 5 extreme transport events at HER in 2024. This is a major limitation of the study, that reduces the support of its utility as a broadly applicable hazard forecasting tool for the region.*

  Thank you for your comment. First of all, we revised these results and realized that we used a wrong threshold for 2024. This did not change the results much because, as you noted, there were no extreme transport events in summer 2024. So in order to have more storm events to validate our regression models, we use now two validation periods, one for autumn (September-November) 2023 ("F2023 autumn") and the second for summer-autumn (July-November) 2024 ("F2024 summer-autumn"). The validation have been extended to the autumn seasons only for the extreme zonal current, not for the sea temperature spells which are cause hazard impacts only during the summer months (see Table 1). The results from these two validation periods are shown in Table 1 (a new version of Table 1 in the manuscript, which will become Table 2; please note that the table includes also regression parameters you requested in the following comments. Note that the Table refers to Figure 3 which will become Figure 4 in the revised version of the manuscript). In order to interpret these results, we show also wind roses and wind vectors in figures R1 and R2.

Table 1: Performance metrics for linear regression models of the zonal (proxy cross-shore) transport ($U/(3.5\sigma)$) evaluated at ocean model grid points around KVA and HER. ~~in summer 2023.~~Position of the points are marked in ~~Fig. 3~~figure 4. The linear regression model is trained for the ocean model output in summer 2023 (In-sample F2023) using the wind stress at HOB as a predictor (Eq.4). The validation of the regression model is done using ocean model output in the ~~subsequent summer~~autumn 2023 (F2023 autumn) and in the summer-autumn 2024 (F2024 summer-autumn). The metrics include the number of identified extreme transport events ($|U| > 3.5\sigma$, see below) in a given model point, percentage of correctly predicted events (relative to the total number of events), the coefficient of determination (R^2), the non-dimensional root-mean-square error (RMSE) and the standard deviation σ [m s^{-1}]. The regression results are statistically significant at 95% confidence level ($p < 0.05$). The number of predicted events can be found by multiplying the values in the row of "Predicted" with the corresponding values of the "Events".

KVA – In-sample (F2023 summer)							
Station	A	B	C	D	F	H	N
Events (count)	12	11	11	12	11	12	11
Predicted (%)	91.7	90.9	100.0	100.0	81.8	83.3	72.7
R^2	0.544	0.565	0.622	0.616	0.483	0.543	0.419
RMSE	0.193	0.188	0.176	0.177	0.205	0.193	0.218
σ [m s^{-1}]	0.036	0.037	0.063	0.046	0.041	0.038	0.039
Intercept A_0	-0.054	-0.039	-0.068	-0.055	-0.030	-0.018	-0.005
Slope A_1 [Pa^{-1}]	-4.064	-4.143	-4.348	-4.325	-3.831	-4.060	-3.568
KVA – Validation (F2023 autumn)							
Events (count)	8	7	7	8	5	6	4
Predicted (%)	100.0	100.0	100.0	100.0	100.0	100.0	75.0
R^2	0.586	0.669	0.653	0.647	0.560	0.589	0.491
RMSE	0.184	0.164	0.168	0.170	0.190	0.183	0.204
σ [m s^{-1}]	0.046	0.052	0.081	0.059	0.054	0.050	0.054
KVA – Validation (F2024 summer-autumn)							
Events (count)	5	18	3	5	2	6	4
Predicted (%)	40.0	33.3	33.3	40.0	0.0	33.3	25.0
R^2	0.507	0.590	0.633	0.604	0.445	0.478	0.323
RMSE	0.207	0.215	0.184	0.182	0.230	0.217	0.265
σ [m s^{-1}]	0.037	0.044	0.067	0.047	0.044	0.040	0.044

HER – In-sample (F2023 summer)							
Station	A	B	C	D	F	I	P
Events (count)	5	2	10	11	13	14	9
Predicted (%)	0.0	100.0	0.0	0.0	7.7	50.0	44.4
R^2	0.034	0.301	0.158	0.168	0.277	0.463	0.370
RMSE	0.281	0.239	0.262	0.261	0.243	0.209	0.227
σ [m s^{-1}]	0.013	0.019	0.031	0.015	0.024	0.042	0.035
Intercept A_0	-0.041	-0.075	0.172	0.150	-0.121	-0.113	-0.139
Slope A_1 [Pa^{-1}]	1.009	3.026	-2.189	-2.259	2.901	3.750	3.353
HER – Validation (F2023 autumn)							
Events (count)	6	5	4	13	3	9	8
Predicted (%)	0.0	20.0	0.0	0.0	0.0	44.4	37.5
R^2	-0.152	0.130	0.265	0.289	0.183	0.446	0.341
RMSE	0.307	0.267	0.245	0.241	0.258	0.213	0.232
σ [m s^{-1}]	0.018	0.019	0.044	0.022	0.029	0.054	0.044
HER – Validation (F2024 summer-autumn)							
Events (count)	42	19	50	62	31	39	25
Predicted (%)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
R^2	-0.080	0.196	0.203	0.217	0.189	0.409	0.288
RMSE	0.349	0.256	0.313	0.310	0.289	0.263	0.294
σ [m s^{-1}]	0.015	0.019	0.038	0.019	0.027	0.050	0.043

