Author's point-by-point response

The revised manuscript includes several new paragraphs of text according to suggestions by the two referees, as well as several new references.

All relevant changes to the manuscript are described in detail in the responses to reviewers, as indicated in the point-by-point response below:

Response to reviewer 1:

Review: "Methane dynamics in the Baltic Sea: investigating concentration, flux, and isotopic composition patterns using the coupled physical, biogeochemical model BALTSEM-CH4 v1.0" by E. Gustafsson, B. G. Gustafsson, M. Hermans, C. Humborg, and C. Stranne, under review for Geoscientific Model Development

As a very short summary, reactions and tracers were added to an ecosystem model that includes the horizontally integrated depth profiles of model tracers in 13 sub-basins of the Baltic Sea to simulate methane concentrations and isotopes and resolve the associated process rates. The baseline simulation shows high CH4 accumulation in the deeper anoxic parts of the Baltic Sea. Redox zonation is strongly affected over time by inflows from the North Sea and stagnation periods. CH4 concentrations above the halocline are generally much lower due to the presence of oxygen. There, seasonal thermal stratification plays an important role due to its effect on O2 availability and CH4 and O2 solubility. This also strongly affects the d13C-CH4 profiles. Qualitative agreement exists between simulated 1-D profiles, representing larger 3-D volumes of the sub-basins, and measured profiles (e.g., Jakobs et al., 2014). The authors also present a preliminary methane budget, which is, however, highly uncertain.

The introduction does a good job of describing the main research questions for studying methane dynamics in the Baltic Sea. It introduces a problem that the 1-D model cannot resolve, viz., lateral CH4 concentrations from the coastal to open waters. Similarly, there is a significant discussion in the manuscript about benthic methane release from shallower parts that cannot be well resolved. The authors could have chosen to add CH4 isotope tracers and reactions to existing 3-D ecosystem models instead of BALTSEM. This would have allowed simulating lateral gradients and also direct comparison of vertical simulated profiles from deeper parts to measured profiles at specific locations. The authors mention that BALTSEM was chosen since it has been calibrated in the past. However, other 3-D ecosystem models have also been calibrated for the Baltic Sea. Given the link between primary production and sedimentary methane production and the importance of shallow point sources of methane, it may seem that a budget including methane emission may only be confidently constrained with a 3-D model. The authors could elaborate on this choice between 1-D and 3-D models.

Response: We have not included a general discussion regarding the pros and cons of 1D versus 3D modelling approaches in the text, since this would derail a bit from the scope of the manuscript. For the purposes of the present study – i.e., parameterizations of processes in the cycling of methane, and a first preliminary budget on a Baltic Sea scale – the use of a horizontally averaged model gives a good idea about the overall methane fluxes. That said, the main issue is not that BALTSEM does not resolve horizontal gradients within the sub-basins, the issue is rather our very limited understanding of how e.g., the benthic methane source and its isotopic composition varies in time and space depending on local conditions

(e.g., organic carbon deposition rate, oxygen concentration, temperature, etc.). A 3D model would produce horizontal gradients of methane concentrations, but model validation would still be limited by knowledge gaps for some of the key processes.

It may however be noted that what appears to be of particular importance here is to improve the knowledge of contributions from ebullitive fluxes, since methane bubbles – as opposed to the diffusive flux – can bypass stratification and thus produce enhanced sea-air fluxes. Considering the very local scales of ebullition events and bubble trains, this is a case where the 1D approach in the present study is likely to be insufficient.

The following text has been added in Section 4.2, line 682-684:

"When it comes to local production of gas bubbles and the transformation and fate of methane in the bubbles, the horizontally averaged approach used in the present study is most likely insufficient, which could be addressed either by 3D modelling or by adding smaller sub-domains to the present model."

One major next step to improve the understanding and close some of the knowledge gaps, is to focus future efforts on the methane dynamics at coastal sites in general, and on ebullitive fluxes in particular – both in terms of measurements and modelling. This is discussed in Section 4.2, line 679-682:

"This indicates that the simulated CH4 outgassing is likely underestimating the real outgassing from the Baltic Sea. Observations of ebullitive fluxes in combination with development of model parameterizations represent important steps to better describe and quantify CH4 emissions from the Baltic Sea."

It is still an open question how important high concentrations and high sea-air fluxes from small shallow areas are compared to low concentrations and fluxes from large open areas.

In the manuscript, there is great uncertainty regarding the sources of methane (acknowledged in the discussion). The paper assumes a constant flux from the sediment to the water column, independent of time or water depth. There are a couple of issues, including some that are not discussed.

