Species richness and functional attributes of fish assemblages across a large-scale salinity gradient in shallow coastal areas

Birgit Koehler[1], Mårten Erlandsson[1], Martin Karlsson[1,a], and Lena Bergström[1]

[1]Institute of Coastal Research, Department of Aquatic Resources, Swedish University of Agricultural Sciences, Skolgatan 6, 74243 Öregrund, Sweden

[a]now at: Department for Marine Management, The Swedish Agency for Marine and Water Management, Gullbergs Strandgata 15, 41104 Gothenburg, Sweden

Correspondence: Birgit Koehler (birgit.koehler@slu.se) and Lena Bergström (lena.bergstrom@slu.se)

Abstract. 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. 1300 km north–south distance and a 12-fold salinity gradient along sub-basins of the Baltic Sea plus the Skagerrak and, depending on the 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 % to 98 %, implying that ca. 2 %–23 % of likely existing fish species had remained undetected. Sample coverage exceeded 98.5 %, suggesting that undetected species represented $<$ 1.5 % of incidences across the sub-basins, i.e. highly rare species. To compare sub-basins, we calculated standardized SR (SR_{std}) and estimated SR (SR_{est}). Sub-basin-specific SR_{est} varied between 35 ± 7 (SE) and 109 ± 6 fish species, being ca. 3 times higher in the most saline (salinity 29–32) compared to the least saline sub-basins (salinity < 3). Analysis of functional attributes showed that differences with decreasing salinity particularly reflected a decreasing SR of benthic and demersal fish, of piscivores and invertivores, and of marine migratory species. 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.

1 Introduction

Biodiversity is essential for ecosystem processes and ultimately for the humans depending on these (IPBES, 2019). Coastal ecosystems are often biologically diverse and highly productive, providing valuable ecosystem services to humans, such as food, water purification and protection against floods (Griffiths et al., 2017; Kraufvelin et al., 2018; Pan et al., 2013). However, threats to coastal biodiversity from, for example, overfishing, habitat loss, pollution, eutrophication and climate change are many and profound (Duncan et al., 2015; Griffiths et al., 2017; Pan et al., 2013). At the same time, the number of species occurring in coastal habitats often remains uncertain (Appeltans et al., 2012), making improved understanding of coastal biodiversity especially important to support conservation and management measures (Pan et al., 2013; Rooney and McCann, 2012).

Taxonomic inventories, or species censuses, are required, for example, for the analysis of biodiversity patterns, delineation of species ranges and prioritization of conservation efforts (Mora et al., 2008). Species richness (SR), i.e. the number of species in an ecosystem, is a classical indicator of biodiversity, also referred to as "alpha diversity" (Gotelli and

Colwell, 2001; Hill, 1973). However, since achieving complete species inventories is often impracticable with realistic sample efforts, most censuses remain incomplete, and many rare species remain unknown. Consequently, it is important to consider the effect of sample size and inventory completeness (IC) on observed SR (SR_{obs}) to avoid biased or misleading comparisons or interpretations (Chao and Chiu, 2016; Chao et al., 2020; Colwell and Coddington, 1994; Mora et al., 2008).

SR is connected to several ecosystem processes, such as productivity (Duffy et al., 2017), and the efficiency of resource use and nutrient cycling. SR may also facilitate the simultaneous provision of several ecosystem processes, i.e. an ecosystem's multifunctionality (Byrnes et al., 2014). However, since species do not contribute equally to ecosystem functioning, the diversity of species functional attributes adds another important dimension to ecosystem understanding (Duncan et al., 2015; Reiss et al., 2009). Functional diversity can enhance long-term stability through functional redundancy and complementarity and can help to buffer ecosystems against disturbances (O'Gorman et al., 2011).

Salinity is a key variable influencing SR in coastal areas as natural differences in salinity among locations function as a threshold or "ecological barrier" for the distribution of freshwater and marine species, for example, in the Baltic Sea (Olenin and Leppäkoski, 1999; Vuorinen et al., 2015). At the same time, an on average intensified water cycle caused by global warming is changing the salinity regimes of marine and coastal ecosystems (Durack et al., 2012; Liblik and Lips, 2019; Meier et al., 2022). It is important to understand how salinity influences species distributions in aquatic ecosystems to be able to better predict how potential changes may affect ecosystem functioning.

The Baltic Sea, one of the world's largest brackish water bodies, exhibits a pronounced, geographically stable salinity gradient that is maintained by sporadic inflows of saline water from the North Sea through the Danish Straits and by freshwater input from large rivers, especially in the north. Hence, the Baltic Sea gradient can serve as a model on the influence of salinity on species distributions (Johannesson and Andre, 2006; Ojaveer et al., 2010). SR_{obs} is often higher at the more saline conditions, e.g. for macroalgae, benthic bacteria, and benthic meio- and macrofauna (Broman et al., 2019; Klier et al., 2018; Middelboe et al., 1997; Schubert et al., 2011). In other studies, SR_{obs} was highest at highest salinity, lowest at intermediate salinity and intermediate at lowest salinity, e.g. for phytoplankton and benthic macrofauna (Bonsdorff, 2006; Olli et al., 2019; Zettler et al., 2014), or there was no clear trend between SR_{obs} and salinity, e.g. for bacterio-, pico- and mesoplankton (Herlemann et al., 2016; Hu et al., 2016).

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). In the Baltic Sea, fish SR_{obs} is generally higher in areas of higher compared to lower salinities (HELCOM, 2020; Hiddink and Coleby, 2012; Lappalainen et al., 2000; MacKenzie et al., 2007; Ojaveer et al., 2010; Pecuchet et al., 2016; Thorman, 1986). Various studies have also reported changes in fish SR_{obs} or species composition over time (e.g. Ammar et al., 2021; Törnroos et al., 2019). However, despite concerns that fish SR may decline in the future due to decreasing upper-layer salinity (e.g. MacKenzie et al., 2007; Pecuchet et al., 2016; Vuorinen et al., 2015), information on how the complete coastal fish assemblage varies spatially in relation to the Baltic Sea salinity gradient, including potential differences across functional groups, is lacking. Hence, there is a need to complement already existing information on the influence of salinity on various Baltic Sea organism groups with more complete information in relation to fish diversity, taxonomically and functionally. This kind of understanding for multiple trophic levels is needed to better understand and predict how changing salinity, in the Baltic Sea and in coastal areas in general (Durack et al., 2012; Liblik and Lips, 2019), may affect ecosystem structure and functioning (MacKenzie et al., 2007). For example, if different species groups are differently affected, this may also change biotic interactions such as benthic–pelagic coupling, with effects on exchanges of energy, mass or nutrients between benthic and pelagic habitats (Griffiths et al., 2017). Moreover, understanding SR at a broader, sub-regional scale is important to support analyses of potential SR and species compositions at more local scales within each sub-basin.

To this aim, we compiled a large dataset on fish species observations in shallow (< 30 m depth) Swedish coastal areas, defined in alignment with the EU Marine Strategy Framework Directive as the area within 1 nmi, that is, 1852 m, beyond the national baseline (EC, 2017). The data compilation was based on multiple existing sources of Swedish mapping and monitoring combined over the years 1975–2021. The extensive dataset covered fish species incidence information from 1638 unique observations/fishing occasions, during which in total 21 670 species incidences were recorded. Geographically, the data covered 12 hydrographically distinct sub-basins and a ca. 12-fold salinity gradient from close to freshwater conditions in the inner Baltic Sea to close to fully marine conditions at the Swedish west coast. The annual mean water temperature varies ca. 2-fold across the same area. Since SR_{obs} is strongly dependent on sample size, which differed between sub-basins, we used statistical rarefaction and extrapolation methods to estimate IC and standardized SR per sub-basin. Further, we categorized each fish species according to origin (marine vs. freshwater), as well as three functional attributes based on coastal habitat preference, vertical preference and feeding habitat, and investigated the influence of salinity (and, for comparison, temperature) on fish SR in total and within the functional attributes. We discuss the results in the context of the regulating influ-

Table 1. Salinity and temperature in Swedish coastal areas, given as mean ($\pm$ SE) annual values per sub-basin across the years 1993–2019. Values represent conditions near the bottom at 0–30 m depth based on data from the EU Copernicus Marine Service Information (CMEMS, 2021).

Sub-basin	Salinity	Temperature (°C)
Bothnian Bay	2.68 ± 0.01	4.53 ± 0.23
The Quark	4.26 ± 0.01	5.38 ± 0.25
Bothnian Sea	5.10 ± 0.01	5.44 ± 0.22
Åland Sea	5.80 ± 0.01	6.44 ± 0.25
N Baltic Proper	6.37 ± 0.01	6.43 ± 0.22
E Gotland Basin	6.85 ± 0.01	7.30 ± 0.24
W Gotland Basin	6.88 ± 0.01	6.48 ± 0.20
Bornholm Basin	7.60 ± 0.02	8.15 ± 0.24
Arkona Basin	10.96 ± 0.07	8.92 ± 0.26
The Sound	23.42 ± 0.14	9.72 ± 0.24
Kattegat	29.02 ± 0.05	9.32 ± 0.21
Skagerrak	32.40 ± 0.03	9.62 ± 0.22

ence of salinity on fish SR and community composition in coastal ecosystems, as well as potential implications for conservation and ecosystem management.

2 Methods

2.1 Study system

The Baltic Sea, an enclosed, essentially non-tidal brackish marine region with maximum and mean depths of 460 and 54 m, respectively, and a water residence time of 25–40 years, is among the world's largest estuaries (area: 415 000 km^2; HELCOM, 2018). Its current brackish conditions were formed by the gradual narrowing of its opening to the North Sea and have been in place for ca. 3000 years (Russell, 1985). Due to its geographically variable but locally relatively stable salinity conditions, the Baltic Sea has been called a "marine-brackish-limnic continuum" (Bonsdorff, 2006). Its surface salinity changes from < 3 (psu) in the inner-most areas in the north and northeast to almost fully marine (ca. 29) in the Kattegat in the southwest (Table 1). Within this gradient, the Baltic Sea can be divided into hydrographically distinct sub-basins, mostly separated by shallow sounds or sills. To strengthen the database with respect to higher-salinity areas we additionally included a North Sea sub-basin adjacent to the Kattegat, i.e. the Skagerrak (salinity ca. 32; Table 1).

Reflecting its salinity conditions the Baltic Sea harbours a unique fish fauna with a mixture of freshwater species (e.g. northern pike (*Esox lucius*), perch (*Perca fluviatilis*), pikeperch (*Sander lucioperca*)) and marine species (e.g. cod (*Gadus morhua*), herring (*Clupea harengus*); Olsson et al., 2012). Further, many marine fish populations have adapted to the brackish conditions from their Atlantic counterparts

(Laikre et al., 2005), for example, Baltic cod and herring populations, and one flounder species is endemic to the Baltic Sea (Momigliano et al., 2018). Hence, the Baltic Sea may also have a unique value as a refuge for evolutionary lineages and constitute an important genetic resource for management and conservation (Johannesson and Andre, 2006).