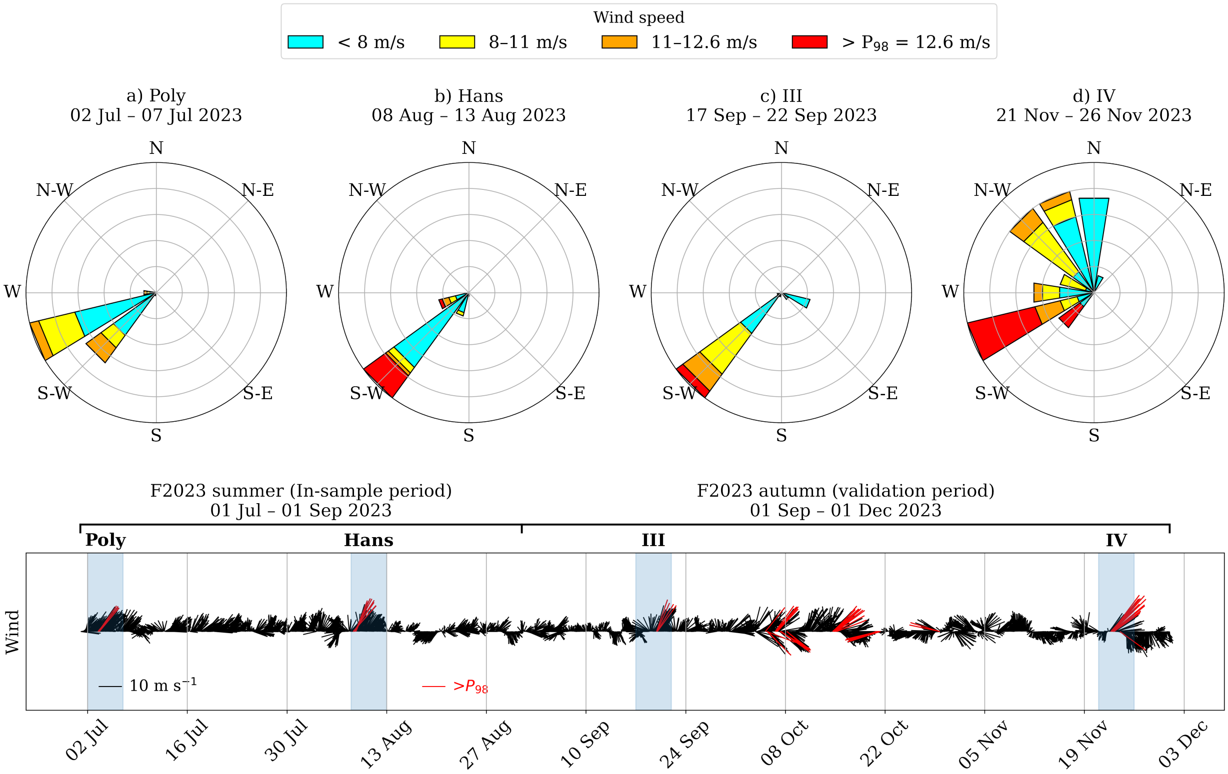

Figure R1: Wind roses (upper row) showing distribution of wind direction and speeds. The range of speeds is *calm* wind conditions< 8 m/s, *breeze* 8 − 11 and *strong breeze* 11-12.6 m/s. The exceedances of the 98th-percentile (P_{98} = 13.6 m/s) are shown in red. The bottom panel shows the stick plot for the wind. The period shown is from 1 July till 1 December 2023.

