Authors' response (AR) to comments by referee 1 (RC)

RC: "Walve et al. (bg-2017-496) present an interesting study where phosphorus dynamics in an estuary in the Baltic Sea are quantified using a box-model approach. The study period spans from 1968 to 2015, and has a seasonal temporal resolution. Box model budgets for water and salt are calculated from measured freshwater inflow and salinity from two stations. These water budgets are further used to infer phosphorus fluxes to and from the study estuary. In general, the study comprises of a straightforward model exercise where abundant monitoring data has been used to build the model. One of the main findings is the identification of the sediment as a sink of phosphorus in spring and as a source during summer and autumn. Overall, the manuscript is well written and easy to follow."

AR: We are pleased that our study was found interesting and easy to follow, and are grateful for the constructive comments. Although the finding of the sediment as seasonal source and sink is important, we consider the finding that legacy P release was clearly demonstrable only for the initial 10-20 years, and that later on seasonal P release is larger than any possible legacy P release, as the main results. Moreover, seasonal P release nearly equals the sink meaning that present annual net P retention is very low.

RC: "The manuscript would benefit from a clear formulation of scientifically relevant research question and/or hypotheses, which would potentially be relevant to other systems as well."

AR: We agree. We have added the requested text in the final paragraph of the introduction. Research questions are added to include:

1. In an estuary that has long been subject to a very high P loading, is there a long-term influence of legacy P once the load has been drastically reduced? What is the current annual P retention of the area?

2. To what extent does oxygen control seasonal P release and annual P retention? Can we expect improvements in oxygen conditions to increase long-term P retention?

3. Can an effect of the temperature increase on oxygen consumption and P release be identified?

RC: "What are the main discrepancies compared to other studies, and why adding one more model study of this system is justified?"

AR: This is the first temporally highly resolved dynamic box model study of the Stockholm IA that follows the long-term P balance, including the period of change from high to low P load, and the first with seasonal resolution. Moreover, we link P retention to oxygen concentrations and oxygen consumption (from a deep-water oxygen budget). We find lower P retention than previous studies, even periods of net P export. We find that oxygen consumption from degradation of organic matter is predominantly directly linked to P release. We also find that low oxygen can promote an additional release of P. We have clarified this justification of our study in the final paragraph of the introduction. Discrepancies from other studies (e.g. in mean annual P retention) are presently dealt with in the discussion, but are now noted also in the introduction.

RC: "The title is very concise, maybe even too much so. The seasonality behind the source and sink actions could be introduced in the title already"

AR: We have expanded the title to "A Baltic Sea estuary as phosphorus source and sink after drastic load reduction: seasonal and long-term mass-balances for the Stockholm inner archipelago 1968-2015". This emphasizes the long-term results, the seasonality aspect, and gives the exact location of the study.

RC: "Abstract P1, L10: In effect, two box models are used in the study, and the box models are vertically layered. This should be made clear here as well."

AR: We now make it clear already in the Abstract that we use more than one model, and that both are vertically layered.

RC: "P1, L10: Here and elsewhere: there is inconsistent use of terms "box-model" and "box model". Authors should choose either one and use the same term throughout."

AR: We now consistently write "box model".

RC: "P1, L20: Replace stores with pools or storages."

AR: Corrected.

RC: "P1, L23: Sentence is incomplete, contributed to what?"

AR: Contributed to improved oxygen conditions. We have removed lines 21-24 that dealt with reasons for oxygen increases in order to shorten and focus the abstract (as also suggested by reviewer 2). We have replaced them with a short sentence about temperature effect on oxygen consumption.

RC: "Introduction P1, L31-32: Sentence is vague and does not read well, please rephrase."

AR: We agree, and have rephrased.

RC: "P3, L1: Remove "coastal", as estuaries are by definition coastal. If estuaries are to be emphasized here, then the sentence should be rephrased."

AR: Corrected and rephrased.

RC: "In general, estuaries have short residence time in the mixed layer, but is that the case in the Stockholm Archipelago? I would assume that the stratification and the relatively deep basins lead to high water residence times in the water masses close to the bottom, compared to the fast water exchange in the surface layer."

AR: due to the estuarine circulation and the relatively large water mass in the upper 10 meters, the retention time of the deep-water is shorter than the surface water. We have added a sentence in section 2.1 with reference to Engqvist and Andrejev (2003), who give estimates of water retention times. This reference is cited elsewhere in the manuscript and was already in the reference list.

RC: "P3, L6: "deteriorated badly" does not read well, rephrase."

AR: Changed to: ... deteriorated seriously ...

RC: "P3, L32: Is this not the case with all box models, that they are dynamic? Or is this a special case of a "dynamic" box model? If so, it should be elaborated here."

AR: We intended to stress the relatively high temporal resolution and the imbalanced instantaneous salt inflow and outflow, i.e. that the model is not just based on annual means assuming steady-state. We have rephrased to clarify this.

RC: "P4, L1-4: See above; here I would expect to see a set of research questions, objectives and/or hypotheses which would clarify the aim of the study and also make it easy to follow throughout the manuscript if the goals are met in the study."

AR: We agree. We will add the requested text. See response to second comment above.

RC: "Methods P4, L7: should be "comprises of"."

AR: We do not agree, and neither does our dictionary.

RC: "P4, L20: Does this mean that temperature was measured from inside the sampling vessels? Please clarify in text."

AR: Thermometer inside water sampler, then in-situ with CTD. Clarified.

RC: "P5, L8-13: Why four layers? What is the justification of this number of layers and the depth ranges of those layers? Please elaborate."

AR: We used a box model with fixed depth layers and need three layers to separate deep water (primarily influenced by inflowing salt water) from a pycnocline depth layer and surface water. We further divided the surface layer since at high fresh water flows there was occasionally a shallower halocline. This was elaborated on page 6, line 20-24. We have moved these lines to the first paragraph in section 2.3 and added further explanation.

RC: "P6, L20: Replace at with during."

AR: Corrected. This sentence is moved to first paragraph of section 2.3.

RC: "Discussion P13, L21: Replace scale with resolution."

AR: Changed.

RC: "P13, L21: The reader should be reminded at this point what is IA."

AR: Agree. Corrected.

RC: "P13, L22: "...and has poor long-term P retention.." does not read well, rephrase."

AR: Rephrased.

RC: "P13, L26-27: When both models yield similar results in P budgets, what is the motivation to use both? If the poorer prediction of water exchange in the mean model is not important (indicating decoupling between surface and bottom water masses), then why include the boundary model at all?"

AR: The boundary model should be more correct if looking at the P balance of the inner area. However, the mean model is more representative regarding the internal dynamics, based on central stations data. We think we have to present all the results to be clear. The similarity of the results of both models is in itself an interesting result that adds credence to the analysis. We added clarification in this paragraph.

RC: "P14, L22: Replace yearly with annual."

AR: Corrected. Replaced throughout the manuscript.

RC: "P14, L18-19: This is an important factor causing bias in sediment studies; in practice, most sediment cores are taken from sedimentation areas (local bathymetric depressions) where sedimentation rates are higher than in their surrounding area in general. This matter should be elaborated here."

AR: We have elaborated and clarified in this paragraph (in response also to comment by reviewer 2). The final sentences of this paragraph is now a separate paragraph.

RC: "P15, L10-11: What is the significant difference between the inner and outer part here? Why do you think that the polychates do not have an effect there? This should be elaborated here."

AR: Factors other than the polychaetes could explain the observations in the cited paper. The inner part of the IA has a complex water circulation. We conclude we cannot see an effect, but lack a solid basis for further speculation.

RC: "P15, L13-21: So where does the TP in the upper layer go? It seems like sedimentation is suggested, but it is vaguely formulated and should be made clearer."

AR: We agree, and have modified the second sentence.

RC: "P16, L14: Iron obviously plays a major role in the P cycling in the sediments, therefore the iron fluxes and concentrations presented on page 17 should be brought up already here. This would better justify the lengthy discussion about the role of Fe in the system."

AR: We now bring this up earlier, and have moved the sentence on Fe loading to second paragraph in section 4.3.

RC: "P17, L15: Is this mean annual concentration? Should be made clear."

AR: Corrected.

RC: "P17, L19-20: Due to salinity differences between the surface and the bottom layers in the system, this is more like a rule than an exception, is it not? Therefore the assumption of water masses being well-mixed seems not justified."

AR: There is hardly any horizontal salinity gradient in the deep water, so this is likely not a problem. The vertical salinity differences are handled by the model.

RC: "P17, L32: Replace yearly with annual."

AR: Corrected.

RC: "P18, L10: Here, P is presumed to be deposited as organic P. Why iron-bound (inorganic) P is not considered here, as it clearly is a significant pathway for P sedimentation as presented before."

AR: The intended point here is that the organic P load was very likely higher in the 1970s than in later years, which should contribute to high oxygen consumption, and mostly low O₂ concentrations, unless O₂ inputs are particularly large. There is probably significant inorganic P sedimentation too, but this will not contribute to oxygen consumption. We have rephrased to clarify.

RC: "Conclusions Overall, this section is too long and would benefit from shortening. Ideally, conclusions are a concise presentation of the key findings, which relate to the research questions and/or hypotheses presented earlier. For instance, the seasonality aspects could be the main concluding remarks."

AR: Yes, we have now shortened the conclusions by moving most of the last two paragraphs to the new section 4.6.

RC: "Tables Table 1: Replace yearly with annual in the caption Table 2: Replace yearly with annual in the caption"

AR: Corrected.

RC: "Figure 1: Sampling points can not be distinguished well enough from their background, so they should be made larger and with better contrast."

AR: Corrected.

RC: "Figure 4: These time series seem to have an increasing trend over the study period. This could be emphasized by plotting a regression line of some kind to the plots. This would also allow the quantification of the annual increase in observed values."

AR: In principle yes. But then we should for consistency add such lines in many of the following graphs too. Many changes are not steady trends but rather short term fluctuations and thus a regression line could be a bit misleading. We prefer mentioning it in the text.

RC: "Figure 5: In panel (c), the net values are hard to read because of the scale issue. For that reason, the net values should be shown only in the separate panel (d). The motivation to show also the values from the mean model in panel (d) is unclear."

AR: We agree that the net values in panel (c) can be removed. The results of the mean model are important for comparisons, but since they are reported in Table 3, they can be removed from fig. 5d.

Authors' response (AR) to comments by referee 2 (RC)

RC: "The manuscript by Walve et al. is a local study investigating whether the Stockholm inner Archipelago acts as a source or sink for phosphorus. They use a four-level box model based on observations to calculate the inflow and outflow of phosphorus to and from the study area as well as to calculate the retention of phosphorus within the area. The study is interesting and the manuscript is overall easy to follow."

AR: We are pleased that our study was found interesting and the manuscript easy to follow, and are grateful for the constructive comments.

RC: "The title is not specific enough, since it is a local study in a limited area and not valid in the entire Baltic Sea."

AR: We have expanded the title to "A Baltic Sea estuary as phosphorus source and sink after drastic load reduction: seasonal and long-term mass-balances for the Stockholm inner archipelago 1968-2015"

RC: "The abstract is too long. It can be more concise. There are some text that would go under study site description and some text that is more of a discussion."

AR: We agree to some shortening. We deleted the sentence starting in the end of row 11 and the following sentence. We modified the sentence starting on row 13. The "discussion" sentences (lines 21-24) about causes of oxygen improvements have been replaced by a single sentence about oxygen consumption, P release and temperature.

RC: "The study is not put into a larger perspective, to what extent is this study interesting for the rest of the Europe or the World?"

AR: We have added perspective in the introduction concerning the general importance this type of studies, citing reviews by Duarte et al. (2015), Testa et al. (2017), Regnier et al. (2013), and Cloern et al. (2014), to better link to general interest and stress the importance of this type of study. We have modified the final part of the introduction to clarify the motivation for the study (see response below) and added text to stress the importance of studies in the Stockholm inner archipelago, which has often been used as an example in the debate about effects of P and N mitigation to coastal waters (Boesch et al. 2006, Schindler and Vallentyne 2008, Schindler et al. 2008). The latter reference is new and has been added to the reference list.

Our study also contributes to clarifying the importance of internal seasonal P cycling versus legacy P loads, and that oxygen control of P release alone is an oversimplification. We have also added a management point in the new section 4.6 about artificial oxygenation of bottom waters.

RC: "Are there any more studies performed in the same way, and what do they show? Weaknesses and strengths in in those and this study?"

AR: We have added more general introduction and further discussion of results of some other studies, e.g. the already cited Testa and Kemp (2008) and Staehr et al. (2017), and the new reference Boynton et al. (2018). Weaknesses and strengths of previous studies in the same area are already discussed in section 4.1, but we have modified the final paragraph of the introduction to bring up the differences already here.

RC: "There are other studies of the retention of phosphorus in the same area as this study, which are mentioned in the manuscript, but what is new with this study. What is the new scientific question?"

AR: This is the first temporally highly resolved dynamic box model study of the study area that follows the long-term P balance, including the period of change from high to low P load, and the first with seasonal resolution. Moreover, we link P retention to oxygen concentrations and oxygen consumption (from oxygen budget). We find lower P retention than previous studies, even periods of net P export. We find that oxygen consumption from degradation of organic matter is predominantly directly linked to P release. We also find that low oxygen can promote an additional release of P. We

have clarified this in the final paragraph of the introduction. Discrepancies from other studies (mean annual P retention) are taken up in the discussion. We have also added text in the final paragraph of the introduction clarifying the scientific questions of the study. Such questions include:

1. In an estuary that has long been subject to a very high P loading, is there a long-term influence of legacy P once the load has been drastically reduced? What is the current annual P retention of the area?

2. To what extent does oxygen control seasonal P release and annual P retention? Can we expect improvements in oxygen conditions to increase long-term P retention?

3. Can an effect of the temperature increase on oxygen consumption and P release be identified?

RC: "It is not necessary to use too many abbreviations since it makes the text unnecessarily complex to read. One example is the abbreviation of the water layers, D, M, LS, US."

AR: We have removed most of the water layer abbreviations from the text and write "deep water " etc, including in methods P7 L1-3. However, we still need to introduce these abbreviations for use in the equations. We have removed the abbreviations introduced on P7 L23-31.

RC: "The Conclusion section is too long and should be more concise. It is too much of a discussion in it. The last section in the conclusion is not even discussed anywhere in the manuscript. From where was this concluded?"

AR: We have moved most of the last two paragraphs of "Conclusions" to a new final section 4.6 and have made other parts of the conclusion more concise.

RC: "P1, L12. what do you mean with the "greater" Stockholm Area. However, it can be questioned if the information about the sewage treatment plants should be in the abstract at all. "

AR: We agree and have deleted this sentence.

RC: "P1,L18. ".... bloom and is exported during winter...." exported to where? From the sediment to the water or to outer areas?"

AR: To outer areas. Rephrased.

RC: "P1, L22-L24 The word "probably" makes it sound as it is a discussion part and not anything that is concluded."

AR: We have removed the section about reasons for improved oxygen conditions from the abstract (lines 21-24), since this is not a main topic of the paper and needs more studies.

RC: "P1, L32 would the word "occurrence" fit better than "incidence"?"

AR: We agree and have rephrased this sentence to make it clearer.

RC: "P3,L5 It is a bit unclear what you mean with small and medium towns. I would assume that it is a question of definition which can be different in different parts of the world. What do you mean?"

AR: Largest "medium" towns are up to 150 000 inhabitants. Clarified.

RC: "P3, L25 Actually this is discussed in the study by Almroth-Rosell et al. (2015), full reference below."

AR: We now discuss and have added this reference.

RC: "P4, L13-L14 What is the reference for the description of the area?"

AR: The reference is the SMHI basin register (SMHI, 2003), used elsewhere. This reference is now added also here.

RC: "P4, L18 DIP is already defined above sections."

AR: Corrected.

RC: "P4, L29 STP is also already defined in above sections."

AR: Corrected.

RC: "P6, L16 Should it be "S in Fig. 1"?"

AR: Yes, clarified.

RC: "P7, L2 "LS" is not defined in the text. "

AR: Defined on P5 L10. However, we have remove layer abbreviations from the running text, keeping them only for the equations.

RC: "P7, L17 K as in K in fig. 1 ? or why have K within parenthesis? "

AR: Yes, modified.

RC: "P7, L27 In PEXP-MM, what does MM mean?"

AR: Mean model. The abbreviation is removed here, and defined and used in table only.

RC: "P7, L31 In PEXP-BM, BM is not defined as well. "

AR: Boundary model. The abbreviation is removed here, and defined and used in tables only.

RC: "P8, L2 The net import should be defined once if it is the import-export or the opposite, but it is by definition a net result, and the word "net" should not be included as well when it is written out as P import-export. This goes for the entire manuscript."

AR: Yes, corrected.