First of all, the text does not elaborate on sources of methane in the water column, which could be particularly important for methane emissions. Studies have shown that methane can be produced in oxic water. Weber et al. (2019) argue that these pathways are needed to explain the general oversaturation of methane in ocean surface waters, and they mention a strong correlation between methane production and net primary production. Could methane production in the water column play an important role in the Baltic Sea? Could the degradation of methylphosphonate form an alternative explanation for higher CH4 emissions (instead of lower CH4 solubility and increased benthic methanogenesis) in exceptionally warm summers, wherein PO4 is more likely a strong limiting factor for primary production? Coccolithophores and zooplankton can also release methane in the water column. Anoxic microzones in sinking particles could harbor methanogenesis both in surface waters and deeper waters. The paper mentions that the model does not reproduce a measured local

minimum d13C-CH4 at 30 m depth in the water column (line 431). Local production of methane could potentially explain this.

Response: The processes behind aerobic CH4 production are still not resolved in the Baltic Sea, although recent evidence suggests links to zooplankton grazing (as briefly mentioned in Section 4). The text in Section 4 has now been replaced by the following expanded discussion on potential unresolved CH4 sources (Section 4, lines 599-609):

"The present study does not include the potential contributions from aerobic CH$_4$ production. There are, however, several potential pathways for CH$_4$ production in shallow oxic waters, including e.g., direct CH$_4$ production by phytoplankton (Lenhart et al., 2016) and cyanobacteria (Bižić et al., 2020), CH$_4$ production as a biproduct of microbial degradation processes (Karl et al., 2008; Damm et al., 2010), and CH$_4$ formation in anoxic microniches within degrading detritus (Karl and Tilbrook, 1994; Holmes et al., 2000). In the Baltic Sea, local CH$_4$ maxima coinciding with $\delta^{13}C$-CH$_4$ minima have been observed in oxic waters just below the summer thermocline (Jakobs et al., 2014; Schmale et al., 2018). These signals can be coupled to zooplankton grazing activities, both directly through CH$_4$ production during digestion, and indirectly via release of methanogenic substrates that can subsequently be degraded to methane by microbes (Schmale et al., 2018; Stawiarski et al., 2019). However, the main pathways as well as magnitude of aerobic CH$_4$ production in the Baltic Sea remain to be resolved in detail. Parameterizations of these processes can then potentially be included in models such as BALTSEM that explicitly include both phytoplankton and zooplankton groups as model state variables."

Secondly, fluxes from sediment to the water are likely not constant. Clearly, the authors are well aware of seasonal variations in CH4 effluxes (lines 458-464) and also that it strongly depends on the oxygen levels in bottom waters (Reed et al., 2011). This makes the simulated temporal patterns less reliable. The discussion and sensitivity analyses sufficiently address this issue and, indeed, show a high sensitivity toward the parameter used for benthic CH4 release.

Response: Yes, as discussed in the manuscript the main uncertainty in this study is the CH4 release from sediment to water: this flux is known to be highly variable both in space and time (e.g., Roth et al., 2022) as a function of organic carbon deposition rate, oxygen and sulfate concentrations, etc. Observed fluxes are very sparse which means that it is a major challenge at this point to parameterize the CH4 effluxes from different sediment areas of the Baltic Sea. The calibrated constant flux gives a very good idea about the long-term mean CH4 release from deep water sediments in stagnant waters where CH4 in the water seems to be only marginally influenced by other processes (e.g., AOM very low in anoxic waters as demonstrated by measurements by e.g., Jakobs et al., 2013).

The following text was added to Section 4, line 565-575:

- *"It is, however, likely that the calibrated CH$_4$ release is mainly representative for present-day conditions (e.g., organic carbon deposition rates, oxygen concentrations, temperature, etc.). Both climate change and nutrient load change are going to affect e.g., oxygen concentrations in the future which means that the benthic CH$_4$ source is likely to change as well. In order to address this, it is necessary to improve the knowledge of CH$_4$ release rates depending on local conditions. One major uncertainty here is what is the contribution from more recent organic carbon deposition, and*

what is the contribution from "old" carbon deeper down in the sediments, i.e., if nutrient loads and organic carbon deposition decreases, and oxygen conditions improve, would this have a major impact on the CH_4 release from sediments, or is the release more heavily dependent on older carbon deposits? This is one of the major remaining open questions regarding CH_4 cycling in the Baltic Sea, but this cannot be addressed by the model at this point."

- *"While the calibrated flux gives a good idea of the present-day CH_4 source in deeper areas, it is more challenging to constrain the sediment source in shallower oxic waters, where the source can be largely compensated by MOX in the water column."*

Furthermore, the uncertainty has previously been addressed in the Section 4 (line 556-562):

- *"However, there are uncertainties in our estimates, in particular regarding the benthic CH4 source. In the standard model run, benthic release is the dominant CH4 source (Table 2). The sediment source is set as constant over time, at all depths, and in all sub-basins. In the real Baltic Sea, however, large spatial and temporal variations are expected (e.g., Roth et al., 2022). Furthermore, the isotopic composition of the sediment source is set either to -80‰ or -60‰ depending on oxygen conditions in the overlying water. This assumption is a simplified representation. The main uncertainty in our present large-scale estimates is that spatial and temporal variations of the sediment source are not well known."*

