Comment on egusphere-2025-5050

This is an interesting study that applies statistical techniques in new ways to predict high frequency (10 min resolution) changes in TA and DIC in an inlet of the Wadden Sea. The methods are instructive, and include time series kriging (with a semivariogram) for uncertainty analysis, Wasserstein distances to reveal hydrological vs biological influences on DIC, and multiple linear regression for the prediction of total alkalinity. I enjoyed reading about the approach that was used to derive high frequency predictions of TA and DIC from the observations of S, T, and pH.

We thank the reviewer for their thoughtful assessment of our manuscript and for their positive feedback on our methodological approach.

All replies below reference changes made in the track-changes version of the manuscript.

Co-reviewer 1:

The physical-hydrological component of this study could be more clearly described.

- Ridderinkhof et al., (1990, 2002) describe inflow to the Wadden Sea occurring primarily on the south side of the Marsdiep channel and outflow from the Wadden Sea occurring on the North side of the channel - where these measurements were made. Therefore, these data would preferentially show Wadden Sea outflowing waters. How does this alter (if appropriate) your interpretation of results?

We thank the reviewer for raising this important point. While the reviewer is correct that our monitoring location on the norther side of the Marsdiep inlet is positioned in the residual outflow area, this does not significantly bias our observations towards Wadden Sea waters in the way the reviewer is suggesting. It is important to make a distinction between residual flow patterns and tidal flow patterns. the residual circulation pattern as described in Ridderinkhof et al.,(1990, 2002), also confirmed by long-term ADCP observations (van der Molen et al., 2022) refers to the net flow that is averaged over many tidal cycles, where there is indeed an outflow on the northern side of the Marsdiep inlet.

In this inlet, the integrated residual flows have decreased from approximately 1000 m^3s^{-1} in 2009 to near zero in recent years, while volume transports at peak tidal currents reach 5-10*10^4 m^3s^{-1} (van der Molen et al., 2022). These residual flows are therefore orders of magnitude smaller compared to the tidal volume transports. Because of the semi-diurnal tides in the Marsdiep, the two high and two low tides per day, creates strong bidirectional flows that dominates the water exchange at our measurement location. Additionally, our data clearly demonstrate this water exchange (for instance the salinity and alkalinity patterns observed during high and low tides).

While there might be some bias towards Wadden Sea in terms of net volume over time, our measurement with a high-frequency measurement of 10-minute capture both North Sea and Wadden Sea waters as they alternate with the tide.

This is now further elaborated in the Introduction (lines 83–87) and in the Materials and Methods Sections 2.1 (lines 108–110).

Suggested reference:

van der Molen, J., Groeskamp, S., and Maas, L. R. M.: Imminent reversal of the residual flow through the Marsdiep tidal inlet into the Dutch Wadden Sea based on multiyear ferry-borne acoustic Doppler current profiler (ADCP) observations, Ocean Sci., 18, 1805–1816, https://doi.org/10.5194/os-18-1805-2022, 2022.

- Hoppema (1990, 1993) is cited to suggest that different Wadden Sea waters become more or less dominant at different times. This seems vague. Different water masses are introduced at the end of the paper - beginning at line 812. I'd recommend including clarifying information on the water masses in 4.1.1.

We appreciate the reviewer's suggestion. We note that the three water masses (Wadden Sea, Ijsselmeer, and the North Sea) are identified in the paragraph immediately following the Hoppema (1990, 1993) citations within section 4.1.1 (see lines 556-558). However, to improve clarity, we have revised the text to introduce these water types where they were initially referred to (see lines 547-549).

Additionally, we have revised the terminology throughout the manuscript from "water masses" to "water types," as "water mass" is not appropriate for surface waters in this context. The revisions can be found in lines 442, 553, and 557.

- minor point: 4.1.1, "endmember" seems to be used both correctly - to refer to the riverine and North Sea "source" waters, and incorrectly, at line 514 to refer to the high (TA or salinity) waters observed in the inlet.