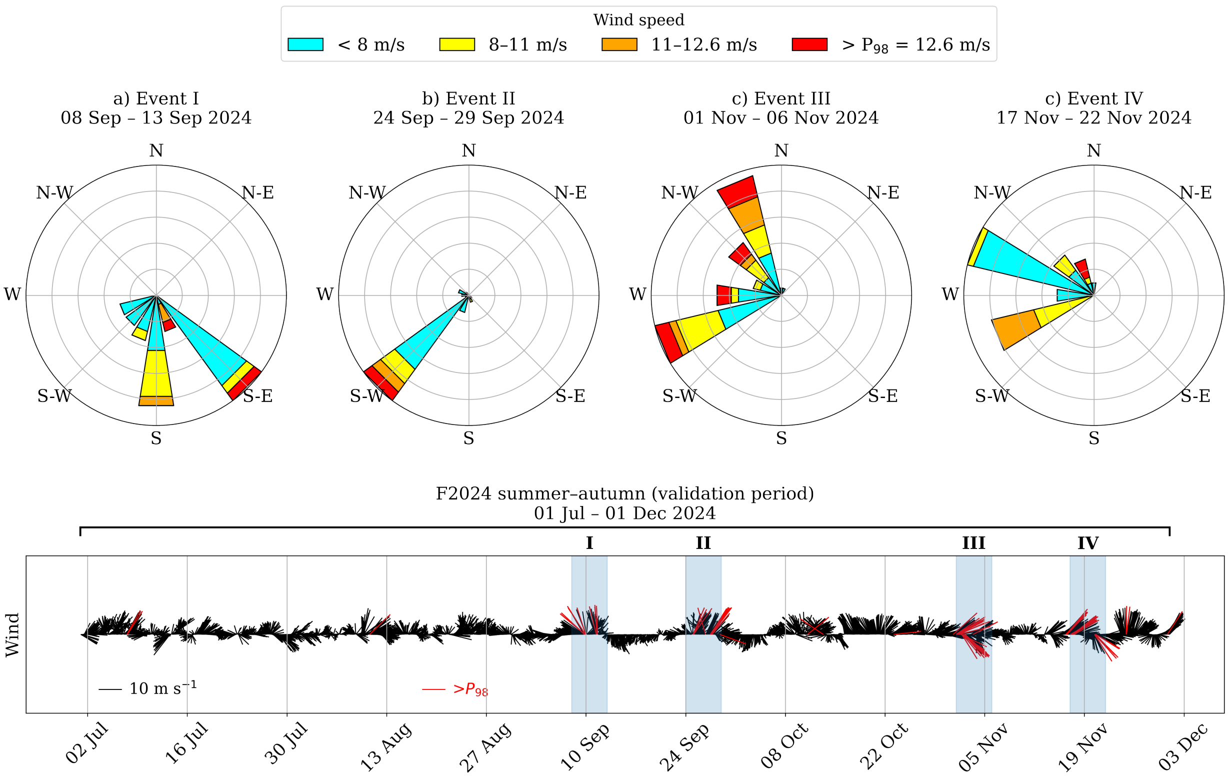

Figure R2: Same as figure above for the period 1 July till 1 December 2024.

Regarding the validation period F2023 autumn, the linear regression model for zonal transport U predicted 100% of the extreme transport events at all the near-by model points for KVA except one (the furthermost point N where the local recirculation weakens off, see Figure 4 and sect. 3.2.1 in the submitted version of the manuscript). At HER, up to about 40% events were predicted at two near-by model points (I and P, south off the desalination plant where the coastal jet associated with SW storms is pronounced, see Figure 4 in the old version of the manuscript), which is similar to the training period F2023 summer, but for point B only 20% events were predicted compared to 100% for the training period. The R^2 scores for the validation F2023 autumn are comparable to the training period F2023 summer at both, KVA and HER.

However, for the validation period of F2024 summer-autumn, the percentage of predicted events was reduced to maximum 40% at two near-by model points for KVA, and none was predicted for HER.

We attribute these results about the regression model performance to the different wind directions characterizing the storms during the two validation periods that in turn determined the coastal circulation. In the validation period F2023 autumn, there are two storms with extreme southwesterly winds similar to Poly and Hans in the training period F2023 summer, hence the similar performance of the regression model (see Sect. 3.1.2). In summer 2024 there are no storms, and in autumn 2024 there are four storms with extreme winds from varying directions and the coastal current response to these was not well captured by our linear regression model with the alongshore wind stress as predictor. We have no impact reports from the desalination plant for 2024 however to test alternative versions of the model for the winds from other directions than southwesterly.

We will incorporate the updated validation results in the following way:

First of all, we will add define the various ocean model data sets for 2023 and 2024 in section 2.2, **line 235**:

"This data set is called F2023 summer. To validate the linear regression model (see sect. 3.2), we also use data for ~~the following~~autumn (September-November) (F2023 autumn) as well as data for summer-autumn (July-November) ~~of the following year~~2024 (F2024 summer-autumn) in order to validate the linear regression model".

and at **line 323**, we will add:

"We validate the regression model using ocean model data for ~~the following summer 2024~~autumm (F2023 autumn) as well as summer and autumn 2024 (F2024 summer-autumn)."

In the Results section 3.2.1, **at line 485**, we will present the results from the validation as in the answer above, but for the updated references to the Table (this will be Table 2 in the new version of the manuscript) as well as figures (Figure 4 will become Figure 5 in the new version, and Figures R3 and R4 will become Figures A6 and A7 in the Appendix):

~~The validation of the linear regression model in the following summer (F2024) is shown in Table 1. We got no impact reports from desalination plants for summer 2024 due to unremarkable summer conditions. There were no extreme transport events at KVA and only a couple of events at HER which were not predicted~~The validation of the linear regression model is shown in Table 1. Regarding the validation period F2023 autumn, the linear regression model for zonal transport U predicted 100% of the extreme transport events at all the near-by model points for KVA except one (the furthermost point N where the local recirculation weakens off, see Figure 5 and section 3.2.1). At HER, about 40% events were predicted at two near-by model points (I and P, south off the desalination plant where the coastal jet associated with SW storms is pronounced, see Figure 5 and section 3.2.1), which is similar to the training period F2023 summer, but for point B only 20% events were predicted compared to 100% for the training period. The R^2 scores for the validation F2023 autumn are comparable to the training period F2023 summer at both, KVA and HER.

However, for the validation period of F2024 summer-autumn, the percentage of predicted events was reduced to maximum 40% at two near-by model points for KVA and none was predicted for HER. We attribute these results to the different wind directions characterizing the storms during the two validation periods (see Appendix Figures A6-A7). In the validation period F2023 autumn, there were two storms associated with SW winds and wind speeds exceeding 98th percentile, similar to Hans, hence the similar performance of the regression model. In summer 2024 there were no extreme storms, and in autumn 2024 there were four storms with extreme winds from different directions and the coastal current response to these was not well captured by our linear regression model with the alongshore wind stress as predictor.