Section 4.2 (line 657-670):

- *"As previously discussed, the main uncertainty in the model simulations lies in our limited understanding of CH4 release from different sediment areas, as well as the isotopic composition of CH4 released into the water column. Both the flux and the isotopic composition depend on the balance between production and oxidation rates in the sediment. A high production could be compensated by high oxidation and thus result in a relatively small CH4 release to the water column in spite of a large production. This would then be evident by a 13C-CH4 enrichment, i.e., comparatively heavy CH4. Alternatively, a relatively small CH4 production could still result in a substantial release to the water column in a case where the oxidation rate is low, which would then also be evident by CH4 depleted in 13C-CH4, i.e., comparatively light CH4. "*
- *"Improved knowledge of properties of CH4 released from sediment to water column in different areas of the Baltic Sea (e.g., the Kattegat and the major gulfs – the Gulf of Bothnia, Gulf of Riga, and Gulf of Finland) would help to improve model parameterizations and thus reduce the main uncertainties of model simulations. This was, however, beyond the scope of the present study because of the missing knowledge concerning both temporal and spatial patterns of the CH4 source. A logical progression at this stage would involve detailed observations combined with modeling studies focused on processes in the sediments, i.e., production and oxidation rates, depending on carbon accumulation rate, oxygen conditions, and the presence of methanotrophs."*

Thirdly, the introduction mentions gas ebullition versus diffusive sources. This part should be expanded. Other workers have shown that hotspots, such as cold seeps, vents, and mud

volcanoes, are significant for methane emissions on the scale of the global ocean (e.g., Hornafius et al., 1999; Weber et al., 2019). There are plenty of studies about cold seeps in the Baltic Sea. To what extent do the authors expect these to dominate emissions to the atmosphere? Methane from these point sources may be laterally transported and affect the CH4 concentrations in surface waters in large parts of the Baltic Sea. There are also other hotspots, such as inundated peat lands. It would be beneficial for readers to know more about the prevalence of cold seeps and methane-rich sediments in both shallow and deeper parts of the Baltic Sea to gain a sense of their importance in the overall budget. It could be very well that methane emissions to the atmosphere are greatly underestimated by the model, as benthic methane release is primarily fitted to CH4 concentrations in the deeper basins.

Response: Already discussed in Section 4.2, but the text has now been expanded and reads as follows (Section 4.2, line 671-682): "A crucial missing link in this study is the formation, transport, and fate of CH_4 bubbles. Estimates by Weber et al. (2019) indicate that ebullitive fluxes contribute a major fraction of CH_4 released to the atmosphere from shallow coastal areas. Ebullition events have been observed in the Baltic Sea, both at coastal sites (e.g., Humborg et al., 2019; Lohrberg et al., 2020; Lehoux et al., 2021; Hermans et al., submitted) and deep water accumulation bottoms (C. Stranne, unpublished data). Ebullition has been included in lake models (e.g., Greene et al., 2014; Stepanenko et al., 2016; Bayer et al., 2019); however, we do not have experimental data to calibrate and validate the large-scale influence of ebullition in the Baltic Sea. The calibrated benthic CH_4 source represents a "bulk" CH_4 release, including in theory both the influences of diffusive flux and bubble dissolution on CH_4 concentrations in the water column. However, CH_4 ebullition might bypass methanotrophy and consequently contribute to higher CH_4 emissions, in particular in shallow-water areas (e.g., Broman et al., 2020). This indicates that the simulated CH_4 outgassing is likely underestimating the real outgassing from the Baltic Sea. Observations of ebullitive fluxes in combination with development of model parameterizations represent important steps to better describe and quantify CH_4 emissions from the Baltic Sea."

More than 90% of the citations appear to be works from scientists who have studied the Baltic Sea. Literature from other parts of the world is largely ignored. Occasionally, the wrong articles are cited. For instance, Broman et al. (2020) did not discover methanotrophy (lines 48-50). Not the latest articles, but the articles that made the initial discoveries should be cited. In Table 1, it would be interesting to compare the values of rate constants to the literature. In the discussion, the reaction kinetics are discussed and compared to lake studies. However, there is also a vast body of literature about methane oxidation in ocean waters (e.g., Chan et al., 2019b, 2019b and Pack et al. 2015 show up in a first search attempt).

