

We would like to thank the reviewer for the thorough and constructive review of our manuscript “Species richness and functional attributes of fish assemblages across a large-scale salinity gradient in shallow coastal areas”.

The assessment and comments are very helpful, and we agree with the suggested changes and needed clarifications/elaborations to improve the manuscript.

Please find our specific responses below, with suggested new or revised text parts in italic font.

### **General comments**

Concerning the general comments, we agree it would benefit the clarity of the manuscript to remove the offshore data, given that only for one sub-basin there was enough data to conduct the statistical (rarefaction-extrapolation) analysis of fish species richness.

Regarding the specific comment concerning the title, i.e. that only coastal but not offshore areas were mentioned, this will be solved when the offshore data is excluded, following the reviewers suggestion.

We agree that the potential factors influencing fish SR should be mentioned already in the introduction. We therefore suggest to revise the following sentence in the introduction to include that aspect:

*“The species composition of fish in the Baltic Sea is regulated by salinity as well (Olsson et al., 2012; Pekcan-Hekim et al., 2016), even though other factors, such as temperature or habitat complexity, might also influence large-scale patterns of fish SR in estuaries (Vasconcelos et al., 2015; Schubert et al., 2011).”*

We further suggest to add two new discussion paragraphs on other regulating factors, specifically:

*“Besides salinity and temperature, which show a pronounced gradient over the large spatial scale of our study (Table 1) and were identified as likely main drivers here (Fig. 2), fish SR might also be influenced by other factors, such as human pressures. The cumulative pressure from human activities in the Baltic Sea, combining factors such as fishing, eutrophication and hazardous substances, is generally higher in the southern and south-western sub-basins. These sub-basins also show both relatively higher salinity and fish SR, compared to for example the northernmost sub-basins with lower cumulative human pressure, salinity and fish SR (Table 2; Korpinen et al., 2012; HELCOM 2018). This indicates that there is no negative relationship between cumulative human pressure and fish SR on the large spatial and temporal scales studied here. It does, however, not contradict well-documented influences that human pressures can have on fish concerning other aspects, or possibly for SR on smaller spatial scales (not studied here). For example, besides direct effects of fish extraction and habitat disturbance on fish species, human-induced depletion of larger predatory fish by Baltic Sea fisheries has likely contributed to an increase in coastal mesopredatory fish abundances (Eriksson et al., 2011), and eutrophication has been connected to increasing abundances of benthic feeding fish in the Baltic Sea (Snickars et al., 2015). Since the rarefaction-extrapolation analyses that we used here are based on species incidence frequencies (Chao et al., 2020), the statistical results could potentially be influenced by human pressures that alter these frequencies, even if fish SR in itself may not be affected. However, given that the rarefied and extrapolated SR (i.e.  $SR_{std}$  and  $SR_{est}$ ) are based on SC,*

*where rare species are not influential, these statistics are rather robust against such effects (unless there would be severe changes in the incidence frequencies of common species). Another potential explanatory factor to consider is habitat complexity, related to e.g. diversity of substrate or habitat-forming macrophytes, which can increase aquatic biodiversity (Soukup et al., 2021). Differences in habitat complexity may play some role in the observed large-scale patterns in fish SR given that macroalgal SR increases with increasing salinity across the Baltic Sea, with a larger share of habitat-forming and perennial species in more marine waters (Middleboe et al., 1997, Schubert et al., 2011). Hence, the greater habitat complexity with increasing salinity may enhance fish SR, further reinforcing any salinity-induced distributional pattern.”*

We agree that further detail, explanation and clarification would benefit section 2.3, “Analysis of species richness data”. In working on the foreseen revisions, we noted an unfortunate error in that the values of sample coverage had been incorrectly termed inventory completeness in the initially submitted manuscript (table and text). We will correct this, better explain the terms (following the reviewers suggestion) and include the values of inventory completeness (besides sample coverage) in Table 2.