The following text will be inserted into the Discussion (**line 571**):

"The linear regression models captured 100% of the extreme transport events at least at one near-by model point for KVA and HER for the training period in summer 2023. For the validation period in autumn 2024 (when two other storms with SW winds exceeding 98th-percentile wind speeds occurred), the linear regression model predicted 100% of the extreme transport events at all but one near-by model points for KVA and about 40% events were predicted at two near-by model points at HER. However, for the validation period of summer-autumn F2024, only about 40% events were predicted near KVA, and none was predicted at HER. This was due to the different wind directions characterizing the storms during this period and the coastal current response to these was not well captured by our linear regression model with the alongshore wind stress as predictor. We have no impact reports from the desalination plant for 2024 however to test alternative versions of the model. .

- *Tables 1 and 2: Can the authors include another row in both training and validation boxes for the number of predicted events, in order to compare against the observed events?*

  The number of predicted events can be found by multiplying the values in the row of "predicted" with the corresponding values of the "events". We will add this information in the captions of both Tables 2 and 3 (that will become Table 3 and 4 in the revised version of the manuscript). Both quantities are already displayed in the current table, so we think adding the number of events as a separate block would be superfluous.

- *Section 3: Can the authors include the fitted coefficients in the Results section, and comment on the predictive power of their chosen variables based on the coefficient sign and values.*

  Thank you for the comment, we think this is very useful. The fitted coefficients (intercepts and slopes) A_0 and A_1 for $U/3.5\sigma$ (Eq.4 in the manuscript) and α_0 and α_1 for $T_b/12$ (Eq.5 in the manuscript) will be added in Table 1 and Table 2, respectively (see Table 1 and 2 in this response).

  Regarding the linear regression model for the zonal transport, in Section 3.2.1, we will add a few comments regarding the linear regression parameters (**line 479**, please note that it refers to Figure 4 in the submitted version which will become Figure 5 in the revised version of the manuscript):

  "At the model points A-D located in a shallow coastal area close to KVA (7a-e), the linear regression has R2 = 0.54-0.62 (corresponding to a linear correlation of about 0.75) and the model was able to predict nearly 90%, i.e., all but one of the 11-12 extreme transport events. There is however a pronounced structure in the residuals (the lower transports tend to be underestimated and the extreme ones tend to be overestimated), which suggest to seek more complex models. Note that the negative sign of the slope ($A_1 < 0$) for all grid points close to KVA indicates the increase of the westward (negative) zonal current with the wind stress, associated with the recirculation seen in Figure 5e,f,h. Note also that the intercept (A_0), which is the value of the predicted variable (zonal current) if the predictor (wind stress) is zero), is slightly positive (up to a few cm/s) which indicates a bias in the model which might arise due to other drivers not included in our simple regression model, limited time resolution, etc. The performance of the model is worse at point F as well as H-N further away from the coast, which have lower magnitudes of the slope, i.e., lower dependence on the wind stress (Table 1)."

  And **in line 483** (for HER):

  "(...) and the linear regression has generally lower performance close to the shore because of the complex cape-like topography making the ocean currents to veer (see Figure 5d,g,i). This is reflected in the regression slope A_1 changing the sign from positive for points located close and south of HER (A,B,F,I,P) to negative for points C-D located north of it. Note also that also here the intercept (A_0) is slightly positive (up to 15 cm/s) marking a bias in the regression model. The best performance is shown by points B and I south of HER (Fig. 7g-h), with R2 = 0.3 (0.4) and 100% (50%) predicted events at B (I), which are the points that have positive and highest value of the slope and relatively small values of the intercept compared to other points.

  And we will add the following sentences in section 3.2.2 to comment on the coefficients and the predictive power of the linear regression model for cold spell in KVA (at **line 501**, please note that figure 5 will become figure 6 in the revised version of the manuscript):

"In KVA, no cold spells events occur during the Poly and Hans storms, but some events are registered in the observations at the intake station as well as at the grid points (F,H and N) during a NW wind event (upwelling favorable on the western Gotland coast) on the 17-21 of August (see time series in figure ~~5b~~6b). Linear regression on observations from intake station yields small determination coefficient (R2 =0.05). Regarding the linear regression of T_b with wind impulse as predictor for the operational model grid points, the slope α_1 is negative for the model grid points A-D close to the coast and positive for model points F,H,N away from it reflecting the recirculation seen in Figure 5f and h. ~~For the model grid points C and N, R2 =0.47 and lower for the other points.~~For the model grid point C, the determination coefficient is highest, R2 =0.47, reflecting the highest magnitude of the slope (dependence of the wind impulse). Note that R2 is also 0.47 at the furthermost grid point N and the slope magnitude is high as well. Point N is about 20 m deep, comparable to the depth of the summer thermocline, and these values reflect sensitivity of the thermocline displacement the wind forcing). "

and at lines **line 503**) for cold spells in HER:

"At HER, the performance of the linear regression models is higher reflecting the coastal upwelling response (sect.3.1.3) with R^2 of about 0.5 and negative coefficients α_1 about 3-4 times larger than for KVA, indicating stronger decrease of bottom temperature with the wind impulse in line with the linear coastal upwelling response. The percentage of predicted events is ~~and~~ over 70-80% ~~predicted events~~ at points A-B and F-P (south and offshore off HER)."