RC: "P8, L2 "closely mirrors", I am not sure that it can be seen so easy from the figure. What is the correlation coefficients? "

AR: This is not a straightforward calculation since stations are often not sampled the same day. Using dates that coincide or are very close we find a linear regression coefficient r^2 of 0.75. The mirroring may be visible also when there are delays, but delays obscure direct regression fits. However, we removed the word "closely".

RC: "P9, L2 Higher in. Higher than what? This can be seen also in other parts of the manuscript. Please go through and check this. "

AR: Here higher after the shift 1989 than in the previous period. Clarified here and elsewhere.

RC: "P10, L18 "closely mirrors", correlates? What is the correlation coefficients?"

AR: r^2 =0.97. We add this number in the text.

RC: "P11, L1 6b, do you mean 6c? Correct also at other places throughout the manuscript."

AR: Yes, it should be 6c (6c should also be 6b in other places).

RC: "P11, L22 change July to August?"

AR: We focus here on the July-October period.

RC: "P11, L23 insert water at the end or the row: surface water"

AR: Modified.

RC: "P11, L29-L30 "Although among years." Can be seen where? "

AR: Variability is indicated in fig 7c but is not shown explicitly for each year. We added this figure reference.

RC: "P12, L19 insert "negatively" in front of "correlated". "

AR: Modified.

RC: "P14, L18-L31 I am not sure what you would like to say with this paragraph. Rewrite."

AR: We have rephrased this paragraph and divided it into two. The first deals with the reasons why there is an apparent discrepancy between accumulation of sediments and low P retention (sometimes even net P release). The reasons are the problem of defining representative P burial from sediment accumulation rates and legacy P release from organic matter degradation. The second paragraph continues with further aspects on legacy P release.

RC: "P17, L1 Blomqvist et al. (2004) should be cited here, full reference below."

AR: Agree. We add this reference here too.

RC: " P18, L26 ". . . . 1990 on. . ." Remove "on"."

AR: Corrected.

RC: "P30, L18-L20 in c) and d) and e) it should be clarified that it is in the inner archipelago and not at a specific station."

AR: Now clarified in (d) and (e).

RC: "P30, L23-L25 Here it should be said something about the two models shown in fig. 4b."

AR: Yes, added.

RC: "P30, L26 Replace the first comma with " of". "

AR: We have corrected text here by removing "water flows of".

RC: "P31, L14 Replace "yearly" with "annual" " AR: Corrected.

RC: "P31, L24 No P-value? "

AR: As stated in Results (P13 L5-6), there is reason to be careful here due to interdependencies of concentrations and model results. This is why no p-value was added. Disregarding this, there is a significant linear regression, y=6.31-0.10x, r^2 =0.23, p<0.001 for all data. (y=6.68-0.10x, r^2 =0.38, p<0.0001 if normalized O₂ input values <0.8 are excluded). We have added p-value for the first regression within parenthesis with reference to results section.

RC: "P32, L1 Replace "yearly" with "annual" "

AR: Corrected.

RC: "Fig. 1 The letters showing the different stations cannot be seen properly. They are too small and the contrast with the background is too bad."

AR: OK, modified.

RC: "Fig.3 It is hard to distinguish the dots in the legend from each other. Make larger, and change color is one suggestion. "

AR: In Fig. 3a symbols have been enlarged and also have changed order to correspond to order of lines in graph. Symbols in Fig 3c are also enlarged.

A Baltic Sea estuary as phosphorus source and sink <u>after drastic load</u> <u>reduction: seasonal and long-term mass-balances for the Stockholm</u> <u>inner archipelago 1968-2015</u>

5 Jakob Walve¹, Maria Sandberg^{1,3}, Ulf Larsson¹, Christer Lännergren^{2,4}

¹Department of Ecology, Environment and Plant Sciences, Stockholm University, SE-106 91 Stockholm, Sweden
 ²Stockholm Vatten AB, SE-106 36 Stockholm, Sweden
 ³Present adress: Stora vägen 16, SE-795 70 Vikarbyn, Sweden
 ⁴Present adress: Träskobacken 10, SE-129 42 Hägersten, Sweden

10 Correspondence to: Jakob Walve (jakob.walve@su.se)

Abstract. Internal phosphorus (P) loading from sediments, controlled by hypoxia, is often assumed to hamper the recovery of lakes and coastal areas from eutrophic-conditionsation. In the early 1970s, the external P load to the inner archipelago of Stockholm, Sweden (Baltic Sea) was drastically reduced by improved sewage treatment, but the internal P loading and its controlling factors have been poorly quantified. We use two slightly different four-a-layer box models to calculate the area's

- 15 seasonal and annual <u>P balance (inputs, _export)</u>, retention and <u>the internal eycling of P exchange with sediments in the inner archipelago of Stockholm, Sweden (Baltic Sea) in 1968–2015. after a large reduction of P load in 1972. The area receives freshwater from Lake Mälaren and treated sewage from the greater Stockholm area. The sewage treatment plants (STPs) have improved their nutrient removal in steps, starting with P in 1972 and nitrogen in 1996. In the firstFor 10-20 years after the main P load reduction, in 1972 76, the models showsthere was, in comparison to the load, a small negative annual P</u>
- 20 balance, <u>small in comparison to the external load, and probably due to release from legacy sediment P storages</u>. <u>Later</u>, <u>T</u>the now stabilized, near neutral P balance indicates no <u>continued-remaining</u> internal loading from legacy P, but P retention is low, despite improved oxygen conditions. Seasonally, sediments are a P sink in spring and a P source in summer and autumn. Most of the deep-water P release from sediments in summer-autumn appears to be derived from the settled spring bloom and is exported to outer areas during winter. Oxygen consumption and P release in the deep water are generally
- 25 tightly coupled, indicating limited control-iron control of by P binding to iron oxyhydroxides under oxic conditionsrelease. However, enhanced P release in years of deep-water hypoxia enhanced P release suggest some contribution from redoxsensitive P pools. The oxygen conditions in the area have generally improved, probably due both to lower sedimentation of organic matter from the 1970s and lower STP ammonium loads from the late 1990s. Increased oxygen inputs to the intermediate and deep waters due to weakened stratification and enhanced vertical mixing have probably also contributed,
- 30 while increased respiration rates due to elevated bottom water temperatures probably explain worsened oxygen conditions during the 1990s. Increasing deep-water temperatures that stimulate oxygen consumption rates in early summer, have counteracted the effect of lowered organic matter sedimentation on oxygen concentrations. Since the P turnover time is short and legacy P minutesmall, measures to bind P in Stockholm inner archipelago sediments would primarily accumulate recent

P inputs, imported from the Baltic Sea and from Lake Mälaren inflow, and management here should focus on reducing external nutrient inputs.

1 Introduction

Bottom water hypoxia (here-with_oxygen concentration <2 mg L⁻¹) and anoxia is a common eutrophication problem in stratified lakes and coastal waters (Diaz and Rosenberg, 2008; Conley et al., 2011). Its root cause is excessive inputs of nutrients resulting in increased settling of organic matter, which increases oxygen (O₂) consumption. The occurrence of hypoxia is also influenced by vVariations in deep-water ventilation, which depend on basin morphology and meteorological forcing-also influence the incidence of hypoxia (Zhang et al., 2010). In addition to its direct adverse effects on benthic fauna, hypoxia also influences nutrient-cycling_processes_cycling_nitrogen (N) and phosphorus (P), with feedback effects on water

10 quality. It is well known that when bottom waters turn anoxic, phosphorus (P) release of phosphate-P from sediments can be drastically enhanced when bottom waters turn anoxic (e.g. Einsele, 1936; Mortimer, 1941; Middelburg and Levin, 2009). Such "internal P loading" is often assumed to hamper the recovery of eutrophicated ecosystems (Vahtera et al., 2007; Schindler, 2006, 2012; Vahtera et al., 2007; Puttonen et al., 2014; Stigebrandt et al., 2014; Puttonen et al., 2014).

In the Baltic Sea, one of the world's largest brackish ecosystems, control of both external and internal P loading is considered crucial for reducing the recurring summer blooms of nitrogen-fixing cyanobacteria (Vahtera et al., 2007). Here, large fluctuations in water mass P inventory are linked to the areal extent of hypoxia (Conley et al., 2002; Jilbert et al., 2011) and numerous experimental studies have documented drastically increased sediment P-release as formerly oxic sediments turn anoxic (e.g. Koop et al., 1990; Gunnars and Blomqvist, 1997). The mechanism usually invoked is the classical redoxsensitive release of <u>phosphate-P (or dissolved inorganic P₁ (DIP)</u> from P bound to iron(III)oxyhydroxides (Fe-P) as Fe is

20 reduced to Fe(II) (Mortimer, 1941; Mort et al., 2010). The importance of O₂ control of P release from sediments is also supported by in-situ flux studies showing that oxygenation of anoxic sediments can drastically decrease the P release (Ekeroth et al., 2016a; Hall et al., 2017).

The long-standing paradigm that "O₂ controls P release from sediments" is however too simple to fully describe the sediment P release processes (Gächter and Müller, 2003; Hupfer and Lewandowski, 2008). On a longer time-scale the P release from sediments is <u>due to anthe</u> imbalance between the P sedimentation and the P binding capacity <u>in-of</u> the deeper, usually anoxic, sediment (Hupfer and Lewandowski, 2008; Carey and Rydin, 2011; Rydin et al., 2011). Even though Fe(III)-bound P can be abundant in surface sediments in the Baltic Sea, it seems to contribute little to the long-term P retention, since P is permanently buried mainly in organic forms (Jensen et al., 1995; Lukkari et al., 2009; Rydin et al., 2011; Mort et al., 2010) and, under some conditions, as persistent Fe(II)-P minerals (Slomp et al., 2013; Reed et al., 2016). In marine sediments sulphate reduction promotes the binding of Fe as sulphides, decreasing the Fe available for P-binding,-and allowing considerable P release from oxic surface sediments (Caraco et al., 1990; Blomqvist et al., 2004). This suggests that the control of sediment P release is tightly coupled to sedimentation and-organic matter decomposition (Caraco et al., 1990)

<u>rather than directly by the</u> and that lower O_2 concentrations are often the result of organic matter mineralization rather than the controlling factor for P release (Hupfer and Lewandowski, 2008).

Estuaries are central in the land-ocean continuum, acting as strong nutrient filters and transforming terrestrial nutrient inputs by complex biogeochemical interactions, involving geological, physical, chemical and biological processes, affected

- 5 by a wide array of forcings, e.g. wind, light, water temperature, freshwater discharge and, where relevant, tides (Regnier et al., 2013). Due to large nutrient inputs and intense nutrient cycling, productivity is usually high, with a 10-fold variability between estuaries (Cloern et al., 2014). In a recent review, Testa et al. (2017) listed a set of challenges and key research questions for estuarine ecosystem science, e.g. maintaining spatially distributed time-series datasets and investigating likely trajectories of ecosystem recovery in response to restoration, to improve mechanistic knowledge of ecosystem change from
- 10 comparative studies. The recovery of estuaries can follow different pathways and are further complicated by climate change, emphasizing the need of comparative studies (Duarte et al 2015). Baltic Sea coastal archipelagos, and bays and estuaries are affected by P-nutrient inputs with local freshwater runoff, import from the open Baltic Sea and sometimes also by direct discharges of treated municipal sewage water. <u>IOften, internal P loading from sediments is often regarded as an important additional P source (e.g. Puttonen et al., 2014; Rydin et al., 2017). Moreover, it is often claimed (e.g. Schindler and</u>
- 15 Vallentyne, 2008), that following high P loading, historical P storages in coastal sediments may leak for many years, preventing the recovery of good water quality. Indeed, a high load of organic matter can increase the sediment organic P fraction (Lukkari et al., 2009), which potentially contributes to internal loading as external loading is relieved. Accumulated organic matter will also elevate oxygen demand, delaying improvement of O₂ conditions (Middelburg and Levin, 2009; Reed et al., 2011). In lakes, recurrent seasonal internal P loading can be an important factor delaying ecosystem recovery (Schindler, 2006; Nürnberg and LaZerte, 2016). In contrast, coastal estuaries usually have higher water-turnover rates than
- <u>lakes</u> and the resulting larger export of released P should facilitate water quality recovery.

The Stockholm archipelago, NW Baltic Sea Proper (Fig. 1), has historically received high nutrient loading <u>as-in the</u> <u>form of</u> sewage from the Stockholm area (Brattberg, 1986) and with fresh_water from Lake Mälaren (at LM in Fig. 1), which drains a large inland area with agriculture, forests and <u>several_small</u> to medium towns <u>(<150 000 inhabitants)</u>. Due to increasing discharges of mostly untreated sewage-water, water quality in the archipelago had deteriorated <u>badly-seriously</u> by the mid-20th century (Brattberg, 1986; Schindler and Vallentyne, 2008). In the early 1970's, effective BOD (biological oxygen demand) reduction and chemical P precipitation was implemented in <u>Stockholm's-the</u> major sewage treatment plants (STPs) in <u>Stockholm</u> and around Lake Mälaren (Karlgren and Ljungström, 1975). Archipelago water quality then improved rapidly, within a few yearswith P concentrations and phytoplankton biomass, particularly of cyanobacteria, <u>decreased</u>

30 <u>decreasing</u> substantially <u>in a few years</u> (Anon, 1975; Karlgren and Ljungström, 1975; Brattberg 1986; Johansson and Wallström, 2001). <u>AfterFollowing</u> Sweden's accession to the European Union<u>in 1995</u>, the two largest STPs discharging to the inner Stockholm archipelago <u>had-installed nitrogen (N)</u> removal <u>installed fromin</u> 1996, which also slightly improved P removal. As a result, water quality improved further, both in the still mostly P-limited inner archipelago and further out, where N now decreased to limiting levels (Boesch et al., 2006; Lännergren and Stehn, 2011).

Over 40 years have passed since the P and BOD loads to the Stockholm inner archipelago were drastically reduced in 1972. Levels of P and chlorophyll have declined and oxygen conditions have improved in the inner archipelago (Boesch et al., 2006; Norkko et al., 2011). Still, the sediments are suggested to be an important P source, judging from the regular late summer increase of P concentrations below the summer pychocline that propagates to the surface waters-layer in late summer to autumn (Boesch et al., 2006). Schindler and Vallentyne (2008) suggest that legacy P may still influence the area although Schindler et al. (2008) find that there are indications that P release from sediments is beginning to decrease in

- response to improved oxygen conditions. However, it It has thus been is unclear how much of the P release of this that is a seasonal recycling of P due to organic matter decomposition, the extent to which the P release is controlled by low O_2 concentrations, and whether historical legacy P is still influencing the area.
- 10

- While numerous studies have dealt with the short-term O₂-dependent release of dissolved inorganic P (DIP) from Baltic Sea coastal sediments (e.g. Ekeroth et al., 2016b), fewer studies have focused on the long-term retention of P, relative to external P inputs (e.g. Almroth-Rosell et al., 2016). Even fewer studies explore the dual role of sediments as sink and source for P-also over seasonal and annual scales (but see Almroth-Rosell et al., 2015). This may partly be a methodological issue. Although direct in-situ measurements of sediment release or uptake of DIP are relatively straightforward (e.g. Hall et 15 al., 2017; Sommer et al., 2017), they require need sophisticated equipment and require a very large sampling effort to reliably representatively cover different the variety of depths, sediment types and seasons. Moreover, these methods do not quantify the settling fluxes of P to the sediments and thus are not sufficient to determine if the sediments is a net P sink or
- source (Nielsen et al., 2001; Hupfer and Lewandowski, 2008). Unfortunately, direct sedimentation measurements are prone to large uncertainties (Blomqvist and Larsson, 1994; Gustafsson et al., 2013).
- 20 Box models can help constrain and integrate net nutrient fluxes over larger coastal areas, for full seasonal cycles and for longer time-periods (e.g. Savchuk, 2005; Testa and Kemp, 2008; Testa et al., 2017). In this study, we use a dynamic box model to analyse a time series of nearly 50 years of water P and O_2 concentrations in the Stockholm inner archipelago. Previous P budgets for the inner Stockholm archipelago (Anon., 1975; Karlgren and Ljungström, 1975; Karlsson et al., 2010; Almroth Rosell et al., 2016) have focussed focused primarily on the long-term P balance and are old (Anon., 1975; Karlgren
- 25 and Ljungström, 1975), based on long-term annual means (Karlsson et al., 2010), or rely on mechanistic benthic models (Almroth-Rosell et al., 2016). In this study, we use a temporally highly resolved dynamic (non-steady state) box model to analyse a time-series of nearly 50 years of water P and O₂ concentrations in the Stockholm inner archipelago. Firstly, longterm changes of the area's annual P balance are used to evaluate whether the previously high P loading, relieved-drastically reduced over 40 years ago, can is still influencing the coastal area due to legacy sediment P release. Our expectation was that
- 30 P retention gradually improved once the high P loading was relieved. We also wanted an improved estimate of the present P retention efficiency and its annual variability, with implications for P export to the open Baltic Sea where excess P may stimulate N-fixing cyanobacterial blooms. Secondly, we evaluate the common assumption that low O_2 concentrations is the main mechanism controlling P release, and the suggestion that improved O_2 conditions have decreased the P release. Alternatively, P release is merely related to degradation rates of organic matter, which would imply that seasonal P release is

mainly controlled by seasonal P deposition. We test these mechanisms by We also calculate calculating the area's seasonal and annual function as source and sink for P relative to the O_2 concentrations and the <u>deep-water</u> O_2 consumption rates. Finally, the temperature in the deep water of the inner archipelago has increased several degrees, although it is unclear whether this is an effect of global warming or of local water circulation patterns. The temperature increase should elevate

5 rates of O_2 consumption and P release and we use our model to test this, accounting for estimated changes in organic matter deposition rates, which also should influence O_2 consumption. We test the sensitivity of our model results by varying some basic assumptions, e.g. of on the origin of saline water flowing into the inner archipelago and of on the estimated errors of accuracy of the P concentrations measured in fresh-water input and STP loads.