2.2 Species richness data

The primary source of fish species data was the Swedish National database of coastal fish (KUL; https://www.slu.se/kul, last access: 27 April 2021), which holds public, quality-assured data from surveys encompassing coastal fish monitoring, mapping projects and surveillance programmes over the entire salinity gradient of the Baltic Sea plus the Skagerrak. Coastal areas were delineated as the zone within one nautical mile beyond the national baseline (EC, 2017), as available from the Water Information System Sweden ("WISS"; https://viss.lansstyrelsen.se/Maps.aspx, last access: 10 November 2020, in Swedish), and KUL data records for 1975 to 2021 were extracted. Subsequently, data from shallow depths < 30 m were selected, corresponding to the main represented sampling methods in the database (Table S1 in the Supplement) and approximately to the photic depth in the concerned coastal habitat types (Kaskela et al., 2012). This selection resulted in 154 172 data entries, i.e. individual fish that had been caught and identified to species by specialists, with the number of years with available data differing between sub-basins (Table S2). Further, quality-assured data were included from (1) a national coastal trawl survey ($n = 4420$), (2) the ICES-coordinated International Bottom Trawl Survey (IBTS, $n = 1969$) and (3) national projects using standardized methodology ($n = 893$), all carried out in the Skagerrak, the Kattegat and the Sound, selecting only hauls from < 30 m depth within the concerned geographical delineations (see above). Corresponding trawl data for the inner Baltic Sea area are not collected at depths < 30 m in Swedish waters. Across databases, only sampling occasions with geographical coordinates and verified sampling references were included. Hence, data collected from multiple types of gear were combined, including gill nets, fyke nets, seines, trap nets, low-impact underwater detonations and trawls, in order to maximize the chance of including different species (Table S1). The ambition to collate information from all available fish surveys implied some differences in predominating data collection methods across the studied geographical range. The main data sources were trawls and trap net surveys in the most saline sub-basins, i.e. the Skagerrak, the Kattegat and the Sound, and gill net surveys in the remaining sub-basins (Table S1). While each gear has a specific selectivity and efficiency, a previous comparison of data from gill and fyke net samplings at the Swedish west coast did not reveal consistent differences in biodiversity metrics, and the statistical approaches were chosen to minimize a potential bias when comparing SR among sub-

basins (Bergström et al., 2013; Chao et al., 2020; see also search of additional data sources below and Sects. 2.3 and 4).

Data on observed SR, SR_{obs}, were available for all 12 sub-basins (Table 2), representing between 4 and 47 years of data, depending on the sub-basin (Table S2). However, subsequent statistical analyses and comparisons were conducted only for the 10 sub-basins containing data from at least 30 sampling/fishing occasions, corresponding to several hundred fish incidences (i.e. Bothnian Bay, the Quark, Bothnian Sea, Åland Sea, northern Baltic Proper, western Gotland Basin, Bornholm Basin, the Sound, the Kattegat and the Skagerrak; see also Sect. 2.3). This dataset is hereafter referred to as "raw data", and it contains in total 160 453 entries (i.e. fish individuals caught and determined to species) from 1638 sampling/fishing occasions at 4571 unique locations. Eastern Gotland Basin and Arkona Basin were not statistically analysed since we considered these sub-basins too under-sampled, with only 13 and 7 samplings, respectively, from 9 and 4 different years, and less than 100 species incidences in total (Tables 2, S2). Additionally, we searched for evidence of fish species that had remained undetected in the fish surveys (i.e. incidence database; see Sect. 2.3) by identifying fish species records from three additional sources using the same criteria for geographical and depth delineations of shallow coastal areas as above, i.e. (1) the national public database for citizens' reporting of species observations hosted by the Swedish University of Agricultural Sciences (SLU Swedish Species Information Centre, https://www.artportalen.se/ (last access: 27 April 2021); $n = 8926$, unreasonable species observations deemed as falsely identified were discarded), (2) the national archive for oceanographic data hosted by the Swedish Meteorological and Hydrological Institute (SHARKweb, https://www.smhi.se/en/services/open-data/national-archive-for-oceanographic-data (last access: 27 April 2021); $n = 1259$), and (3) published inventory data from the Skagerrak, the Kattegat and Bornholm Basin (Pihl and Wennhage, 2002; Pihl et al., 1994; Wikström A., 2009). These "additional data sources" were used as complementary information on SR_{obs} but could not be used in the statistical analysis since they did not have complete sampling and species incidence information. Further, our SR results were compared with the HELCOM (2020) checklist on macro-species containing information for all of the sub-basins and depths in the Baltic Sea region.

2.3 Analysis of species richness data

The raw data were first summarized into a dataset of *unique fish species* caught in presence–absence format, per fishing/sampling occasion in each sub-basin, and then further aggregated to an incidence frequency format (Chao et al., 2014), in which incidence is occurrence/presence of a species. The resulting dataset gives the observed total incidence of each species over the number of fishing/sampling occasions per sub-basin and 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, 2020). Calculations were performed using the R package *iNEXT* (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 diversities 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 and Coddington, 1994; Hill, 1973), and it 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 and Jost, 2015; Menegotto and 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 resident as well as migrating and 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 and 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 *stan-*

Table 2. Summary information for the fish incidence database and additional data sources, covering 12 sub-basins in Swedish shallow coastal areas (< 30 m depth). Inventory completeness (IC), sample coverage (SC), standardized (SR_{std}) and estimated (SR_{est}) values were statistically estimated for sub-basins with a sample size of > 30 fishing/sampling occasions. Of these, SR_{std} was calculated for a SC of 98.5 %, which was the lowest SC among sub-basins with sufficient data (i.e. Åland Sea). For comparison, the last two columns give SR_{obs} for Swedish shallow coastal areas when also including species records from additional data sources (not included in the statistical analyses; see Sect. 2.2) and for each Baltic Sea sub-basin according to HELCOM (2020), i.e. across depths and Baltic countries. NA: not applicable; n.d.: not determined.

Sub-basin	Species incidence dataset					Statistical estimations				Including additional data sources	
	Observations										
	Sample size	Species incidences[a]	SR_{obs} (incidence data alone)	Singletons	Doubletons	IC (%)[b]	SC (%)[c]	SR_{std} (with CI)	SR_{est} ($\pm$ SE)	Total SR_{obs} (incidence data plus additional data sources)	SR_{obs} (across depths and Baltic countries; HELCOM (2020))
Bothnian Bay	70	553	29	4	0	82.9	99.3	24^{28}_{25}	35 ± 7	34	51
The Quark	71	754	30	4	2	88.2	99.5	26^{28}_{24}	34 ± 5	35	56
Bothnian Sea	194	2 222	42	2	0	97.7	99.9	37^{38}_{36}	43 ± 2	50	74
Åland Sea	31	394	32	6	3	84.2	98.5	32^{35}_{28}	38 ± 6	45	71
N Baltic Proper	77	1 046	42	6	3	87.5	99.4	36^{38}_{34}	48 ± 6	56	67
E Gotland Basin	13	94	25	8	5	n.d.	n.d.	n.d.	n.d.	37	82
W Gotland Basin	411	5 123	53	6	2	85.5	99.9	39^{40}_{38}	62 ± 10	60	67
Bornholm Basin	68	837	46	5	1	79.3	99.4	42^{44}_{40}	58 ± 17	59	TS1
Arkona Basin	7	67	23	7	7	n.d.	n.d.	n.d.	n.d.	37	110
The Sound	119	1 373	61	8	4	88.4	99.4	54^{57}_{51}	69 ± 7	70	144
Kattegat	353	6012	101	11	8	92.7	99.8	78^{80}_{76}	109 ± 6	114	178
Skagerrak	230	3195	69	13	3	76.7	99.6	52^{53}_{50}	90 ± 16	106	NA[d]

[a] Sum of the number of unique species observed across all sampling occasions. Please note that this does not correspond to "entries" in Sect. 2.2, which is individual fish caught and determined to species.
[b] Percentage of species detected from the estimated "true" (i.e. observed plus undetected) SR, i.e. 100 · SR_{obs}/ SR_{est} (Chao et al., 2020).
[c] Percentage of incidences detected from the estimated "true" (i.e. observed plus undetected) incidences (Chao et al., 2020).
[d] Not included in the Baltic Sea region.

dardized values, i.e. SR_{std}, ShD_{std} and SiD_{std}) and (2) the 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).

2.4 Fish functional attributes

All observed fish species were classified by origin (i.e., marine or freshwater) and assigned three functional attributes based on ecological and behavioural traits (Kullander, 2002). Specifically, the affinity of each species to different parts of the coastal habitat, or habitat preference, was assigned based on Elliott and Dewailly (1995) and Pihl and Wennhage (2002), however with certain adaptations to suit both marine and brackish conditions (Table S3–S4). Applied categories were as follows: catadromous or anadromous migrant (CA), using coastal habitats only when migrating between marine and freshwaters for spawning and feeding; marine juvenile migrant (MJ), using coastal habitats primarily as nursery or feeding grounds; marine visitor (MV), occurring irregularly in the coastal area, having their primary habitat in deeper waters; marine seasonal migrant (MS), making regular seasonal visits to coastal habitats, usually as adults; and coastal resident (CR), spending almost their complete life cycle in coastal habitats or the littoral coastal zone. The main vertical distribution of each species in the water column, considering the adult stage, was assigned based on Elliott and Dewailly (1995) and Koli (1990) as follows: pelagic (P), living mainly in the water column; demersal (D), mainly associated with the bottom substrate; demersal–pelagic (DP), alternating between the water column and bottom substrate; and benthic (B), staying close to the seabed. Main feeding habits were assigned by combining information on feeding guild (Elliott and Dewailly, 1995) with trophic level (TL) and principal diet composition (Froese and Pauly, 2021) as follows: piscivores (Pi; TL 3.6–4.4), invertivores and piscivores (IP; TL 2.9–3.9), invertivores (I; TL 2.8–3.9), planktivores (PL; TL 3.1–3.2), and omnivores (O; TL 2.8–3.5).

2.5 Seawater salinity and temperature

For each sub-basin, data on ambient salinity and temperature were extracted from the "Baltic Sea Physics Reanalysis" product, as calculated by the Swedish Meteorological and Hydrological Institute (SMHI) with the coupled physical–biochemical model system NEMO-SCOBI and available from the year 1993 onwards (CMEMS, 2021). This encompassed full coverage layers with a $4\,km \times 4\,km$ grid. Monthly mean values near the bottom for all grid cells representing areas $< 30\,m$ depth were used for calculating long-term means and standard deviations for the years 1993–2019.

2.6 Statistical analyses

Simple linear regressions were used to analyse the relationships between salinity or temperature and observed, standardized and estimated SR, ShD and SiD. To test for any additional explanatory effect of temperature, after accounting for the effect of salinity, we used ANOVA to compare models with salinity as the only explanatory factor with models with salinity plus temperature as explanatory factors. Furthermore, relationships were tested between the different functional attributes and salinity. To reduce skewness and approximate normality, left-skewed response variables were $\log_{10}$ transformed prior to analysis or, in two cases where the response variable included zero values, Yeo–Johnson transformed (Yeo and Johnson, 2000). All analyses were conducted using R version 4.0.4 (R Core Team, 2021).

3 Results

3.1 Salinity and temperature

The annual mean salinity varied more than 12-fold in the shallow coastal areas across the studied sub-basins, from 2.7 in the northernmost Baltic Sea to 32.4 in the Skagerrak. Across the same geographical range, the annual mean water temperature varied from $4.5\,°C$ in the north to ca. $9–10\,°C$ in the Sound and outwards (Table 1).

