Responses to Review#1

Comment 1*: How have the authors proposed uncertainty be represented in the analysis product? For example, the model was only run for one year and the maps provided average over significant lengths of time for such a widely varying system. The issue surrounding temporal uncertainty- how representative 2021 is, and if the current simulation should be considered that -is important to consider when using these results to decide regualtion. Could the standard deviation over each seasonally averaged period be showcased as a map? Can the authors do the same with salinity from the same embayment and compare that to a longer run they have already? This would at least speak to the interannual variation expected in this region over time. This point is critical for the evaluation of event-based discharges as was asked of us in the review process. The products are currently insufficient as provided to evaluate events as they are largely time-averaged currently.*

Responses: The modeling of the wastewater treatment plant (WWTP) was performed by tracking the concentration of dye released from the outfall. This effort focused initially on the North River, where scenarios were considered using both seasonally averaged and maximum forcing conditions. The dilution maps generated account for both normal and extreme conditions. The dispersion of the outfall tracer is primarily governed by flow advection and mixing. The uncertainty in these processes includes both physical and numerical errors. Physical errors stem from the accuracy of meteorological forcing, while numerical errors are associated with model resolution and the parameterization of mixing and bathymetry.

The meteorological data employed to drive the model were derived from the regional NECOFS reanalysis product, which has assimilated all available observed data to minimize model uncertainty. We conducted extensive tests to evaluate the convergence of the tracer model through comparisons with dye releases and experiments demonstrating the sensitivity of the spatiotemporal variability of the diluted tracer to model resolution. At the regional scale, such as Georges Bank, convergence can be achieved with model resolutions approaching 100 meters (as detailed in our provided paper). In our research on predicting the initial spread of radionuclides from the Fukushima Dai-ichi Nuclear Power Plant, we determined that a model resolution of approximately 5 meters is necessary to accurately resolve the infrastructure of the plant and align with observations. Our WWTP model has been employed by FVCOM users in other states (as referenced), where it was shown that configurations around 100 meters can effectively reproduce the dilution of WWTP effluent.

The bathymetric data used in our WWTP configuration were obtained from high-resolution 1-meter Lidar data and the 5-meter resolution CUDEM (Continuously Updated Digital Elevation Model) dataset, both of which have been thoroughly validated.

We acknowledge that we are not the first to utilize FVCOM to evaluate the impacts of WWTP effluent on New Bedford and Fairhaven. An initial assessment, supported by MIT/WHOI Sea Grant, investigated the influence of wastewater effluent on coastal ocean acidification in Buzzard Bay, led by PIs Scott Doney, Jennie Rheuban, Jim Churchill, and Geoff Cowles. This study involved measurements of dissolved inorganic carbon and total alkalinity, along with nutrients such as NO3- + NO2, NH4+, PO4-3, SiO4-, TN, PON, and POC. The early version of FVCOM used in that study operated at a horizontal resolution of approximately 100 meters, and even with this relatively coarse resolution, the model successfully reproduced the spatial and temporal variability of wastewater influences in alignment with observed data.

For the 2021 simulation, the FVCOM was configured with a horizontal resolution of up to 4 meters based on convergence tests of tracer simulations. Unlike previous models, this update integrates high-resolution satellite sea surface temperature data to enhance predictions of water stratification and is informed by reanalysis meteorological forcing as well as validated boundary conditions from NECOFS. These improvements have significantly reduced the uncertainties within the model. Notably, flow in Buzzard Bay is predominantly affected by tidal currents, and the simulated tides have undergone extensive validation against observational data.

We concur with the reviewer on the necessity of incorporating an assessment of uncertainty within the simulation. This can be accomplished without technical challenges, but it does require personnel support. We can categorize the primary sources of model uncertainty and integrate them into the simulation process. While we don't believe this would result in significant changes to the outcomes, it is valuable to provide stakeholders with explicit ranges of uncertainty.

Regarding interannual variability, we think that the model should be operated in forecast or nowcast mode to adequately track the dilution of wastewater from WWTPs, given the significant spatiotemporal variability observed. In the North River case, we focused on presenting seasonal means and extreme weather scenarios to assess averaged and worst-case conditions, which inadvertently filtered out the short-term (hourly to daily) temporal and spatial variability of the wastewater influent. For outfalls in New Bedford and Fairhaven, we conducted real simulations utilizing discharge records, enabling us to capture both short-term and monthly or seasonal variability. While we did calculate the standard deviation, it was not included in the report. However, we have provided the Division of Marine Fisheries (DMF) with the hourly simulation results to facilitate straightforward access.