2 Methods

10 **2.1 Study site**

15

20

The Stockholm archipelago comprises ~24000 islands (Hedenstierna, 1948) (Fig. 1). Its dominant freshwater source is Lake Mälaren (drainage basin 22000 km²) that empties through Norrström in central Stockholm (LM in Fig. 1) to the relatively enclosed inner archipelago (IA). This connects to the middle archipelago through one major and three shallow minor straits. The 18 m deep Oxdjupet sound (O in Fig. 1) has the only significant deep-water inflow and ~84 % of the surface water outflow (Engqvist and Andrejev, 2003). Water exchange between the inner and middle archipelago is governed mainly by estuarine circulation, generated by the large freshwater flow, while wind and barotropic forcing are less important (Engqvist and Andrejev, 2003), making the IA suitable for mass-balance calculations. Due to the estuarine circulation and large relative volumes of surface water, the deep-water retention time is relatively short (Engqvist and Andrejev, 2003). The IA local drainage area is 323 km², its water surface area 103 km², its mean depth 14 m and its maximum depth 57 m (SMHI, 2003).

2.2 Data sources

We used water chemistry data from the <u>Stockholm Vatten</u> monitoring program of <u>Stockholm Vatten</u> (1968–2015) and the investigations of <u>studies</u> by Mats Waern in 1968–1976 (Nordling et al., 1984). Data from 1968–1981 were digitized from printed reports. Sampling frequency is detailed in Supplement 1. <u>Dissolved inorganic phosphorus</u> (<u>DIP</u>) and <u>T</u>total
phosphorus (TP) and <u>DIP</u> were analysed by standard colorimetric methods (molybdate reactive P), TP after persulphate digestion. Due to analytical problems, nutrient data from 2011 were not used (Walve, 2012). Temperature was measured with-in the water samplers until 1991, and <u>afterwards later</u> at the same depths-with an in-situ CTD probe at the same depths. Salinity was calculated from conductivity of water samples in 1968-1991 and 2011-2015, and <u>from conductivity</u> measured in-situ by CTD in 1992-2010. Oxygen was measured by Winkler titration (SS-EN 25813-1), hydrogen sulphide (H₂S) by spectrophotometer (SS 028115-1) and converted to negative oxygen. On a The few times when few occasions, surveys by

Stockholm Vatten and the Mats Waern group sampled the same station on the same day, with good agreement was good (salinity $r^2=0.99$, slope=0.99, n=19, TP $r^2=0.99$, slope=0.98, n=15).

Data on the daily Lake Mälaren freshwater discharge from Lake Mälaren, calculated by Stockholms Hamnar AB (Ports of Stockholm), were provided by Stockholm Vatten and have an estimated error of <5 % (SMHI, 2006).
Concentrations of TP and DIP in the discharge of Lake Mälaren were measured monthly (1968–1973) or weekly (from 1974) by Stockholm Vatten. Sewage treatment plant (STP) <u>DSTP</u> ddischarges of water and TP to the inner archipelago from STPs-were obtained from Stockholm Vatten and Käppalaförbundet (Henriksdal STP quarterly data 1971–1987, weekly 1988–2013, monthly 2014-15, Käppala_STP quarterly 1973–1985, weekly 1986–2015, Loudden STP quarterly 1971–1988, weekly 1989–2003), Bromma STP weekly 1989–2013, monthly 2014-15 and Hemmesta STP weekly 2004–2006).
For 1968–1970, we only found summed annual TP load from all STPs discharging into the inner archipelago (Anon., 1975).

DIP load, analysed on unfiltered water with an unknown fraction of Fe-bound P, was available from 1989. The Swedish Meteorological and Hydrological Institute (SMHI) provided data on-precipitation data (station Stockholm 98210) and local land runoff 1977-2015 (S-HYPE model). We used the monthly means for 1977-2015 to estimate local runoff for 1968-1976. Since local runoff is <2 % of total annual fresh-water inputs, resulting errors are small.

15 **2.3 Water and salt box model**

We used a box model (Fig. 2) with four <u>fixed_depth</u> layers to calculate the water, salt and P budgets of the IA, <u>with two</u> <u>different approaches (see below)</u>. The model assumes an estuarine circulation (Engqvist and Andrejev, 2003), with outflowing brackish surface water and a counter-current of inflowing saltier water (SW) from the middle archipelago to the deep water of the IA, <u>-creating a corresponding upwelling of deep water in the IA. We also included mixing between water</u>

- 20 layers. The deep-water layer (D, 20-57 m), which has spatially and vertically relatively homogenous salinity in the IA, is separated from the surface water (S, 0-10 m) by an intermediate pycnocline layer (M, 10-20 m). Since there is often a shallower pycnocline during high freshwater input, the surface layer was divided into an upper surface layer (US, 0-4 m), representing outflowing water, and a lower surface layer (LS, 0-4 m). On most occasions the upper water mass is well-mixed to 10 m, suggesting a deeper outflowing layer, but the 0-4 m layer salinity is still representative due to the weak vertical salinity and nutrients gradients on these occasions. We performed sensitivity tests of model results to different depths of the
- outflowing layer (see section 2.6). Model box volumes (392 Mm³ for the upper surface layer US 0-4 m, 465 Mm³ for the lower surface layer LS 4-10 m, 465 Mm³ for the intermediate layer M 10-20 m, and 214 Mm³ for the deep water layer D 20-57 m) were calculated from hypsographic curves of IA basins (SMHI, 2003), excluding three relatively small and isolated small, peripheral basins (Brunnsviken, Edsviken and Kyrkfjärden).
- 30

We used total freshwater inflow (FW_t), salinity of the IA and salinity of inflowing SW to constrain water flows, with two different approaches. Firstly, we used volume-weighted mean salinity of the IA to constrain water flows, typical for conventional one-dimensional box models. Thus, this "*mean model*" yields a SW inflow matching the salt amount in the IA and the salinity of outflowing water (as volume-weighted mean for surface water, i.e. Sal_{out} equals Sal_{US} , Fig. 2). We used

this approach to estimate representative mean internal P fluxes in the IA. For some analyses, we also used a secondary "boundary model". It adds an extra boundary SW inflow needed to match the salinity in the surface water near the boundary, which is higher than the volume-weighted mean surface water salinity of the IA. Thus, the boundary model accounts for the surface water salinity gradient by adding to the water turnover in the outer part of the IA, not represented in the mean model.

5 We give detailed calculations below. We used the boundary model to correct the P input-output budget but do not regard it as representative for the mean internal P budget, since it is mainly a "short-circuit" flow occurring in the outermost sub-

Equations for water exchange used Knudsen's theorem for estuarine circulation (Knudsen, 1900), which assumes constant water volume (Eq. 1) and equilibrium of salt transport (Eq. 2),

basins of the IA. We made several sensitivity tests to check the effects of various assumptions on the results.

15

$$Q_{out} = Q_{in} + Q_{FWt} \tag{1}$$

$$Q_{out} Sal_{out} = Q_{in} Sal_{in} \qquad (2)$$

where Q_{out} is the outflowing volume of surface water (with the relatively low salinity $Sal_{out} = Sal_{US}$ in the mean model), Q_{in} is the inflowing volume of deep water (with the relatively high salinity Sal_{in}), and Q_{FWt} is the total fresh-water input to the estuary, i.e. the sum of Lake Mälaren discharge, local land runoff, precipitation and waste water discharge. Qin is calculated from equations (1) and (2) according to

(3)

(4)

$$Q_{in} = Q_{FWt} Sal_{out}/(Sal_{in} - Sal_{out})$$

Since the Knudsen equations assume steady-state, which for salt is a valid approximation only for a period longer than the turn-over time of the estuary, where volume-weighted salinity can shift with up to one unit within a month, we modified the salt mass-balance equation to include changes in total salt amount, ΔS_{IA} , of the IA,

$$\Delta S_{IA} = Q_{in} Sal_{in} - Q_{out} Sal_{out}$$

 Q_{in} is calculated from Eq. (1) and (4) according to

$$Q_{in} = (\Delta S_{IAtotal} + Q_{FWt}Sal_{out})/(Sal_{in} - Sal_{out})$$
(5)

This is A similar to the box model approach was developed by Hagy et al. (2000) and used by Testa and Kemp (2008). Testa et al. (2008) and Boynton et al. (2008) to calculate salt and water transport in the Patuxent River estuary of Chesapeake bay.

25

The model interpolates the daily changes in salt amount ΔS_{IA} from volume-weighted salinity data for each depth layer, calculated from measured salinity profiles (4 m depth resolution) at the central stations (A, AV, H, L, K in Fig. 1), and water volumes for 1 m depth layers (SMHI, 2003). We tested the representativeness of the central stations with data from all subareas 1998–2005. The central stations slightly underestimated salinity and DIP, but we found no systematic difference for TP

(Suppl. 2). We used these relationships to correct data for each depth layer. A model sensitivity test with non-adjusted data 30

showed smaller uncertainty than for other sensitivity tests (see below). In some years absent winter (February) data for the IA central stations were estimated from Solöfjärden (S in Fig. 1) data using correlations to the central stations, and correlations between IA surface water and deeper layers for years with full February data sets (see sampling frequency in Suppl. 1). December data were generally lacking but model sensitivity tests, run with extrapolated November or February data, rather than from linear interpolation, showed less divergence than for other sensitivity tests (see below).

Since there is often a shallow pycnocline during high fresh water input, the volume-weighted mean salinity of the 0.4 meter layer was used as the salinity of the out flowing water. On most occasions the upper water mass is well mixed to 10 m, suggesting a deeper out flowing layer, but the 0.4 m layer salinity is still representative due to the weak vertical salinity and nutrients gradients on these occasions. The salinity at 20–30 m depth in Trälhavet (T in Fig. 1), just outside the boundary strait Oxdjupet (O in Fig. 1), was used to represent inflowing deep water (Lännergren and Stehn, 2011). This salinity is in good agreement with the bottom water salinity in Solöfjärden (S in Fig. 1), the IA sub-basin closest to Oxdjupet. Oxdjupet data also support the use of this layer (Lännergren and Stehn, 2011), but were not used in the model, due to large data gaps.

The wind-driven and diffusive vertical mixing (Fig. 2) was calculated according to the salt mass-balance equations for the bottom-deep layer D (Eq. 6), intermediate layer M (Eq. 7), and the upper surface layer US (Eq. 8):

15
$$\Delta S_D = Q_{in} Sal_{in} - Q_{in} Sal_D + Q_{Dmix} Sal_M - Q_{Dmix} Sal_D$$
(6)
$$\Delta S_M = Q_{in} Sal_D - Q_{in} Sal_M + Q_{Dmix} Sal_D - Q_{Dmix} Sal_M + Q_{Mmix} Sal_{LS} - Q_{Mmix} Sal_M$$
(7)
$$\Delta S_{US} = Q_{in} Sal_{LS} - Q_{out} Sal_{US} + Q_{Smix} Sal_{LS} - Q_{Smix} Sal_{US}$$
(8)

where ΔS is the change in salt amount in each layer for the time step used, Q_{Dmix} is the vertical mixing of the intermediate (10-20m) water and deep layers D and M (10-20m), Q_{Mmix} is the mixing of the lower surface and intermediate layers M and LS, and Q_{Smix} is the mixing of upper and lower surface layers LS and US (Fig. 2). Mixing terms were calculated by rearranging the equations.

The boundary model adds an extra SW inflow Q_{inxD} according to

5

10

20

$$Q_{inxD} = (Sal_b Q_{out} - Sal_{out} Q_{out})/(Sal_{in} - Sal_b)$$
⁽⁹⁾

where Sal_{h} is surface water salinity at the boundary. This inflow generates a corresponding increase in outflow.

- The salinity profile 20 57 m of the IA central stations is relatively uniform and similar to the deep water in Solöfjärden, indicating that salinity in the deep water is determined mainly by inflowing water from the middle archipelago and that mixing with the overlying water mass is small. The main vertical salinity gradient is generally found in the intermediate 10 20 m layer (*M*). In the innermost part of the IA, discharged sewage water creates a 4-layered water flow, with an outflowing sewage-enriched layer at 8–20 m, below inflowing shallow and an outflowing surface layers and above
- 30 inflowing deep water (Boesch et al., 2006; Lännergren and Stehn, 2011). These intermediate layers are restricted to the innermost part of the IA and do not affect the water and salt budgets of the box model. Moreover, directing STP flows to

either the US-upper surface or M-intermediate layers has no effect on our results as presented, since we-they are aggregated results for the upper 0-20 m.

The modelling used the software ExtendSim8 (Imagine That Inc.), which is suitable for sequential computations and interpolations of large data sets. We used a time step of 1 day and computation $dt = 100 \text{ day}^{-1}$.

5

We correlated model results with water temperature data from the representative station Koviksudde (K<u>in Fig. 1</u>) with the most complete sampling. Data were compiled for depth intervals with consistent data for all sampling occasions. Means were calculated for depth intervals 0-8 meters (data mostly from 0, 4 and 8 m), 12-20 m (mostly 12, 16 and 20 m), and 24-32 m (24, 28 and 32 m). Since some sampling occasions were close to the end or the beginning of months, temperature data were interpolated between occasions before calculations of seasonal means.

10 2.4 Phosphorus model

Phosphorus inputs to the IA come from Lake Mälaren (P_{LM}) , local land runoff (P_{LR}) , precipitation on the sea surface (P_P) , STPs (P_{STP}) -and inflowing saltier water from the middle archipelago (P_{SW}) . P import (P_{SW}) -from the middle archipelago was calculated from the volume of inflowing water Q_{in} and the mean P concentration at 20–30 m depth in Trälhavet (T in Fig. 1). Atmospheric P deposition on the IA is ~15 kg km⁻² yr⁻¹ (Savchuk et al., 2008) – minute compared to other P sources, and therefore disregarded

15 therefore disregarded.

In the The P export of the mean model (defined above), P export (P_{EXP-MM})-is a function of the mean model water outflow and the volume-weighted P concentration in the upper surface layer. The P concentration in this and all other layers was continuously adjusted to fit interpolated observations, by additions or removal from the model. The summed corrections (P_t)-were used to as estimates of net internal losses from, and or inputs to, the water column.

20 In the boundary model, The P export of the boundary model (*P_{EXP-BM}*)-was calculated using the out-flow of the mean model and the P concentration at 0–4 m depth near the boundary (defined by stations S, O and TF, Fig. 1), and then adding the extra flow across the boundary that was needed to match the boundary salinities. The TP export calculated by the boundary model also had the advantage of larger data availability of TP concentrations at the boundary than at the central stations of the IA (Suppl. 1). Using the boundary model, monthly and annual net external P import–export balance was calculated. Thus, a positive P balance means that import from external sources are larger than the export. The extra flows across thegenerated in the boundary model generally resulted in additional annual net P loss from the inner archipelago, due to mostly higher P concentrations in the outflowing surface water than in the inflow (defined by concentration in deep water

at station T).

2.5 Oxygen budget for the deep water

30 The An O_2 budget was calculated only for the deep layer (>20 m). In the model, O_2 is imported to the deep water by SW inflow and mixing with overlying water (10-20 m). Oxygen is exported by upwelling and mixing into overlying water. Consumption of O_2 in the deep water was calculated as for P, i.e. with a function continuously adjusting the estimated

concentration to fit observations. The O_2 consumption was converted to a theoretical mineralization of P assuming Redfield composition (1 mol P per 106 mol O_2) of organic matter. The complete oxidation of ammonium, generated from organic matter, to nitrate would give a 30 % lower P yield (1 mol P per 138 mol O_2). We used the higher P yield (1:106) because of observed_ammonium_periodically accumulated of ammonium in the deep water, and because unknown_ammonium

5 transfer to intermediate water layers was unknown, as was also the useand unknown contribution of nitrate into deep-water organic matter oxidation (denitrification).

