This manuscript aims to study the effects of beach morphological changes, hydrodynamic boundary conditions, and hydrogeologic properties on flow regimes, salt distribution, and the potential for mixing controlled chemical reactions in high-energy beach aquifers. The authors achieve this objective by building 2-D transient density-driven groundwater flow and transport numerical models that systematically explore boundary conditions and hydrogeological parameters in a changing beach morphology setting. The authors then conclude that changing beach morphology causes the migration of infiltration and exfiltration locations along the beach transect, leading to moving flow and salt patterns in the subsurface and enhancing mixing-controlled reactions.

I have read the manuscript with great interest. My overall opinion is that the manuscript is well-written but needs some moderate revisions and clarifications. Below, I have listed comments, hoping they may help improve the manuscript's quality.

Dear Reviewer #1 we appreciate your time and thank you for thoroughly reading and reviewing our manuscript. We appreciate the overall positive evaluation. Your comments are very valuable and improve the quality of our manuscript. We have responded to your comments point by point as indicated by Author Comments (AC) in blue, planned changes to the manuscript are indicated by "speech mark".

Kind regards, Rena Meyer

Specific Comments

1. Some clarification is needed regarding the 2-D numerical model approach.
   1. The beach boundary, where freshwater inflow occurs, is close to where flow, salinity, and mixing variations occur. Have the authors considered moving this boundary further inland to remove possible boundary effects?

   AC: We agree with the reviewer that the mixing effect would probably be enhanced if the boundary of fresh water inflow would lie further inland. We will add this point to the limitation section of the study.

   "The location of the beach boundary to which freshwater inflow is assigned is in close proximity to where flow, transport and mixing variations occur. Consequently, it is likely to exert a boundary effect where increased freshwater inflow reduces TDS concentrations. Conversely, if the boundary were to lie further inland, TDS concentrations below the upper beach would be enhanced, as water is allowed to flow further inland during possible flow reversals. However, this only happens during extreme storm flood events, and as these are infrequent and short lived (1-2 days), there is not enough time for additional salt to accumulate as a result of the flow reversal. Furthermore, as this only occurs along the southern vertical boundary of the models, we expect it to have an insignificant effect on the overall flow, transport and mixing patterns in the intertidal zone."

   2. The description of the freshwater inflow boundary is missing from the methods section. Also, why is the value of 0.5 m^3/d/m decided for most cases? Is this value related to the groundwater discharge in the area where the model is based?

AC: The value for the fresh groundwater inflow of 0.5 m^3/d/m will be added to the methodology section. It relates to the subsurface freshwater inflow from the islands' freshwater lens. The value was estimated by Beck et al. (2017) and relates to the fresh groundwater entering from the South with a flow rate of 0.5m^3 per day and meter of shoreline and was estimated from the approximate distance to the groundwater divide and the approximate recharge rate.

"Freshwater (salt concentration = 0 g/l) entering from the islands' inland (freshwater lens) along the vertical Southern boundary was prescribed using a specified flux of 0.5 m^3/per day per meter shoreline as estimated by Beck et al. (2017) and was uniformly distributed across the cells of the first column (Fig. 1)."

3. The authors mention that meteoric groundwater recharge was applied at the upper beach slope above the mean high water line (MHWL) (lines 123 to 124). What is the value of this recharge (350 or 400 mm/a)? These values are missing in Table 1.

AC: The meteoric groundwater recharge at the upper beach is 400 mm/a. We will add the value to the methodology section.

"Meteoric groundwater recharge of 400 mm/a was applied at the upper beach above the MHWL"

4. What is the initial concentration distribution in the model? Did the authors run a warm-up period before calculating the variations?

AC: We understand the concern of the reviewer about the effect of the initial distribution of TDS. The initial TDS concentrations were 0 g/l landwards from the mean low water line and 35 g/l towards the seaside. This is explained in the methodology section.

The total simulation time was 20a. The calculation of the variability of TDS (SD of TDS) was based on the last 10a of simulation time, when the initial concentrations have no significant influence on the dynamic flow and transport patterns any more. Thus, the first 10a simulation time serve as model spin-up period to avoid any impact of the initial conditions on the final model results. We will extent the explanation in the methodology sections accordingly.