3.2 Fish species observations and distribution

SR_{obs} varied from 23 (Arkona Basin) to 101 (the Kattegat; Table 2, which also contains related information on, for example, sample size and species incidences per sub-basin) and amounted to 125 across sub-basins and years. Since IC was $< 100\,\%$ (see Sect. 3.3), this can be assumed to be a lower-bound estimate of the true SR. Indeed, the additional data sources contained 19 more species that were not represented in the fish incidence database, resulting in a total fish SR_{obs} of 144 (Tables S3, S4). Of the species in the fish incidence database, 49 % occurred only in the higher-salinity Skagerrak–Kattegat region including the Sound, 15 % occurred only in the Baltic Sea region (i.e. inside the Sound) and 36 % occurred in both these regions. The species ranging most widely were herring (*Clupea harengus*), brown trout (*Salmo trutta*), European sprat (*Sprattus sprattus*) and eelpout (*Zoarces viviparous*), with incidences reported from all 12 sub-basins (Tables S3, S4).

3.3 Inventory completeness and sample coverage

The fish species IC in Swedish shallow coastal areas varied from 76.7 % in the Skagerrak to 97.7 % in the Bothnian Sea for the assessed sub-basins (see Sect. 2.2), suggesting that ca. 2 %–23 % of species statistically likely to be present remained undetected (Table 2). The SC exceeded
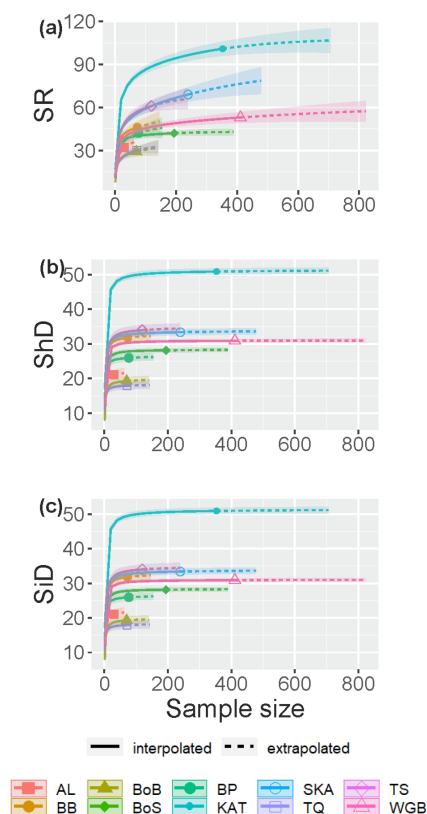
Figure 1. Sample-size-based species accumulation curves showing rarefaction/interpolation (solid curve segments) and extrapolation up to twice the actual sample size (dashed curve segments; Chao et al., 2020) with 95 % confidence intervals (shaded areas) for **(a)** species richness (SR), **(b)** Shannon diversity (the effective number of frequent species in the assemblage, ShD) and **(c)** Simpson diversity (the effective number of very frequent species in the assemblage, SiD) of fish in shallow coastal areas of the 10 assessed sub-basins. The intersection points between solid and dotted lines represent the observed values. Legend acronyms are AL: Åland Sea; BB: Bornholm Basin; BoB: Bothnian Bay; BoS: Bothnian Sea; BP: north Baltic Proper; KAT: Kattegat; SKA: Skagerrak; TQ: the Quark; TS: the Sound; and WGB: western Gotland Basin.

98.5 % in all sub-basins (Table 2), suggesting that these undetected species were highly rare, likely representing $< 1.5\%$ of incidences. The species accumulation curves (SACs) show the SR_{obs} at the conducted sample sizes and SR estimated for hypothetical smaller and larger sample sizes, including 95 % confidence intervals. According to these, the steepest increase in accumulated species occurred with the first ca. 20 samplings in all sub-basins, and coastal fish SR was highest in the Kattegat, followed by the Skagerrak and the Sound, and lowest in the other seven sub-basins (Fig. 1a).

The SACs also visualize differences in IC between sub-basins. For the three most saline sub-basins, the Skagerrak, the Kattegat and the Sound, the SACs were still clearly increasing with increasing sample size even when extrapolating to double the actual sample size. Hence, the SR_{est} val-

ues for these sub-basins are more uncertain and more likely biased low than for sub-basins where the curve flattened, illustrating a more complete inventory, e.g. western Gotland Basin and the Bothnian Sea (Fig. 1a). The SR_{est} values, estimated based on extrapolation of the information in the fish incidence database, were similar to SR_{obs} if complementing the incidence data with records from the additional data sources (Table 2).

3.4 Shannon and Simpson diversities

Rarefaction and extrapolation SACs carried out for Shannon diversity (ShD) show that the effective number of frequently recorded fish species was quite well captured by the samplings in all statistically assessed sub-basins, illustrated by SACs with small remaining slopes at extrapolated higher sample size. As for SR_{obs}, ShD was highest in the Kattegat, while the remaining nine sub-basins clustered in two separate groups. The lowest ShD was noted for the Åland Sea, the Quark and Bothnian Bay (Fig. 1b, Table S5). The effective number of highly frequent species, i.e. Simpson diversity (SiD), was also well captured in all sub-basins, being highest in the Kattegat, while SiD in the remaining sub-basins clustered in four groups (Fig. 1c, Table S5).

3.5 Standardized and estimated species richness

To compare coastal fish SR, ShD and SiD across sub-basins, we estimated their standardized values against the minimum observed SC in any of the sub-basins. This represented a standardization to the SC of the Arkona Basin data (98.5 %; Tables 2 and S3). SR_{std} was ca. 3 times higher in the relatively more saline Kattegat ($SR_{std} = 78$) compared to the least saline Bothnian Bay ($SR_{std} = 24$), as also confirmed by comparing the respective SR_{est} values (Table 2). The differences were smaller for ShD and SiD. For example, based on SiD_{std} and SiD_{est}, the effective number of highly frequent species was ca. 2 times higher in coastal areas of the Kattegat compared to Bothnian Bay (Table S3). This implies, as also seen from the SACs (Fig. 1), that the frequent and most frequent fish species were captured quite well by the samplings for all sub-basins and that remaining uncertainties in differences across the salinity gradient are mostly due to uncertainty in the numbers of highly rare fish species. The estimated SR per sub-basin (SR_{est}, statistical extrapolation of the fish incidence data) was very similar to the total SR_{obs} when also including species records from additional data sources, with a mean ratio of 1.07 ± 0.03 (calculated based on data presented in Table 2).

3.6 Relationships of SR with salinity and temperature

Species richness increased with increasing mean water salinity, which explained 37 %–55 % of the variance in the data based on SR_{obs}, ShD_{obs} and SiD_{obs}. Using the standardized or estimated values, i.e. values corrected for sample size, re-
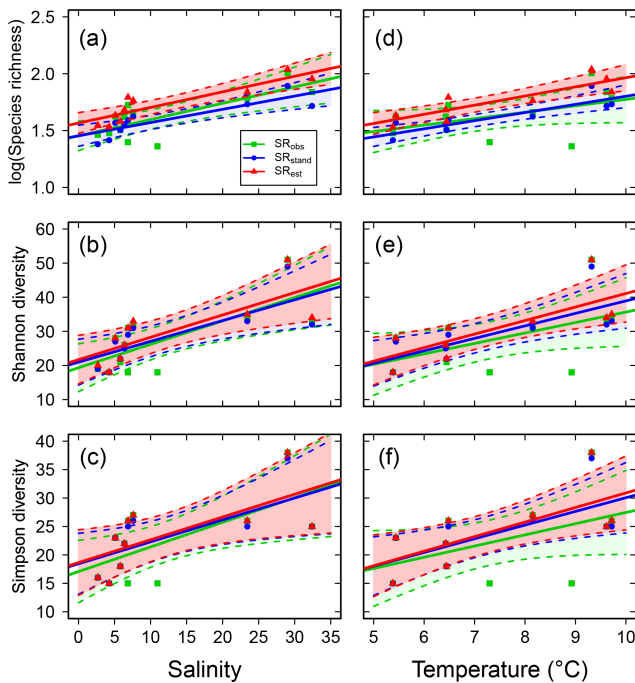
Figure 2. Fish species richness estimates in relation to mean salinity (left column) and mean water temperature (right column), with total species richness ($\log_{10}$ transformed; **a** and **d**), Shannon diversity (effective number of frequent species; **b** and **e**) and Simpson diversity (effective number of highly frequent species; **c** and **f**). Each plot shows the observed, standardized and estimated values and, when significant ($P < 0.05$), the linear regression lines (solid) and 95 % confidence intervals (shaded areas surrounded by dashed lines). Please note that data points (Table S7), lines and shaded confidence intervals are overlying each other within the panels in some cases. For regression equations and statistics, see Table 3.

sulted in stronger correlations, i.e. higher explained variance (40 %–77 %; Fig. 2a–c, Table S6). SR_{obs}, ShD_{obs} and SiD_{obs} were not correlated with mean water temperature, but, using the standardized and estimated values, correlations with temperature were also significant (explaining 48 %–77 % of the variance; Fig. 2d–f, Table S6). The slope estimates of the linear regressions differed more across observed, standardized and estimated values for SR than for ShD and SiD (Table S6). Adding temperature as an explanatory variable to the regressions with salinity did not improve the model in any case (all $P > 0.14$).

3.7   Fish functional attributes

Of the observed fish species, 74 % and 26 % were of marine and freshwater origin, respectively (based on the incidence data, i.e. SR_{obs} of 92 vs. 33 species; Table S3). In the most saline sub-basins, i.e. the Skagerrak and the Kattegat, the SR_{std} of marine fish species was 7 to 10 times higher than that of freshwater fish species. The SR_{std} values of marine vs. freshwater fish were rather similar in the

central Baltic Sea, while in the northernmost and least saline sub-basins, i.e. Bothnian Sea, the Quark and Bothnian Bay, the SR_{std} of freshwater fish species exceeded the SR_{std} of marine fish species by 2 to 3 times. In total, the marine fish SR_{std} decreased by a factor of 8–11 along the salinity gradient, from 39 and 57 marine species (SR_{std}) in the Skagerrak and the Kattegat to 5 in Bothnian Bay. Freshwater fish SR_{std} increased by a factor of 2–4 along the same gradient (Fig. 3, Table S3). These distributional patterns of freshwater vs. marine fish species were also reflected by negative univariate correlations of freshwater SR (obs, std and est) with salinity and positive univariate correlations of marine SR with salinity (Fig. S1 in the Supplement, Table 3).

Concerning habitat preference, half of the fish species in Swedish shallow coastal areas were classified as being coastal resident species (CR; based on the incidence database only, SR_{obs}: 63 species; Table S3). This group dominated coastal fish assemblages in all sub-basins, with CR SR_{std} of 19–30 across sub-basins (Fig. S2, Table S7), and was not linearly related to salinity (Fig. 4a, Table 3). A similar result was noted for catadromous or anadromous fish species, with SR_{std} of 2–6 across sub-basins and CA SR_{std} not related to salinity (Fig. 4b, S2, Tables 3, S7). In the more saline sub-basins, fish species classified as marine visitors or as marine juvenile or seasonal migrants contributed significant numbers to the SR_{std}, while these species groups did not exist or contributed only little to the SR_{std} in the Baltic Sea region (Fig. S2, Table S7). Reflecting this pattern, the SR of marine migrating or visiting fish species (i.e. MJ, MS and MV) was significantly positively related to salinity in most cases, with the strongest correlations for marine juvenile visitors (MJs; Fig. 4c–e, Table 3).