2.6 Sensitivity analysis

We tested the robustness of the P budgets by 1) varying the fresh-water flow from Lake Mälaren by ± 5 %, 2) applying assuming different recruitment depths of for the inflowing SW (12–20, 16–24 and 24–32 m compared to the default depth

10 20–30 m)-from the middle archipelago, 3) assuming a ± 10 % error in the P discharge from the STPs, 4) assuming a ± 10 % error in TP and DIP measurements, and 5) changing the depth layer of the outflowing surface water to 0-10 meters.

3 Results

3.1 Salinity and temperature

- The salinity of the deep water (>20 m) in the IA elosely mirrors the salinity at 20–30 meters depth in Trälhavet (T in Fig. 1),
 the assumed recruitment depth for inflowing salt water (SW) (Fig. 3a). The SW salinity peaked around 1980 at 6.0-6.5 and has since slowly declined and stabilized at ~5.5. The volume-weighted mean salinity of the IA shows-showed considerable seasonal variation (Fig. 3a), mainly due to variable freshwater input affecting the surface water salinity (Fig. 3b). The long-term trend of IA mean salinity mirrors that of the deep water but shows-showed a more abrupt decline, influenced by the large FW inflows in 1977-1988 (Fig. 4a). From the late 1980's, a less distinct halocline and a deeper surface mixed layer
 (Fig. 3b) accompanied the salinity decrease in the deep water. The pycnocline depth (depth of maximum density change per meter) increased from ~10 to ~14 m and its strength (density change per meter) decreased (data not shown but indicated in Fig. 3b). From 1989, there was an abrupt increase (~2°C) of mean July to October temperature in the intermediate and deep layers (Fig. 3c). There is a seasonal temperature rise from July to October (of by ~3 °C), but this showed little changed little, and the higher mean July-October temperature from 1989 on, relative to the previous period, was mostly the result of larger
- 25 seasonal temperature rises earlier in the year. From After 1989, early May temperatures were ~0.5°C and ~1°C higher than previously in the intermediate layer and ~1°C higher in the deep layers, respectively, and with early July temperatures ~2°C and ~1°C higher, respectively. After 2007, temperatures appear to have increased even further (Fig. 3c).

3.2 Phosphorus concentrations

Following the reduction of phosphorus (P) load from the STPs in the early 1970's (see below), annual mean total P concentrations (TP) in the IA decreased from >100 μ g L⁻¹ to around 49 μ g L⁻¹ in 1976-85 and 39 μ g L⁻¹ in 1986-1995 (Table

1, Fig. 3d, e). After the P-treatment upgrade in 1996, TP declined further to 30-32 μ g L⁻¹ (Table 1). In 1968-75, TP was lower in the deep than in the surface water, but this was reversed after the P load reduction (in the following, P refers to TP unless otherwise stated). The P decline was less prominent in the deep water than in the surface water and the deep water varied more among years (Fig. 3d, e). In the early 1970's, substantial amounts of dissolved inorganic phosphorus (DIP) remained in the upper 20 m all year (Fig. 3d). From the mid-1970's, summer (June–August) DIP concentrations declined to values close to or below the analytical quantification limit (3 μ g L⁻¹). In the deep water, DIP was consistently the main P fraction (Fig. 3e). After the P load reduction in the 1970's, the P concentration in the inflowing salt water (SW) from the

3.3 Water budget

5

Total fresh-water input (FW_t) to the IA 1968-2015 (5296 \pm 1418 Mm³ yr⁻¹, or 168 \pm 45 m³ s⁻¹, mean \pm SD among years) was dominated by Lake Mälaren (94 \pm 2.1 %), followed by treated sewage (3.6 \pm 1.2 %), local runoff (1.7 \pm 0.6 %) and direct precipitation on the water<u>surface</u> (1.1 \pm 0.4 %). FW_t input shows no trend, but the early 1970's were relatively dry (1976 extremely dry), followed by a period of high discharge in the late 70's to the late 80's (Fig. 4a, Table 1). In the 90's there was again relatively low discharge (1996 and 1997 were especially dry), followed by high discharges around year 2000 and

middle archipelago (Trälhavet) has declined from ~ 35 to $\sim 25 \mu g L^{-1}$ (Table 2).

- at the end of the time series (Fig. 4a, Table 1). On average, the annual salt water inflow from the middle archipelago (Q_{sw}) was 74 ±22 % of the FW_t input when calculated by the mean model, and 136 ±54 % of the FW_t input, calculated by the mean and by the boundary models, respectively (Fig. 4b, Table 2). Salt-water inflow <u>varied</u> showed less inter annual variation<u>among years</u> than the FW_t discharge (Fig. 4), and on an annual basis, there was no correlation between these flows. The SW inflow was higher at-in the second compared to the first half end-of the time series, especially when calculated by
- 20 the boundary model (Fig. 4b, Table 2). The sensitivity analyses with variable recruitment depth of SW, or variable discharge from Lake Mälaren, changed the SW inflow only ~6 % relative to default models, except for the scenario with shallowest (12-20 m) recruitment depth (+15 %), and the scenario with export depth 0-10 m (+36 %). The mean model gives an annual mean water turnover time of 2.2 ±0.6 months. The seasonal variation is large, ranging from <u>an average of 1.7</u> months in April to 5.2 months in August.

25 3.4 External P loads

In the late 1960's and early 1970's, the direct P discharges from STPs to the inner archipelago were very high, peaking at over 800 tons in 1970 (Fig. 5a). In 1976, when P treatment had been implemented, P loads from <u>Stockholms's</u> STPs had decreased by a factor of ten (Fig. 5a). <u>During this period</u>, <u>R</u>reduced P loads from other <u>STPs</u> to Lake Mälaren <u>also</u> substantially decreased P inputs substantially also FW (FWt minus IA STPs) P inputs during this period. From After 1976,

30 the annual P load from STPs has mostly been considerably smaller than the P inputs from FW or SW inflow (Fig. 5b, Table 1). From 1976 to 1995, the P load from STPs decreased gradually decreased further, from ~80 to 45 tons yr⁻¹ and after the 1996 P-treatment upgrade the load was reduced to ~30 tons yr⁻¹, i.e. to approximately 10 % of the total external (STPs, FW)

and SW) P load to the inner archipelago. From 1976, the mean annual P loading with FW has been 170 \pm 62 tons yr⁻¹, most of whichpredominantly from Lake Mälaren (162 \pm 62 tons yr⁻¹, Table 1). Although there is a significant increase of the modeled inflowing SW volume with time, there is no equivalent increase in the SW P input, because of the decreasing SW P concentrations in Trälhavet (in Trälhavet, T in Fig. 1, Table 2). Atmospheric P deposition is negligible in the P budget of the inner archipelago, contributing only about 1.5 tons yr⁻¹ (using 15 kg P km⁻² yr⁻¹ from Savchuk et al., 2008).

5

10

The Lake Mälaren and SW DIP inputs (data not shown) follow the same pattern as the TP load. From 1989 (when data on DIP loads from STPs are-become available) on average 15 tons of DIP per year originate from STPs, 10 % of the total external DIP inputs to the area, or 17 % of DIP load from land. In contrast to the decrease in TP load, DIP load from STPs remained virtually unchanged after the upgraded-P-treatment upgrade in 1996, indicating that mainly particulate P was removed.

3.5 Annual and long-term P input-export budget

The annual P export from the inner archipelago closely mirrored (linear regression r²=0.97) the summed annual external P loading (including P in SW inflow) (Fig. 5c). Thus, the annual P loading mainly governed the P export-(linear regression r²=0.97). The input-export difference shows the net P balance of the area, with positive values indicating that the area is a net P sink (export < input, i.e. positive P balance and [positive] net P retention), while negative values indicate it is a net P source (export > input, i.e. negative P balance). The annual input-export difference was generally a small fraction (mostly <10 %) of the inputs and exports and showed considerable variation among years (Fig. 5c, d). The boundary model sensitivity analysis indicated an uncertainty of yearlyannual P balance of approximately ±25 tons (Fig. 5d). Following the major relaxation of external P loads in the 1970's, the long-term mean P balance was initially negative (1976-1985- and 1986-1995), but later became slightly positive (1996-2005 and 2006-2015) (Table 3). Sensitivity tests show largest influence sensitivity toof assumed SW recruitment depth, followed by potential errors in measured TP concentrations. Excluding the most extreme scenario with the shallowest SW recruitment depth, the and in 1996-2005 and 2006-2015 mean P balance in 1996-2015 ranged from -19 to +15 tons/year, excluding the most extreme scenario with the shallowest SW recruitment depth for symplement of increasing P retention (0.9 tons year⁻¹).

25 $r^2=0.16$, p<0.02), but there was no trend for 1996-2015 there was no trend ($r^2<0.01$, p>0.8) (Fig. 5d).

In 1968-74, before full-effective P reduction in STPs, P retention was highly variable (Fig. 5d). As STPs in this period were the main source of P, assuming 10 % uncertainty in the STP load will mean a proportionally much-larger uncertainty in the retention value. In 1969 and 1970, years with relatively good O_2 conditions (Fig. 6b6c), a positive P retention was relatively robust. From 1976, when the P loads had decreased substantially, our results indicate a 9-year-period with mostly

30 negative P balance, i.e. the-<u>larger</u> export was larger than input. The discrepancy between the boundary model and the mean model for some of these years (Fig. 5d) was due to better data availability <u>for the boundary model</u> in winter, the season with most of the P export. This was followed by a 5-year-period (1985-1989) with a balanced P input and export which coincided with better O₂ conditions both in the deep (>20 m) and intermediate (10-20m) water. As O₂ conditions deteriorated again in

the 1990's, there was clear net P loss from the IA. This may partly have been release of P deposited in the previous (1985-89), more O₂-rich period (1985-89). In 1992 and 1993, the high P loss from the IA (Fig. 5d) was associated with relatively wide-spread hypoxia both in the deep water and the 10-20 m layer (Fig. 6cb), and exceptionally poor P retention <20 m (see next section and Fig. 6a), suggesting much-high P release from shallow bottoms shallower than 20 meters.

5 3.6 Internal phosphorus P fluxes

10

Annual P budgets were also derived separately for the deep water (>20 m) and for the combined surface and intermediate water layers (0-20 m), using the mean model. For each layer, we calculated internal losses or sources of P from the sum of daily adjustments of the P modelled from external inputs that were required to match the observed P. In the upper 0-20 m layer, the net loss of P needed to balance the annual P budget was interpreted as due to net P sedimentation (Fig. 6a). Conversely, a source of P was needed in the deep_-water (Fig. 6a), mostly in the form of DIP (results not shown, but see Fig. 3e), and was interpreted as a net sediment P release. The deep-water sediments released on average 67 ± 19 tons P yr⁻¹ (2.1 g m⁻² yr⁻¹, i.e. 5.8 mg m⁻² day⁻¹, or 0.19 mmol m⁻² day⁻¹ for 32.3 km² bottom area below 20 m depth) in 1976–2015, with no trend (Fig. 6a). The annual net P loss from 0-20 m was similar, on average 61 ± 29 tons for 1976–2015. The estimated deep-water P-release was most sensitive to assumed potential errors in P concentrations and SW recruitment depth (Table 4).

Most net P sedimentation from the upper 20 m layer (~56 % of annual) occurs in spring (March-May), most of the deep-water net P release (~57 %) in summer and autumn (July-October, Fig. 6be, and Fig. 7 and Suppl. 3). Maximum P release usually occurs in September (13 tons per month, i.e. 13 mg m⁻² day⁻¹ or 0.43 mmol m⁻² day⁻¹). In spring, the net P loss from surface water layers decreases the P pool and removes some of the external P input, and the P export is smaller than the input (Fig. 7). In late summer, when the P pool starts to rebuild, there is net P input in the deep_-water (partly also advected to surface layers) and a_small net P loss from surface layers (Fig. 7). Also contributing to increasing P concentrations in summer and autumn are the larger external P inputs compared to P exports. The yearlyannual net P sedimentation in spring (March-May) correlated with deep-water sediment P release in summer and autumn (July-October) (Fig. 8a), which is consistent with a predominantly internal recycling of P within the plankton growth season March-October. Although there is on average a net P loss from the 0-20 m layer also in summer, it varies considerably among years (Fig. 7c). In some a few years, there is a net internal P source in the upper layer 0-20 m in summer, suggesting larger P release from shallow sediments was larger than gross P sedimentation from the surface water layer.

The combined annual net release or loss for all water layers indicates if the inner archipelago sediments acts as a source or sink for P relative to the water column (Fig. 6a). The pattern was similar to the net input-export budget (Fig. 5d), i.e. in years with net P release to the water, there was often, but not always, an <u>annual larger-negative net-P exportbalance for</u> 30 <u>the IA</u>. Differences are partly due to the different approaches used (mean and boundary models, respectively), but delays between years also contribute, as P accumulated in the water late in the year is partly exported early in the following year.

3.7 Oxygen concentrations

5

10

Oxygen (O₂) concentrations in the deep water vary seasonally, being highest in January–May and lowest usually in September–October (Fig. 7). Anoxia in the central inner basin has been rare since 1976, and hydrogen sulphide was only recorded in 1993 (data not shown). It should, however, be noted that seasonal occurrence of hydrogen sulphide is common in enclosed, peripheral areas of the inner archipelago (Lännergren and Stehn, 2011). The <u>yearlyannual</u> minimum volumeweighted O₂ concentration in the deep water shows a general improving trend (+0.07 mg L⁻¹ yr⁻¹, r²=0.46, p<0.0001, Fig. 6cb). There are also decadal variations, with bad conditions in the 1970's, better in the 1980's, worse again in the 1990's and finally better from 2001. The <u>yearlyannual</u> maximum extent of hypoxic (O₂ <2 mg L⁻¹) bottom areas was estimated using the hypsographic curve (bottom area per depth layer) of the IA. In the period 1968-1980, hypoxic areas varied greatly, from <10 to 100 % of deep-water bottoms and 0 to 100 % of 10-20 m bottoms (Fig. 6cb). In the 1980's only minor deep areas were

hypoxic (except for 1984), followed in the 1990's by larger hypoxic areas and then a return to only occasional deep-water hypoxia in the new millennium. From 1994, there has been no hypoxia in the 10-20 m depth interval.

3.8 Relationship between O2 and P release in the deep water

The deep-water P release in July-October <u>negatively</u> correlated <u>negatively</u> with the <u>annual</u> minimum O₂ concentration. 15 <u>usually occurring in September or October</u> (Fig. 8b). This may partly be due to the classic mechanism of enhanced P release at low O₂ concentrations, but is not necessarily a cause and effect relationship. Both P release and O₂ depletion may be a consequence of organic matter degradation, and thus primarily linked to organic matter sedimentation, as is indicated by the correlation between the net P loss in spring and the net P release in summer and autumn (Fig. 8a). To shed light on the cause and effect relationships we also modelled deep-water O₂ consumption, which was a function of the seasonal decrease of

20 volume-weighted O₂ concentration, the SW O₂ inflow (mean model), and the net O₂ input to the deep water from mixing with the 10-20 m layer. The O₂ consumption rate (Fig. 7) was low in spring, increased through the summer and peaked in September-October (at 800 tons per month, i.e. 0.8 g m⁻² day⁻¹ or 26 mmol O₂ m⁻² day⁻¹ for 32.3 km² bottom area below 20 m depth).

Theoretically expected <u>The</u> P release rates (P₀₂) theoretically expected from mineralization of organic matter were calculated from deep-water O₂ consumption, assuming mineralization of organic matter of Redfield elemental composition (see methods). <u>In Mm</u>ost years, the annual P₀₂ calculated from organic matter decomposition was close to the modelled annual TP and DIP release rates (Fig. 6a-b), <u>but tended to be lower than</u>, <u>Mm</u>odelled TP and DIP release tended to be higher than P₀₂-in years with low minimum O₂ concentrations (e.g. 1984, 1992-93, 1996, 2000, 2008 and 2012) (Fig. 6a and Fig. 9a-b). When the comparison was restricted to July-October, P₀₂ explained less of the P release (Fig. 6<u>be</u> and Fig. 9-c-d), with

30 largest deviation for September (Fig. 7). Conversely, the <u>modelled</u> P release was generally somewhat lower than P_{O2} in May-June (Fig. 7, <u>Suppl. 3</u>).

3.9 Factors affecting oxygen conditions

The observed O_2 minimum correlates with modelled O_2 consumption in July-October (Fig. 8c, here converted to P release, see above). Since the calculation of O_2 consumption depends on the observed O_2 minimum concentration, a correlation is expected and should be interpreted with caution. We find a similar correlation between the O_2 minimum, which is usually in

5 September, and the O₂ consumption in July (data not shown). Although our model results indicate that the O₂ consumption rate influences <u>the O₂</u> concentrations of the deep water, hypoxia <u>usually-primarily</u> developed in years with low O₂ inputs to the deep_-water, i.e. <u>when</u> the sum of O₂ inputs from boundary inflow and mixing with intermediate water <u>was low</u> (Fig. 8c).