"The decision to assess SD TDS and RP$_c$ based on the final 10 years of the simulation period was taken in order to circumvent the potential influence of the initial distribution of TDS, R$_s$ and R$_f$. Hence, the first 10 years of the simulation period serve as model spin-up."

5. In my mind, when the beach morphology changes, part of the model geometry also changes. How do the authors estimate the variations in salinity and mixing patterns with varying geometry?

AC: The topographical changes were not applied to the model grid, this remained constant over time. The topographical changes were used to calculate the hydraulic

heads that are applied to the sea boundary above the LWL (Fig.1). This approach was also used by Greskowiak and Massmann (2021).
We will extent the explanation in the methodology section accordingly, see AC to the next point.

1. Throughout the manuscript, the authors highlight the importance of changes in beach morphology in the dynamics of upper salinity plumes (USPs). However, the modeling choice to account for these changes needs further clarification.

   1. It is unclear how the surface interpolation was made, and given that the authors express the importance of this feature in the modeling choice, more information should be provided on this matter. Thus, I recommend adding more information on the methods of how the interpolation was made and refining the plot in Figure 1 so that the reader can picture how beach morphology changes over time.

   AC: We understand the necessity for clarification of the interpolation method of the beach topography.
   The interpolation of the beach surface was performed as proposed by Greskowiak and Massmann (2021) and their supporting information. The five cross-shore LIDAR scan profiles (as shown in Fig. 1b, obtained for Feb-July), were interpolated to daily increments over a six-month period. Subsequently, the topography was varied in daily steps over half a year and then reversed for the second half of the year. The resulting annual topography was then employed recursively over the 20a simulation period. The resulting time series of daily topography was then used to calculate the hydraulic heads that were applied to the GHB in the intertidal zone.

   We will extent the explanation of the interpolation procedure in the manuscript by:

   "In the intertidal zone (between the MLWL and the upper beach affected by storm floods) the beach surface was interpolated according to the methodology described by Greskowiak and Massmann (2021). Five cross-shore LIDAR scan profiles (as shown in Fig. 1b, obtained for Feb-July), were sampled at 1m resolution and interpolated to daily increments over a six-month period. The topography was then varied in daily increments for six month, and reversed for the second half of the year. The resulting annual topography was applied recursively over the 20a simulation period. Daily topography time series were then used to calculate the hydraulic heads using the tide-average head approach, which were assigned to the sea boundary in the intertidal zone (Fig. 1a, grey box)."

   As suggested by the reviewer, Figure 1 will be updated to better demonstrate how the intertidal topography changes over time:
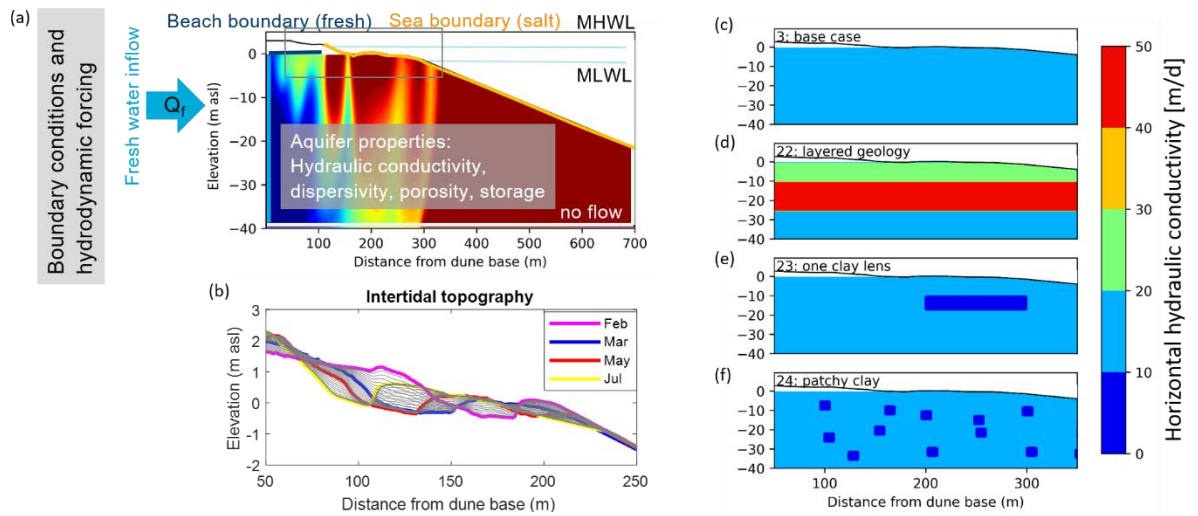
   New Figure 1