We will also specify for clarity at **line 507** at the end of Section 3.2.2:

"Note that both training and validations consider only the summer months (July-August) when the cold spell hazard can occur."

Table 2: Performance metrics for models of the bottom temperature ($T_b/12°C$) evaluated at ocean model grid points around KVA and HER~~in summer 2023~~. Position of the points are marked in Fig.6. The top block shows the linear correlation between the ocean model (F2023 summer) output in the grid points and temperature measurements at the KVA water intake station. Summer is defined as July and August, when the cold spell hazard can occur. The remaining blocks show results from the linear regression model for the bottom temperature around KVA and HER trained for the ocean model output in summer 2023 (In-sample F2023 summer) using the wind impulse as a predictor (Eq.5). The validation of the regression model is done using ocean model output in ~~the subsequent~~ summer (F2024 summer). The metrics include the number of identified cold spells (count) in a given model point, percentage of correctly predicted events (relative to the total number of cold spells), the coefficient of determination (R^2), and the non-dimensional root-mean-square error (RMSE). The correlation and regression are statistically significant at 95% confidence level ($p < 0.05$).

Correlation with observations at KVA water intake							
F2023	A	B	C	D	F	H	N
	0.48	**0.46**	**0.21**	**0.45**	**0.70**	**0.69**	**0.50**
KVA — In-sample (F2023)							
Cold spells (count)	0	0	0	0	109	112	801
Predicted (%)	-	-	-	-	0	0	92.6
R^2	0.18	0.32	0.47	0.17	0.08	0.11	0.47
RMSE	0.10	0.08	0.08	0.10	0.22	0.22	0.25
Intercept α_0	1.57	1.60	1.65	1.56	1.33	1.29	0.71
Slope $\alpha_1 \cdot 10^3$ [Pa^{-1} s^{-1}]	-0.004	-0.005	-0.006	-0.004	0.005	0.006	0.020
KVA — Validation (F2023)							
Cold spells (count)	0	0	0	0	80	81	818
Predicted (%)	-	-	-	-	0	0	91
R^2	-0.11	-0.29	-0.10	-0.13	0.15	0.16	-0.12
RMSE	0.12	0.11	0.11	0.10	0.19	0.18	0.35
HER — In-sample (F2023)							
Station	A	B	C	D	F	I	P
Cold spells (count)	275	312	184	459	355	355	506

HER							
Predicted (%)	73	72	60	52	80	76	70
R^2	0.54	0.52	0.54	0.27	0.43	0.46	0.45
RMSE	0.20	0.21	0.12	0.28	0.22	0.22	0.24
Intercept α_0	1.44	1.42	1.51	1.28	1.32	1.35	1.24
Slope $\alpha_1 \cdot 10^3$ [Pa^{-1} s^{-1}]	-0.018	-0.018	-0.015	-0.014	-0.015	-0.017	-0.017
HER — Validation (F2024)							
Cold spells (count)	114	123	64	159	195	158	364
Predicted (%)	74	65	0	71	59	70	57
R^2	0.23	0.23	0.27	-0.25	0.35	0.35	0.43
RMSE	0.20	0.20	0.17	0.28	0.21	0.20	0.23

- *Sections 4 and 5: The authors' conclusions regarding the linear dependence as evidenced by the results of linear regression model are not well supported given the weak regression fits. I recommend the authors revise the language to better reflect the fit, and/or indicate that a linear dependence model is not sufficiently well suited for this hazard forecasting.*

  Both sections will be edited. See response to your previous comments.

- *Line 475: What is the false positive rate of predictions from the linear regression fits?*

  Thank you for your comment. For the zonal transport, the false positive rate of predictions defined as POFD=FP/(FP+TN), with FP being false positives, and TN the true negatives, is shown in Table 3 for KVA, and in Table 4 for HER, together the percentage of predicted events (Predicted=TP/(TP+FN)) shown already in the submitted manuscript (Table 1 and Table 2 in this answer, which will become Table 2 and Table 3 in the revised version). Please note that we do not round the percentage of the predicted events but keep the second decimal point for consistency with (very small) false positive rates. For completeness, we also include the success ratio (or precision) SR=TP / (TP + FP).
  Across all grid points and validation periods, POFD remains below 1% (0.00–0.79 %).

  The low false positive rate of predictions values are expected because extreme events are rare, resulting in a much larger number of true negatives than false positives.

  For the bottom temperature, POFD is reported in Table 5 and 6 for KVA and HER respectively. Across all stations and validation periods, POFD remains below 0.1%.

Table 3: KVA: Predicted events (as in Table 1), the false positive rate of predictions (Probability of False Detection, POFD) and Success Ratio (SR) for each model grid point for zonal current (FP = False Positives, TN = True Negatives, TP= True Positives).