To respond to this reviewers request to better elaborate and clarify the statistical terms and methods, we suggest to expand and revise section 2.3 to:

*“The raw data was first summarized to a dataset of unique fish species caught per fishing/sampling occasion in presence/absence format, and then further aggregated to an incidence frequency format, giving the observed total incidence of each species over the number of fishing/sampling occasions. This dataset is referred to as “fish incidence database”. Each unique combination of a fishing/sampling location per date was defined as one sampling unit, and these were summed per sub-basin to obtain the sample sizes. Subsequently, incidence-based Hill diversity numbers of three “orders”, which differ in their propensity to include or exclude relatively rarer species (Hill, 1973), were calculated to quantify the species diversity of each assemblage, i.e. 1) species richness (SR), which counts all species equally irrespective their incidence frequency, 2) Shannon diversity (ShD), which considers the incidence frequency and can be interpreted as the effective number of frequent species, and 3) Simpson diversity (SiD), which can be interpreted as the effective number of highly frequent species (Chao et al., 2014; Chao et al., 2020; Hill, 1973). Calculations were performed using the R package iNEXT and the functions ChaoRichness, ChaoShannon and ChaoSimpson (Chao et al., 2020; Hsieh et al., 2016), and the values are hereafter referred to as observed SR, ShD and SiD, respectively. It should be noted that, using these methods, Shannon and Simpson diversity are expressed in terms of richness, i.e. number of species, which differs from other known formats. Specifically, ShD is the exponential of Shannon’s entropy index, and SiD is the inverse of Simpson’s concentration index (Chao et al., 2014).  $SR_{obs}$  is highly dependent on “sample completeness” (Colwell & Coddington, 1994; Hill, 1973) and may typically underestimates the “true” SR due to undetected species, an aspect referred to as under-sampling, sampling bias or sampling problem (Chao et al., 2014; Chao & Jost, 2015; Menegotto & Rangel, 2018). Similar to Hill numbers, “sample completeness” can be calculated for different “orders” (Chao et al., 2020). The zero-order sample completeness is hereafter referred to as inventory completeness (IC). It is calculated as the ratio of  $SR_{obs}$  to the estimated “true” SR (i.e. observed plus undetected SR, see “estimated SR” below), hence giving the proportion of detected species without considering the species incidence frequencies. We calculated IC for the data merged over time, and including both resident and migrating/visiting fish species. The first-order sample completeness, hereafter referred to as “sample coverage” (SC), is a measure where species are weighted by their*

detection probabilities, giving the proportion of incidences detected from the estimated “true” incidences (Chao et al., 2020).

To correct for the effect of differing sample completeness on  $SR_{obs}$ , and allow accurate, unbiased comparisons between sub-basins, we used a coverage-based rarefaction and extrapolation method implemented for incidence data in the R package iNEXT (Chao et al., 2014, 2020; Hsieh et al., 2016). A coverage-based method was chosen because more traditional sample size-based corrections can introduce a systematic bias, since the number of samples needed to fully characterize a community depends on its SR (Chao & Jost, 2012). For each sub-basin, we obtained 1) the rarefied SR,  $ShD$  and  $SiD$ , which were standardized to the minimum observed SC across all included sub-basins (hereafter referred to as standardized values, i.e.  $SR_{std}$ ,  $ShD_{std}$  and  $SiD_{std}$ ), and 2) the actual, asymptotic fish SR extrapolated to twice the actual sample size (hereafter referred to as estimated values, i.e.  $SR_{est}$ ,  $ShD_{est}$  and  $SiD_{est}$ ; Chao et al., 2014, 2020; Hsieh et al., 2016). Similar analyses were also conducted for SR of fish with different functional attributes (see Sect. 2.4). All calculations were conducted using R version 4.0.4 (R Core Team, 2021).”.

We agree as well that a more detailed discussion of the fish functional characteristics and changes across the geographic gradient is warranted. We suggest to expand and revise the respective discussion section, starting L449, to:

“Our study also revealed changes in fish SR for different functional groups across the studied salinity gradient. As the different functional groups represent variability among species in e.g. use of resources or level of connectivity with other areas, this may also translate to taxonomic-driven differences in coastal ecosystem functioning across the different sub-basins (Elliott et al., 2007; Franco et al., 2008). Clearly migrating fish species are typically of marine origin (here classified as marine juvenile migrants, marine seasonal migrants or marine visitors) and cannot tolerate low salinity, explaining their predominance at higher salinities (Fig. 4), and in agreement with known patterns in European estuaries in general (Elliott & Dewailly, 1995; Franco et al., 2008). This pattern of marine fish species temporarily using coastal areas may be related to comparatively higher prey densities and to food types not encountered in marine areas, as well as to typically more turbid waters providing better protection from predators (Franco et al., 2008). Moreover, the high migratory fish SR at higher salinity is likely relevant for the ecological connectivity between ecosystems, e.g. by transport of local “coastal” production to open sea and vice versa (Franco et al., 2008), and emphasizes the important role of higher salinity coastal areas as nursery grounds, migration routes and refuge areas for marine fish species (Elliott et al., 2007). Connectivity is also maintained in the less saline sub-basins, though the concerned functional groups are represented by only a few species (Fig. 1, S2; Berkström et al., 2021).