We want to clarify that the dilution simulation results serve as an alternative reference for DMF, but the final decision, to our understanding, is based on integrating this data with additional analyses conducted by DMF.

Comment 2: There are several other forms of uncertainty – performance based in the hydrodynamic simulation (which the authors have tried to address by providing substantial model evaluation). This one is the most constrained – but process-based uncertainty remains. Here, the vertical mixing described in the results of the simulation whereby bottom waters were readily *mixed and dispersed at the surface seems like a process that while simulated should also be evaluated somehow. Maybe salinity observations could help? Model based mixing errors are a notorious issue for modelers everywhere and this pathway seems critical to this issue.*

Another process-based uncertainty emerges with the design of the inert tracer. It seems likely that the tracer was allowed to behave passively, when in reality the wastewater would degrade or be naturally broken down as it is exposed to light (Dizer et al. 1993, others) as well as it exists in the natural environment. This timescale was not included in the considerations for the residence times or dilution metrics provided.

In addition, only two WWTP discharge outfalls were considered providing additional uncertainty as to how the region would be impacted by more outfalls. A simulation with 26 locations assuming the discharge is like that of the ones they have data from could be useful in the interim to just understand where and when the outfall has the largest impact.

Finally, the analysis currently largely lack of consideration for the habitat or organism at the center of concern and management. While the authors provide some output on the bottom, which seems like the natural location for them to reside, there is not much attention made to the organism's phenology and or their potential ability to clear the material from its system?

Responses: We completely agree that validating mixing versus stratification with observations is essential. By assimilating high-resolution satellite-derived sea surface temperature (SST) data into the model, we observed improvements in vertical stratification at a regional scale. We acknowledge that the contribution of salinity to stratification requires justification with observational data. In our simulation, we included freshwater discharges and found that the dispersion of wastewater influent is primarily driven by tidal and wind-induced flows, which are orders of magnitude stronger horizontally than vertically. Although stratification can affect the vertical dispersion of the tracer, its influence is of secondary importance. We believe that both temperature and salinity measurements around the outfalls can assist in further model validation.

Regarding the conservative tracer simulations, we did not account for the degradation of wastewater concentration due to light exposure. To address this, it would be necessary to incorporate a water quality model that considers both biological and chemical processes. We have developed the Northeast Coastal Ocean Acidification Model (NeBEM), which was initially validated in Massachusetts Bay, incorporating the Boston Harbor outfall. In comparisons with UG-RCA, a water quality model used to assess the impact of the MWRA outfall in Boston Harbor, we found that NeBEM yielded more promising results. We could utilize NeBEM to address the reviewers' comments if additional funding is available to support the necessary personnel.

While the New Bedford and Fairhaven outfalls are two significant WWTP influents, there are 26 wastewater outfall sites along the coast. Unfortunately, we lacked comprehensive discharge data from these 26 outfalls at the time of the simulation. The data we received indicated that discharges from these outfalls mainly occur during wet weather events with significant rainfall. Despite considering only two major outfalls, our model predicts a significant influence of WWTP influents on local marine environmental conditions. The situation could become even more severe temporarily if all 26 outfalls were added. Technically, it would not pose a problem to include these additional outfalls in the model simulation, provided the relevant data is available.

As mentioned in our response to the major comment above, the dilution simulation results serve as a reference point for DMF (Division of Marine Fisheries), but we understand that the final decision will be based on integrating this model data with additional analyses conducted by DMF, which consider the habitat and organisms involved. If we wish to enhance the model's role in assessing the impact of WWTP influents, we should incorporate water quality modeling, like what we did for the MWRA outfall in Boston Harbor. However, this integration requires personnel support.