Mean O_2 consumption in July-October (converted to P mineralization of from organic matter, P_{O2}) did not correlate with water temperature (Fig. 10a). However, there was a tendency of higher P_{O2} at high temperatures when normalized to the

- 10 net organic matter sedimentation in spring (as modelled P loss from upper water mass, Fig. 10b), suggesting a higher degradation rate per amount of settled organic matter at higher temperatures. The results were clearer when the same analysis (P_{02} per P sedimentation as a function of temperature) was done for each month separately (data not shown), with similar slopes in July, August and September, and with decreasing correlation coefficients decreasing from July on (r^2 = 0.41, 0.30, 0.18, <0.01, for each month July-October, respectively), and with no correlation in October (r^2 <0.01) (data not shown).
- 15 There was no long-term trend in Θ_2 consumption-July-October \underline{O}_2 consumption (Fig. 11). The <u>A</u> decrease from the 1970's to the 1980's was soon-followed by increased O_2 consumption in the 1990's (Fig. 11), possibly influenced by increased temperatures.

4 Discussion

- Our study highlights the dual function of sediments in a coastal area as sink and source for P and shows that this function strongly depends on temporal scaleresolution. Sediments of the Stockholm inner archipelago (IA) change from a P sink in spring to a P source in summer and autumn, are an annual net source in some years, especially when there is hypoxia, and has a poor retains little of the long term P retention relative to P inputs to the IA_over the long term. We used two different model approaches to evaluate the internal P fluxes and the P input-export budget of the Stockholm inner archipelago (IA). By accounting for the surface salinity gradient from the IA central stations to the boundary, the boundary model gives the mosta more realistic boundary flow, considerably higher than in the mean model, where water turnover will balance mean IA salinity. Despite this difference, both models result in similar input-export P budgets, with differences for most years within the uncertainty estimates of the sensitivity tests. We consider the boundary model the most reliable for the input-export budget due to the more realistic boundary flow and the larger data availability at the boundary compared to IA central stations, especially in winter, when net P export is highest. In contrast, water turnover rates of the mean model should be
- 30 <u>most-more</u> representative, <u>since it uses</u> when using the central stations to estimate mean internal fluxes in surface and deep water. Fortunately, the period with highest internal flux rates (April to October) has the best data availability. The similarity of the input-export budgets of the two models suggest that internal nutrient fluxes in the outer IA section are not

fundamentally different from that represented by the central stations. More winter data for the boundary model in the period from the late 1970s to the early 1980s may have contributed to the larger P export of the boundary model in this period (Table 3).

4.1 Legacy P release and Long-term-phosphorus P retention

- 5 Our study highlights the dual function of sediments in a coastal area as sink and source for P and shows that this function strongly depends on temporal resolution. Sediments of the IA change from a P sink in spring to a P source in summer and autumn, are an annual net source in some years, especially when there is hypoxia, and retains little of the P inputs to the IA over the long term. Two previous studies have estimated long-term P budgets for the Stockholm IA. Using a mass-balance P budget based on long-term annual means, Karlsson et al. (2010) calculated a P retention of 51 tons per year for the period
- 10 1982-95 and 53 tons yr⁻¹ for 1996-2007. They also calculated a P retention of 70 tons yr⁻¹ based on sediment accumulation rates. Almroth-Rosell et al. (2016) used coupled physical and biogeochemical models to calculate a P retention of 30 tons yr⁻¹ for 1990-2012, corresponding to 18 % of the land-derived P loading. In their biogeochemical model, sediment DIP release was O₂ sensitive and a fraction of P was assumed to be permanently buried, apparently a nearly constant amount per year. Thus, the retention estimate was strongly dependent on the <u>chosen-assumed</u> model parameters. <u>Compared to other coastal</u>
- 15 areas a relatively low P retention efficiency (<20 %) is expected for the Stockholm IA, based on its low water residence time (Almroth-Rosell et al., 2016; Nixon et al., 1996; Nielsen et al., 2001).

<u>Overall, o</u>Our results indicate a near balance between input and export of P., with a tendency to net export <u>D</u>during the first two decades following the main P treatment upgrade in the 1970's, there was a tendency to net P export ("negative retention"), indicating legacy P release. For the last 20-year period, our sensitivity tests suggest a long-term mean P retention

- 20 of at most 15 tons per year (Table 3), corresponding to 8 % of summed FW and STP inputs, half the value found-calculated by Almroth-Rosell et al. (2016). Compared to other coastal areas, a relatively low P retention efficiency (<20 %) is should be expected for the Stockholm IA, based on its low water residence time (Almroth Rosell et al., 2016; Nixon et al., 1996; Nielsen et al., 2001; Almroth-Rosell et al., 2016;). It is also-difficult to reconcile a sediment net burial of 50-70 tons P per year (Karlsson et al., 2010) with our internal flux estimates of net P sedimentation losses from the 0-20 m water layer in</p>
- 25 March-May of usually less than 50 tons. There is, however, _an apparent contradiction between a-the_very low P retention and the presence of accumulation bottoms with a net sediment deposition in the inner archipelago, with a net deposition of sediments (Jonsson et al., 2003). The difficulty of assessing estimating representative area-integrated net sediment accumulation rates from sediment cores that are sampled in the deepest areas may contribute to this discrepancy. A major partMost of the dry matter sedimentation in the archipelago areas is probably resuspended sediments from shallow areas,
- 30 with a large fraction having the elemental signature of clay (Blomqvist and Larsson, 1994). Persson and Jonsson (2000) found large variation in annual sediment accumulation rates, related to the <u>frequency occurrenceoccurrence</u> of strong wind events, and concluded that erosion of shallow bottoms contribute much <u>to-of</u> the sediment accumulation in the deeper areas. <u>It may therefore be misleading to assume that s</u>Sediment accumulation rates in deep areas therefore do not represent net loss

of P. Moreover, <u>mobilization of accumulated P from deeper sediment layers may also contribute to the apparent</u> <u>contradiction between net P release and net accumulation of sediments in deep areas.</u> Tehe net P export that seem to have <u>followedafter</u> the external P load reduction in the 1970's indicates that sediments were <u>then</u> a net source of P, <u>possibly</u> <u>presumably</u> as a legacy of the previous high P loading. Degradation of organic P, combined with limited retention of Fe-P in surface sediments during poor O₂ conditions, may have promoted driven this excessive high P release.

- However, <u>ILegacy P release</u>, observed for up to 20 years after the main P load reduction, iswas small relative to the previously very high P loading, with accumulated inputs from <u>STPs alone</u> <u>1968</u> 75 as high as 3900 tons in <u>1968-75</u> from <u>STPs only</u> (see Table 1), legacy P release is small compared to a net-gative P balance indicating net P export of 450 tons in 1976-1995 (260 tons in 1976-1985). Conditions favouring large P export also at high P loading seem to have limited the
- 10 build-up of a large sediment legacy P storages. (Fig. 5c). Sediment cores from the 1990's indicate-provide no evidence of a remaining legacy of carbon in IA sediments (Jonsson et al., 2003). Similarly, DIP mass-balances for a shallow Danish estuary indicated no substantial legacy P release after reduced external P loading, but rather a close link between P loading and summer DIP release (Staehr et al., 2017). Boynton et al. (2018) reviewed global estuarine and coastal sediment-water flux data and found only a few time series long enough to evaluate remedy actions. In Boston Harbor and Chesapeake Bay
- 15 sediment processes responded quickly to load reductions and did not exhibit long-term legacy effects.

From 1996, the The now stabilized, near neutral P balance of the Stockholm IA means that there was still is little P retention, in spite of the improved O₂ conditions. Although oxidation of surface sediments can increase short-term P retention due to accumulation of Fe-oxyhydroxides, many studies indicate little effect of this P pool on long-term P retention, which depends on the amount of P permanently buried in the deeper anoxic parts of the sediments (Jensen et al., 1995; Lukkari et al., 2009; Rydin et al., 2011). Recent results indicate that a persistent Fe(II)-P mineral can form in sediments of the Bothnian Sea (Slomp et al., 2013) and to some extent also in the Baltic proper (Reed et al., 2016). To which extent formation of Fe(II)-P minerals have contributed to long-term retention in the Stockholm IA is unknown, but the low overall P retention suggest this process is currently not very significant. Possibly, preferential Fe(II) sulphide formation in IA sediments competes with Fe(II)-P formation at the current organic loading (Reed et al. 2016).

25

5

Norkko et al. (2011) have suggested that the recent (post-2005) increase of the invasive polychaetes *Marenzelleria* spp. in the IA may have enhanced the oxygenation of the sediment and hence increased the binding of P-in the sediments, presumably as Fe-oxyhydroxides. We found neither lower P release rates nor increased P retention in these years. However, our results are not directly comparable to those of Norkko et al. (2011), who used data from the inner<u>most-part of the</u> IA only.

30 4.2 Seasonal phosphorus budget in surface water <u>P budget</u>

In spring, there is net internal TP loss (largely as DIP) from the whole water mass of the IA, due to larger <u>net</u> losses from the upper layers (0-20m) than net P inputs to the deep water >20m (Fig. 7). <u>The-This</u> DIP and TP loss, <u>presumably to the</u> sediment, —was associated with the phytoplankton spring bloom, which typically contributes to nutrient uptake and

sedimentation in temperate coastal areas (Blomqvist and Larsson, 1994; Testa and Kemp, 2008; Staehr et al., 2017). In the IA, diatoms, known to have high sedimentation rates, and dinoflagellates, known to have high sedimentation rates, dominate the spring bloom (Lännergren and Stehn, 2011). The incorporation in higher trophic levels probably contributes only to a small extent to the P loss, since zooplankton are included in the TP analysis and the build-up of fish biomass occurs mainly occurs in summer (Hjerne and Hansson, 2002). P uptake by benthic vegetation likely contributes little to the overall P budget, since the deep, turbid IA has small bottom areas suitable for macrophytes, and while large algae, such as bladderwrack (*Fucus vesiculosus*), are absent from this low-salinity area.

In summer and autumn, net P losses from the 0-20 m surface-upper layer decrease while net P inputs, mainly of DIP, increase in the deep water. This requires a net internal P source, which we interpret as net sediment P release, and which

10

5

increases the P concentration in all water layers of the IA (Fig. 7). However, there is no net export of P from the IA until the winter months.

Our model approach can only give net rates for each compartment. That the annual net P loss from the upper 0-20 meters is relatively small in some years (Fig. 6a) probably does not reflect lower gross sedimentation rates, but substantial regeneration of P also from shallow sediments. Such P release in summer can explain the much larger variability of 0-20 m

- 15 layer net P loss for the full year than for the spring only, when P recycling from sediments should generally be low. Bottoms above 0-20 m depth of which only a small proportion are accumulation bottoms (Jonsson, 2003), account for 69 % of the area in the inner archipelago, but only a small proportion of them are accumulation bottoms (Jonsson, 2003). These shallow bottoms and are probably particularly important for P regeneration of P in years when transport to deeper accumulation bottoms is less effective, or when hypoxia affect also shallow areas, e.g. in 1992 and 1993. Sedimented matter apparently is
- 20 probably <u>normally</u> efficiently transported to deep accumulation bottoms, as most of the net P sedimentation in spring seems to be regenerated in the deep water in summer and autumn (Fig. 6). However, it remains difficult to estimate the contribution of gross sedimentation P losses from the surface layer to the deep water in summer.

4.3 Seasonal deep-water Pphosphorus and oxygen O2 budgets in deep water

Our monthly P-budgets for the deep_water indicate a more or less continuous internal DIP source, with a clear increase in summer and a peak in September, similar to what Jensen et al. (1995) found in Aarhus Bay. The budget for the whole water column shows a net loss of P to the sediments in spring, suggesting deep-water sediments are a net sink for P in spring and early summer. The delayed net P input to the whole water column in summer and autumn suggests that most of this internal P source in the deep water is regeneration from seasonal storage of P in deep<u>water</u> sediments, but a P redistribution by transport of shallow sediments to the deep water can also contribute. On an annual basis, there was a strong correlation between deep_water O₂ consumption rates and correlated strongly to TP and DIP release rates, in close to Redfield proportions (Fig. 9a), indicating that P release is mainly short-term (within-year) recycling of sedimented organic P. Moreover, net P release in the deep water correlated with net P spring sedimentation (Fig. 8a, Fig. 7), suggesting that a large share of the net P release in the deep water in summer and autumn is a seasonal recycling of P from the spring bloom. Staehr

et al. (2017) came to the same conclusion based on a DIP mass-balance for a shallow Danish estuary, where summer net DIP production coincided with changes in TP inputs indicating that P inputs in spring was the primary source of P release in summer.

- The summer increase in P flux from sediments to the deep_-water starts in June (Fig. 7), despite O₂ concentrations in June and July well above hypoxic levels, even in the deepest bottom water. In August and September, P release continues to increase also in years with relatively good O₂ conditions. These observations, and the close coupling between deep-water P release and respiration rates, indicate limited classical Fe-P trapping (Einsele, 1936; Mortimer, 1941) in the oxic surface sediments. Similarly, field studies in the Finnish archipelago have shown high DIP release also under oxic conditions, with limited Fe-P formation as the suggested explanation (Lehtoranta and Heiskanen, 2003; Lehtoranta et al., 2009). Monitoring
- 10 data from the Lake Mälaren outlet Norrström show that the Fe transport is substantial, ~400 tons per year, corresponding to an annual mean concentration of ~70 μ g L⁻¹ (data from the Swedish University of Agricultural Sciences, www.slu.se/institutioner/vatten-miljo/datavardskap/). However, much of this Fe will be transported out of the IA by the surface water and much of the Fe that reaches the sediment , but Iin relatively sulphate rich coastal sediments, Fe can be immobilized as solid, unreactive FeS or FeS₂ in anoxic layers of these sulphate-rich coastal sediments, where sulphate-
- 15 reducing bacteria produce hydrogen sulphide (Blomqvist et al., 2004). As a result of low Fe availability relative to the production of <u>dissolved-DI</u>P from organic matter decomposition, the rate of P release to the water mass is then mainly determined by the rate of remineralisation (Caraco et al., 1990; Hupfer and Lewandowski, 2008). We found that P release was somewhat lower than expected from O₂ consumption in spring and higher in summer. This suggests some early season trapping of P in Fe-P complexes, which partly dissolve in late summer, when the redox-cline in the sediment becomes
- 20 shallower and higher temperatures increase reduction rates of Fe-oxyhydroxides (Jensen et al., 1995; Jilbert et al., 2011). Jensen et al. (1995) found that P sedimented in spring was largely retained as Fe-bound P until September-October, when half the annual P release occurred. For the Arkona basin of the Baltic Sea, Reed et al. (2011) estimated that sediment DIP release at oxic conditions approximately doubled from winter to summer (0.05 0.1 mmol m⁻² day⁻¹), recycling much of the P mainly deposited as organic matter in spring. Aerobic degradation of organic matter was the dominant P mineralization
- 25

process, but P release was largely mediated via rapid Fe-oxyhydroxide turnover. Elevated temperatures in the IA after 1990 may have increased anaerobic respiration and sulphide production (Bågander, 1977), promoting P release by decreasing the available Fe due to Fe-sulphide formation.

Lehtoranta et al. (2009) suggested that the Fe-dependent P-binding capacity of marine sediments can be negatively affected-reduced by previous anoxic events, if much of the Fe pool has becomes permanently bound as sulphides and is

30 unavailable for P-binding upon reoxidation, as described by (Blomqvist et al. (-2004). Effective P-binding capacity will then be regained only slowly, upon replenishment of as the Fe-pool is replenished. Early investigations show that IA sediments were anoxic long before the rapidly increasing loads of organic matter and nutrients in the mid-20th century (Brattberg, 1986). Hydrogen sulphide was recorded as early as 1909 and was present in the inner parts of the IA (Norrström – Höggarnsfjärden) on 10 out of all 10 sampling occasions in 1909–1913 and oin 40 out of 50 occasions in 1930–1944 (Anon., 1945). Dredging of the strait Oxdjupet, from 8 to 10.5 m in 1919 and to 12 m in 1929 apparently had little effect, but dredging in 1950 to the present 18 m <u>did_improved</u> O_2 conditions. Sediment cores show <u>increasing</u>-lamination frequency <u>increased_until</u> 1950, followed by decreased frequency, indicative of improved O_2 _conditions <u>and-with_more</u> benthic fauna (Jonsson et al., 2003). In 1970–1980, lamination frequency increased again, remaining high at least until the mid-1990's

5 (Jonsson et al., 2003). In 2008, surface sediments were <u>again</u> oxidized (Karlsson et al. 2010). How <u>the this</u> history of anoxic events has affected later Fe- and P-cycling in the Stockholm IA is not well understood, but the short water turnover time of the IA<u>and high Fe loading</u> suggests that continuous inputs of Fe should be relatively more important for formation of sediment Fe(III)-P complexes than the <u>long-term cycling of the</u> internal <u>Fe</u> pool. <u>Monitoring data from the Lake Mälaren outlet Norrström show that the Fe transport is substantial, ~400 tons per year, corresponding to a mean annual concentration</u>

10 of --70 μg L⁻⁺ (measurements by Sveriges Lantbruksuniversitet, www.slu.se/institutioner/vatten miljo/datavardskap/).