We have revised the text, as reflected in lines 544-545.

- minor point: Line 588, Salt et al., (2014) was cited for showing similar nutrient patterns for the Marsdiep channel, but their study reports pH and CO2 for the North Sea.

We were referring to Chapter 3 of Leslie Salt's PhD thesis, which analysed nutrient concentrations in the Marsdiep basin of the Wadden Sea where our data were collected. We have corrected the reference to cite the thesis chapter rather than the 2014 article (see lines 624, 625 628 in section 4.13, and lines 1143-1142 in the References lists).

Figure 4. Please include Kw, to understand how much of flux is driven by wind.

We thank the reviewer for this suggestion. As Figure 4 already contains substantial information, we have added the gas transfer velocity data to the supplementary materials (Fig. S5) to maintain clarity. This new supplementary figure is now referenced in the main text in lines 604-606. We have also renumbered all subsequent supplementary figures accordingly in both supplementary material and the main text.

Figure 4. I found the legend misleading. High frequency TA and DIC were both predicted, not "continuously measured." Perhaps the color used to represent them could be changed to help clarify this.

We thank the reviewer for this comment. We have revised the Figure 4 caption to explicitly distinguish between measured and calculated variables, while maintaining the blue colour scheme for consistency. The caption now provides more details to clearly indicate which parameters are continuously measured and which are calculated (see lines 394-401)

4.1.3. The ΔTA/ΔDIC discussion is interesting, but it also seems under-developed. Please include references here that could provide more confidence in interpreting these results. How do we know that the winter-to-summer decrease in the ratio isn't simply driven by the reduced rainfall and river input in summer?

We thank the reviewer for this insightful comment. We have revised this section to include additional references and to discuss the potential role of freshwater input.

We acknowledge that without direct measurements of freshwater endmember TA and DIC, we cannot completely rule out a freshwater contribution to the seasonal δTA/δDIC pattern. However, the fact that TA and DIC co-vary seasonally rather than diverging suggests that mixing alone is not the primary driver. Furthermore, the negative September value (-0.08) means DIC is increasing while TA stays flat or decreases, which only makes sense if respiration is happening. September therefore represents a transition zone between the photosynthesis-dominated summer and the sediment-driven winter signal.

The seasonal pattern is instead consistent with known biogeochemical stoichiometry: the winter δTA/δDIC = 0.89 aligns with sedimentary processes such as denitrification (Chaillou et al., 2024), similar to values reported in other coastal systems (Yau et al., 2022; Miao et al., 2025), while the September ratio is consistent with aerobic respiration stoichiometry (Chaillou et al., 2024).

We have revised and elaborated on this section, see lines 635-638 and 643-65. We have also revised the terminology throughout, replacing ΔTA/ΔDIC with δTA/δDIC. We added the following references in the references lists:

Chaillou, Gwénaëlle, Gwendoline Tommi-Morin, and Alfonso Mucci. "Production and fluxes of inorganic carbon and alkalinity in a subarctic subterranean estuary." *Frontiers in Marine Science* 11 (2024): 1323463.

Yau, Yvonne YY, Pei Xin, Xiaogang Chen, Lucheng Zhan, Mitchell Call, Stephen R. Conrad, Christian J. Sanders, Linwei Li, Jinzhou Du, and Isaac R. Santos. "Alkalinity export to the ocean is a major carbon sequestration mechanism in a macrotidal saltmarsh." *Limnology and Oceanography* 67 (2022): S158-S170.

Miao, Yanyi, Bin Wang, Jacob Carstensen, Dewang Li, Xiao Ma, Qianwen Sun, Zhongsheng Xu et al. "Diverging relationships between acidification and hypoxia off the Changjiang Estuary." *Journal of Geophysical Research: Oceans* 130, no. 11 (2025): e2025JC022675.