Response:

- *The sentence has been simplified and the Broman reference removed (line 46-48): "Methane formation in sediments can be substantial, but aerobic and anaerobic oxidation processes can efficiently remove CH_4 both in the pore water and water column"*
- *The text on lines 89-91 has been updated and now reads: "Methane cycling has previously been investigated in both lake (e.g., Lopes et al., 2011; Greene et al., 2014; Tan et al., 2015; Stepanenko et al., 2016; Bayer et al., 2019) and ocean (e.g., Nihous and Masutani, 2006; Wåhlström and Meier, 2014; Malakhova and Golubeva, 2022) modeling studies."*

- *The reference list has now been expanded to include studies from wetlands, lakes, and ocean areas. The following text has been added in Section 4, line 583-589: "Studies from wetlands (Segers, 1998), lakes (Martinez-Cruz et al., 2015; Tan et al., 2015), and oceanic sites (Kessler et al., 2011; Crespo-Medina et al., 2014; Pack et al., 2015; Rogener et al., 2018; Chan et al., 2019a) show that MOX rates can vary by several orders of magnitude. For example, observed deep water MOX in the Gulf of Mexico increased from a background rate of around 60 pM d-1 to a peak rate of 5900 nM d-1 after the Deepwater Horizon oil spill (Rogener et al., 2018). The observed MOX rates from the central Baltic Sea (approximately 0.1-4 nM d-1; Schmale et al., 2012; 2016; Jakobs et al., 2014) are in the same range as MOX rate observations from the eastern tropical North Pacific Ocean (Pack et al., 2015), but typically lower than MOX rates observed in lakes (Martinez-Cruz et al., 2015; Tan et al., 2015)."*
- *The following text has been added in Section 2.3.6, line 440-441: "Observations indicate a wide range of fractionation during CH_4 oxidation (e.g., $\varepsilon \sim 4\text{-}30‰$, Whiticar (1999); $\varepsilon \sim 16\text{-}54‰$, Chan et al. (2019b))".*

Overall, I think the strength of the paper is the simulated vertical structure and temporal variability of CH4 concentrations, which is representative of locations with greater water depth. It identifies the interesting role of thermal stratification in surface water, which can affect methane emissions, and the dynamics related to inflows from the North Sea. The authors discuss uncertainties in the overall budget, which could further be improved by considering methane production in the water column and by elaborating on the role of cold seeps and other point sources. Beyond simulating some interesting dynamics, I am currently not convinced that the model will be able to constrain a methane budget for the entire Baltic Sea in the future, as it will always be difficult to represent methane dynamics in coastal areas, and model output cannot be directly compared to measured profiles from particular locations. Maybe the authors can elaborate on whether switching to a 3-D model will be necessary.

Response: It is indeed likely that the calibrated CH4 flux is mainly representative for present-day conditions. This is now discussed in Section 4, line 565-575 – see major point above.

Specific comments:

Line 15: "land loads"

I do not understand what land loads could mean, since CH4 is not a solid but a gas. Based on the text, I think river runoff is meant. However, sometimes river runoff and land loads are mentioned in a single sentence as separate sources (e.g., line 594). In the text river runoff, river load, and land loads may denote the same source.

Response: The "river runoff" indicates the freshwater supply to the various sub-basins, whereas "land load" includes both river loads and point sources of organic and inorganic carbon and nutrients. For methane, the contributions from point sources and river loads, respectively, are not known. For that reason, we have only included guessed river loads of methane in the model simulations. This has now been clarified in the text; "land load" of methane has been replaced by "river load" of methane.

Lines 20-21: "to our knowledge this is the first time that CH4 isotopes have been included in a physical-biogeochemical model"

What about Nihous and Masutani (2006)? The full reference is listed below.

Response: Thanks. Their model only includes fractionation during oxidation in the water column (and not air-sea exchange, etc.), but this is still a relevant reference that has now been included. The sentence has been modified and now reads (line 19-20): "Modeling of stable CH$_4$ isotopes can help to constrain process rates."

Line 53-55: "This notion... Humborg et al. 2019)"

Could there not be an alternative explanation, such as increased methane production in the water column?

Response: Potentially yes, but observations of gas flares indicate that outgassing from the sediments was the major source in this particular case. The sentence has nevertheless been rephrased and now reads (line 51-53): "This notion was qualitatively supported by acoustic observations of outgassing from the sediments during a recent field study, where exceptionally high CH4 emissions were reported from the coastal Baltic Sea at the end of a summer heat wave (~250 μmol m-2 day-1, Humborg et al., 2019)."

Line 354: "calibrated"

This word is used several times in the text. However, due to the scarcity of data, I think it cannot be called a calibration. Also, it is annoying that the data is not shown. This shows the disadvantage of not using a 3-D resolved model.

Response: The scarcity of data is a major issue for sure. For most of the parameters in the methane modelling we do not have observed rates to rely on, which means that the different rate constants are calibrated to produce simulated methane concentrations and isotope signatures as closely as possible to the few available observations. Perhaps this word can be used a little bit differently depending on context, but I still believe it is the most appropriate word in this case. I also doubt that using a 3D resolved model would make calibration easier considering the massive computational cost.

Lines 430-431: "Furthermore, a local... model run."

This could indicate a local source of methane in the water column.

Response: Yes. The potential influence from local CH$_4$ sources were already briefly addressed in Section 4 and Section 4.2. However, as suggested in one of the major comments, this has now been elaborated in some detail in Section 4, line 599-609 (see major comment above).

Lines 482-486: "The rate constant... lakes cited above."