Concerning vertical distribution, benthic fish species (B) were important contributors to SR_{std} in the sub-basins of higher salinity, but only a few or no fish species belonged to this group in the less saline sub-basins (Fig. S3; Table S7). A similar though less pronounced distribution pattern was also found for demersal fish species (D). Accordingly, the SR of these groups were positively related to salinity in all cases (i.e. for SR_{obs}, SR_{std} and SR_{est}; Fig. 5a, b, Table 3). The SR of demersal–pelagic (DP) fish species varied between sub-basins with a SR_{std} of 6–16 not related to salinity (Fig. 5c, Table 3). A similar picture was found for pelagic fish species (P), in which SR_{std} varied between 5 and 12 across sub-basins (Fig. S3, Table S7) and was not related to salinity (Fig. 5d, Table 3).

The two most common feeding groups, across all sub-basins, were invertivore and piscivore (IP) species, as well as invertivores (I). The third-most represented feeding group was piscivores (Pi), followed by planktivores (PL) and omnivores (O) with lower and often similar SR_{std} (Fig. S4, Table S7). SR_{std} of Pi and IP increased with increasing salinity, and SR_{std} and SR_{est} of I increased with increasing salinity (Fig. S5, Table 3).

Table 3. Simple linear regressions between salinity and SR for the different fish functional attributes (observed (obs), standardized (std) and estimated (est) SR) in Swedish shallow coastal areas. YJ(y) denotes that the response variable was Yeo–Johnson transformed (Yeo and Johnson, 2000). n.s. = not significant.

Response variable			Parameters ($\pm$ SE)	n	R^2	P value
Origin	Marine	obs	$\log_{10}(y) = 1.0\ (\pm 0.1) + 0.03\ (\pm 0.01) \cdot x$	12	0.74	< 0.001
		std	$\log_{10}(y) = 0.9\ (\pm 0.1) + 0.03\ (\pm 0.01) \cdot x$	10	0.71	0.001
		est	$\log_{10}(y) = 1.1\ (\pm 0.1) + 0.03\ (\pm 0.01) \cdot x$	10	0.66	0.003
	Freshwater	obs	$y = 22.0\ (\pm 3.0) - 0.6\ (\pm 0.2) \cdot x$	12	0.39	0.018
		std	$y = 19.0\ (\pm 1.1) - 0.4\ (\pm 0.1) \cdot x$	10	0.84	< 0.001
		est	$y = 25.4\ (\pm 2.5) - 0.5\ (\pm 0.2) \cdot x$	10	0.47	0.017
Habitat preference	CR	obs	$y = 25.3\ (\pm 4.6) + 0.1\ (\pm 0.3) \cdot x$	12	−0.08[a]	n.s. (0.661)
		std	$y = 26.5\ (\pm 2.0) - 0.1\ (\pm 0.1) \cdot x$	10	−0.04[a]	n.s. (0.538)
		est	$y = 32.5\ (\pm 3.0) + 0.1\ (\pm 0.2) \cdot x$	10	−0.10[a]	n.s. (0.665)
	CA	obs	$\log_{10}(y) = 0.4\ (\pm 0.1) + 0.01\ (\pm 0.01) \cdot x$	12	−0.10[a]	n.s. (0.351)
		std	$\log_{10}(y) = 0.5\ (\pm 0.1) + 0.005\ (\pm 0.004) \cdot x$	10	−0.10[a]	n.s. (0.314)
		est	$\log_{10}(y) = 0.5\ (\pm 0.1) + 0.006\ (\pm 0.005) \cdot x$	10	−0.10[a]	n.s. (0.287)
	MJ	obs	$\log_{10}(y) = 0.14\ (\pm 0.13) + 0.04\ (\pm 0.01) \cdot x$	12	0.60	0.002
		std	$\log_{10}(y) = 0.57\ (\pm 0.09) + 0.016\ (\pm 0.004) \cdot x$	6	0.69	0.025
		est	$\log_{10}(y) = 0.43\ (\pm 0.07) + 0.018\ (\pm 0.004) \cdot x$	6	0.84	0.007
	MS	obs	$y = 0.7\ (\pm 0.6) + 0.3\ (\pm 0.04) \cdot x$	12	0.79	< 0.001
		std	$y = 2.4\ (\pm 1.0) + 0.2\ (\pm 0.1) \cdot x$	6	0.67	0.029
		est	$y = 3.0\ (\pm 1.4) + 0.2\ (\pm 0.1) \cdot x$	6	0.55	n.s. (0.056)
	MV	obs	$YJ(y) = -1.0\ (\pm 0.2) + 0.09\ (\pm 0.01) \cdot x$	12	0.81	< 0.001
		std	$y = -3.4\ (\pm 7.1) + 0.7\ (\pm 0.3) \cdot x$	4	0.65	n.s. (0.123)
		est	$y = -6.5\ (\pm 10.8) + 1.3\ (\pm 0.4) \cdot x$	4	0.72	n.s. (0.100)
Vertical distribution	B	obs	$YJ(y) = -1.0\ (\pm 0.2) + 0.09\ (\pm 0.01) \cdot x$	12	0.76	< 0.001
		std	$\log_{10}(y) = 0.3\ (\pm 0.1) + 0.031\ (\pm 0.004) \cdot x$	7	0.91	0.001
		est	$\log_{10}(y) = 0.4\ (\pm 0.1) + 0.039\ (\pm 0.004) \cdot x$	7	0.95	< 0.001
	D	obs	$\log_{10}(y) = 1.0\ (\pm 0.1) + 0.02\ (\pm 0.01) \cdot x$	12	0.44	0.011
		std	$\log_{10}(y) = 0.97\ (\pm 0.04) + 0.013\ (\pm 0.013) \cdot x$	10	0.70	0.002
		est	$\log_{10}(y) = 1.1\ (\pm 0.1) + 0.014\ (\pm 0.004) \cdot x$	10	0.54	0.010
	DP	obs	$y = 13.1\ (\pm 2.1) - 0.1\ (\pm 0.1) \cdot x$	12	−0.10[a]	n.s. (0.679)
		std	$y = 13.9\ (\pm 1.3) - 0.1\ (\pm 0.1) \cdot x$	10	0.16	n.s. (0.136)
		est	$y = 15.5\ (\pm 1.8) - 0.03\ (\pm 0.11) \cdot x$	10	−0.12[a]	n.s. (0.792)
	P	obs	$\log_{10}(y) = 0.8\ (\pm 0.1) + 0.006\ (\pm 0.004) \cdot x$	12	0.07	n.s. (0.203)
		std	$\log_{10}(y) = 0.80\ (\pm 0.05) + 0.004\ (\pm 0.003) \cdot x$	10	0.06	n.s. (0.243)
		est	$\log_{10}(y) = 0.9\ (\pm 0.1) + 0.008\ (\pm 0.004) \cdot x$	10	0.27	n.s. (0.072)
Feeding habit	Pi	obs	$\log_{10}(y) = 0.70\ (\pm 0.06) + 0.018\ (\pm 0.004) \cdot x$	12	0.63	0.001
		std	$\log_{10}(y) = 0.69\ (\pm 0.06) + 0.014\ (\pm 0.004) \cdot x$	10	0.64	0.004
		est	$\log_{10}(y) = 0.76\ (\pm 0.06) + 0.020\ (\pm 0.004) \cdot x$	10	0.77	0.001
	IP	obs	$y = 8.7\ (\pm 1.8) + 0.7\ (\pm 0.1) \cdot x$	12	0.75	< 0.001
		std	$y = 9.9\ (\pm 1.1) + 0.4\ (\pm 0.1) \cdot x$	10	0.81	< 0.001
		est	$y = 10.7\ (\pm 1.9) + 0.8\ (\pm 0.1) \cdot x$	10	0.84	< 0.001
	I	obs	$\log_{10}(y) = 0.8\ (\pm 0.2) + 0.02\ (\pm 0.01) \cdot x$	12	0.16	n.s. (0.108)
		std	$\log_{10}(y) = 0.85\ (\pm 0.04) + 0.015\ (\pm 0.003) \cdot x$	10	0.79	< 0.001
		est	$\log_{10}(y) = 0.93\ (\pm 0.04) + 0.018\ (\pm 0.002) \cdot x$	10	0.86	< 0.001
	PL	obs	$\log_{10}(y) = 0.5\ (\pm 0.1) + 0.007\ (\pm 0.004) \cdot x$	12	0.12	n.s. (0.142)
		std	$\log_{10}(y) = 0.4\ (\pm 0.1) + 0.008\ (\pm 0.003) \cdot x$	10	0.36	0.041
		est	$\log_{10}(y) = 0.5\ (\pm 0.1) + 0.007\ (\pm 0.004) \cdot x$	10	0.20	n.s. (0.108)
	O	obs	$y = 3.6\ (\pm 0.9) - 0.1\ (\pm 0.1) \cdot x$	12	−0.03[a]	n.s. (0.419)
		std	$y = 3.2\ (\pm 0.5) + 0.1\ (\pm 0.04) \cdot x$	8	0.23	n.s. (0.131)
		est	$y = 3.5\ (\pm 0.5) + 0.1\ (\pm 0.04) \cdot x$	8	0.33	n.s. (0.080)

[a] Adjusted R^2 can turn negative for multiple R^2 close to zero.
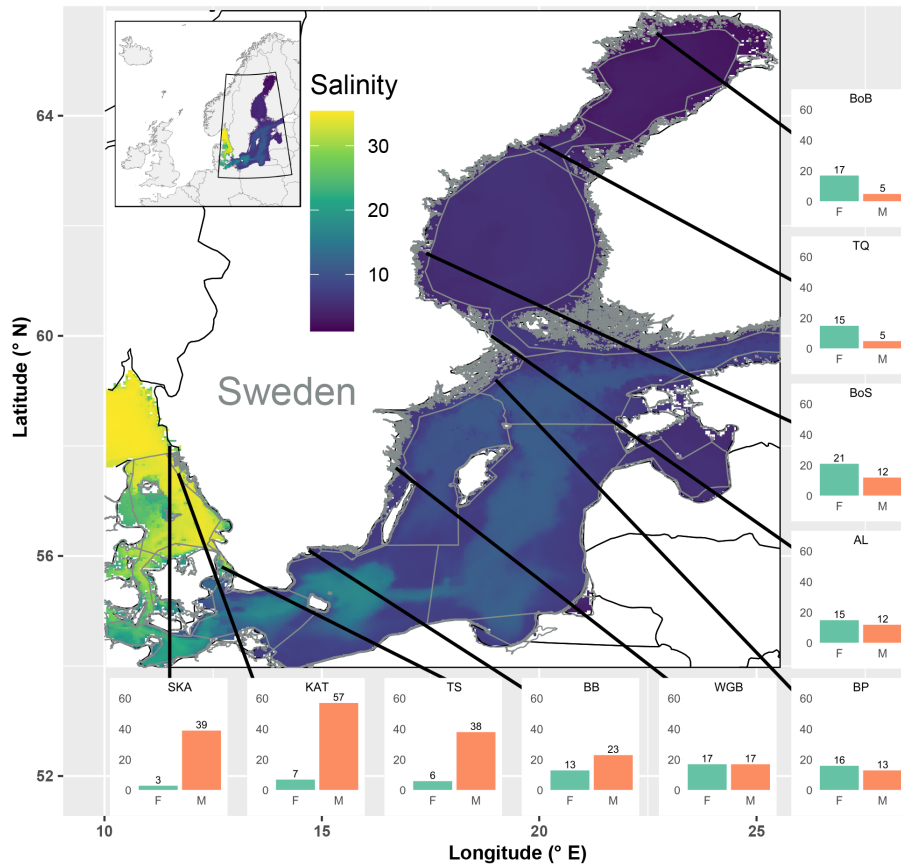
Figure 3. TS2 Map of the study area covering the Baltic Sea and the Skagerrak, colour-coded by mean salinity. Grey lines delineate sub-basins and shallow coastal areas. Bar plots show standardized fish species richness for shallow coastal areas in each of the 10 assessed sub-basins, separately for species of freshwater (F) and marine (M) origin. SR was standardized across sub-basins to similar sample coverage (Table S5). Black lines indicate the positions of the sub-basins, but the exact sampling sites were spread across the shallow areas of each of the sub-basins. SKA: Skagerrak; KAT: Kattegat; TS: the Sound; BB: Bornholm Basin; WGB: western Gotland Basin; BP: northern Baltic Proper; AL: Åland Sea; BoS: Bothnian Sea; TQ: the Quark; and BoB: Bothnian Bay.