		KVA						
Dataset	Metric [%]	A	B	C	D	F	H	N
Summer 2023	Predicted	91.67	90.91	100.00	100.00	81.82	83.33	72.73
	POFD=FP/(FP+TN)	0.20	0.27	0.41	0.34	0.07	0.14	0.07
	SR =TP / (TP + FP)	78.57	71.43	64.71	70.59	90.00	83.33	88.89
Autumn 2023	Predicted	100.00	100.00	100.00	100.00	100.00	100.00	75.00
	POFD=FP/(FP+TN)	0.32	0.37	0.74	0.55	0.23	0.23	0.18
	SR=TP/(TP+FP)	53.33	46.67	30.43	40.00	50.00	54.55	42.86
Summer–Autumn 2024	Predicted	40.00	33.33	33.33	40.00	0.00	33.33	25.00
	POFD=FP/(FP+TN)	0.55	0.44	0.79	0.74	0.38	0.49	0.11
	SR=TP/(TP+FP)	9.09	27.27	3.33	6.90	0.00	10.00	20.00

Table 4: HER: Predicted events (as in Table 1), the false positive rate of predictions (Probability of False Detection, POFD) and Success Ratio (SR) for each model grid point for zonal current (FP = False Positives, TN = True Negatives, TP= True Positives).

		HER						
Dataset	Metric [%]	A	B	C	D	F	I	P
Summer 2023	Predicted	0.00	100.00	0.00	0.00	7.69	50.00	44.44
	POFD=FP/(FP+TN)	0.00	0.07	0.00	0.00	0.00	0.14	0.07
	SR=TP/(TP+FP)	-	66.67	-	-	100.00	77.78	80.00
Autumn 2023	Predicted	0.00	20.00	0.00	0.00	0.00	44.44	37.50
	POFD=FP/(FP+TN)	0.00	0.23	0.00	0.00	0.14	0.09	0.14
	SR=TP/(TP+FP)	-	16.67	-	-	0.00	66.67	50.00
Summer–Autumn 2024	Predicted	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	POFD=FP/(FP+TN)	0.00	0.00	0.00	0.00	0.00	0.06	0.03
	SR=TP/(TP+FP)	-	-	-	-	-	0.00	0.00

Table 5: KVA: Predicted events (as in Table 2), the false positive rate of predictions (Probability of False Detection, POFD) and Success Ratio (SR) for each model grid point for bottom temperature (FP = False Positives, TN = True Negatives, TP= True Positives).

		KVA						
Dataset	Metric [%]	A	B	C	D	F	H	N
Summer 2024	Predicted	-	-	-	-	0.00	0.00	92.63
	POFD=FP/(FP+TN)	0.00	0.00	0.00	0.00	0.00	0.00	0.40
	SR=TP/(TP+FP)	-	-	-	-	-	-	73.76
Summer 2024	Predicted	-	-	-	-	0.00	0.00	90.95
	POFD=FP/(FP+TN)	0.00	0.00	0.00	0.00	0.00	0.00	0.70
	SR=TP/(TP+FP)	-	-	-	-	-	-	62.26

Table 6: HER: Predicted events (as in Table 2), the false positive rate of predictions (Probability of False Detection, POFD) and Success Ratio (SR) for each model grid point for bottom temperature (FP = False Positives, TN = True Negatives, TP= True Positives).

		HER						
Dataset	Metric [%]	A	B	C	D	F	I	P
Summer 2024	Predicted	73.09	72.12	60.33	52.29	79.15	75.77	77.87
	POFD=FP/(FP+TN)	0.0412	0.0469	0.0078	0.1045	0.0505	0.0568	0.1608
	SR=TP/(TP+FP)	80.40	80.65	91.74	69.57	83.38	81.02	71.90
Autumn 2024	Predicted	73.68	65.04	0.00	71.07	58.46	69.62	57.42
	POFD=FP/(FP+TN)	0.0237	0.0380	0.0000	0.0552	0.0394	0.0345	0.0764
	SR=TP/(TP+FP)	72.41	61.07	-	61.08	69.51	70.97	71.33

- *Line 78: The parenthesis around (cold water) are unnecessary.*

  The sentence will be corrected:

  "at the intake pipe during summer (cold spells)"

- *Line 102: Sentence is unclear.*

  The sentence will be edited (please note that Fig. 2 will become Fig. 3 in the revised version of the manuscript):

  "~~An i~~During Poly and Hans, an intense upwelling occurred along the western Baltic coast leading to a ~~drop of~~decrease in the sea surface temperature (SST), which for Hans ~~were extending over~~ extended over 100 km offshore (Fig. ~~2~~3a,d)."