Rate constants could be compared to kinetic studies of methane oxidation in ocean water instead of lakes. It should be noted that microbes that oxidize methane may have more than

one trick on their sleeve (Rogener et al., 2018), allowing them to survive on other energy sources.

Response: Text added in Section 4, line 583-589 (see major comment above).

Line 599: "References"

The reference list is incomplete. At least, Weber et al. (2019) and Roth et al. (2022) are missing.

Response: Thanks. These references (and several new references) have been added to the list.

Minor comments:

Line 52-53: "In shallow... Borges et al. (2016)"

The sentence would improve by replacing "emissions" with "seafloor ebullition" and removing the last part.

Response: Ok, updated according to suggestion (line 50-51):

"In shallow, organic-rich sediments, seafloor ebullition will increase in response to ocean warming due to increased biogenic CH_4 production and decreased CH_4 solubility (Borges et al., 2016)."

Line 354: "intension"

Response: Corrected.

Line 388-389: "The δ13C-CH4 in water... temperature stratification."

The transition from the previous sentences is not smooth. It would be good to point the readers to the right figure here.

Response: The text has been rewritten and now reads (483-487): "This temperature dependence on oxidation rates also has an impact on the isotopic composition of CH_4 – the $\delta^{13}C$-CH_4 in water above the top of the halocline is strongly influenced by the seasonality of temperature stratification (Fig. S3-S4, supporting information). However, the variations of isotopic composition in surface waters are significantly smaller than the variations at depth where $\delta^{13}C$-CH_4 mainly depends on transitions between oxic and anoxic conditions (Fig. 4-5)."

Figure 6: Abbreviations should be explained, as figures should be understandable without reading the text. "ASE" is also not defined in the text.

Response: Agreed. The figure legend has been updated; abbreviations have been replaced by words, and "land load" has been replaced by "river load".

The term "redox zone" appears to be wrongly used at various locations in the text, where the authors intent to mean "redoxcline".

Response: In this case we used the same wording as e.g., Schmale et al. (2012) and Jakobs et al. (2013;2014) for the oxic-anoxic transition zone. But anyway, "redox zone" has now been replaced by "redoxcline" in the text.

References:

Chan, E. W., Shiller, A. M., Joung, D. J., Arrington, E. C., Valentine, D. L., Redmond, M. C., ... & Kessler, J. D. (2019a). Investigations of aerobic methane oxidation in two marine seep environments: Part 1—Chemical kinetics. Journal of Geophysical Research: Oceans, 124(12), 8852-8868.

Chan, E. W., Shiller, A. M., Joung, D. J., Arrington, E. C., Valentine, D. L., Redmond, M. C., ... & Kessler, J. D. (2019b). Investigations of aerobic methane oxidation in two marine seep environments: part 2—isotopic kinetics. Journal of Geophysical Research: Oceans, 124(11), 8392-8399.

Hornafius, J. S., Quigley, D., & Luyendyk, B. P. (1999). The world's most spectacular marine hydrocarbon seeps (Coal Oil Point, Santa Barbara Channel, California): Quantification of emissions. Journal of Geophysical Research: Oceans, 104(C9), 20703-20711.

Nihous, G. C., & Masutani, S. M. (2006). A model of methane concentration profiles in the open ocean. Journal of Marine Research, 64(4), 629-650.

Pack, M. A., Heintz, M. B., Reeburgh, W. S., Trumbore, S. E., Valentine, D. L., Xu, X., & Druffel, E. R. (2015). Methane oxidation in the eastern tropical North Pacific Ocean water column. Journal of Geophysical Research: Biogeosciences, 120(6), 1078-1092.

Rogener, M. K., Bracco, A., Hunter, K. S., Saxton, M. A., & Joye, S. B. (2018). Long-term impact of the Deepwater Horizon oil well blowout on methane oxidation dynamics in the northern Gulf of Mexico. Elem Sci Anth, 6, 73.

Weber, T., Wiseman, N. A., & Kock, A. (2019). Global ocean methane emissions dominated by shallow coastal waters. Nature communications, 10(1), 4584.

Response to reviewer 2:

Gustafsson et al. developed the model BALTSEM-CH4 v1.0, where methane isotopes are included to provide more information about the methane dynamics in the Baltic Sea. They reported their study focusing on three objectives: *1. identify and roughly quantify key CH4 fluxes, 2. set up a preliminary CH4 budget on a Baltic Sea scale, and 3. perform sensitivity experiments on CH4 concentration and isotopic composition depending on transport and transformation processes.*

While I would agree that the authors have by and large achieved their objectives, I think the manuscript in its current form won't meet the curiosity of interested readers, particularly for

those who are technical detail oriented and who are likely a significant fraction of gmd's readership. In the least, I think the authors should provide a thorough technique note that details the model structure with all governing equations and their supporting assumptions, as well as instructions on how initial and boundary conditions are set, and how the numerical solution is obtained.