4.4 Low O2 and P Phosphorus release at low oxygen concentrations

15

In our model, all-inflowing SW is well-fully mixed into the deep water and the estimates of deep-water P release are a direct function of the deep-water turnover rates. Thus, if the deep_water becomes partly stagnant and the SW inflow partly mixes into shallower (10-20 m) water layers, the model will somewhat overestimate deep-water P release rates. Since stagnation may contribute to hypoxia, this risk is higher in years with hypoxia and therefore the differences in P release between hypoxic and oxic years are more likely overestimated than underestimated. However, variable deep-water turnover rates will have similar effect on deep-water P release and O_2 consumption. Thus, the finding of high P release rates relative to theoretical P remineralization (based on O_2 consumption) in years with low O_2 concentrations (Fig. 9) is not affected by this. Likewise, the P retention estimate for the entire IA is independent of modelled deep-water P cycling.

- 20 The particularly relatively high P release in years with deep-water hypoxia (Fig. 9) suggests that classical redox-dependent release of Fe(III)-bound P contributed significantly to P release in those years. We also found particularly high P net export in years when hypoxia affected intermediate water depths (e.g. in 10-20 m layer in 1992-1993), suggesting unusually large P release from sediment Fe-P p-in these water layers. This is in contrast to the net retention in more O₂-rich years, when Fe-P pools probably were restored. Puttonen et al. (2014) estimated the reserve of potentially mobile P in Baltic Sea archipelago and Stockholm inner archipelago surface sediments to 3.5 g m⁻², stored predominantly in redox-sensitive Febound forms on accumulation bottoms. This P reserve is similar to our estimated annual net P release from deep IA bottoms (up to 3.1 g m⁻² yr⁻¹ and a mean of 2.1 g m⁻² yr⁻¹) and to the P release of 2.7 g m⁻² yr⁻¹ estimated by Rydin et al. (2011) in Torsbyfjärden in the outer IA. Our estimate of the P amount released in the deep water (67 tons yr⁻¹) is not sensitive to the defined depth of the deep -water. However, with 15-57 m instead of 20-57 m as deep-water layer the sediment area increases
- 30 by 40 % and the area-specific mean P release is <u>only</u> 1.5 g m⁻² yr⁻¹. Still, Fe-P pools (Puttonen et al. 2014) are not very large relative to <u>yearlyannual</u> P recycling from deep-water sediments of the IA and the notion that such mobile P pools must have accumulated over decades of high anthropogenic P loads is <u>questionable</u>-untenable for the Stockholm IA.

4.5 Factors affecting oxygen conditions

5

10

Our model indicates that although O_2 consumption is an important factor for deep-water O_2 concentrations in the Stockholm IA, hypoxia is particularly prone to develop in years with low O_2 inputs to the deep water, i.e. $\underline{O_2}$ inputs with boundary inflow and mixing of deep water with intermediate water (Fig. 8c). This was evident already in the 1970's. Despite <u>relatively</u> high organic matter loads to the deep water in these years (judging from the P budget and high surface water chlorophyll

concentrations), our model results indicate that high O₂ inflow prevented widespread hypoxia in 1969 and 1970.

We find no long-term trend (1968-2015) in July-October O_2 inflow to the IA deep water from the middle archipelago (data not shown). However, there is a trend of increasing O_2 input to the deep water from mixing with intermediate water layers (10-20m), where O_2 concentrations have increased. STP discharges of O_2 consuming substances, mostly ammonium, was-were ~15000 tons O_2 -equivalents year⁻¹ before sewage treatment was improved in the mid-1990's and decreased to ~3000 tons by 2000 (Lännergren and Stehn, 2011). We made no O_2 budget for the intermediate water, but considering mixing rates, the improvement-decreased discharges werewas large enough to contribute to the disappearance of hypoxia in the 10-20 m depth layer. In addition, the weakening of the stratification that occurred around 1990 promoted vertical mixing and thus O_2 transport to the intermediate layer.

15 Organic matter inputs ultimately control the total O₂ consumption in the deep water. However, temperature-dependent aerobic and anaerobic microbial processes should influence O_2 consumption rates. We found seasonally increasing O_2 consumption rates from May to September or October, when the temperature increased from ~2 to 9 °C. The maximum O_2 consumption rate, in September (26 mmol O_2 m⁻² d⁻¹), corresponds well with summer O_2 consumption rates measured in Danish estuaries (Ærtebjerg et al., 2003) and Finnish archipelago areas (Kauppi et al., 2017). Ærtebjerg et al. (2003) found 20 that temperature explained most of the within-year variation in O₂ consumption. In the Stockholm IA, decomposition rates of organic material should have increased due to higher elevated temperatures from 1990 on (Fig. 3c). We find that O_2 consumption rate in July-October normalised to organic matter input in spring increased as postulated (Fig. 10a). As expected, the correlation based on monthly data was strongest in early summer, when most organic matter from the spring bloom is still undegradedundecomposed. We find no long-term trend in total O₂ consumption in July-October. Lowered 25 organic matter deposition probably explains the decreased O_2 consumption in the late 1970's. However, in the 1990's increased O₂ consumption due to higher temperatures probably counteracted this improvement and contributed to the renewed spread of hypoxia.

4.6 Effects and potential management of internal P load

The effect of hypoxia-controlled internal P loading on long-term ecosystem productivity depends on the limiting nutrient

30 and, if P, on the size of the internal P load relative to external P loads and the P inventory in the water mass. In the Baltic Sea as a whole, hypoxia clearly affects the P inventory of the water mass over long periods (Conley et al., 2002). The Baltic Sea P inventory of ~400,000 tons fluctuates by 100,000 tons between years (Jilbert et al., 2011), much more than the current

annual P loading from land of ~28,000 tons (HELCOM, 2011). In the well-flushed Stockholm IA, the annual mean P inventory in the water mass is currently ~46 tons, fluctuating by only a few tons between years and 30 tons seasonally (Fig. 7), much less than the annual P loading from land of 190 tons and the total annual P inputs of almost 400 tons. Release of P during hypoxic events will largely be exported in winter and will have very limited long-term local effects.

- 5 Measures to artificially increase P trapping in sediments, for example by precipitation with aluminium salts, can be successful in archipelago bays with restricted water exchange (Rydin et al., 2017). The short P turnover time in the Stockholm IA suggests that for long-term effects on local water quality, sediment treatment here must enhance long-term annual P retention, probably by repeated treatments. Such measures would primarily bind P imported from the Baltic Sea and Lake Mälaren and to a minor extent the relatively small current P inputs from STPs, rather than legacy P. However, not
- 10 <u>only management of P should be considered, since intensified P limitation in the IA will lead to a larger N export to the N-limited middle and outer archipelago (cf. Brattberg, 1986; Paerl, 2009; Lännergren and Stehn, 2011).</u>

Our study shows that the often-observed correlation between water O_2 concentrations and P (e.g. Norkko et al., 2011) is only partly a cause-and effect relationship, since both O_2 consumption and P release are driven by organic matter degradation (cf. Hupfer and Lewandowski, 2008; Reed et al., 2011). Thus, P binding by Fe compounds can apparently at

15 most give some temporary sediment P retention, but do not prevent P release into the overlying oxic water in the longer term. These results have implications for expected outcomes of artificial methods to oxygenate anoxic deep water. Such efforts may precipitate P in the short term but a corresponding long-term P storage cannot be expected.

5. Conclusions

25

- 20 <u>In accord with several recent studies our result indicate that estuarine and coastal sediments can respond rapidly to</u> load reductions and exhibit little long-term legacy effects.
 - We conclude that tIndeed, tThe seasonal increase of water column P concentrations in the Stockholm inner archipelago in summer and autumn is due mainly to internal P loading from the sediments. However, Mmost of this P release is simply a recycling of P deposited by the sedimentation of the spring bloom. Legacy P release from historical P-deposits, presumably organic P, was apparently probably-restricted to the first 10-20 years after the major P load reduction, when model results suggest sediments were net, but small, P sources. The long-term P balance has now stabilized, with means of our sensitivity analyses close to neutral balance (i.e. export equals inputs) in the last two-ten-20-year-periods, indicating that internal loading from legacy P has essentially ceased. The sensitivity analysis suggests a retention <15 tons per year, or 8 % of land-based inputs. The eMost external P inputs are either exported without seasonal delay, or are efficiently recycled from sediments in summer and autumn despite the improved O₂ conditions, and then exported in winter. The stabilized P balance means that fFurther water quality improvements cannot be expected forin the Stockholm IA unless will require either further decrease in nutrient load is further decreased, or improved that long term sediment P retention is improved.
 - 22

• The coupled P release and O₂ consumption, even under oxic conditions, indicate that Fe binding of P cannot efficiently prevent release of P mineralized from organic matter degradation. The tightly coupled P release and O₂ consumption processes show that (<u>The often observed correlation between water O₂ concentrations and P (e.g. Norkko et al., 2011) is only partly a cause and effect relationship, since both O₂ consumption and P release are linked to<u>driven by</u> organic matter degradation (cf. Hupfer and Lewandowski, 2008; Reed et al., 2011). Thus, <u>P</u> binding by Fe compounds can apparently at most give some temporary sediment P retention, but cannot prevent P release into the overlying oxic water in the longer term. In winter and spring, when O₂ concentrations peak, some temporary storage of Fe P probably occurs, which is released when Fe is increasingly bound to sulphides in summer, as the oxic upper sediment layer gets thinner because of increasing O₂-consumption rates at the higher temperatures. <u>However, C</u>consistent with the classical view of redox-dependent P release, summer and autumn hypoxia triggered elevated P release, often resulting in a-low annual P retention of the area, especially when sediments at intermediate depth were affected by hypoxia.</u>

5

10

- We found that hHypoxia in the IA deep water was influenced by the size of the spring bloom, and but also by temperature affecting O₂ consumption rates, but also by deep-water renewal rates, and the O₂ concentrations in the water ventilating the deep_-water. The generally improving O₂ conditions of the IA are probably linked_due_to lowered sedimentation of organic matter (especially in-since_the 1970s), to lowered STP loading of ammonium (from the late 1990's) and enhanced vertical mixing (from the 1990's). The enhanced mixing has also contributed to higher deep-water temperature and respiration rates. More dDetailed studies are needed to clearly separate the complex factors influencing the O₂ conditions in the Stockholm IA.
- 20 The impact <u>effect</u> of hypoxia controlled internal P loading on long term ecosystem productivity depends on <u>the</u> <u>limiting nutrient and, if P, on the size of this P pool load</u> relative to external P loading <u>loads</u> and the P inventory in the water mass. In the Baltic Sea as a whole, hypoxia clearly affects the P inventory of the water mass over long periods (Conley et al., 2002). The Baltic Sea P inventory of ~400,000 tons fluctuates by 100,000 tons between years (Jilbert et al., 2011), considerably <u>much</u> more than the <u>current</u> yearly<u>annual</u> P loading from land of ~28,000 tons
 25 (HELCOM, 2011). In the well flushed Stockholm IA, the annual mean P inventory in the water mass is currently ~46 tons, fluctuating by only a few tons between years and 30 tons seasonally, much less than the yearly<u>annual</u> P loading from land of 190 tons. Release of P during hypoxic events will largely be exported in winter and will have very limited long term <u>local</u> effects locally.
- Further water quality improvements in the Stockholm IA will require either further decrease in nutrient load, or improved long-term sediment P retention. Since the P turnover is rapid and legacy P small in the Stockholm IA, measures to-binding in-P in sediments would primarily accumulate P recently imported from the Baltic Sea, Lake Mälaren and local STPs, and would need to be repeated regularly for enduring effect. This situation should be typical of well-flushed Baltic Sea coastal areas and bays.

Measures aimed at artificially increasing P trapping in sediments, for example by aluminium treatment, can be successful in archipelago bays with restricted water exchange (Rydin et al., 2017). The short P turnover time in the Stockholm IA, suggests that here sediment treatment must enhance long term annual P retention efficiency, probably by recurrent treatment, to have long term effects on local water quality. Even so, treatment of IA sediments primarily would not bind legacy P, but rather accumulate P from the Baltic Sea and Lake Mälaren and to minor extent the relatively small current P inputs from STPs, suggesting control of external P inputs is preferable. However, not only management of P should be considered, since further intensifying P limitation in the IA would lead to larger N export to N-limited parts of the archipelago (cf. Brattberg, 1986; Paerl, 2009; Lännergren and Stehn, 2011).

Data availability

5

15

20

25

10 Monitoring <u>and runoff</u> data are <u>deposited_available fromat</u> the Swedish Meteorological and Hydrological Institute (www.smhi.se). <u>SMHI also provides runoff data.</u>

Acknowledgements

Data were kindly provided by Stockholm Vatten AB (now Stockholm Vatten och Avfall). Model development was initiated in the EU-project SPICOSA, and was also-financially supported <u>also</u> by Sydvästra Stockholmsregionens va-verksaktiebolag (SYVAB). Data compilation and model analysis was supported by grants from Svealands kustvattenvårdsförbund (SKVVF). Data digitizing was financed by the Swedish EPA, Stockholm Vatten AB, and SKVVF. We thank Ragnar Elmgren for valuable comments on the manuscript.

References

- Ærtebjerg, G., Andersen, J. H., and Hansen, O. (eds): Nutrient and eutrophication in Danish marine waters. A challenge for science and management, National Environmental Research institute, 126 pp, 2003.
- Almroth-Rosell, E., Eilola, K., Kuznetsov, I., Hall, P. O. J., and Meier, H. E. M.: A new approach to model oxygen dependent benthic phosphate fluxes in the Baltic Sea, J. Marine Syst., 144, 127-141, 10.1016/j.jmarsys.2014.11.007, 2015.

Almroth-Rosell, E., Edman, M., Eilola, K., Meier, H. E. M., and Sahlberg, J.: Modelling nutrient retention in the coastal zone of an eutrophic sea, Biogeosciences, 13, 5753-5769, 10.5194/bg-13-5753-2016, 2016.

Anon.: Bilaga till Berättelse över gatukontorets verksamhet år 1944 – Om vattenbeskaffenheten i Stockholm hamn och inre skärgård åren 1909-1944 samt dess samband med föroreningen genom stadens avloppsvatten, K. L. Beckmans Boktryckeri, Stockholm, 33 pp. (in Swedish), 1945.

- Anon.: Recipientutredning Stockholms skärgård. Stockholm läns landsting regionalplanekontoret, rapport 10, 250 pp. (in Swedish), 1975.
- Blomqvist, S., and Larsson, U.: Detrital bedrock elements as tracers of settling resuspended particulate matter in a coastal area of the Baltic Sea, Limnol. Oceanogr., 39, 880-896, 1994.
- 5 Blomqvist, S., Gunnars, A., and Elmgren, R.: Why the limiting nutrient differs between temperate coastal seas and freshwater lakes: A matter of salt, Limnol. Oceanogr., 49, 2236-2241, 2004.
 - Boesch, D., Hecky, R., O'Melia, C., Schindler, D., and Seitzinger, S.: Eutrophication of Swedish seas, Swedish Environmental Protection Agency, Report 5509, 70 pp, 2006.
- Boynton, W. R., Hagy, J. D., Cornwell, J. C., Kemp, W. M., Greene, S. M., Owens, M. S., Baker, J. E., and Larsen, R. K.:
 Nutrient budgets and management actions in the Patuxent River estuary, Maryland, Estuar. Coast., 31, 623-651,
 - Boynton, W. R., Ceballos, M. A. C., Bailey, E. M., Hodgkins, C. L. S., Humphrey, J. L., and Testa, J. M.: Oxygen and Nutrient Exchanges at the Sediment-Water Interface: a Global Synthesis and Critique of Estuarine and Coastal Data, Estuar. Coast., 41, 301-333, 10.1007/s12237-017-0275-5, 2018.
- 15 Brattberg, G.: Decreased phosphorus loading changes phytoplankton composition and biomass in the Stockholm archipelago, Vatten, 42, 141-153, 1986.
 - Bågander, L. E.: In Situ studies of bacterial sulfate reduction at the sediment-water interface. Ambio Special Report 5, 147–155, 1977.
 - Caraco, N., Cole, J., and Likens, G. E.: A comparison of phosphorus immobilization in sediments of fresh-water and coastal
- 20 marine systems, Biogeochemistry, 9, 277-290, 1990.