4 Discussion

Data from species censuses have been called "probably the most basic data in ecology" as they are widely useful, for example, to define species ranges and biodiversity patterns and support conservation efforts (Gaston and Blackburn, 2000). A limitation for the use of taxonomic inventory data for biodiversity purposes, however, is their completeness, i.e. the fraction of species in a given location that has been sampled (Mora et al., 2008). In this study, the coastal fish taxonomic inventory completeness (IC) varied between 77 % and 98 % for the 10 assessed sub-basins, exceeded 80 % in eight sub-basins, and averaged 86 % (Table 2). The comparatively low IC in the Skagerrak (77 %; Table 2) can be related to, in this context, a rather small sample size for a relatively high number of marine migrating and/or visiting species which are only more rarely present in this sub-basin and hence harder to detect during sampling (Fig. S2, Table S7). The average IC of 86 % in shallow Swedish coastal waters is high compared to

an assessment of marine fish species census data worldwide, in which global IC averaged 79 %, indicating that ca. 21 % of fish species still remained to be described. Marine fish IC exceeded 80 % in less than 2 % of marine areas worldwide, with the highest IC of 92 % found for reef-associated species and the lowest of 56 % for bathydemersal species (Mora et al., 2008). Similarly, a global assessment published in 2012 concluded that ca. 77 % of marine fish were known.

The SR of frequent and very frequent species (i.e. Shannon and Simpson diversities, ShD and SiD) were well described by the sample sizes available to date in the assessed sub-basins, with calculated ("observed") ShD and SiD being similar to both standardized and estimated values (when effects of differing sample sizes are considered; Table S5). This indicates that the remaining uncertainty in fish SR_{obs} is caused by a number of highly rare species, corroborated by a high sample coverage (SC) across sub-basins (Table 2). This is a typical pattern since well-known species are usually common and have large geographical ranges, whereas
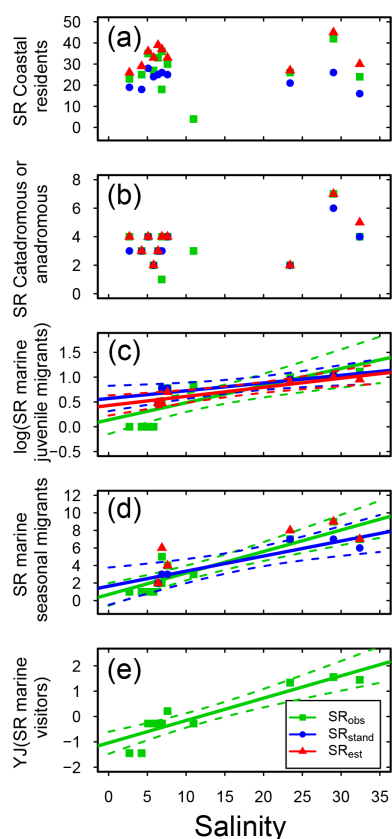
Figure 4. Fish species richness (SR) in relation to salinity in Swedish shallow coastal areas for different types of habitat preference, with **(a)** coastal residents, **(b)** catadromous or anadromous migrants, **(c)** marine juvenile migrants ($\log_{10}$ transformed), **(d)** marine seasonal migrants, and **(e)** marine visitors (Yeo–Johnson transformed). Each plot shows the observed, standardized and estimated SR and, when significant ($P < 0.05$), the linear regression lines (solid) and 95 % confidence intervals (shaded areas surrounded by dashed lines). Please note that data points (Table S7), lines and shaded confidence intervals are overlying each other within the panels in some cases. For regression equations and statistics see Table 3. For marine visitors **(e)**, for clarity following transformation, only the observed SR is shown (no transformation was needed for the standardized and estimates values, and there were no significant relationships; Table 3).

newly discovered species are usually (more) locally rare and geographically concentrated (Appeltans et al., 2012; Mora et al., 2008; Pimm et al., 2014). Specifically, in terms of species numbers and depending on the sub-basin, 1–37 statistically likely existing and highly rare fish species remained undetected in the samplings (Table 2, SR_{est} minus SR_{obs}). However, the high similarity between SR_{est} and total SR_{obs} (i.e. SR_{obs} based on incidence data plus species records from additional data sources; Table 2) suggests that the total observed species lists for fish in Swedish shallow coastal areas were close to complete for all assessed sub-basins, i.e. including the highly rare species undetected in the fish surveys

(i.e. incidence database). This comparison between SR_{est} and total SR_{obs} also independently confirmed/validated that the SR values estimated based on the fish incidence database (SR_{est}) were realistic.

The most recent checklist of Baltic Sea macrospecies, i.e. containing fish species reported across Baltic countries at both shallow and deeper water depths but excluding the Skagerrak, contains 242 fish species (HELCOM, 2020). In our analyses of Swedish shallow coastal areas the total fish SR_{obs} amounted to 144 (i.e. SR_{obs} from incidence data plus additional data sources), also if the Skagerrak is excluded. Comparing the sample-size-corrected estimates of SR in coastal areas (SR_{est}) with the sub-basin-specific SR_{obs} of HELCOM (2020) suggests that ca. 50 %–90 % of all reported Baltic Sea fish species are found in Swedish shallow coastal areas (i.e. $100 \cdot SR_{est} / SR_{obs}$ all countries and depths; Table 2).

Our study reinforces that SR_{obs} is strongly dependent on SC, as relatively rare species are more likely to be missed at lower sample size/sample coverage, and that comparisons of SR_{obs} in species assemblages without accounting for this effect can lead to biased or misleading conclusions (Chao and Chiu, 2016; Colwell and Coddington, 1994; Gotelli and Colwell, 2001; Hill, 1973; Menegotto and Rangel, 2018; Pimm et al., 2014). When sample sizes are not uniform among sites or over time, SR_{obs} values need to be corrected for SC before valid comparisons can be made. However, such methods have so far only rarely been used for coastal and estuarine fish assemblages (Waugh et al., 2019).

Besides the effects of sample size, SR and IC might to some extent also have been differentially influenced by differences in predominating sampling gear across sub-basins. Multi-mesh gill nets dominated in seven of the statistically assessed sub-basins, while trap nets and trawls dominated in the other three (Table S1). As one "sample" in each of these cases represents a different effort due to differences in gear selectivity and sampling approach, this strictly does not allow for direct comparisons (Bergström et al., 2013; Waugh et al., 2019). For example, at the Swedish west coast, gill nets typically sample a wider range of species than fyke nets, which are more selective towards demersal and demersal–pelagic species (Bergström et al., 2013). Merging data from multiple gear types into one analysis may have caused a certain bias in this regard. However, we argue that our approach was feasible given that the gear types used in the different sub-basins are optimized for the locally prevailing conditions, i.e. aiming to sample the existing assemblages as completely as possible (Bergström et al., 2013), as additional data from relevant trawl surveys were also included, and considering the long time horizon of data collection. Further supporting our approach, biodiversity metrics that were standardized against catch size revealed no consistent differences when comparing gill and fyke net samplings at the Swedish west coast (Bergström et al., 2013). Our assumption also appears justified given that SR_{est} was similar to SR_{obs} from incidence
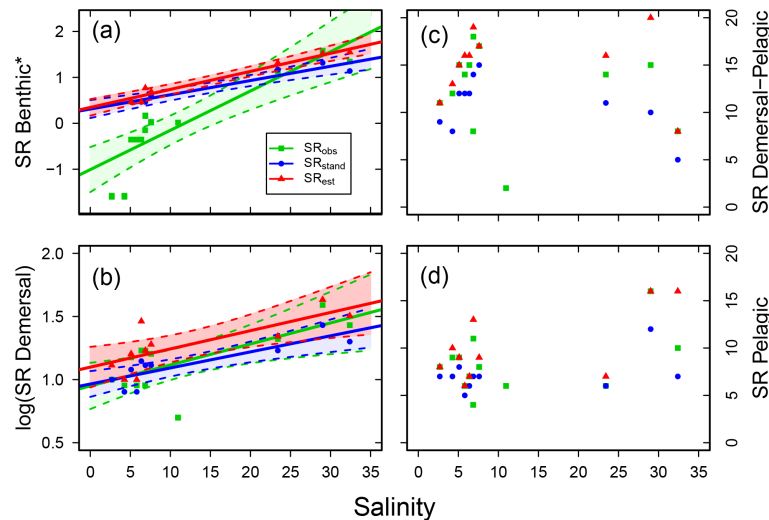
Figure 5. Fish species richness (SR) in relation to salinity in Swedish shallow coastal areas for different types of main vertical distribution, with **(a)** benthic (*Yeo–Johnson transformed for SR_{obs} and $\log_{10}$ transformed for SR_{std} and SR_{est}), **(b)** demersal ($\log_{10}$ transformed), **(c)** demersal–pelagic and **(d)** pelagic fish species. Each plot shows the observed, standardized and estimated SR and, for cases with a significant linear relationship ($P < 0.05$), also the regression line (solid) and 95 % confidence intervals (shaded areas surrounded by dashed lines). Please note that data points (Table S7), lines and shaded confidence intervals are overlying each other within the panels in some cases. For regression equations and statistics see Table 3.

data plus species records from additional data sources (Table 2), giving confidence that the potentially introduced bias due to differing fishing gear and methods did not strongly influence the general patterns and results of this comparative and large-scale statistical analysis.

As anticipated based on earlier Baltic Sea studies on fish (e.g. Hiddink and Coleby, 2012; Ojaveer et al., 2010; Olsson et al., 2012) and other organism groups (e.g. Broman et al., 2019; Zettler et al., 2014), coastal fish SR was positively correlated with salinity (Fig. 2a, Table S6), with fish SR increasing ca. 3-fold together with the ca. 12-fold increase in salinity (Table 2). The clear predominance of marine species in the most saline sub-basins compared to freshwater species in the inner parts of the Baltic Sea (Fig. 3, Table S7) is in agreement with the fact that salinity functions as threshold or "ecological barrier" for the distribution of many freshwater and marine species (Olenin and Leppäkoski, 1999; Vuorinen et al., 2015). It also corroborates patterns reported earlier for fish SR_{obs} in three Baltic sub-basins (Hiddink and Coleby, 2012) and estuaries in general (Whitfield, 2015). The relatively small number of freshwater fish species incidences observed in the higher-salinity sub-basins in our study (Fig. 3) likely stems from sampling close to freshwater tributaries and reflects that many freshwater fish species can stand extended exposure to certain salinity levels ($<$ ca. 9) and tolerate brief exposure to higher salinities ($>$ ca. 15; Peterson and Meador, 1994).