For example, the current paper leaves me with many questions like:

- How is the reactive-transport problem being formulated?
- Does the model explicitly represent diagenesis?
- Is the sediment represented with explicit biogeochemistry?
- How diagenesis and biogeochemistry are coupled with temperature dynamics and vertical mixing?

Response to the four points above: This model does not include a reactive-transport model for sediment diagenesis. Instead, we use a simplified parameterization where the model accounts for the depth dependent sediment pools of organic C, N, P, and Si at a resolution of 1 meter water depth. The total sediment areas at different water depths are described by the hypsographic functions for the respective basins. Sinking organic material from the water column feeds into these sediment pools at different depths, where the accumulated organic material is subjected to burial as well as temperature dependent mineralization. Mineralized carbon and nutrients can then again be released to the water column. Oxygen concentration in the water overlying the sediments controls the release of ammonium and nitrate, respectively, as well as the denitrification loss term. Oxygen concentration further controls phosphate sequestration, representing phosphate bound to iron oxides, as well as an additional release of phosphate to the water column during transitions from oxic to anoxic conditions, representing reduction of iron oxides.

The sediment dynamics in the model is described in some detail on line 212-228 in the revised manuscript.

- In what way is river load applied?

Response: River loads (as well as point sources and atmospheric depositions) of organic and inorganic carbon and nutrients are applied as monthly mean loads to the respective sub-basins (see line 232-233 in the revised manuscript).

- And how the lateral and vertical transition of water depth and sediment thickness are handled?
- How is the lateral exchange formulated from the shallow water zone to deep water zone?

Response to the two points above: The depth dependent areas of both water and sediment are described by the hypsographic functions for the respective sub-basins. Sediment thickness is

not explicitly modelled, instead, the model accounts for the depth dependent pools of organic C, N, P, and Si (see further in the response above, as well as the description on 212-228 in the revised manuscript). The model includes a parameterization for lateral transports of organic material from sediments at shallow depths toward deeper areas, representing resuspension and redeposition processes.

- From some part of the paper, it seems the sediment is not explicitly represented. Then How should this be justified if the model is used for long-term projection, where active accumulation/degradation of sediment organic matter will be significant?

Response: As discussed in responses above and also described in some detail in the updated model description (line 212-228), the sediment is indeed explicitly represented in the model. Active accumulation/degradation of sediment organic matter occurs in the model as a function on one hand on deposition rate, and on the other hand on burial as well as temperature dependent mineralization of organic material accumulated in the sediments. The model is a very useful tool for long-term projections and has in several prior publications been used both to do hindcast simulations and different future scenario simulations depending on climate change and nutrient load change.

Besides, although the authors mentioned calibration in the paper, the results do not show any comparison with observations. (They did say the model more or less agree with some measurement in the text, but I think this is insufficient.) Since there is no differential equation of the reactive-transport system described, I cannot judge how well the model is performing, even I may trust the authors are making confident statement.

In all, I expect the authors do a major revision to present a more convincing manuscript to the readers.

Response: Observed methane profiles were unfortunately not available. The "model description" section has been completely rewritten and now includes a lengthy qualitative description of hydrodynamic and biogeochemical processes in the model. The description also includes detailed references to studies where all governing equations and model parameterizations are formulated, but we have not repeated all the differential equations in the present paper – instead the focus here is on the new features of the model, i.e., the process parameterizations for stable isotope methane cycling, which is outlined in detail in Section 2.3.1-2.3.6.

The new and expanded model description now reads as follows, line 119-241 in the revised manuscript:

2.2 The model

BALTSEM is a horizontally averaged, but vertically resolved process-oriented model that couples hydrodynamic and biogeochemical modules in time-dependent, numerical

simulations. In the model, the Baltic Sea is divided into thirteen coupled sub-basins (Figure S1), with geometric characteristics as summarized in Table S1. The hydrodynamic module has been described in detail by Gustafsson (2000; 2003), whereas the biogeochemical module has been described in detail by Savchuk (2002). Below, the two modules are qualitatively recapped (Sect. 2.2.1-2.2.2).

In this study, a new expanded version of the model, BALTSEM-CH4 v1.0, with state variables representing both ^{12}C-CH$_4$ and ^{13}C-CH$_4$ is presented for the first time. Figure 1 illustrates processes involved in CH$_4$ cycling that are included in the model. This study focuses on the modeling of stable CH$_4$ isotopes: The CH$_4$ sources (i.e., river load and sediment release), boundary conditions (i.e., atmospheric CH$_4$ and CH$_4$ at the open ocean boundary), transport and transformation processes (i.e., CH$_4$ oxidation and air-sea exchange), as well as the isotopic fingerprints associated with these processes are described in Sect. 2.3. The model parameterizations for both hydrodynamic and biogeochemical processes (prior to the inclusion of CH$_4$) have been described in detail in earlier publications (e.g., Gustafsson 2000; 2003; Gustafsson et al., 2012; 2014; 2017; Savchuk, 2002; Savchuk et al., 2012); this will not be repeated here. A list of all state variables in the model is included in Appendix A (Table A1).