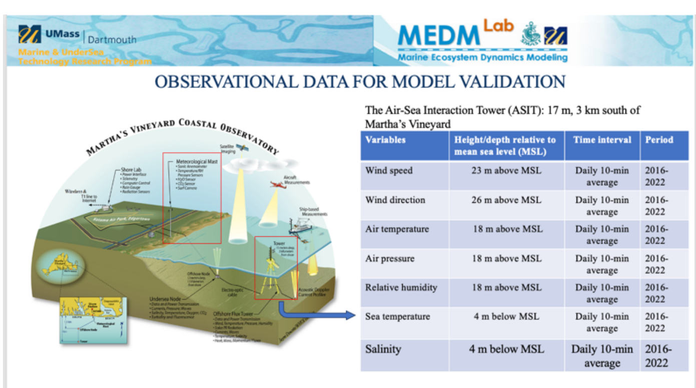
General Questions for the report for further consideration:

QS. 1: Spinup time 1-2 weeks is short. Do the authors begin that simulation from rest or from a semi-spun up state based on the forecast model at coarser resolution?

Answer: The model was initialized using our 39-year hindcast simulation results from NECOFS, which encompasses Buzzard Bay with a horizontal resolution of approximately 100 meters. Thus, it is not spun up from a rest state. The primary justification for the spin-up period lies in the water exchanges between the bay and wetlands, which allows for the system to reach an equilibrium state over a few days (approximately 5-6 semi-diurnal tidal cycles). Therefore, the 1-2 weeks of spin-up time is adequate given the hindcast-generated initial and boundary conditions.

QS. 2: What is the resolution of the atmospheric forcing from WRF and does it sufficiently cover the nooks and crannies of the complicated coastline simulated?

Answer: The atmospheric forcing from WRF operates at a resolution of 3 km. We have compared the WRF-predicted wind with observed data along the coast in previous years. The WRF model employed in these comparisons is the same one used for the WWTP project, though we increased the resolution to 1 km in the offshore wind farm development region. The simulated wind at the observational sites is approximately 3 km, and the results are very promising. We have included some slides from our conference presentations to illustrate the model's performance.

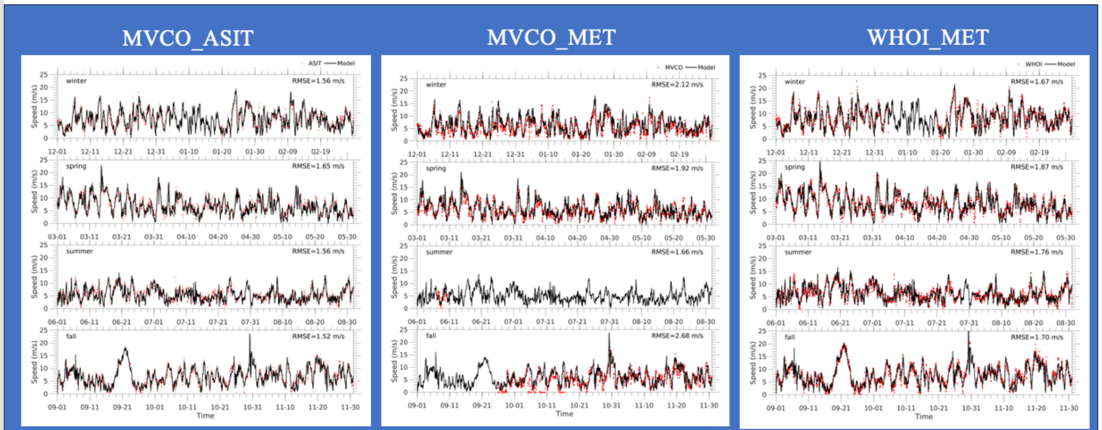


Variables	Height/depth relative to mean sea level (MSL)	Time interval	Period
Wind speed	23 m above MSL	Daily 10-min average	2016-2022
Wind direction	26 m above MSL	Daily 10-min average	2016-2022
Air temperature	18 m above MSL	Daily 10-min average	2016-2022
Air pressure	18 m above MSL	Daily 10-min average	2016-2022
Relative humidity	18 m above MSL	Daily 10-min average	2016-2022
Sea temperature	4 m below MSL	Daily 10-min average	2016-2022
Salinity	4 m below MSL	Daily 10-min average	2016-2022