While temperature did not significantly correlate with observed SR, ShD or SiD, it was positively correlated with the standardized and estimated values (Fig. 2d–f, Table S6),

which may indicate a temperature effect on fish biodiversity. In previous studies, temperature has shown positive correlations with SR_{obs} in North Atlantic demersal and benthopelagic fish assemblages (Gislason et al., 2020), as well as with fish SR_{obs} in the coastal Norwegian Skagerrak (Lekve et al., 2002), and in estuaries worldwide (Vasconcelos et al., 2015), all being examples of the often found general pattern that broader-scale SR co-varies with climatic variables such as temperature (Currie et al., 2004). However, given the clear relationship between salinity and the incidences of freshwater vs. marine fish species across the studied sub-basins (Fig. 3), we consider the studied salinity gradient to represent a case to which the "physiological tolerance hypothesis" applies strongly, i.e. that SR in a particular area is limited by the number of species that can tolerate the local salinity conditions (Currie et al., 2004). In accordance, the regression models with salinity alone did not improve by adding temperature as an additional explanatory variable. This conclusion is in agreement with observations from estuaries that fish SR is influenced by the broader distributions and habitat preference patterns of marine and freshwater species that can colonize these areas (Vasconcelos et al., 2015).

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, for example, 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 southwestern 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 (Tables 1, 2; Korpinen et al., 2012; HELCOM, 2018). Hence, no negative relationship between cumulative human pressure and fish SR is expected on the large spatial and temporal scales studied here. This does, however, not contradict that variation in levels of human pressures can have effects on fish concerning other aspects, such as population sizes or growth rates (Olsson, 2019; Bastardie et al., 2021; Reckermann et al., 2022), or could possibly affect SR on smaller spatial scales than those studied here. Evidently, since the rarefaction and 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}) values are based on SC, in which 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, for example, diversity of substrate or habitat-forming macrophytes, which can promote higher 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, a greater habitat complexity with increasing salinity may enhance fish SR, further reinforcing any salinity-induced distributional pattern.

In compiling data we have assumed that salinity changes have been minor during the time period from which samples were obtained compared to the pronounced spatial salinity gradient (observations over 17 to 47 years in the statistically assessed sub-basins, during 1975–2021; Table S2). According to monitoring data from the Baltic Sea, temporal changes in surface salinity varied between an increase of 3 psu in the Kattegat and a decrease of 1 psu in the Bothnian Sea during the past decades (1980–2015; Ammar et al., 2021), which can be considered small compared to the spatial salinity gradient ranging from 3 to 29. Considering temporal patterns in SR and community composition, it was earlier reported that the observed fish SR increased in the Kattegat, Arkona Basin and the central Baltic during 2001–2008 (Hiddink and Coleby, 2012) and that the observed SR of demersal fish increased in the Baltic Proper and the Bothnian Sea during ca. 1971–2013 (Törnroos et al., 2019). First-time observations of known fish species in sub-basins where they were not previously caught have been particularly related to increasing spring temperatures ($+3$–$6\,°C$ during 1980–2015; Ammar et al., 2021). Such potential temporal patterns were not analysed in our study, in which we focused on large-scale spatial patterns and merged the fish incidence data across years to provide as accurate as possible SR estimates at the sub-basin scale.

Along with the changes in species, our study revealed changes in fish SR for different functional groups across the studied salinity gradient. As the functional groups represent differences among species groups in, for example, use of resources or level of mobility, these changes may also lead to differences in coastal ecosystem functioning across the different sub-basins (Elliott et al., 2007; Franco et al., 2008). Migrating fish species are typically of marine origin (here, marine juvenile migrants, marine seasonal migrants or marine visitors) and cannot tolerate low salinity, explaining their predominance at higher salinities (Fig. S2) and being in agreement with known patterns in European estuaries in general (Elliott and Dewailly, 1995; Franco et al., 2008). This pattern, with marine open-sea fish species periodically using coastal areas, may be related to enhanced prey availability, hiding places, and the typically more turbid waters providing protection from predators (Franco et al., 2008). Moreover, the higher SR of migratory fish at higher salinity is likely relevant for the ecological connectivity between ecosystems, i.e. by transport of local "coastal" production to the open sea and vice versa (Franco et al., 2008). It also emphasizes the important role of 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. S2; Berkström et al., 2021).

Further, benthic and demersal fish SR decreased with decreasing salinity (Fig. 5a, b, Table 3), corroborating a previously documented decrease in demersal fish SR_{obs} from the saline Kattegat to the less saline northern Baltic Proper (Pecuchet et al., 2016) and in accordance with the high benthic preference of marine fish species in European estuaries (Elliott and Dewailly, 1995; Franco et al., 2008). This pattern further corresponds with the observed SR of benthic meio- and macrofauna, which are a dominating prey for benthic and demersal fish, also decreasing with decreasing salinity in the Baltic Sea (Broman et al., 2019; Zettler et al., 2014). Taken together, these patterns suggest that 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 and Dewailly, 1995). Also, the higher piscivorous fish SR in the more saline sub-basins (Fig. S5e) and the low omnivorous fish SR which was unrelated to salinity (Fig. S5c) were in agreement with findings from European estuaries (Franco et al., 2008). In summary, the identified 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.

5   Conclusions

Since fish SR and a number of functional attributes changed along the salinity gradient, respective changes in the coastal fish communities may be foreseen if climate change further alters salinity conditions in the Baltic Sea. While the confidence in future salinity projections remains low (HELCOM, 2021), recent ensemble simulations estimate that the two main drivers of climate-related changes in salinity in the Baltic Sea region, increasing river runoff (leading to lower salinity) and sea level rise (leading to higher salinity), approximately compensate for each other and may result in no net salinity changes (Meier et al., 2022). Mean (depth-integrated) observed Baltic Sea salinity did not change during 1982–2016; however, vertical changes were observed with freshening trends of the upper layer down to 40–50 m depth in most sub-basins and increasing salinity below the halocline in the deep layer of some sub-basins (Liblik and Lips, 2019). Hence, if not considering potential phenotypical acclimation or genetic adaptation, an upper-layer freshening would, based on the results from this and earlier studies (e.g. Hiddink and Coleby, 2012; MacKenzie et al., 2007; Pecuchet et al., 2016), likely lead to reduced native fish SR in shallow coastal areas, where more and more marine species are excluded. Further, successful recovery of marine overfished species may become less probable, while certain freshwater fish species may be favoured (MacKenzie et al., 2007; Peterson and Meador, 1994). Indeed, marine fish species were negatively affected by a period of freshened conditions in the Baltic Sea during the ca. 1970s–1990s (Ojaveer and Kalejs, 2005). Benthic fish species, being mostly of marine origin, may be especially vulnerable to freshening in the inner Baltic Sea (i.e. Arkona Basin and inwards) where their proportion in the fish assemblage is already relatively low to date.

Besides salinity changes, fish SR and distribution may also be influenced by other climate-change-related processes, including warming and resulting higher deep-water oxygen consumption rates, or changes in the Baltic Sea circulation (HELCOM, 2021; MacKenzie et al., 2007). Increasing water temperatures have already been linked to increased observed fish SR in the adjacent North Sea (Hiddink and Ter Hofstede, 2008) and in the Kattegat (Hiddink and Coleby, 2012). Further ecosystem-based assessments are needed to obtain realistic predictions of the net effect of such ongoing environmental changes on future fish SR and community composition and on how they may interact with human activities such as fishing patterns, as well as with conservation needs for biodiversity management.

Code and data availability. The data used in this study are publicly available via the SLU Database for Coastal Fish – KUL (https://www.slu.se/en/departments/aquatic-resources1/databases/database-for-coastal-fish-kul/, Swedish University of Agricultural Sciences, 2021a), the SLU Species Observation System (https://www.artportalen.se/ViewSighting/SearchSighting, Swedish University of Agricultural Sciences, 2021b), the SMHI SHARKweb (https://www.smhi.se/en/services/open-data/national-archive-for-oceanographic-data, Swedish Meteorological and Hydrological Institute, 2021), the Fish-Base (https://fishbase.se/search.php, Froese and Pauly, 2021), the EU Copernicus Marine Service Information (https://resources.marine.copernicus.eu/product-detail/BALTICSEA_REANALYSIS_PHY_003_011/INFORMATION, CMEMS, 2021) and the Water Information System Sweden (https://viss.lansstyrelsen.se/Maps.aspx, The Competent Authorities of the Swedish Water Districts et al., 2020). The SLU Trawl Survey data (Department of Aquatic Resources) and the R code used for analysis and plotting are available from Birgit Koehler upon request.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-19-1-2022-supplement.

Author contributions. LB, MK and BK designed the study. LB, ME and MK compiled the data. MK and LB assigned the functional characteristics to the fish species. BK analysed the data. BK and ME visualized the data. All co-authors discussed and validated the results. BK prepared the manuscript with contributions from all co-authors.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgements. We thank two anonymous reviewers for their constructive comments, Malin Werner (SLU) who kindly contributed with the compilation of trawl data, and the many colleagues involved in the Swedish national programme for long-term coastal fish monitoring, data hosted by SLU, which made this study possible.

Financial support. This research has been supported by The Swedish Environmental Protection Agency (grant no. NV-03421-10, WATERS project: Waterbody Assessment Tools for Ecological Reference conditions and status in Sweden). Long-term monitoring of coastal fish in Sweden is mainly financed by the Swedish Agency for Marine and Water Management.

Review statement. This paper was edited by Emilio Marañón and reviewed by two anonymous referees.

References

Ammar, Y., Niiranen, S., Otto, S. A., Möllmann, C., Finsinger, W., and Blenckner, T: The rise of novelty in marine ecosystems: The Baltic Sea case, Glob. Change Biol., 27, 1485–1499, https://doi.org/10.1111/gcb.15503, 2021.

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., and Berta, A.: The magnitude of global marine species diversity, Curr. Biol., 22, 2189–2202, https://doi.org/10.1016/j.cub.2012.09.036, 2012.

Bastardie, F., Brown, E. J., Andonegi, E., Arthur, R., Beukhof, E., Depestele, J., Döring, R., Ritzau Eigaard, O., García-Barón, I., Llope, M., Mendes, H., Piet, G. J., and Reid, D.: A review characterizing 25 ecosystem challenges to be addressed by an ecosystem approach to fisheries management in Europe, Front. Mar. Sci., 7, 629186, https://doi.org/10.3389/fmars.2020.629186, 2021.

Berkström, C., Wennerström, L., and Bergström, U.: Ecological connectivity of the marine protected area network in the Baltic Sea, Kattegat and Skagerrak: Current knowledge and management needs, Ambio, 51, 1485–1503, https://doi.org/10.1007/s13280-021-01684-x, 2021.

Bergström, L., Karlsson, M., and Pihl, L.: Comparison of gill nets and fyke nets for the status assessment of coastal fish communities, Deliverable 3.4-2, WATERS Report no. 2013:7, Havsmiljöinstitutet, Sweden, https://gupea.ub.gu.se/handle/2077/37083 (last access: 21 April 2021), 2013.

Bonsdorff, E.: Zoobenthic diversity-gradients in the Baltic Sea: continuous post-glacial succession in a stressed ecosystem, J. Exp. Mar. Biol. Ecol., 330, 383–391, https://doi.org/10.1016/j.jembe.2005.12.041, 2006.