Figure 1: (a) Model setup with dimensions, boundary conditions and parameters, (colors indicate TDS distribution: red = saline, blue = fresh). The thin grey box encompasses the part of the sea boundary where changes in beach morphology occur. (b) Four LIDAR scans (Grünenbaum et al., 2020a) of intertidal topography (colored lines) and respective interpolated topographies (grey lines in 10-days increments). (c) – (f) Four different geological settings, note K = 0.005 m/d for the clay lens and clay patches (dark blue).

2. This reviewer understands the study's limitations. Between lines 351 and 355, the authors reference the limitations of linearly interpolating the beach morphology. However, how do the authors reconcile the idea of cycling beach morphology when changes in the hydrologic forcings and hydraulic properties occur? Have the authors considered changing the order of the interpolation in the beach profiles to see if different patterns emerge? Also, the authors only considered a stable case with and without storm floods. Have they considered exploring the variations in hydrogeologic parameters and boundary conditions with stable cases to compare them with the changes in beach morphology?

AC: As we outline in the limitation section we recognized the issue about temporally linear interpolation and cycled morphology which most probably underestimates the mixing in the STE sections. The suggestion of the reviewer to change the order of the interpolation is a good idea and we will consider it for future studies. Given that we already have an extended set of many model cases, we refrain from further extending our model suite in this study. However, we will bring up the reviewers point in the outlook that it would be a valid step for future studies.

"The implementation of real data, including a high-resolution beach topography, as, for example, currently collected from a beach observatory on Spiekeroog (Massmann et al., 2023), in model calibration will further help to better constrain and understand the biogeochemical functioning of high energy STEs. "

The variations in hydrogeological parameters and boundary conditions with stable morphologies have been studied before, e.g. Michael et al., (2016). Therefore, we decided to use the model of Greskowiak and Massmann (2021) as base case and studied the effect of single changes to boundary conditions and parameters to evaluate their individual effect on the flow, transport and mixing dynamics and refrain from

considering combinations of the parameters and boundary conditions. This aspect was mentioned in the limitation section:

"In our generic modelling approach, we systematically changed single aquifer parameters and boundary conditions though some of these parameters may be correlated. At the same time, boundary conditions can be inter-dependent or affected by seasons. We chose this approach to better disentangle single effects and refrain from analyzing combined cases."

2. Regarding mixing-controlled reactions, the authors used the initial salinity patterns for the Rf and Rs concentration patterns. This modeling choice could create misleading mixing patterns because these chemical constituents were assumed conservative, and the model needs to stabilize first. So, did the authors consider warming up the model before setting the initial concentration patterns?

AC: The reviewer is right that the initial distribution of Rf and Rs could have had an effect on the final patterns of their mixing product (Mp). In order to resolve this issue, we have now subtracted the accumulated mixing concentration after 10a (when the initial concentrations of TDS did not influenced final salinity patterns any longer) from the final (20a) simulation concentration. The general mixing patterns look very similar to the previous results, as presented in the new Figure 5 (see below). However, slight changes in normalized concentrations are visible (Fig. 5). Therefore, we decided to follow the reviewers recommendations and update Figure 5 and also re-calculate the cluster analysis (Fig. 6) which also shows slight differences. Moreover, we will modify the corresponding text about the cluster analysis. However, the overall changes were only minor and the general picture and conclusion driven are still valid.