UMass Dartmouth
Marine & UnderSea Technology Research Program

MEDM Lab
Marine Ecosystem Dynamics Modeling

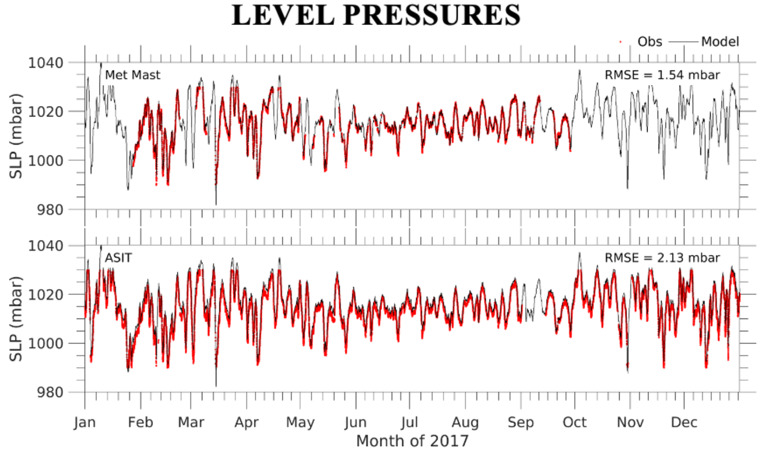
WIND SPEED COMPARISONS AT18-23 M ABOVE MEAN SEA LEVEL (MSL)



UMass Dartmouth
Marine & UnderSea Technology Research Program

MEDM Lab
Marine Ecosystem Dynamics Modeling

COMPARISONS BETWEEN MODEL-SIMULATED AND OBSERVED SEA LEVEL PRESSURES

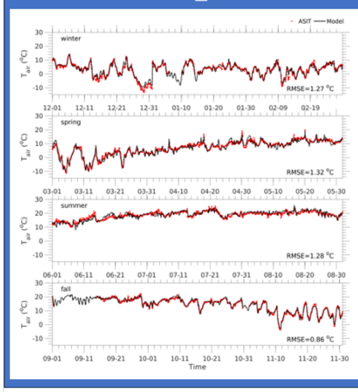


UMass | Dartmouth
Marine & UnderSea Technology Research Program

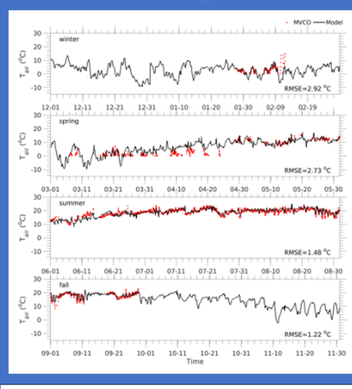
MEDM Lab
Marine Ecosystem Dynamics Modeling

AIR TEMPERATURE COMPARISONS AT 18 M ABOVE MSL

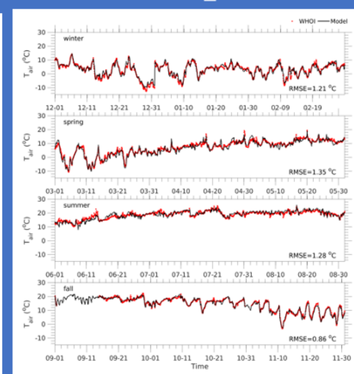
MVCO_ASIT



MVCO_MET



WHOI_MET

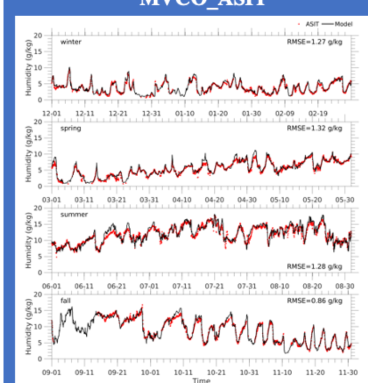


UMass | Dartmouth
Marine & UnderSea Technology Research Program

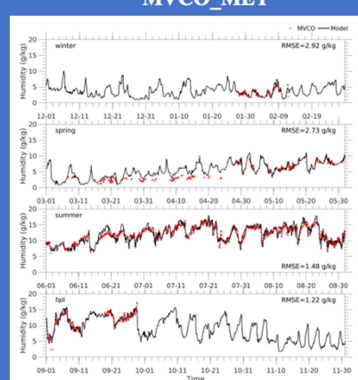
MEDM Lab
Marine Ecosystem Dynamics Modeling

AIR HUMIDITY COMPARISONS AT 18 M ABOVE MSL

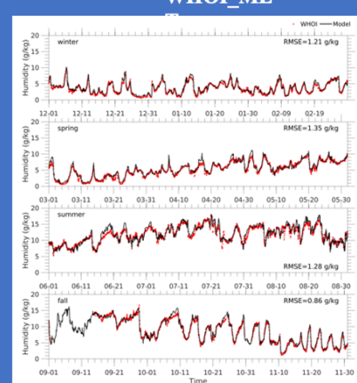
MVCO_ASIT



MVCO_MET



WHOI_ME



QS. 3: Where did they get the WWTP data to force the rivers and are other variables available like temperature and salinity? Were those properties included for the WWTP discharge in the simulation?