Figure 8. It would be helpful to show the tidal height, or to at least indicate the approximate locations of low tide and high tide.

We have added tidal height information to Figure 8 and explained it in the figure caption, see lines 798.

Co-reviewer 2:

Fig 1 – the + sign is quite small & hard to see (esp. if printed in black & white), so it would be helpful if you could label the Marsdiep Channel and maybe make the jetty indicator more clear.

We have added a label for the Marsdiep channel on the map and increased the size of the cross symbol used to indicate the jetty location. Additionally, we have improved the map by adding a distinct colour for the IJsselmeer, which was not clearly distinguishable from the surrounding land in the previous version (See lines 118-119).

Line 120 - The authors say they removed all data that was not collected directly after the cleaning process which happened every 10 minutes. Was there a significant effect of biofouling in the 10 minutes between cleans? Was this necessary?

This approach was implemented as a precautionary measure to minimize any potential biofouling effect. However, the 10-minute interval is short enough that significant biofouling is unlikely to develop. Additionally, a 10-minute resolution is sufficient for our study, as the processes of interest (tidal cycles, diel variations) occur on timescales of hours rather than minutes. This sampling strategy also ensures consistency, as the time between cleaning and measurement remains constant throughout the dataset. Furthermore, this is a standard procedure developed based on decades of experience with other types of sensors (eg., temperature, salinity, and fluorescence) deployed at the jetty.

Line 120 - By placing the equipment at a fixed depth with respect to NAP, the depth of the instruments is changing with respect to the water surface, and the samples are taken at the surface. How can you be sure there is no stratification, which would mean that (a) your sensors are measuring different water masses at different points in the tidal cycle, and (b) your samples are not measuring the same water mass as your sensor? The authors cite Buijsman and Ridderinkhof, 2008 to support the claim that the water column can be assumed to be fully mixed but there is no clear supporting evidence of that in this paper. The other references (Otto 1990, Postma 1954) also don't support this – Otto 1990 explicitly excludes the Wadden Sea in its analysis and Postma 1954 is from over half a century ago. Also, the non-zero coefficient for water level and water level difference in equation (4) (TA-salinity reference) suggests that the water is not fully mixed. If the authors have any data that would better support the assumptions of mixing, that should be included; if not, they will need to address and quantify the uncertainties that this creates.

We thank the reviewer for this important and detailed comment. To address the concern about potential stratification, we analysed vertical profiles of temperature, salinity, and density (both measured and recalculated using the GSW toolbox with the TEOS-10 equation) from CastAway CTD casts collected at the study site. Measurements were conducted over multiple days across different seasons, through the years, with several casts per day, providing a comprehensive view of the vertical structure of the water column under varying conditions.

The profiles revealed that the water column rarely exceeded 3.5 m depth, and virtually no systematic variation in temperature, salinity, or density was observed between the surface and bottom across the vast majority of casts. The highest density difference observed was 1.7 kg m^{-3}, recorded on February third, 2021 (between 2.2 and 3.2 meters), which represented an exceptional case. Apart from this day, the density difference between surface and bottom across all years and seasons was less than 0.1 kg m^{-3}.

For reference, Groeskamp et al.(2011) documented stratification in the Marsdiep channel, a much deeper system (~30 m) with density anomalies ranging from 16 to 24 kg/m^3 and noted that mixing due to wind and tidal currents is often so vigorous that the water is usually vertically well-mixed. In that study, the upper ~5 m of the water column was generally well mixed and no stratification was observed (see their Fig. 4), with tidal mixing due to strong flood and ebb currents leading to a vertically well-mixed water column. Even in the deeper part of the water column, stratification was a transient phenomenon occurring primarily during slack tide, lasting between

20 minutes and 3 hours, before strong tidal currents re-established well-mixed conditions (see their Fig. 4). The density differences and water depths at our shallow jetty location are substantially smaller, making sustained stratification far less likely. While brief episodes of weak stratification cannot be entirely excluded, particularly during slack tide, the water column at our study site was predominantly well-mixed, and any stratification would have been very short and of limited magnitude.