- *Line 106: personal → personnel*
  The text will be changed as suggested.
- *Figure 1: In *(b)* and (d), red sticks indicate…*
  The text will be changed as suggested.
- *Figure 2: Superimposed on panels (a,*d*)*
  The text will be changed as suggested.
- *Line 135: Please clarify whether the symbol should be strictly less than () or less than equal to (≤).*

  The symbol should be ≤, it will be specified.
- *Line 142: The last sentence by itself is unclear. Perhaps the authors are contrasting their linear regression approach with more complex regression models? It will be helpful to expand on this sentence, and provide examples of higher complexity models and alternative approaches. Inclusion of additional literature review summarizing other common methods used for similar hazard assessment will also be helpful, along with their current limitations to serve as additional motivations of this study.*

  Thank you for your comment. This sentence (which is in the Introduction section) is actually pertaining to the Discussion, so we will remove it from the Introduction. In the Discussion, we already discuss the possible extensions of our approach in the submitted version of the manuscript **lines 601-624**.
- *Line 167: …might lead *to* the accumulation of…*

  The text will be changed as suggested:

  "The inner shelf circulation might lead to the accumulation"
- *Line 193: accumulation of → its accumulation*

  The text will be changed as suggested:
  "and the convergence associated with downwelling and inner shelf (see [1] and [2]) will cause its accumulation ~~of~~ at the coast "
- *Line 218: Can the authors add information about how "quality-marked observations" are marked/determined?*

  Quality flags are determined by the data provider (SMHI), which specify that data with no "good quality" can be due to the station being out of service for a period, or the value needing manual review before it can be approved. The information about the flags are specified by the provider (SMHI) at the website where the data can be downloaded and the website is already specified in **line 215-216**: `https://www.smhi.se/data/hitta-data-for-en-plats/ladda-ner-vaderobservationer/airtemperatureInstant/68560`. The website is already give in the text. So we think that no extra information is needed in the text, as the reader can refer to the website for specification about how quality flags are determined.
- *Line 224: Can the authors add why the data cannot be used directly?*

  The reason is because the observations of temperature and salinity are not taken at the sea, but inside the intake station on land. We will reformulate the sentence **at line 220**:

  "In-situ temperature (ITS-90) and conductivity measurements were conducted by SMHI inside the intake station of the desalination plant in Kvarnåkershamn using Sea-Bird Electronics Inc. SBE 37-SM at sampling rate of 600 seconds in the period 29 June-13 September 2023. The raw data were processed using SeatermV2 2.8.0.119 software. These data are used to describe the temperature and salinity conditions ~~at~~inside the intake station on land and ~~related to the ocean model results~~are used to compare with the ocean model results (sect. 3.1.3), but they ~~cannot be used directly to~~do not represent the oceanographic conditions in the sea outside the intake station~~or~~, so they cannot be used to evaluate the performance of the ocean model."

- *Line 241: Sentence is unclear.*

Thank you for catching this, we swallowed some verbs and information in this sentence. We will rephrase the sentence:

"These products employ the biogeochemical model ERGOM (Neumann et al., 2022), which is one way coupled to the aforementioned ~~to the~~ physics' forecast and reanalysis models~~mentioned above~~. The output variable "chlorophyll" from ERGOM was used as a proxy for phytoplankton bloom intensity in a general overview of regional response (Fig. 2, sect. 1.2). ~~This product is no~~The biogeochemical model is not validated close to the coast~~for~~, so we do not use it for detailed analyses close to the desalination plants. "

- *Eqs 1, 2: Can the authors include complete descriptions of notations, e.g., Tr, Cov, Var, etc.?*

Yes, the whole section will be expanded for clarity and we will add references that used the vector correlation previously:

"The correspondence between the wind observations and the wind forcing can be evaluated through the squared vector correlation ρ_v^2 (Crosby et al., 1993):

$$\rho_v^2 = \mathrm{Tr}\left[(\Sigma_{11})^{-1}\Sigma_{12}(\Sigma_{22})^{-1}\Sigma_{21}\right] \tag{1}$$

where ~~$\vec{u^w}_{(obs)} = (u^w_{(obs)}, v^w_{(obs)})$ are~~ Tr denotes the trace operator. The matrix blocks Σ_{ij} are defined in the following for the observed wind vector time series at Hoburg or Östergarnsholm $\vec{u^w}_{(obs)} = (u^w_{(obs)}, v^w_{(obs)})$, and the AROME wind vector time series at ~~every~~ each ocean model grid point $\vec{u^w}_{\mathrm{AROME}} = (u^w_{\mathrm{AROME}}, v^w_{\mathrm{AROME}})$:

$$\Sigma = \begin{pmatrix} \mathrm{Var}(u^w_{\mathrm{obs}}) & \mathrm{Cov}(u^w_{\mathrm{obs}}, v^w_{\mathrm{obs}}) & \mathrm{Cov}(u^w_{\mathrm{obs}}, u^w_{\mathrm{AROME}}) & \mathrm{Cov}(u^w_{\mathrm{obs}}, v^w_{\mathrm{AROME}}) \\ \mathrm{Cov}(v^w_{\mathrm{obs}}, u^w_{\mathrm{obs}}) & \mathrm{Var}(v^w_{\mathrm{obs}}) & \mathrm{Cov}(v^w_{\mathrm{obs}}, u^w_{\mathrm{AROME}}) & \mathrm{Cov}(v^w_{\mathrm{obs}}, v^w_{\mathrm{AROME}}) \\ \mathrm{Cov}(u^w_{\mathrm{AROME}}, u^w_{\mathrm{obs}}) & \mathrm{Cov}(u^w_{\mathrm{AROME}}, v^w_{\mathrm{obs}}) & \mathrm{Var}(u^w_{\mathrm{AROME}}) & \mathrm{Cov}(u^w_{\mathrm{AROME}}, v^w_{\mathrm{AROME}}) \\ \mathrm{Cov}(v^w_{\mathrm{AROME}}, u^w_{\mathrm{obs}}) & \mathrm{Cov}(v^w_{\mathrm{AROME}}, v^w_{\mathrm{obs}}) & \mathrm{Cov}(v^w_{\mathrm{AROME}}, u^w_{\mathrm{AROME}}) & \mathrm{Var}(v^w_{\mathrm{AROME}}) \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{2}$$

where Var, Cov denote the variance and the covariance operators, respectively. The vector correlation quantifies how much two time series of vectors are interdependent, see Davini et al. (2012); Pelland et al. (2016); Ferrari et al. (2017) for example applications. ~~In other words, t~~The perfect vector correlation ($\rho_v = 2$), between two time series of vectors (...)"