The methodology section will be extended by:

"The decision to assess SD TDS and RP_c based on the final 10a of the simulation period was taken in order to circumvent the potential influence of the initial distribution of TDS, R_s and R_f."
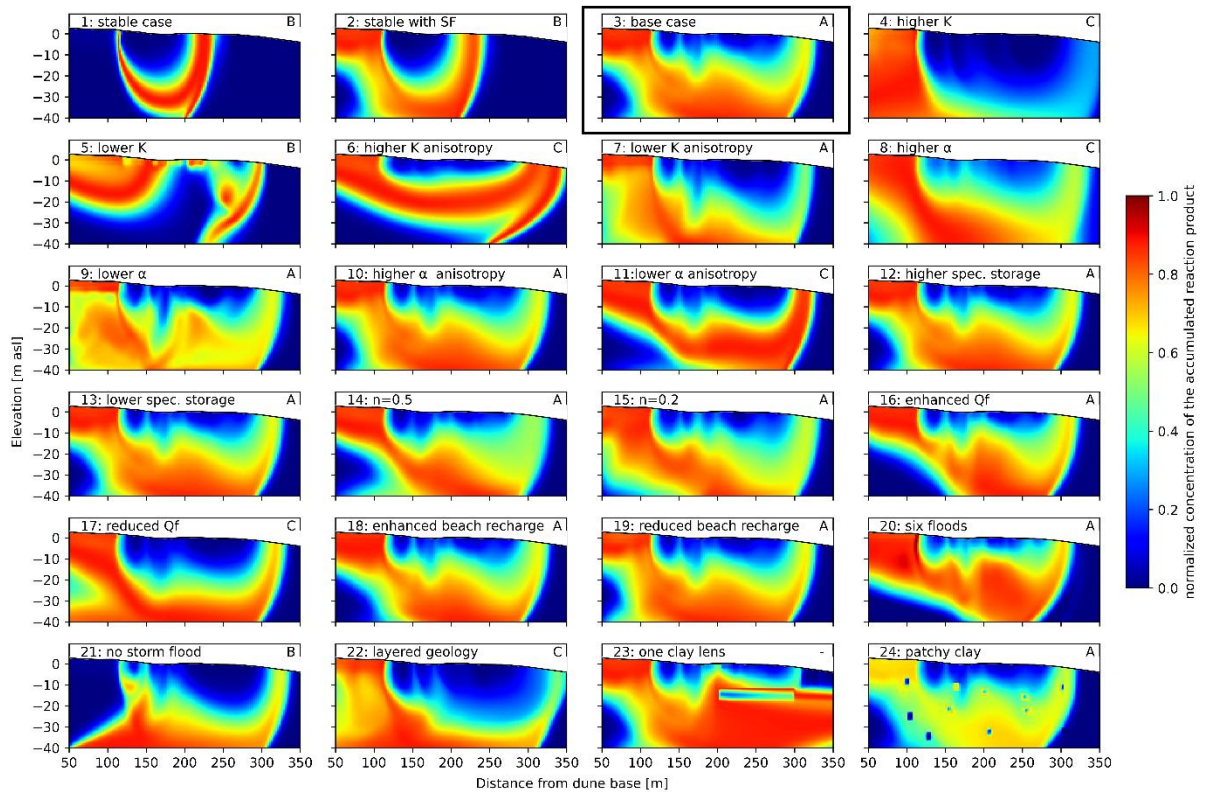
Update to Figure 5:

Figure 5: Normalized concentration of the accumulated reaction product (Mp) indicating the reaction potential for all 24 model cases after 20 years of simulation time. The base case is framed in black. The letter in the upper right corner refers to the cluster group (cf. Tab. 1, Fig. 6).