Broman, E., Raymond, C., Sommer, C., Gunnarsson, J. S., Creer, S., and Nascimento, F. J.: Salinity drives meiofaunal community structure dynamics across the Baltic ecosystem, Mol. Ecol., 28, 3813–3829, https://doi.org/10.1111/mec.15179, 2019.

Byrnes, J. E., Gamfeldt, L., Isbell, F., Lefcheck, J. S., Griffin, J. N., Hector, A., Cardinale, B. J., Hooper, D. U., Dee, L. E., and Emmett Duffy, J.: Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions, Methods Ecol. Evol., 5, 111–124, https://doi.org/10.1111/2041-210X.12143, 2014.

Chao, A. and Chiu, C. H.: Nonparametric estimation and comparison of species richness, in: eLS, edited by: John Wiley and Sons, Ltd, Chichester. https://doi.org/10.1002/9780470015902.a0026329, 2016.

Chao, A. and Jost, L.: Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size, Ecology, 93, 2533–2547, https://doi.org/10.1890/11-1952.1, 2012.

Chao, A. and Jost, L.: Estimating diversity and entropy profiles via discovery rates of new species, Methods Ecol. Evol., 6, 873–882, https://doi.org/10.1111/2041-210X.12349, 2015.

Chao, A., Gotelli, N. J., Hsieh, T., Sander, E. L., Ma, K., Colwell, R. K., and Ellison, A. M.: Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies, Ecol. Monogr., 84, 45–67, https://doi.org/10.1890/13-0133.1, 2014.

Chao, A., Kubota, Y., Zelený, D., Chiu, C. H., Li, C. F., Kusumoto, B., Yasuhara, M., Thorn, S., Wei, C. L., and Costello, M. J.: Quantifying sample completeness and comparing diversities among assemblages, Ecol. Res., 35, 292–314, https://doi.org/10.1111/1440-1703.12102, 2020.

CMEMS: EU Copernicus Marine Service Information [data set], https://resources.marine.copernicus.eu/product-detail/BALTICSEA_REANALYSIS_PHY_003_011/INFORMATION, last access: 10 November 2021.

Colwell, R. and Coddington, J.: Estimating terrestrial biodiversity through extrapolation, Philos. T. R. Soc. B, 345, 101–118, https://doi.org/10.1098/rstb.1994.0091, 1994.

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guégan, J. F., Hawkins, B. A., Kaufman, D. M., Kerr, J. T., Oberdorff, T., and O'Brien, E.: Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness, Ecol. Lett., 7, 1121–1134, https://doi.org/10.1111/j.1461-0248.2004.00671.x, 2004.

Duffy, J. E., Godwin, C. M., and Cardinale, B. J.: Biodiversity effects in the wild are common and as strong as key drivers of productivity, Nature, 549, 261–264, https://doi.org/10.1038/nature23886, 2017.

Duncan, C., Thompson, J. R., and Pettorelli, N.: The quest for a mechanistic understanding of biodiversity–ecosystem services relationships, P. Roy. Soc. B-Biol. Sci., 282, 20151348, https://doi.org/10.1098/rspb.2015.1348, 2015.

Durack, P. J., Wijffels, S. E., and Matear, R. J.: Ocean salinities reveal strong global water cycle intensification during 1950 to 2000, Science, 336, 455–458, https://doi.org/10.1126/science.1212222, 2012.

EC (European Commission): Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU, http://data.europa.eu/eli/dec/2017/848/oj (last access: 10 November 2021), 2017.

Elliott, M. and Dewailly, F.: The structure and components of European estuarine fish assemblages, Neth. J. Aquat. Ecol., 29, 397–417, https://doi.org/10.1007/BF02084239, 1995.

Elliott, M., Whitfield, A. K., Potter, I. C., Blaber, S. J. M., Cyrus, D. P., Nordlie, F. G., and Harrison, T. D.: The guild approach to categorizing estuarine fish assemblages: a global review, Fish Fish., 8, 241–268, https://doi.org/10.1111/j.1467-2679.2007.00253.x, 2007.

Franco, A., Elliott, M., Franzoi, P., and Torricelli, P.: Life strategies of fishes in European estuaries: the functional guild approach, Mar. Ecol. Prog. Ser., 354, 219–228, https://doi.org/10.3354/meps07203, 2008.

Froese, R. and Pauly, D. (Eds.): FishBase, World Wide Web electronic publication, version 06/2021 [data set], https://fishbase.mnhn.fr/search.php, last access: 17 May 2021.

Gaston, K. and Blackburn, T. (Eds.): Pattern and Process in Macroecology, Blackwell Science Ltd, Oxford, England, https://doi.org/10.1002/9780470999592, 2000.

Gislason, H., Collie, J., MacKenzie, B. R., Nielsen, A., Borges, M. d. F., Bottari, T., Chaves, C., Dolgov, A. V., Dulčić, J., and Duplisea, D.: Species richness in North Atlantic fish: Pro-

cess concealed by pattern, Global Ecol. Biogeogr., 29, 842–856, https://doi.org/10.1111/geb.13068, 2020.

Gotelli, N. J. and Colwell, R. K.: Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness, Ecol. Lett., 4, 379–391, https://doi.org/10.1046/j.1461-0248.2001.00230.x, 2001.

Griffiths, J. R., Kadin, M., Nascimento, F. J., Tamelander, T., Törnroos, A., Bonaglia, S., Bonsdorff, E., Brüchert, V., Gårdmark, A., and Järnström, M.: The importance of benthic–pelagic coupling for marine ecosystem functioning in a changing world, Glob. Change Biol., 23, 2179–2196, https://doi.org/10.1111/gcb.13642, 2017.

HELCOM: State of the Baltic Sea – Second HELCOM holistic assessment 2011–2016, Baltic Sea Environment Proceedings no. 155, ISSN 0357-2994, 2018.

HELCOM: HELCOM Checklist 2.0 of Baltic Sea Macrospecies, Baltic Sea Environment Proceedings no. 174, ISSN 0357-2994, 2020.

HELCOM: Climate Change in the Baltic Sea, 2021 Fact Sheet, Baltic Sea Environment Proceedings no. 180, ISSN: 0357-2994, 2021.

Herlemann, D. P., Lundin, D., Andersson, A. F., Labrenz, M., and Jürgens, K.: Phylogenetic signals of salinity and season in bacterial community composition across the salinity gradient of the Baltic Sea, Front. Microbiol., 7, 1883, https://doi.org/10.3389/fmicb.2016.01883, 2016.

Hiddink, J. and Ter Hofstede, R.: Climate induced increases in species richness of marine fishes, Glob. Change Biol., 14, 453–460, https://doi.org/10.1111/j.1365-2486.2007.01518.x, 2008.

Hiddink, J. G. and Coleby, C.: What is the effect of climate change on marine fish biodiversity in an area of low connectivity, the Baltic Sea? Global Ecol. Biogeogr., 21, 637–646, https://doi.org/10.1111/j.1466-8238.2011.00696.x, 2012.

Hill, M. O.: Diversity and evenness: a unifying notation and its consequences, Ecology, 54, 427–432, https://doi.org/10.2307/1934352, 1973.

Hsieh, T., Ma, K., and Chao, A.: iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers): Methods Ecol. Evol., 7, 1451–1456, https://doi.org/10.1111/2041-210X.12613, 2016.

Hu, Y. O., Karlson, B., Charvet, S., and Andersson, A. F.: Diversity of pico-to mesoplankton along the 2000 km salinity gradient of the Baltic Sea, Front. Microbiol., 7, 679, https://doi.org/10.3389/fmicb.2016.00679, 2016.

IPBES: Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, edited by: Brondizio, E. S., Settele, J., Díaz, S., and Ngo, H. T., IPBES secretariat, Bonn, Germany, 1148 pp., https://doi.org/10.5281/zenodo.3831673, 2019.

Johannesson, K. and Andre, C.: Invited review: Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea, Mol. Ecol., 15, 2013–2029, https://doi.org/10.1111/j.1365-294X.2006.02919.x, 2006.

Kaskela, A., Kotilainen, A., Al-Hamdani, Z., Leth, J., and Reker, J.: Seabed geomorphic features in a glaciated shelf of the Baltic Sea, Estuar. Coast. Shelf S., 100, 150–161, https://doi.org/10.1016/j.ecss.2012.01.008, 2012.

Klier, J., Dellwig, O., Leipe, T., Jürgens, K., and Herlemann, D. P.: Benthic bacterial community composition in the oligohaline-marine transition of surface sediments in the Baltic Sea based on rRNA analysis, Front. Microbiol., 9, 236, https://doi.org/10.3389/fmicb.2018.00236, 2018.

Koli, L.: Suomen kalat (translation: "Fishes of Finland"), Werner Söderström Osakeyhtiö, Helsinki, 357 p., 1990 (in Finnish).

Korpinen, S., Meski, L., Andersen, J. H., and Laamanen, M.: Human pressures and their potential impact on the Baltic Sea ecosystem, Ecol. Indic., 15, 105–114, https://doi.org/10.1016/j.ecolind.2011.09.023, 2012.

Kraufvelin, P., Pekcan-Hekim, Z., Bergström, U., Florin, A.-B., Lehikoinen, A., Mattila, J., Arula, T., Briekmane, L., Brown, E. J., and Celmer, Z.: Essential coastal habitats for fish in the Baltic Sea, Estuar. Coast. Shelf Sci., 204, 14–30, https://doi.org/10.1016/j.ecss.2018.02.014, 2018.

Kullander, S. O.: Förteckning över svenska fiskar (that is: List of Swedish fish), https://www.yumpu.com/es/document/read/19727234/forteckning-over-svenska-fiskar-pdf-naturhistoriska-riksmuseet, last access: 10 November 2020 (In Swedish).

Laikre, L., Palm, S., and Ryman, N.: Genetic population structure of fishes: implications for coastal zone management, Ambio, 34, 111–119, https://doi.org/10.1579/0044-7447-34.2.111, 2005.

Lappalainen, A., Shurukhin, A., Alekseev, G., and Rinne, J.: Coastal-Fish Communities along the Northern Coast of the Gulf of Finland, Baltic Sea: Responses to Salinity and Eutrophication, Int. Rev. Hydrobiol., 85, 687–696, https://doi.org/10.1002/1522-2632(200011)85:5/6<687::AID-IROH687>3.0.CO;2-4, 2000.

Lekve, K., Stenseth, N. C., Gjøsæter, J., and Dolédec, S.: Species richness and environmental conditions of fish along the Norwegian Skagerrak coast, ICES J. Mar. Sci., 59, 757–769, https://doi.org/10.1006/jmsc.2002.1247, 2002.

Liblik, T. and Lips, U.: Stratification has strengthened in the Baltic Sea–an analysis of 35 years of observational data, Front. Earth Sci., 7, 174, https://doi.org/10.3389/feart.2019.00174, 2019.

MacKenzie, B. R., Gislason, H., Möllmann, C., and Köster, F. W.: Impact of 21st century climate change on the Baltic Sea fish community and fisheries, Glob. Change Biol., 13, 1348–1367, https://doi.org/10.1111/j.1365-2486.2007.01369.x, 2007.