10.1007/s12237-008-9052-9, 2008.

- Carey, C. C., and Rydin, E.: Lake trophic status can be determined by the depth distribution of sediment phosphorus, Limnol. Oceanogr., 56, 2051-2063, 10.4319/lo.2011.56.6.2051, 2011.
 - Cloern, J. E., Foster, S. Q., and Kleckner, A. E.: Phytoplankton primary production in the world's estuarine-coastal ecosystems, Biogeosciences, 11, 2477-2501, doi:10.5194/bg-11-2477-2014, 2014.
- 25 Conley, D. J., Humborg, C., Rahm, L., Savchuk, O. P., and Wulff, F.: Hypoxia in the Baltic Sea and basin-scale changes in phosphorus biogeochemistry, Environ. Sci. Technol., 36, 5315-5320, 10.1021/es025763w, 2002.
 - Conley, D. J., Carstensen, J., Aigars, J., Axe, P., Bonsdorff, E., Eremina, T., Haahti, B. M., Humborg, C., Jonsson, P., Kotta, J., Lännergren, C., Larsson, U., Maximov, A., Medina, M. R., Lysiak-Pastuszak, E., Remeikaite-Nikiene, N., Walve, J., Wilhelms, S., and Zillén, L.: Hypoxia Is Increasing in the Coastal Zone of the Baltic Sea, Environ. Sci. Technol., 45, 6777-6783, 10.1021/es201212r, 2011.
- 30
- Diaz, R. J., and Rosenberg, R.: Spreading dead zones and consequences for marine ecosystems, Science, 321, 926-929, 10.1126/science.1156401, 2008.
- Duarte, C. M., Borja, A., Carstensen, J., Elliott, M., Krause-Jensen, D., and Marba, N.: Paradigms in the Recovery of Estuarine and Coastal Ecosystems, Estuar. Coast., 38, 1202-1212, 10.1007/s12237-013-9750-9, 2015.

- Einsele, W.: Über chemische und kolloidchemische Vorgänge in Eisen-Phosphat-Systemen unter limnochemischen und limnogeologischen Gesichtpunkten, Arch. Hydrobiol., 33, 361–387, 1936.
- Ekeroth, N., Blomqvist, S., and Hall, P. O. J.: Nutrient fluxes from reduced Baltic Sea sediment: effects of oxygenation and macrobenthos, Mar. Ecol. Prog. Ser., 544, 77-92, 10.3354/meps11592, 2016a.
- 5 Ekeroth, N., Kononets, M., Walve, J., Blomqvist, S., and Hall, P. O. J.: Effects of oxygen on recycling of biogenic elements from sediments of a stratified coastal Baltic Sea basin, J. Marine Syst., 154, 206-219, 10.1016/j.jmarsys.2015.10.005, 2016b.
 - Engqvist, A., and Andrejev, O.: Water exchange of the Stockholm archipelago a cascade framework modelling approach, J. Sea Res., 49, 275-294, 10.1016/s1385-1101(03)00023-6, 2003.
- 10 Gächter, R., and Muller, B.: Why the phosphorus retention of lakes does not necessarily depend on the oxygen supply to their sediment surface, Limnol. Oceanogr., 48, 929-933, 2003.
 - Gunnars, A., and Blomqvist, S.: Phosphate exchange across the sediment-water interface when shifting from anoxic to oxic conditions an experimental comparison of freshwater and brackish-marine systems, Biogeochemistry, 37, 203-226, 10.1023/a:1005744610602, 1997.
- 15 Gustafsson, O., Gelting, J., Andersson, P., Larsson, U., and Roos, P.: An assessment of upper ocean carbon and nitrogen export fluxes on the boreal continental shelf: A 3-year study in the open Baltic Sea comparing sediment traps, Th-234 proxy, nutrient, and oxygen budgets, Limnol. Oceanogr.-Meth., 11, 495-510, 10.4319/lom.2013.11.495, 2013.
 - Hagy, J. D., Sanford, L. P., and Boynton, W. R.: Estimation of net physical transport and hydraulic residence times for a coastal plain estuary using box models, Estuaries, 23, 328-340, 10.2307/1353325, 2000.
- 20 Hall, P. O. J., Almroth Rosell, E., Bonaglia, S., Dale, A. W., Hylén, A., Kononets, M., Nilsson, M., Sommer, S., van de Velde, S., and Viktorsson, L.: Influence of natural oxygenation of Baltic Proper deep water on benthic Recycling and removal of phosphorus, nitrogen, silicon and carbon, Front. Mar. Sci., 4 (27), 10.3389/fmars.2017.00027, 2017.
 - Hedenstierna, B.: Stockholms skärgård. Kulturgeografiska Undersökningar i Värmdö Gamla Skeppslag, Geografiska Annaler, 30, 1–444, (in Swedish), 1948.
- HELCOM: Fifth Baltic Sea pollution load compilation (PLC-5), Baltic Sea Environment Proceedings, 128, 1–217, 2011.
 Hjerne, O., and Hansson, S.: The role of fish and fisheries in Baltic Sea nutrient dynamics, Limnol. Oceanogr., 47, 1023-1032, 2002.
 - Hupfer, M., and Lewandowski, J.: Oxygen Controls the Phosphorus Release from Lake Sediments a Long-Lasting Paradigm in Limnology, Int. Rev. Hydrobiol., 93, 415-432, 10.1002/iroh.200711054, 2008.
- 30 Jensen, H. S., Mortensen, P. B., Andersen, F. O., Rasmussen, E., and Jensen, A.: Phosphorus cycling in a coastal marine sediment, Aarhus Bay, Denmark, Limnol. Oceanogr., 40, 908-917, 1995.
 - Jilbert, T., Slomp, C. P., Gustafsson, B. G., and Boer, W.: Beyond the Fe-P-redox connection: preferential regeneration of phosphorus from organic matter as a key control on Baltic Sea nutrient cycles, Biogeosciences, 8, 1699-1720, 10.5194/bg-8-1699-2011, 2011.

- Johansson, L., and Wallstrom, K.: Urban impact in the history of water quality in the Stockholm archipelago, Ambio, 30, 277-281, 10.1639/0044-7447(2001)030[0277:uiitho]2.0.co;2, 2001.
- Jonsson, P., Persson, J., and Holmberg, P.: The seafloor of the Baltic Sea archipelagoes, Swedish Environmental Protection Agency Report No 5212, ISBN 91-620-52512-8, 112 pp (in Swedish), 2003.
- 5 Karlgren, L., and Ljungström, K.: Nutrient budgets for Inner archipelago of Stockholm, J. Water Pollut. Con. F., 47, 823-833, 1975.
 - Karlsson, O. M., Jonsson, P. O., Lindgren, D., Malmaeus, J. M., and Stehn, A.: Indications of Recovery from Hypoxia in the Inner Stockholm Archipelago, Ambio, 39, 486-495, 10.1007/s13280-010-0079-3, 2010.
 - Kauppi, L., Norkko, J., Ikonen, J., and Norkko, A.: Seasonal variability in ecosystem functions: quantifying the contribution
- 10 of invasive species to nutrient cycling in coastal ecosystems, Mar. Ecol. Prog. Ser., 572, 193-207, 10.3354/meps12171, 2017.
 - Knudsen, M.: Ein hydrographischer Lehrsatz, Annalen der Hydrographie und Maritimen Meterologie, 28, 316–320, 1900.
 - Koop, K., Boynton, W. R., Wulff, F., and Carman, R.: Sediment-water oxygen and nutrient exchanges along a depth gradient in the Baltic Sea, Mar. Ecol. Prog. Ser., 63, 65-77, 10.3354/meps063065, 1990.
- 15 Lehtoranta, J., and Heiskanen, A. S.: Dissolved iron: phosphate ratio as an indicator of phosphate release to oxic water of the inner and outer coastal Baltic Sea, Hydrobiologia, 492, 69-84, 10.1023/a:1024822013580, 2003.
 - Lehtoranta, J., Ekholm, P., and Pitkanen, H.: Coastal Eutrophication Thresholds: A Matter of Sediment Microbial Processes, Ambio, 38, 303-308, 2009.
 - Lukkari, K., Leivuori, M., and Kotilainen, A.: The chemical character and behaviour of phosphorus in poorly oxygenated
- 20 sediments from open sea to organic-rich inner bay in the Baltic Sea, Biogeochemistry, 96, 25-48, 10.1007/s10533-009-9343-7, 2009.
 - Lännergren, C., and Stehn, A.: Undersökningar i Stockholms skärgård 2010, Stockholm Vatten report 2011-04-11, Dnr 10SV299 (in Swedish), 2011.

Middelburg, J. J., and Levin, L. A.: Coastal hypoxia and sediment biogeochemistry, Biogeosciences, 6, 1273-1293, 2009.

- 25 Mort, H. P., Slomp, C. P., Gustafsson, B. G., and Andersen, T. J.: Phosphorus recycling and burial in Baltic Sea sediments with contrasting redox conditions, Geochim. Cosmochim. Ac., 74, 1350-1362, 10.1016/j.gca.2009.11.016, 2010.
 - Mortimer, C. H.: The exchange of dissolved substances between mud and water in lakes: I and II, J. Ecol., 29, 280–329, 1941.

Nielsen, K., Risgaard-Petersen, N., Somod, B., Rysgaard, S., and Bergo, T.: Nitrogen and phosphorus retention estimated

- 30 independently by flux measurements and dynamic modelling in the estuary, Randers Fjord, Denmark, Mar. Ecol. Prog. Ser., 219, 25-40, 10.3354/meps219025, 2001.
 - Nixon, S. W., Ammerman, J. W., Atkinson, L. P., Berounsky, V. M., Billen, G., Boicourt, W. C., Boynton, W. R., Church, T. M., Ditoro, D. M., Elmgren, R., Garber, J. H., Giblin, A. E., Jahnke, R. A., Owens, N. J. P., Pilson, M. E. Q., and Seitzinger, S. P.: The fate of nitrogen and phosphorus at the land sea margin of the North Atlantic Ocean,

Biogeochemistry, 35, 141-180, 10.1007/bf02179826, 1996.

- Nordling, C., Brattberg, G., Ek, K., and Engström, G.: Analysresultat från Näringsämnesundersökningen i Stockholms Skärgård 1969–1976. Primärdataredovisning, SNV PM 1855, 136 pp. (in Swedish), Statens Naturvårdsverk, 1984.
- Norkko, J., Reed, D. C., Timmermann, K., Norkko, A., Gustafsson, B. G., Bonsdorff, E., Slomp, C. P., Carstensen, J., and
- 5 Conley, D. J.: A welcome can of worms? Hypoxia mitigation by an invasive species, Global Change Biol., 18, 422-434, 10.1111/j.1365-2486.2011.02513.x, 2012.
 - Nürnberg, G. K., and LaZerte, B. D.: More than 20 years of estimated internal phosphorus loading in polymictic, eutrophic Lake Winnipeg, Manitoba, J. Great Lakes Res., 42, 18-27, 10.1016/j.jglr.2015.11.003, 2016.

Paerl, H. W.: Controlling Eutrophication along the Freshwater-Marine Continuum: Dual Nutrient (N and P) Reductions are

10 Essential, Estuar. Coast., 32, 593-601, 10.1007/s12237-009-9158-8, 2009.

- Persson, J., and Jonsson, P.: Historical development of laminated sediments an approach to detect soft sediment ecosystem changes in the Baltic Sea, Mar. Pollut. Bull., 40, 122-134, 10.1016/s0025-326x(99)00180-0, 2000.
- Puttonen, I., Mattila, J., Jonsson, P., Karlsson, O. M., Kohonen, T., Kotilainen, A., Lukkari, K., Malmaeus, J. M., and Rydin,E.: Distribution and estimated release of sediment phosphorus in the northern Baltic Sea archipelagos, Estuar. Coast.
- 15 Shelf S., 145, 9-21, 10.1016/j.ecss.2014.04.010, 2014.

- Reed, D. C., Slomp, C. P., and Gustafsson, B. G.: Sedimentary phosphorus dynamics and the evolution of bottom-water hypoxia: A coupled benthic-pelagic model of a coastal system, Limnol. Oceanogr., 56, 1075-1092, 10.4319/lo.2011.56.3.1075, 2011.
- Reed, D. C., Gustafsson, B. G., and Slomp, C. P.: Shelf-to-basin iron shuttling enhances vivianite formation in deep Baltic
 Sea sediments, Earth Planet. Sc. Lett., 434, 241-251, 10.1016/j.epsl.2015.11.033, 2016.
 - Regnier, P., Arndt, S., Goossens, N., Volta, C., Laruelle, G. G., Lauerwald, R., and Hartmann, J.: Modelling Estuarine Biogeochemical Dynamics: From the Local to the Global Scale, Aquat. Geochem., 19, 591-626, 10.1007/s10498-013-9218-3, 2013.

- Rydin, E., Kumblad, L., Wulff, F., and Larsson, P.: Remediation of a Eutrophic Bay in the Baltic Sea, Environ. Sci. Technol., 51, 4559-4566, 10.1021/acs.est.6b06187, 2017.
 - Savchuk, O. P.: Resolving the Baltic Sea into seven subbasins: N and P budgets for 1991-1999, J. Marine Syst., 56, 1-15, 10.1016/j.jmarsys.2004.08.005, 2005.
- 30 Savchuk, O. P., Wulff, F., Hille, S., Humborg, C., and Pollehne, F.: The Baltic Sea a century ago a reconstruction from model simulations, verified by observations, J. Marine Syst., 74, 485-494, 10.1016/j.jmarsys.2008.03.008, 2008.
 - Schindler, D. W.: Recent advances in the understanding and management of eutrophication, Limnol. Oceanogr., 51, 356-363, 2006.
 - Schindler, D. W., and Vallentyne, J. R.: The Algal Bowl. Overfertilization of the World's freshwaters and Estuaries,

Rydin, E., Malmaeus, J. M., Karlsson, O. M., and Jonsson, P.: Phosphorus release from coastal Baltic Sea sediments as estimated from sediment profiles, Estuar. Coast. Shelf S., 92, 111-117, 10.1016/j.ecss.2010.12.020, 2011.

Earthscan, London, 2008.

20

- Schindler, D. W., Hecky, R. E., Findlay, D. L., Stainton, M. P., Parker, B. R., Paterson, M. J., Beaty, K. G., Lyng, M., and Kasian, S. E. M.: Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year wholeecosystem experiment, Proc. Natl. Acad. Sci. USA, 105, 11254-11258, 10.1073/pnas.0805108105, 2008.
- 5 Schindler, D. W.: The dilemma of controlling cultural eutrophication of lakes, P. Roy. Soc. B-Biol. Sci., 279, 4322-4333, 10.1098/rspb.2012.1032, 2012.
 - Slomp, C. P., Mort, H. P., Jilbert, T., Reed, D. C., Gustafsson, B. G., and Wolthers, M.: Coupled Dynamics of Iron and Phosphorus in Sediments of an Oligotrophic Coastal Basin and the Impact of Anaerobic Oxidation of Methane, Plos One, 8, 10.1371/journal.pone.0062386, 2013.
- 10 SMHI: Djupdata för havsområden 2003, Swedish Meteorological and Hydrographical Institute, Oceanography Report No 73, ISSN 0283-7714, Norrköping, 69 p (in Swedish), 2003.
 - SMHI: Mätning av vattenflöde från Mälaren till Saltsjön 2006-05-03, Swedish Meteorological and Hydrographical Institute, Report 2006/1204/183 (in Swedish), 2006.
 - Sommer, S., Clemens, D., Yücel, M., Pfannkuche, O., Hall, P. O. J., Almroth-Rosell, E., Schulz-Vogt, H. N., and Dale, A.
- 15 W.: Major bottom water ventilation events do not significantly reduce basin-wide benthic N and P release in the eastern Gotland Basin (Baltic Sea), Front. Mar. Sci., 4 (18), 10.3389/fmars.2017.00018, 2017.
 - Staehr, P. A., Testa, J., and Carstensen, J.: Decadal Changes in Water Quality and Net Productivity of a Shallow Danish Estuary Following Significant Nutrient Reductions, Estuar. Coast., 40, 63-79, 10.1007/s12237-016-0117-x, 2017.
- Stigebrandt, A., Rahm, L., Viktorsson, L., Odalen, M., Hall, P. O. J., and Liljebladh, B.: A New Phosphorus Paradigm for
- Testa, J. M., and Kemp, W. M.: Variability of biogeochemical processes and physical transport in a partially stratified
 - estuary: a box-modeling analysis, Mar. Ecol. Prog. Ser., 356, 63-79, 10.3354/meps07264, 2008.

the Baltic Proper, Ambio, 43, 634-643, 10.1007/s13280-013-0441-3, 2014.