Benthic and demersal fish SR also decreased with decreasing salinity, corroborating previous results where demersal fish  $SR_{obs}$  decreased from the saline Kattegat to the less saline northern Baltic Proper (Pecuchet et al., 2016), and in accordance with high benthic preference of marine fish species in European estuaries (Elliott & Dewailly, 1995; Franco et al., 2008). This pattern further corresponds with that the observed SR of benthic meio- and macrofauna, which are generally the dominating prey for benthic and demersal fish, also decreased with decreasing salinity in the Baltic Sea (Broman et al., 2019; Zettler et al., 2014). Taken together, these patterns suggest that the process of benthic-pelagic coupling through fish predation likely involves a lower number of species links, or functional redundancy, towards lower salinity sub-basins. Concerning feeding habits, the general composition of feeding guilds noted in the higher-salinity sub-basins was similar to that reported on a larger European scale (Elliott & Dewailly, 1995). Also our findings of higher

*piscivorous fish SR in the more saline sub-basins (Fig. 6) and the pattern showing a low proportion of omnivorous fish that was unrelated to salinity levels was in agreement with findings from European estuaries (Franco et al., 2008). In summary, the found differences in functional traits of fish along the salinity gradient were largely related to the respective changes in the predominating fish origin, i.e. freshwater vs. marine species.”*

### **Specific comments**

We thank the reviewer for pointing out the needed technical corrections and suggestions for improving tables, and will follow all of them.

Reflecting the revisions conducted to the manuscript we further suggest to slightly revise/adapt the Abstract to:

*“Coastal ecosystems are biologically productive and their diversity underlies various ecosystem services to humans. However, large-scale species richness (SR) and its regulating factors remain uncertain for many organism groups, owing not least to the fact that observed SR ( $SR_{obs}$ ) depends on sample size and inventory completeness (IC). We estimated changes in SR across a natural geographical gradient using statistical rarefaction and extrapolation methods, based on a large fish species incidence dataset compiled for shallow coastal areas (<30 m depth) from Swedish fish survey databases. The data covered a ca. 1,300 km north-south distance and a 12-fold salinity gradient along sub-basins of the Baltic Sea plus Skagerrak and, depending on sub-basin, 4 to 47 years of samplings during 1975–2021. Total fish  $SR_{obs}$  was 144, and the observed fish species were of 74% marine and 26% freshwater origin. In the 10 sub-basins with sufficient data for further analysis, IC ranged from 77–98%, implying that ca. 2–23% of likely existing fish species remained undetected. Sample coverage ranged from 98.5–99.9%, suggesting that the undetected species represented <1.5% of incidences across the sub-basins, i.e. highly rare species. To compare between sub-basins, we calculated standardized SR ( $SR_{std}$ ) and estimated SR ( $SR_{est}$ ). Sub-basin specific  $SR_{est}$  varied between  $35 \pm 7$  (SE) and  $109 \pm 6$  fish species, being ca. three times higher in the most saline (salinity 29–32) compared to the least saline sub-basins (salinity <3). Additional information on functional attributes showed that differences with decreasing salinity particularly reflected a decreasing SR of benthic and demersal fish, piscivorous and invertebrate-eating fish, and marine migratory fish. We conclude that, if climate change continues causing an upper-layer freshening of the Baltic Sea, this may influence the SR, community composition and functional characteristics of fish, which in turn may affect ecosystem processes such as benthic-pelagic coupling and connectivity between coastal and open sea areas.”*

We thank the reviewer and Editors for their efforts with this work, and are looking forward to hear from you about our manuscript.

Yours sincerely,

Birgit Koehler and co-authors