The cluster analysis in the result section will be modified to:

"A k-means analysis based on γ and RP_M of each model case normalized to the same statistics of the base case, resulted in the clustering of the model cases into three main groups (Fig. 6, Table 1). Cluster A (red circles) had a γ (+/- 20%) and RP_M (+/- 20%) similar to the base case (located at the coordinates 1,1 in the diagram in Fig. 6). Cluster B included the less dynamic and more stable cases with a lower γ, reduced by 40-98%, and lower RP_M, reduced by 30-70%, compared to the base case. Cluster B contained the cases with either changing topography only (and no storm floods, case 21) or only storm floods (and no changing topography, case 2) or neither (case 1) and the low K case (case 5). Cluster C was characterized by a lower γ reduced by 40-80%, while keeping a RP_M (+/- 20%) similar to the base case. One outlier, case 23 with one clay layer, showed a significantly higher γ and higher RP_M compared to the base case."
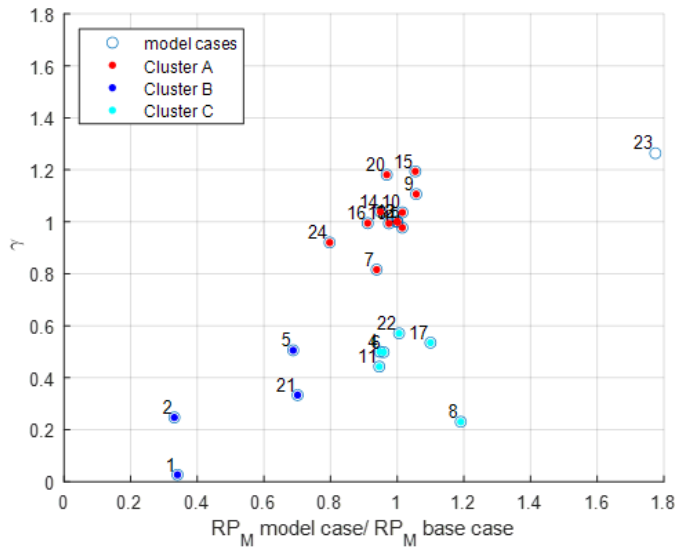
Update to Figure 6:

Figure 6: Cluster analysis (k-means,3) of all model cases based on the models' variation in salinity (γ) and the sum of reaction product of each model (RP_M). Individual model values were normalized to the base case (located at 1,1 in the diagram).

Technical Corrections

Besides the comments described above, I have a few technical recommendations for the manuscript.

1. In line 17 (in the abstract), I recommend being more specific in the sentence: "The objective was to investigate their individual effects…," whose individual effects?

   AC: We will modify the sentence in the abstract to be more specific to:

   "The objective was to investigate the individual effects of boundary conditions and hydrogeological parameters on flow regime, salt distribution, and potential for mixing controlled chemical reactions in a system with a temporally-variable beach morphology."

2. I recommend rewording the sentence between lines 55 and 58. It is difficult to understand what the other studies found and their limitations.

   AC: We will modify the sentence for the sake of clarity to:

   "These studies provided field evidence supported by numerical models for the - at least temporal - occurrence of more than one USP for different beach slopes (Abarca et al., 2013) or typical sandy beach surfaces like runnel-ridge (Grünenbaum et al., 2020b) and through-berm structures (Robinson et al., 2006)."

3. A comma is missing between "…model Greskowiak…" in line 62.

   AC: We will add a comma.

4. Review the sentence between lines 67 and 69. It sounds redundant.

   AC: The sentence will be rephrased to:

   "Greskowiak et al. (2023) concluded that redox zone dynamics in the STE are strongly affected by beach morphodynamics. While some redox reactions take place in the USPs and storm flood affected area, mixing-controlled reactions driven by mixing of two solutes in different end members occur in the fringes of the USP and at the SW interface (Heiss et al., 2017)."

5. There is a double comma in line 86.

   AC: We will remove it.

6. The term **variability** of total dissolved solutes (TDS) (in line 88) is ambiguous. Are there variations in the concentration values of TDS, the spatial distribution or which variability?

   AC: We understand the reviewer that clarification is needed for the term variability of TDS and will change the wording to be more specific to:

   "[…] the aims are to systematically evaluate the effect of […] on (1) the flow regime, (2) the distribution of total dissolved solids (TDS) and their temporal concentration variability calculated as the standard deviation of TDS, [...]"

7. The statement **generic modelling approach** (in line 93) is also ambiguous. What do the authors mean by "generic modelling"?