We have added this discussion in lines 110–114 and we have deleted the sentence the reviewer refers to in lines 138–139.

We have also added the following reference (see lines 1019-1020):

Groeskamp, S., Nauw, J. J., & Maas, L. R. (2011). Observations of estuarine circulation and solitary internal waves in a highly energetic tidal channel. *Ocean Dynamics*, *61*(11), 1767-1782.

Section 2.4. – the language is a bit confusing overall here. First, section 2.4.1 is entitled "pH sensor calibration" but there doesn't seem to be any actual calibration happening here, just the implementation of the Nernst equation. Figure 2 shows raw & "corrected" pH sensor data, but it's not clear what the correction process is. It might be that the "uncertainties" calculated in section 2.4.2 were used as corrections, but this is not explicitly stated anywhere in the text and this needs to be clarified and perhaps re-named to reflect that it is a correction process, not a calculation of uncertainty.

We thank the reviewer for pointing out this confusion. We have revised sections 2.4.1 (see lines 171-173) and 2.4.2 (see lines: 209, and 218-219) to clarify the distinction between the calibration and uncertainty estimation processes. We have also revised Figure 2 caption for clarity, see lines 205-206.

Section 2.4.1 describes the calibration of the continuous pH sensor voltage (EMF) using discrete reference measurements from spectrophotometry. At each calibration point, the Nernst equation is used to back-calculate the reference potential (EMF0), which drifts over time due to sensor aging and fouling. EMF0 is then interpolated between calibration points using PCHIP, and the Nernst equation is applied to convert all continuous sensor voltages into corrected pH values.

Section 2.4.2 is entirely separate and quantifies the measurement uncertainty of the corrected pH values based on their temporal distance from the nearest calibration point using a modified Kriging approach. No additional corrections are applied in this section.

Eq. 4 – what is the physical rationale behind having both a sine and a cosine term in the model predicting TA from S?

We thank the reviewer for this question. The combination of sine and cosine terms allows the model to capture seasonal TA variations with timing of the seasonal maximum.

A single sinusoidal function would constrain the seasonal peak to occur at a fixed point in the calendar year. By including both terms, the model can simultaneously optimize both the amplitude (strength of seasonal variation) and phase (timing of the seasonal maximum) through least-squares fitting, rather than requiring these to be specified before-hand.

Line 414 – "In contrast, months with limited sampling (n=1-2), showed unreliable RMSD estimates, as insufficient observations cannot accurately capture the real monthly variability needed by the model to predict TA. These results reflect a sampling limitation rather than model deficiency." I'm not sure that you can meaningfully decouple a sampling vs model weakness, since you're using the samples to create the model.

The reviewer is correct that we cannot meaningfully separate sampling limitations from model performance. We have revised the text, see lines 439-442.

Line 450 – I'm not sure what is meant by the phrase "However, the measured DIC values exhibited comparable values and patterns" – it sounds like it is comparing DIC_measured to DIC_calc to justify the calculations, but this seems contradictory to the previous two sentences. Note there is a minor typo on line 450, "form" instead of "from".

We have revised the text to better explain our validation approach (see lines 477-480). The key point we wanted to convey is that DIC_measured (obtained from independent laboratory analysis of discrete samples) and DIC_calc (calculated from TA_pred and continuous pH sensor data) represent two completely independent methods. The strong agreement between these independent approaches validates both our TA prediction model and pH sensor calibration.

Abstract - Is it reasonable to make claims about the entire Wadden Sea net CO2 source based on one point measurement?

The reviewer raises a valid point. We have revised the abstract to clarify that our CO_2 flux estimate applies specifically to our measurement location rather than the entire Wadden Sea (see lines 26-27).