Answer: It appears there may be some misunderstanding regarding our methodology. We did not utilize WWTP data to directly force the river discharges. Instead, we implemented the WWTP discharge through the FVCOM river discharge numerical module. All river discharges incorporated into the model included temperature and salinity data, which differ significantly from the passive tracer used for the WWTP discharge.

The temperature for river discharges is derived from a hydrological model known as the "Water Balance Model" (WBM). The simulated temperatures from WBM have been validated against observed river temperature measurements. Salinity, on the other hand, is set to zero, as the input sources are located far inland from the coast.

QS. 4: Why are monthly average maps the appropriate product to use to address this issue?

Answer: We have provided the Division of Marine Fisheries (DMF) with hourly model output data, including animations that illustrate the temporal and spatial variability of the WWTP-diluted waters. However, DMF expressed a preference for monthly averaged dilution maps. They require these averages as an alternative reference to determine which months should be designated for closures or openings.

While the monthly average maps offer a simplified overview, the temporal and spatial variability of diluted waters is still considered in DMF's analysis. It's important to note that this report serves as a summary of the modeling activities. We also hold regular meetings with scientists and managers at DMF to discuss the model results in detail.

We just want to clarify again that the dilution simulation results serve as an alternative reference for DMF, but the final decision, to our understanding, is based on integrating this data with additional analyses conducted by DMF

QS. 5: What do the authors consider for year-to-year variation in the dilution value simulated?

Answer: Yes, we do take interannual variability into account, primarily based on the seasonal variability of winds, even though we conducted a single-year simulation for the outfalls in New Bedford and Fairhaven. For the outfall in the North River, we analyzed seasonal averages and extreme weather conditions over a 40-year period from 1978 to 2017.

In our findings for the outfalls in New Bedford and Fairhaven, we observed that the distribution of WWTP-diluted water varied significantly with changing wind patterns, in addition to tidal influences. This underscores the importance of conducting a real-time simulation over a year. We propose to continue running simulations in successive years to better examine interannual variability; however, this effort is constrained by funding limitations.

QS. 6: The simulation was only for one year – how was 2021 chosen and why is it representative?

Answer: That's a good question. The modeling was conducted under a contract with DMF, and the selection of the year 2021 was made through a review process led by DMF.

2021 was chosen because it experienced a range of representative weather conditions, including Nor'easter storms (e.g., the Nor'easter on January 4), hurricanes (e.g., Hurricane Henri on August 22), and heatwaves. Additionally, DMF has comprehensive WWTP discharge records for both New Bedford and Fairhaven outfalls, which indicate significant variability in discharge rates, ranging from 15 MGD to 60 MGD at the New Bedford outfall. This variability is particularly pronounced during the summer months when water stratification occurs.

QS. 7: *Are all the WWTP sites that exist in the region considered? Or are some not included in the simulation*?

Answer: See our responses to Comment 2.

QS. 8: *Could Figure 5 be compared to the WRF model used? How well does it perform*?

Answer: See our answer to QS. 2.

QS. 9: How was dilution determined in the simulations? What definition of residence time was considered?

Answer: Dilution in the simulations was determined by analyzing the concentration of WWTP water over the simulation period. This can be represented as a snapshot or through hourly, daily, monthly, or seasonal averaging of the concentration.

Since the WWTP water is injected continuously throughout the year, based on the discharge records, rather than a single point injection at a time, the residence time is calculated based on the total influent of the WWTP water in the system at any given moment. This approach allows for a more accurate representation of how the WWTP discharge interacts with bay water over time.

QS. 10: How are the modelers conveying uncertainty in their estimates? There are several estimate of various kinds of uncertainty provided – performance estimates as well as variance of the target variable over a year. Can the modelers combine these somehow or translate these into a confidence interval somehow? There are some examples of this in Kessouri et al (year?).

Answer: See our responses to Comment 2.