   AC: We understand that some clarification about the term generic modelling is needed. We consider our modelling approach "generic" in contrast to a field-site model, given that our objective is not to resemble the hydrogeological situation at a specific location in detail. It is not our intention in this study to utilize any field observations, for example for the purpose of calibrating the models. In contrast to a site-specific approach, our methodology entails a systematic alteration of boundary conditions and hydrogeological parameters, thereby facilitating a more comprehensive understanding of the physical processes occurring in STEs. However, we chose the model geometry and hydro-meteorological forcings loosely based on Spiekeroog's conditions because it is a well-studied site and representative for barriers islands under similar high energy settings (e.g., in terms of tidal amplitude and significant wave height), for example those occurring along the Wadden Sea (Netherlands, Germany and Denmark).

   We will add a short explanation of our understanding of a "generic modelling approach" to the last part of the introduction:

   "The objective of the present study is to investigate the interplay of morphological changes and hydrodynamic boundary conditions paired with aquifer properties in the subsurface of high energy beaches in a 2-D density-dependent generic modeling approach. Our model is considered 'generic' because it doesn't aim to replicate specific site conditions using field data or calibration. Instead, boundary conditions and parameters a varied to explore physical processes in STEs. While based on

Spiekeroog's conditions, the model aims to represent barrier islands in high-energy environments like the Wadden Sea, rather than a specific location."

8. Line 95 is missing a space in "…conditions(Hayes, 1979)…."

    AC: The space will be added.

9. Consider rewording the sentence and citation between lines 100 and 101.

    AC: We will rephrase the respective sentence to:

    "A slightly adapted version of the model by Greskowiak and Massmann (2021) serves as the base case for the 24 simulation cases in this study."

10. In Figure 1,
    1. There is a small box in (a) that is not described in the caption. Is this the area where changes in beach morphology occur? If so, I recommend stating it in the caption.

    AC: The reviewer is right that the box encompasses the area where the variable beach morphology was applied. We will extend the caption of Figure 1 respectively. We will include an update to Figure 1 (see above)

    "Figure 1: (a) Model setup with dimensions, boundary conditions and parameters, (colors indicate TDS distribution: red = saline, blue = fresh). The thin grey box encompasses the part of the sea boundary where changes in beach morphology occur. (b) Four LIDAR scans (Grünenbaum et al., 2020a) of intertidal topography (colored lines) and respective interpolated topographies (grey lines show 10-day increments). (c) to (f) Four different geological settings, note $K_h = 0.005$ m/d for the clay lens and clay patches (dark blue)."

    2. Is the hydraulic conductivity presented in (c) through (f) the horizontal (K_h) or the vertical (K_v) component?

    AC: The horizontal hydraulic conductivity is presented in Fig.1 c-f . We will add it to the figures' colorbar (see updated Figure 1 above).

    3. There is a space missing between "…scans(Grünenbaum et al., 2020a)…" in the caption.

    AC: We will add the space.

11. The concentration of saltwater is missing in line 134.

    AC: The concentration of saltwater of 35g/l will be added.

12. This might conflict with the journal's requirements, but I suggest changing the equations to a math format as it is easier to read.

    AC: We will consider this in consultation with the Journal requirements.

13. In Table 1,
    1. What do the bold values mean? Are they to highlight the changing parameters? If so, I recommend stating this in the caption and keeping the bold values consistent throughout the rows.

    AC: The reviewer is right, these are the values that were changed compared to the base case. We will add this explanation to the caption.

    "Table 1: Aquifer properties and boundary conditions for the 24 model cases. Numbers in bold highlight the changes compared to the base case. Note that the base case is a case with a dynamic topography resembling the model by Greskowiak and Massmann (2021)."

    2. What is the `-"-` symbol in α_L in cases 4 and 5?

    AC: The symbol was wrong, and the value of the longitudinal dispersivity of 2m will be added instead.

    3. What is the acronym SF? Is it Storm Flood? If so, please state it somewhere in the text.

    AC: The reviewer is right, SF is the acronym for storm flood. We will add it to the text.