- *Line 247: Does the equation assume normal distribution?*

No, the definition of vector correlation does not assume normal distribution. The vector correlation proposed in can be used to evaluate the linear dependence between two time series of vectors $\vec{W}_1$, $\vec{W}_2$, in the case the distribution of $\vec{W}_1$ or $\vec{W}_2$ is unknown or non-standard. This is explained in the method's reference (Crosby et al., 1993) that is cited in the submitted manuscript.

- *Line 254: Why is the vector correlation represented by a single scalar value = 2?*

The vector correlation between two time series of vectors $\vec{W}_1$, $\vec{W}_2$ is defined in the range $\rho_v = [0, 2]$ $\rho_v = 2$ represents the case the two vectors are linearly dependent. It comes from its definition. All this information can be found in (Crosby et al., 1993) that we refer to.

- *Line 254: Typically correlation ranges from 0 to 1, so I am unfamiliar with how a value of 2 is derived and its interpretation. Can the authors add more information about the basis of the equations and interpretation of results?*

As mentioned above, this is explained in the method's reference (Crosby et al., 1993). We do not introduce a new method, the vector correlation is a quite established metrics. We will add a few references as examples for applications of vector correlation (see also the above response): Davini et al. (2012); Pelland et al. (2016); Ferrari et al. (2017).

- *Eq 3, 4: Can the authors include complete descriptions of notations?*

  Thank you for the catch. **Line 295** will be modified as:

  "The ocean currents are decomposed to the zonal component (u, in the zonal direction x, positive eastward) and the meridional component (v, in the meridional direction y, positive northward), and z is the vertical coordinate, positive upward. The material (volume) transport (U, V) is estimated from the depth-integrated ocean model velocity field:

$$(U, V) = \frac{1}{H}\left(\int_{-H(x,\not{z}y)}^{0} u dz', \int_{-H(x,\not{z}y)}^{0} v dz'\right), \tag{3}$$

  where $H(x, y)$ is the local sea depth."
  Please note that we found a mistake in the notation in Eq.3 ($H(x, \not{z}y)$) that has been corrected above.

- *Line 326: We chose *the* winds...*
  The text will be corrected:

  "We choose ~~to~~the winds at HOB station instead of (...)"

- *Line 329: os -¿ is*
  The text will be corrected:

  "(...) while the OST station ~~os~~is located further off-shore."

- *Line 432: missing closing parenthesis*
  Thanks, it will be corrected.

- *Line 483: The data in parenthesis - R2=0.4 and 50 % does not match Figure 7h or any of the other HER panels.*

  Point I is not shown in Figure 7. In the Table 1 we reported $R^2 = 0.46$ for point I in HER, which should be $= 0.5$, not 0.4 (rounding error) so the value must be corrected in the text. The 50 % value is correctly reported in the Table, matching the value reported in the text. The final sentence will be edited as:

  "The best performance is shown by points B and I south of HER (Fig. 7g-h), with $R^2 = 0.3$ (~~0.4~~0.5) and 100% (50%) predicted events at B (I)".

- *Line 489: The authors mention couple of events at HER but list 5 events in Table 1. Can they clarify?*
  The table will be substituted with the new version discussed in previous response.

References

Crosby, D. S., Breaker, L. C., and Gemmill, W. H.: A Proposed Definition for Vector Correlation in Geophysics: Theory and Application, Journal of Atmospheric and Oceanic Technology, 10, 355 – 367, https://doi.org/10.1175/1520-0426(1993)010⟨0355:APDFVC⟩2.0.CO;2, 1993.

Davini, P., Bechini, R., Cremonini, R., and Cassardo, C.: Radar-Based Analysis of Convective Storms over Northwestern Italy, Atmosphere, 3, 33–58, https://doi.org/10.3390/atmos3010033, 2012.

Ferrari, R., Artana, C., Saraceno, M., Piola, A. R., and Provost, C.: Satellite Altimetry and Current-Meter Velocities in the Malvinas Current at 41°S: Comparisons and Modes of Variations, Journal of Geophysical Research: Oceans, 122, 9572–9590, https://doi.org/https://doi.org/10.1002/2017JC013340, 2017.

Pelland, N. A., Eriksen, C. C., and Cronin, M. F.: Seaglider surveys at Ocean Station Papa: Circulation and water mass properties in a meander of the North Pacific Current, Journal of Geophysical Research: Oceans, 121, 6816–6846, https://doi.org/https://doi.org/10.1002/2016JC011920, 2016.