QS. 11: Doesn't fecal coliform bacteria degrade with UV? Did the modelers consider adding this kind of decay to their tracer design? A sensitivity to this kind of forcing – an example can be found in these works (Kragh et al. 2022; Delre et al 2023) could be implemented in the longer run and help contribute to the process based evaluation or the uncertainty/confidence interval discussion.

Answer: See our responses to Comment 2.

QS. 12: Could the modeling team add the WWTP sites that served as sources for the dye experiments to Figures 6-9?

Answer: Yes, we can add those WWTP sites to Figures 6-9. We have re-conducted the experiments with improved forcing conditions and refined coastal geometry in the Fairhaven bridge area and along the coast. Additionally, we have included the WWTP discharge from the Dartmouth outfall in our analysis. These updates will be reflected in the revised report.

QS. 13: Even if dye was not released from all the WWTP, was the freshwater delivery considered from all the FW sources depicted in the earlier figure?

Answer: In our simulation, the WWTP discharge was treated as dye rather than as freshwater. While we accounted for the contribution of WWTP discharges, our focus was specifically on their role as sources of dye. Freshwater discharges from rivers can alter water density and subsequently affect water currents; however, the WWTP discharge was regarded solely as dye for the purposes of this study.

QS. 14: What about human engineered barriers that have been inserted into the system -like the storm/flood walls around the harbor etc? are these physical features included in the simulation's high resolution LIDAR based bathymetric forcing? Are they an issue for the WWTP dispersal?

Answer: Yes, the model incorporates human-engineered barriers, including storm and flood walls around the harbor. The model grid covers both land and ocean areas using dry and wet treatment to accurately simulate storm-induced coastal inundation. Additionally, the grid is configured based on high-resolution LIDAR bathymetry data at a 1x1 meter resolution, allowing for precise representation of these physical features and their potential effects on WWTP dispersal.

Recommendations –

1. *Consider expanding the simulation beyond one year and expanding the uncertainty estimates into a confidence interval that includes several aspects of modeled uncertainty surrounding the dilution metric. This could be achieved either through expanding the current run, or by looking in the current run at the extent of a particular isohaline associated with the dilution metric and then using that isohaline in longer simulations to explore uncertainty, its not clear what simulations this study has access to be this group has a lot of experience and model fields already to draw from.*

Response: That is exactly what we propose to do. However, this will require sufficient personnel support.

2. *Consider augmenting the evaluation with some other evaluation of the WWTP outfall – like salinity or temperature observations from the discharge to instill confidence in the simulated fine scale mixing implicated in the distribution of the pollutants by the modeling team in their explanation in the report.*

Response: Temperature and salinity measurements could provide valuable information for evaluating the hydrodynamic model, particularly regarding the accuracy of water stratification. This can be accomplished with support to conduct hydrographic surveys.

3. *Work to further define this threshold with the biology in mind. For example, the WWTP discharge is highest in the autumn in New Bedford but the model suggests that the dilution thresholds experience seasonality and are most severe in winter through spring in 2021. If you were to bring in the biology of the oysters and other shellfish – what months are they harvested, when are they growing the fastest and eating the most? How long does the bacteria from the WWTP remain a threat to humans within the shellfish after they ingest it? Could the farms remain open but harvest delayed in certain months?*

Response: his can be addressed by incorporating a water quality model simulation alongside the current dye-dilution-based WWTP model, which also requires sufficient personnel support.

4. *In line with this last question above – the model estimates are largely provided at the surface with some attention paid to the bottom conditions, but where are the farms mostly? Are the organisms growing at the surface in the intertidal or are they mostly at the bottom and what depths typically do they reside? Can the dilution estimates be expanded in the subsurface to better align with their habitat? Is bottom appropriate or do the farms allow the cages to reside higher up in the water column?*

Response: The WWTP model estimations involved dye being injected throughout the water column rather than solely at the surface. The model output includes a comprehensive 3-D spatiotemporal distribution of the diluted water. The figures in the report are illustrative examples depicting distribution at the sea surface. We can analyze conditions at any depth in the water column. Based on our findings, the largest spread of the WWTP-diluted water occurred at the sea surface.

5. *The forecast recommendation from the author team at the end seems useful to help build out some of the products described above in the future. If this was to proceed, then convening regular meetings with managers and the forecasters to co-produce and co-design the products expected and build trust in the model simulation seem warranted.*

Response: If the forecast system is established, it is essential to hold regular meetings with managers to ensure that the products meet management requirements.