    4. I recommend capitalizing the "based case" in the description of case 3 to be consistent with the rest of the table.

    AC: We will change it as recommended by the reviewer.

    5. I recommend correcting the spacing in the description of case 15 to be consistent.

    AC: The spacing will be changed as recommended.

14. I consider the clustering analysis a neat exercise for examining similar models. However, switching between Figure 6 and Figures 3 through 5 can be cumbersome. I recommend adding boxes around the subplots in Figures 3 to 5 that represent the colors of the clusters; that way, it is easier for the reader to visualize which models are clustered together.

    AC: We understand the need for a better recognition of clusters in the Figures 3, 4 and 5. Instead of adding colored boxes we will add the cluster groups name "A, B, C", in the upper right corner for each subfigure in Figures 3 to 5. These cluster groups are also referred to in Table 1. We will extend the Figure captions respectively by

    "The letter in the upper right corner refers to the cluster group (cf. Tab. 1, Fig. 6)."

References

Abarca, E., Karam, H., Hemond, H.F., Harvey, C.F., 2013. Transient groundwater dynamics in a coastal aquifer: The effects of tides, the lunar cycle, and the beach profile. Water Resour. Res. 49, 2473–2488. https://doi.org/10.1002/wrcr.20075

Beck, M., Reckhardt, A., Amelsberg, J., Bartholomä, A., Brumsack, H.J., Cypionka, H., Dittmar, T., Engelen, B., Greskowiak, J., Hillebrand, H., Holtappels, M., Neuholz, R., Köster, J., Kuypers, M.M.M., Massmann, G., Meier, D., Niggemann, J., Paffrath, R., Pahnke, K., Rovo, S., Striebel, M., Vandieken, V., Wehrmann, A., Zielinski, O., 2017. The drivers of biogeochemistry in beach ecosystems: A cross-shore transect from the dunes to the low-water line. Mar. Chem. 190, 35–50. https://doi.org/10.1016/j.marchem.2017.01.001

Greskowiak, J., Massmann, G., 2021. The impact of morphodynamics and storm floods on pore water flow and transport in the subterranean estuary. Hydrol. Process. 35, 1–5. https://doi.org/10.1002/hyp.14050

Greskowiak, J., Seibert, S.L., Post, V.E.A., Massmann, G., 2023. Redox-zoning in high-energy subterranean estuaries as a function of storm floods , temperatures , seasonal groundwater recharge and morphodynamics. Estuar. Coast. Shelf Sci. 290, 108418. https://doi.org/10.1016/j.ecss.2023.108418

Grünenbaum, N., Ahrens, J., Beck, M., Gilfedder, B.S., Greskowiak, J., Kossack, M., Massmann, G., 2020a. A Multi-Method Approach for Quantification of In- and Exfiltration Rates from the Subterranean Estuary of a High Energy Beach. Front. Earth Sci. 8, 1–15. https://doi.org/10.3389/feart.2020.571310

Grünenbaum, N., Greskowiak, J., Sültenfuß, J., Massmann, G., 2020b. Groundwater flow and residence times below a meso-tidal high-energy beach: A model-based analyses of salinity patterns and 3H-3He groundwater ages. J. Hydrol. 587, 124948. https://doi.org/10.1016/j.jhydrol.2020.124948

Heiss, J.W., Post, V.E.A., Laattoe, T., Russoniello, C.J., Michael, H.A., 2017. Physical Controls on Biogeochemical Processes in Intertidal Zones of Beach Aquifers. Water Resour. Res. 53, 9225–9244. https://doi.org/10.1002/2017WR021110

Michael, H.A., Scott, K.C., Koneshloo, M., Yu, X., Khan, M.R., Li, K., 2016. Geologic influence on groundwater salinity drives large seawater circulation through the continental shelf. Geophys. Res. Lett. 43, 10,782-10,791. https://doi.org/10.1002/2016GL070863

Robinson, C., Gibbes, B., Li, L., 2006. Driving mechanisms for groundwater flow and salt transport in a subterranean estuary. Geophys. Res. Lett. 33, 3–6. https://doi.org/10.1029/2005GL025247