2.2.1 Hydrodynamic module

The vertical stratification in each sub-basin is resolved by a variable number of horizontally homogenous layers. The numbers of layers in the respective sub-basins increase over time because of both inflows from adjacent basins and instances of pycnocline retreat, as described below, but are kept below maximum values by mixing of the two layers that require the least amount of energy to be merged (Gustafsson, 2000).

Flow dynamics through the straits that connect different sub-basins depends on the width of the strait compared to the internal Rossby radius, determining whether or not earth rotation influence the water exchange. In general, lateral exchange between sub-basins is forced by barotropic pressure gradients across the straits that depend on sea level difference and wind set-up, as well as baroclinic pressure gradients caused by differences in stratification. In narrow straits, the water flow is influenced by frictional resistance and dynamical contraction due to the Bernoulli effect, while the transport through wider straits is further controlled by earth rotation effects (Gustafsson, 2000; 2003).

Dynamics of the mixed surface layer in each sub-basin is forced by wind stress and buoyancy fluxes, but also depends on earth rotation, following Stigebrandt (1985). The pycnocline is eroded whenever the buoyancy flux is negative (e.g., if surface water density increases because of net evaporation, or by cooling when the water temperature is above the temperature for maximum density), or when the buoyancy flux is positive but the power generated by wind stress is sufficient to do work against the buoyancy forces. Pycnocline erosion means that the mixed surface layer becomes thicker and denser as a result of deep-water entrainment into the surface layer. If the power is not sufficient, the turbulent mixing becomes limited either by earth rotation or by buoyancy fluxes, leading to a pycnocline retreat and the formation of a new and shallower mixed surface layer. The thickness of the

new surface layer will be determined either by the Ekman or Monin-Obukov length-scale – whichever is shorter (Stigebrandt, 1985).

Entrainment flows are further modified by the presence of sea-ice (Gustafsson, 2003). Ice dynamics is based on a sea-ice model by Björk (1997), but adapted to the Baltic Sea following Nohr et al. (2009). Calculations for heating/cooling and evaporation at the sea, ice or snow surface follow Björk (1997). About half of the incoming short-wave radiation is absorbed at the surface while the remaining fraction attenuates exponentially using constant attenuation factors for water, ice and snow, respectively.

Turbulent vertical diffusion in deeper layers below the mixed surface layer is parameterized as a function of stratification and mixing wind (Stigebrandt, 1987; Axell, 1998), representing the energy inputs from inertial currents and breaking internal waves. The model further includes dense gravity currents (i.e., deep-water inflows along the seafloor), where entrainment of surrounding deep water into the gravity currents depends on bottom slope and friction as well as density difference between the gravity current and the surrounding water (Stigebrandt, 1987). Entrainment of surrounding water into gravity currents has the effect that the volume flow increases while at the same time density decreases, influencing at what depth the gravity current will be interleaved, i.e., the depth of neutral buoyancy. Deep-water inflows cause an uplift of the entire water column above the intrusion depth.

2.2.2 Biogeochemical module

Biogeochemical processes are calculated using a nutrient-phytoplankton-zooplankton-detritus model setup that closely follows Savchuk (2002), but that has been expanded with state variables representing e.g., dissolved organic compounds and the inorganic carbon system (Gustafsson et al., 2014).

The biogeochemical module includes pelagic state variables for oxygen (O_2), hydrogen sulfide (H_2S), total alkalinity, dissolved inorganic carbon, nitrate + nitrite, ammonium, phosphate, dissolved silica, labile and refractory fractions of dissolved organic carbon (C), nitrogen (N), and phosphorus (P), particulate organic C, N, P, and silicon (Si), three functional groups of phytoplankton (representing diatoms, 'summer species', and diazotrophic cyanobacteria), and one bulk state variable for heterotrophs that represents zooplankton and other organisms that consume and mineralize phytoplankton and detrital matter. All pelagic state variables are subject to transport processes (vertical mixing and horizontal advection) as well as various biological and chemical transformation processes; source and sink terms for each state variable are computed in all water layers in each sub-basin. BALTSEM further includes sediment pools of C, N, P, and Si that are subject to mineralization and burial. The pelagic and benthic realms are coupled by sedimentation of organic matter and sediment-water exchange of dissolved inorganic compounds. Oxygen, CO_2, and CH_4 are exchanged at the air-sea boundary depending on solubilities, wind speed, and gradients between sea surface and air of the respective gases.

Phytoplankton growth depends on water temperature and is further limited by light and nutrient availability (Savchuk, 2002). Light penetration in water in the biogeochemical module is calculated as a function of the biogeochemical state. The phytoplankton groups assimilate dissolved inorganic C, N, and P according to fixed Redfield ratios while at the same time producing oxygen, but also take up an excess of dissolved inorganic carbon which

is transformed into dissolved organic carbon, representing extracellular production (Gustafsson et al., 2014). The cyanobacteria group is able to fix atmospheric N when ammonium and nitrate become limiting. The diatom group is the only phytoplankton group that requires dissolved silica. Loss terms for phytoplankton include natural mortality, grazing by zooplankton, and sinking. Dead phytoplankton are converted into detrital C, N, P, and Si according to their elemental stoichiometry.