- Testa, J. M., Kemp, W. M., Boynton, W. R., and Hagy, J. D.: Long-Term Changes in Water Quality and Productivity in the Patuxent River Estuary: 1985 to 2003, Estuar. Coast., 31, 1021-1037, 10.1007/s12237-008-9095-y, 2008.
- 25 <u>Testa, J. M., Kemp, W. M., Harris, L. A., Woodland, R. J., and Boynton, W. R.: Challenges and Directions for the</u> Advancement of Estuarine Ecosystem Science, Ecosystems, 20, 14-22, 10.1007/s10021-016-0004-0, 2017.
 - Vahtera, E., Conley, D. J., Gustafsson, B. G., Kuosa, H., Pitkanen, H., Savchuk, O. P., Tamminen, T., Viitasalo, M., Voss, M., Wasmund, N., and Wulff, F.: Internal ecosystem feedbacks enhance nitrogen-fixing cyanobacteria blooms and complicate management in the Baltic Sea, Ambio, 36, 186-194, 10.1579/0044-7447(2007)36[186:iefenc]2.0.co;2, 2007.
- 30 Walve, J.: Värdefull tidsserie riskerar att förstöras, Svealandskusten 2012, Svealands Kustvattenvårdsförbund, Stockholm, Sweden, pp. 14–15 (In Swedish), 2012.
 - Zhang, J., Gilbert, D., Gooday, A. J., Levin, L., Naqvi, S. W. A., Middelburg, J. J., Scranton, M., Ekau, W., Pena, A., Dewitte, B., Oguz, T., Monteiro, P. M. S., Urban, E., Rabalais, N. N., Ittekkot, V., Kemp, W. M., Ulloa, O., Elmgren, R.,

Escobar-Briones, E., and Van der Plas, A. K.: Natural and human-induced hypoxia and consequences for coastal areas: synthesis and future development, Biogeosciences, 7, 1443-1467, 10.5194/bg-7-1443-2010, 2010.

Table 1. Period means of annual water discharge (Q), total phosphorus (TP) load and volume_-weighted input TP concentrations (±standard deviation). The volume_-weighted TP concentration in the whole water mass and in the surface layers of the model area are also shown. LM=Lake Mälaren, LR=local runoff, P=precipitation, STP=sewage treatment plants, IA=inner archipelago.

		1968–1975	1976–1985	1986–1995	1996-2005	2006-2015
QLM	Mm ³ yr ⁻¹	4027 ±884	5414 ±1717	5004 ±939	4711 ±1682	5543 ±1189
Qlr	Mm ³ yr ⁻¹	83 ±0	88 ±16	85 ±21	76 ±20	88 ±17
Q _P	Mm ³ yr ⁻¹	50 ±10	58 ±8	58 ±8	54 ±7	59 ±10
QSTP	Mm ³ yr ⁻¹	158 ±7	162 ±10	189 ±18	182 ±14	195 ±14
TPLM	tons yr-1	228 ±108	223 ±69	143 ±43	138 ±59	143 ±27
TP _{LR} **	tons yr-1	8 ±0	9 ±2	8 ±2	8 ±2	9 ±2
TP _{STP}	tons yr-1	485 ±272	74 ±12	61 ±12	28 ±3	33 ±6
TP _{LM}	$\mu g L^{-1}$	55 ±18	42 ±6	28 ±4	29 ±4	26 ±1
TP _{LM + LR + STP}	$\mu g L^{-1}$	167 ±69	57 ±15	40 ±5	35 ±3	32 ±2
TP _{IA} 0-57 m	μg L ⁻¹	118 ±28*	49 ±4	39 ±3	32 ±2	30 ±2*
TP _{IA} 0-10 m	μg L ⁻¹	123 ±38*	45 ±4	36 ±3	30 ±3	28 ±2*
TP _{IA} 0-20 m	μg L ⁻¹	122 ±32*	48 ±3	38 ±3	31 ±2	29 ±2*
TP _{IA} 20-57 m	μg L ⁻¹	93 ±11*	55 ±7	43 ±5	40 ±5	39 ±3*

5 * Years 1975 or and 2011 excluded.

**A mean local runoff TP concentration of $100 \,\mu$ g/l is assumed, see methods.

Table 2. Period means of modeled annual water inflow (Q), total phosphorus (TP) input and volume_-weighted input TP concentrations (±standard deviation). SW=inflowing salt water from middle archipelago, MM=mean model, BM= boundary model.

		1968-1975	1976–1985	1986–1995	1996-2005	2006-2015
QSW-MM	Mm ³ yr ⁻¹	3126 ±443*	3189 ±611	3684 ±643	3973 ±554	4109 ±419
Qsw-bm	Mm ³ yr ⁻¹	5608 ±828*	5441 ±935	7095 ±960	7717 ±890	6963 ±767
TP _{SW-MM}	tons yr ⁻¹	110 ±20*	90 ±15	93 ±15	100 ±12	99 ±18*
TP _{SW-BM}	tons yr ⁻¹	199 ±41*	159 ±32	186 ±33	201 ±28	175 ±26*
TP _{SW-MM}	μg L-1	35 ±4*	28 ±2	25 ±2	25 ±3	24 ±2*
TP _{SW-BM}	μg L-1	35 ±4*	29 ±2	26 ±2	26 ±3	25 ±2*

* Years 1975 or-and 2011 excluded.

Table 3. Effects of sensitivity tests of net P export (tons per year) from the inner archipelago (IA) with boundary model (0-6), and mean model (7). Positive values indicate that the IA accumulates (is net a sink for) P, negative values that it is a net source of P. Year 2011 is excluded. Sensitivity tests included 1. Variable freshwater flow with Lake Mälaren (Q_{LM}), 2. Recruitment depth of inflowing salt water (SW) from the middle archipelago (affecting salinity and P), 3. Sewage treatment plant (STP) discharge of P, 4. Water P concentrations in the IA and Lake Mälaren (LM), 5. Deeper layer for water export

- 5
 - from the IA, 6. Non-adjusted data from central IA stations (see methods), 7. Mean model.

Sensitivity test	1976–1985	1986–1995	1996-2005	2006-2015	1976-2015
0. Default boundary model	-25.8	-19.6	0.1	3.8	-10.7
1a. Q _{LM} -5 %	-20.7	-15.7	1.8	5.3	-7.7
1b. Q _{LM} +5 %	-30.9	-23.4	-1.5	2.2	-13.8
2a. SW 12–20 m	-51.0	-53.6	-32.9	-23.8	-40.7
2b. SW 16–24 m	-39.1	-39.9	-19.3	-13.8	-28.4
2c. SW 24–32 m	-18.3	-7.2	12.3	15.1	0.1
3a. STP -10 %	-33.2	-25.7	-2.7	0.5	-15.7
3b. STP +10 %	-18.4	-13.4	2.9	7.0	-5.8
4a. TP conc. IA, LM -10 %	-8.0	-1.3	13.8	15.1	4.7
4b. TP conc. IA, LM +10 %	-43.6	-37.8	-13.6	-7.6	-26.1
5. Export depth 0-10 m	-48.1	-28.0	-5.4	4.9	-19.8
6. Non-adj. IA Salinity/ TP	-23.5	-19.0	1.0	4.0	-9.7
7a. Mean model with boundary P conc. in export calculation.	7.9	3.5	10.3	10.1	7.9
7b. Mean model	1.2	-16.2	-8.4	-7.9	-7.8

Table 4. Effects of sensitivity tests on deep layer (>20 m) P release (tons per year) in the inner archipelago (IA) using the mean model. Year 2011 is excluded. Sensitivity tests included 1. Variations in freshwater flow with Lake Mälaren (Q_{LM}), 2. Recruitment depth of inflowing salt water (SW) from the middle archipelago (affecting salinity and P), 3. Sewage treatment plant (STP) discharge of P, 4. Water P concentrations in the IA and Lake Mälaren (LM), 5. Deeper layer for water export from the IA, 6. Non-adjusted data from central IA stations (see methods), 7. Use of boundary P concentration for export

5

from the IA.

	1976–1985	1986–1995	1996-2005	2006-2015	1976-2015
0. Default mean model	70.4	59.5	62.2	71.6	66.2
1a. QLM -5 %	67.4	57.3	59.7	68.9	63.6
1b. Q _{LM} +5 %	73.4	61.7	64.7	74.2	68.8
2a. SW 12–20 m	83.9	75.7	72.9	82.8	79.0
2b. SW 16–24 m	78.0	69.1	69.4	79.6	74.3
2c. SW 24–32 m	66.0	53.5	57.3	66.0	61.0
3a. STP -10 %	70.4	59.6	62.2	71.6	66.3
3b. STP +10 %	70.3	59.4	62.2	71.5	66.2
4a. TP conc. IA, LM -10 %	62.9	53.5	55.9	64.3	59.4
4b. TP conc. IA, LM +10 %	77.8	65.5	68.5	78.8	73.0
4c. TP conc. IA -10 %	54.7	45.1	46.9	55.4	50.7
4d. TP conc. IA +10 %	86.1	73.9	77.6	87.7	81.8
5. Export depth 0-10 m	89.2	71.6	74.3	85.0	80.3
6. Non-adj. IA Salinity/ TP	62.6	54.0	55.0	65.7	59.5
7. Boundary P conc. in export	70.4	59.5	62.2	71.5	66.2

Figure legends

25

Figure 1. Maps showing the location of the inner Stockholm archipelago (IA), in the NW Baltic Proper. Lake Mälaren with drainage area is shown in the top left map and its outlet Norrström in Stockholm (LM) in the lower map. The small local drainage area of the IA is shown in the lower map. The central stations (A, AV, H, L, and K) were used to represent the IA.

- 5 The <u>major-main</u> water exchange between the IA and middle archipelago is through Oxdjupet (O). Data from Trälhavet (T) were used to represent inflowing water. Surface water data from Solöfjärden (S), Oxdjupet (O) and Trälhavet Fiskare (TF, weekly station for surface water) were used to represent outflowing water of the boundary model. Depths (m) in maps are spline interpolated from nautical chart data (scale 1:25000, published 1984).
- 10 Figure 2. Schematic illustration of the box model for water flows (*Q*), salt amounts (*S*) and salinities (*Sal*). Salt is transported from the middle archipelago through Oxdjupet strait ($Q_{in} Sal_{in}$ according to Eq. 2) and forms the inner archipelago deep water, with salt content S_D . Upwelling water (equivalent to Q_{in}) transports salt to the above layers and mixing (Q_{D_mix}, Q_{M_mix} and Q_{S_mix}) relocates salt between the layers.
- Figure 3. Time series 1968-2015 of (a) Volume_-weighted mean salinity of the whole water mass and bottom water (20–57 m) of the inner archipelago (IA) and salinity of the assumed inflowing water (SW) from the middle archipelago at 20-30 m depth just outside the boundary strait (Trälhavet, T in Fig.1). (b) Salinity stratification in the inner archipelago. The black contours show integer values 2-6. The colour scale has a resolution of 0.25. (c) Seasonal (July-October) mean temperature at 0-8 m, 12-20 m and 24-32 m depth at station Koviksudde (K in Fig. 1). (d) IA Volume-volume_-weighted concentrations of total phosphorus (TP) and dissolved inorganic phosphorus (DIP) 0-20 m. (e) IA Volume_-weighted concentrations of TP and DIP in the deep layer, >20 m.

Figure 4. (a) Annual mean fresh-water input (including STPs) and (b) modelled inflow from Middle archipelago (Baltic Sea) to the inner archipelago <u>calculated with boundary and mean models (see methods)</u>. Lake Mälaren contributes 93.6 ±2.1 % (mean and SD for all years) and STPs 3.6 ±1.2 % of the total freshwater inputs.

Figure 5. Annual water flows_of, total phosphorus (TP) inputs, TP export and net TP input exportbalance. (a) Annual TP load from sewage treatment plants (STP), from fresh-water runoff from Lake Mälaren plus the local catchment (FW), and (b) from modelled inflow of middle archipelago deep water <u>(SW-(Trälhavet), calculated with boundary and mean models (see methods)</u>. (c) Total annual external TP input and export, and net input export, using the boundary model. Positive values indicate that the inner archipelago accumulates P and negative values that it is a net source of P. (d) Net input-export according to boundary model (enlargement of difference of values in c) with error bars showing maximum and minimums

from sensitivity test. <u>Positive values indicate that the inner archipelago accumulates P and negative values that it is a net</u> source of P. Dots show net input-export using the mean model.

Figure 6. Internal total phosphorus (TP) fluxes and oxygen (O₂) conditions. (a) Annual net TP release in deep layer (>20 m),
annual net TP loss from 0-20 m, net TP balance for whole water mass, and calculated P remineralization from organic matter (from modelled O₂ consumption). Positive net TP balance indicates net TP release from sediments to water mass and negative balance indicates net TP loss from water to sediments. (b) As in (a) but net TP release in deep layer (>20 m) for July-October and net TP loss from 0-20 m in March-May (c) Minimum volume-weighted O₂ concentration in deep water and maximum percent hypoxic (O₂ <2 mg L⁻¹) bottom area in deep water and 10-20 m layer (July-October).

10

15

Figure 7. Monthly means and changes 1996-2015 (2011 excluded) of (**a**) total phosphorus (TP) stock in whole water mass and deep water and the monthly change (**b**) TP input and export with boundary model, (**c**) net TP release in deep water (>20m) and net TP losses in upper water layers 0-20m and calculated P remineralization (from modeled O_2 consumption), with mean model, (**d**) deep-water temperature, O_2 consumption and minimum O_2 concentration. Error bars show standard deviation.

Figure 8. (a) Yearly-Annual net total phosphorus (TP) release (tons) in deep water (>20m) in July-October, as a function of net TP loss (tons) from the layers 0-20 m in Mar-May 1969-2015 (1970 excluded as outlier [P loss 222 tons], 1975 excluded due to lack of data and 2011 due to unreliable data). Colour-code shows mean bottom water temperature July-October (station K, 24-32m, see Fig. 3). There is a significant linear regression, y=19.8 + 0.55x, r²=0.38, p<0.0001. Only the 1:1 line is shown in the graph. (b) YearlyAnnual net TP release in deep water July-October (tons) as a function of the minimum oxygen (O₂) concentration (mg L⁻¹). Colour-code as in (a). There is a significant linear regression, y=73.2 - 8.50x, r²=0.48, p<0.0001. (c) Minimum O₂ concentration (mg L⁻¹) as a function of O₂ consumption rate, converted to expected P release (see methods). Colour-code shows model-calculated O₂ input to the deep water normalized to the mean O₂ input for all years. A value >1 indicates a higher O₂ input than the mean value for all years, and value <1 indicates a lower O₂ input than the mean value for all years, and value <1 indicates a lower O₂ input than the mean for all years. The linear regression is y=6.31-0.10x, r²=0.23 (p<0.001, however, see Results).

Figure 9. <u>Deep-water (>20 m) n</u>Net total phosphorus (TP, **a**, **c**) and dissolved inorganic phosphorus (DIP, **b**, **d**) release in tons per year in the deep water layer (>20 m) versus expected P remineralization (calculated from oxygen consumption) for

30 full years (a, b) or July-October (c, d) 1969-2015 (1975 and 2011 excluded). The 1:1 line is shown for comparison. The data set is colour-coded according to minimum oxygen (O₂) concentration July-October. Linear regressions were all significant (not shown in graphs), a: y=8.18+1.19x, r²=0.42, p<0.0001 for all values, y=11.1+0.91x, r²=0.73, p<0.0001 for O₂ minimums >4; b: y=4.79+1.17x, r²=0.30, p<0.0001 for all values, y=7.71+0.88x, r²=0.64, p<0.0001 for O₂ minimums >4; c:

y=-2.69+1.70x, $r^2=0.46$, p<0.0001 for all values, $r^2=10.6+0.91x$, $r^2=0.20$, p<0.05 for O₂ minimums >4; **d:** y=-8.55+1.89x, $r^2=0.48$, p<0.0001 for all values, y=6.27+1.04x, $r^2=0.20$, p<0.05 for O₂ minimums >4).

Figure 10. (a) <u>Yearly Annual</u> oxygen consumption July-October 1969-2015 (1975 and 2011 excluded) recalculated to
theoretical deep-water P remineralization, as a function of temperature. The data set is colour-coded according to minimum oxygen concentration July-October: (y=28.3-0.31x, r²=0.006, p=0.60).

(b) <u>A</u>es in (a) but <u>with</u> theoretical P remineralization is-normalized to P loss from 0-20 m layer in March-May- (y=-0.030+0.13x, r²=0.25, p<0.001).

10 **Figure 11.** <u>July-October o</u>Oxygen consumption <u>July October</u>-in the Stockholm inner archipelago deep_-water 1968-2015, calculated with the mean model.



Figure 1.

Figure 2.

Figure 4.

Figure 5.

Figure 6.

Figure 8.

Figure 9.

Figure 10.

Figure 11.