Meier, H. E. M., Kniebusch, M., Dieterich, C., Gröger, M., Zorita, E., Elmgren, R., Myrberg, K., Ahola, M. P., Bartosova, A., Bonsdorff, E., Börgel, F., Capell, R., Carlén, I., Carlund, T., Carstensen, J., Christensen, O. B., Dierschke, V., Frauen, C., Frederiksen, M., Gaget, E., Galatius, A., Haapala, J. J., Halkka, A., Hugelius, G., Hünicke, B., Jaagus, J., Jüssi, M., Käyhkö, J., Kirchner, N., Kjellström, E., Kulinski, K., Lehmann, A., Lindström, G., May, W., Miller, P. A., Mohrholz, V., Müller-Karulis, B., Pavón-Jordán, D., Quante, M., Reckermann, M., Rutgersson, A., Savchuk, O. P., Stendel, M., Tuomi, L., Viitasalo, M., Weisse, R., and Zhang, W.: Climate change in the Baltic Sea region: a summary, Earth Syst. Dynam., 13, 457–593, https://doi.org/10.5194/esd-13-457-2022, 2022.

Menegotto, A. and Rangel, T. F.: Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness, Nat. Commun., 9, 1–6, https://doi.org/10.1038/s41467-018-07217-7, 2018.

Middelboe, A. L., Sand-Jensen, K., and Brodersen, K.: Patterns of macroalgal distribution in the Kattegat-Baltic region, Phy-

cologia, 36, 208–219, https://doi.org/10.2216/i0031-8884-36-3-208.1, 1997.

Momigliano, P., Denys, G. P., Jokinen, H., and Merilä, J.: *Platichthys solemdali* sp. nov. (Actinopterygii, Pleuronectiformes): a new flounder species from the Baltic Sea, Front. Mar. Sci., 5, 225, https://doi.org/10.3389/fmars.2018.00225, 2018.

Mora, C., Tittensor, D. P., and Myers, R. A.: The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes, P. Roy. Soc. B-Biol. Sci., 275, 149–155, https://doi.org/10.1098/rspb.2007.1315, 2008.

O'Gorman, E. J., Yearsley, J. M., Crowe, T. P., Emmerson, M. C., Jacob, U., and Petchey, O. L.: Loss of functionally unique species may gradually undermine ecosystems, P. Roy. Soc. B-Biol. Sci., 278, 1886–1893, https://doi.org/10.1098/rspb.2010.2036, 2011.

Ojaveer, E. and Kalejs, M.: The impact of climate change on the adaptation of marine fish in the Baltic Sea, ICES J. Mar. Sci., 62, 1492–1500, https://doi.org/10.1016/j.icesjms.2005.08.002, 2005.

Ojaveer, H., Jaanus, A., MacKenzie, B. R., Martin, G., Olenin, S., Radziejewska, T., Telesh, I., Zettler, M. L., and Zaiko, A.: Status of biodiversity in the Baltic Sea, PloS ONE, 5, e12467, https://doi.org/10.1371/journal.pone.0012467, 2010.

Olenin, S. and Leppäkoski, E.: Non-native animals in the Baltic Sea: alteration of benthic habitats in coastal inlets and lagoons, Hydrobiologia, 393, 233–243, https://doi.org/10.1023/A:1003511003766, 1999.

Olli, K., Ptacnik, R., Klais, R., and Tamminen, T.: Phytoplankton species richness along coastal and estuarine salinity continua, Am. Nat., 194, E41–E51, https://orcid.org/0000-0003-2895-1273, 2019.

Olsson, J.: Past and current trends of coastal predatory fish in the Baltic Sea with a focus on perch, pike, and pikeperch, Fishes, 4, 7, https://doi.org/10.3390/fishes4010007, 2019.

Olsson, J., Bergström, L., and Gårdmark, A.: Abiotic drivers of coastal fish community change during four decades in the Baltic Sea, ICES J. Mar. Sci., 69, 961–970, https://doi.org/10.1093/icesjms/fss072, 2012.

Pan, J., Marcoval, M. A., Bazzini, S. M., Vallina, M. V., and Marco, S.: Coastal marine biodiversity: Challenges and threats, in: Marine Ecology in a Changing World, edited by: Arias, A. H. and Menendez, M. C., CRC Press, Boca Raton, United States, 43–67, https://doi.org/10.1201/b16334-3, 2013.

Pecuchet, L., Törnroos, A., and Lindegren, M.: Patterns and drivers of fish community assembly in a large marine ecosystem, Mar. Ecol. Prog. Ser., 546, 239–248, https://doi.org/10.3354/meps11613, 2016.

Pekcan-Hekim, Z., Gårdmark, A., Karlson, A. M., Kauppila, P., Bergenius, M., and Bergström, L.: The role of climate and fisheries on the temporal changes in the Bothnian Bay foodweb, ICES J. Mar. Sci., 73, 1739–1749, https://doi.org/10.1093/icesjms/fsw032, 2016.

Peterson, M. S. and Meador, M. R.: Effects of salinity on freshwater fishes in coastal plain drainages in the southeastern US, Rev. Fish. Sci., 2, 95–121, https://doi.org/10.1080/10641269409388554, 1994.

Pihl, L. and Wennhage, H.: Structure and diversity of fish assemblages on rocky and soft bottom shores on the Swedish west coast, J. Fish Biol., 61, 148–166, https://doi.org/10.1111/j.1095-8649.2002.tb01768.x, 2002.

Pihl, L., Wennhage, H., and Nilsson, S.: Fish assemblage structure in relation to macrophytes and filamentous epiphytes in shallow non-tidal rocky-and soft-bottom habitats, Environ. Biol. Fish., 39, 271–288, https://doi.org/10.1007/BF00005129, 1994.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., and Sexton, J. O.: The biodiversity of species and their rates of extinction, distribution, and protection, Science, 344, 1246752, https://doi.org/10.1126/science.1246752, 2014.

R Core Team: R: A language and environment for statistical computing. Version 4.0.4, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (last access: 27 March 2022), 2021.

Reckermann, M., Omstedt, A., Soomere, T., Aigars, J., Akhtar, N., Bełdowska, M., Bełdowski, J., Cronin, T., Czub, M., Eero, M., Hyytiäinen, K. P., Jalkanen, J.-P., Kiessling, A., Kjellström, E., Kuliński, K., Larsén, X. G., McCrackin, M., Meier, H. E. M., Oberbeckmann, S., Parnell, K., Pons-Seres de Brauwer, C., Poska, A., Saarinen, J., Szymczycha, B., Undeman, E., Wörman, A., and Zorita, E.: Human impacts and their interactions in the Baltic Sea region, Earth Syst. Dynam., 13, 1–80, https://doi.org/10.5194/esd-13-1-2022, 2022.

Reiss, J., Bridle, J. R., Montoya, J. M., and Woodward, G.: Emerging horizons in biodiversity and ecosystem functioning research, Trends Ecol. Evol., 24, 505–514, https://doi.org/10.1016/j.tree.2009.03.018, 2009.

Rooney, N. and McCann, K. S.: Integrating food web diversity, structure and stability, Trends Ecol. Evol., 27, 40–46, 2012.

Russell, G.: Recent evolutionary changes in the algae of the Baltic Sea, Brit. Phycol. J., 20, 87–104, https://doi.org/10.1080/00071618500650111, 1985.

Schubert, H., Feuerpfeil, P., Marquardt, R., Telesh, I., and Skarlato, S.: Macroalgal diversity along the Baltic Sea salinity gradient challenges Remane's species-minimum concept, Mar. Pollut. Bull., 62, 1948–1956, https://doi.org/10.1016/j.marpolbul.2011.06.033, 2011.

Soukup, P.R., Näslund, J., Höjesjö, J., and Boukal, D.S.: From individuals to communities: Habitat complexity affects all levels of organization in aquatic environments, Wiley Interdiscip. Rev. Water, 9, e1575, https://doi.org/10.1002/wat2.1575, 2021.

Swedish Meteorological and Hydrological Institute: National archive for oceanographic and marine biological data, SHARKweb [data set], https://www.smhi.se/en/services/open-data/national-archive-for-oceanographic-data, last access: 27 April 2021.

Swedish University of Agricultural Sciences: Database for Coastal Fish – KUL [data set], https://www.slu.se/en/departments/aquatic-resources1/databases/database-for-coastal-fish-kul/ (last access: 27 April 2021), 2021a.

Swedish University of Agricultural Sciences: Species Observation System [data set], https://www.artportalen.se/ViewSighting/SearchSighting (last access: 27 April 2021), 2021b.

The Competent Authorities of the Swedish Water Districts, the County Administrative Boards, and the Swedish Agency for Marine and Water Management: Water Information System Sweden [data set], https://viss.lansstyrelsen.se/Maps.aspx, last access: 10 November 2020.

Thorman, S.: Seasonal colonisation and effects of salinity and temperature on species richness and abundance of fish

of some brackish and estuarine shallow waters in Sweden, Ecography, 9, 126–132, https://doi.org/10.1111/j.1600-0587.1986.tb01201.x, 1986.

Törnroos, A., Pecuchet, L., Olsson, J., Gårdmark, A., Blomqvist, M., Lindegren, M., and Bonsdorff, E.: Four decades of functional community change reveals gradual trends and low interlinkage across trophic groups in a large marine ecosystem, Glob. Change Biol., 25, 1235–1246, https://doi.org/10.1111/gcb.14552, 2019.

Vasconcelos, R. P., Henriques, S., França, S., Pasquaud, S., Cardoso, I., Laborde, M., and Cabral, H. N.: Global patterns and predictors of fish species richness in estuaries, J. Anim. Ecol., 84, 1331–1341, https://doi.org/10.1111/1365-2656.12372, 2015.

Vuorinen, I., Hänninen, J., Rajasilta, M., Laine, P., Eklund, J., Montesino-Pouzols, F., Corona, F., Junker, K., Meier, H. M., and Dippner, J. W.: Scenario simulations of future salinity and ecological consequences in the Baltic Sea and adjacent North Sea areas–implications for environmental monitoring, Ecol. Indic., 50, 196–205, https://doi.org/10.1016/j.ecolind.2014.10.019, 2015.

Waugh, A., Elliott, M., and Franco, A.: Debunking paradigms in estuarine fish species richness, Mar. Ecol. Prog. Ser., 613, 125–138, https://doi.org/10.3354/meps12883, 2019.

Whitfield, A.: Why are there so few freshwater fish species in most estuaries?, J. Fish Biol., 86, 1227–1250, https://doi.org/10.1111/jfb.12641, 2015.

Wikström, A. and Magnusson, M.: Provfiske med nordiska kustöversiktsgarn utmed två kustområden I Hanöbukten: Pukavik/Elleholm och Skräbeåns mynning/Valjeviken, Marine Monitoring AB, ISBN: 978-91-86461-01-0, 2009 (in Swedish).

Yeo, I. K. and Johnson, R. A.: A new family of power transformations to improve normality or symmetry, Biometrika, 87, 954–959, https://doi.org/10.1093/biomet/87.4.954, 2000.

Zettler, M. L., Karlsson, A., Kontula, T., Gruszka, P., Laine, A. O., Herkül, K., Schiele, K. S., Maximov, A., and Haldin, J.: Biodiversity gradient in the Baltic Sea: a comprehensive inventory of macrozoobenthos data, Helgoland Mar. Res., 68, 49–57, https://doi.org/10.1007/s10152-013-0368-x, 2014.

Remarks from the typesetter

TS1   Please give an explanation of why this needs to be changed. We have to ask the handling editor for approval. Thanks.

TS2   Please confirm the cut figure.