Heterotroph/zooplankton growth depends on grazing rate which is regulated by water temperature and food concentration (phytoplankton and detritus) as well as the respective availability of different food sources (Savchuk, 2002). Grazing is in addition strongly inhibited at low oxygen concentrations. Fractions of each food source that are not digested are instead assigned to detritus pools in accordance with stoichiometry of the food sources. Zooplankton have elemental stoichiometry that differ from their food sources; growth thus becomes limited by the element in relative shortage, while carbon and nutrients in excess compared to zooplankton stoichiometry are excreted. Zooplankton biomass decreases by natural mortality and excretion; dead zooplankton are converted into detrital C, N, and P according to elemental stoichiometry.

Phytoplankton and detritus sink through the water column; phytoplankton that are not lost by grazing or natural mortality in the water column settle on the seafloor where their constituents are assigned to sediment pools of C, N, P, and Si according to elemental composition. Temperature dependent leaching converts a fraction of the detritus into dissolved organic C, N, and P, as well as dissolved silica in the water column, while the remainder is either consumed by zooplankton in the water column or settles on the seafloor where it is assigned to the respective sediment pools. Organic carbon and nutrients in the water column are mineralized either by means of zooplankton respiration (dissolved inorganic carbon) and excretion (ammonium and phosphate) or by temperature dependent oxidation of dissolved organic compounds; these processes also consume oxygen. Nitrification converts ammonium into nitrate while consuming oxygen. Heterotrophic and chemolithoautotrophic denitrification processes represent loss terms for nitrate. In the absence of both oxygen and nitrate, organic matter is instead oxidized by sulfate, which also leads to hydrogen sulfide production. Sulfide can be oxidized by either oxygen or nitrate (i.e., chemolithoautotrophic denitrification); sulfide oxidation thus represents loss terms for either oxygen or nitrate.

The sediment compartment in each sub-basin can be described as a series of horizontal terraces with a resolution of one terrace per one meter water depth; the area of each terrace is a function of the hypsographic curve for the respective sub-basins. Sediment state variables are not vertically resolved on the individual terraces, but instead formulated as pools of bioavailable C, N, P, and Si that have been deposited on the different terraces – representing the "active" (i.e., not permanently sequestered) top layer of sediments (Savchuk et al., 2012). The carbon and nutrients in phytoplankton and detritus that settles on the terraces are added to the respective sediment pools. A fraction of the sediment pools is permanently sequestered and thus removed from the biogeochemical cycling, while the remaining fraction undergoes temperature dependent mineralization into inorganic carbon and nutrients that can again be released to the water column.

Nutrient cycling and release from the sediments is strongly coupled to oxygen concentration in the overlying water. During oxic conditions, mineralized N is released in the form of nitrate, but an oxygen dependent fraction of the nitrate is lost by denitrification. A fraction of the mineralized P is retained in the sediments during oxic conditions, representing phosphate bound to e.g., iron oxides. P retention capacity is further regulated by salinity, representing a proxy for both sulfate concentration and iron availability (Savchuk et al., 2012). During anoxic conditions in the overlying water, mineralized N is released in the form of ammonium. At the same time, mineralized P cannot be retained in the sediments during anoxic conditions; instead, previously sequestered phosphate is released to the water column, representing reduction of metal oxides that are thus unable to bind phosphate. During oxic conditions, sediment mineralization consumes oxygen in the overlying water; during anoxic conditions, the sediments release hydrogen sulfide to the overlying water, representing sulfate reduction.

2.2.3 Model forcing, boundary conditions and initial conditions

The meteorological forcing includes three-hourly wind data, air temperature, cloudiness, air pressure, and precipitation. Model forcing for the hydrodynamic module also includes observed daily mean sea level in the Kattegat as well as monthly mean river runoff to each sub-basin. Further, the model forcing includes monthly mean loads of inorganic and organic carbon and nutrients, and alkalinity from land (point sources and river loads) and atmosphere. Daily profiles of salinity and temperature (i.e., stratification), as well as concentrations of all biogeochemical state variables (Table A1) define the conditions at the open boundary between the Northern Kattegat (sub-basin 1; Figure S1) and the Skagerrak (open ocean). Monthly mean atmospheric partial pressures of CO_2 and CH_4 comprise the atmospheric boundary conditions for the respective gases. The model forcing is further detailed in Appendix B.

An initial model run over the period 1970-2000 started with initial profiles for the different state variables based on observations when possible or else calibrated values. The initial model run was then used as a spin-up for a series of model runs covering the period 2001-2020 that are performed to examine the sensitivity of e.g., CH_4 concentration and isotopic composition depending on process parameterizations (Sect. 4.1).