ICES Journal of Marine Science, 2025, Vol. 82, Issue 4, fsaf044
https://doi.org/10.1093/icesjms/fsaf044
Received: 8 November 2024; **revised:** 10 March 2025; **accepted:** 11 March 2025
Original Article

ICES CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Using the Newcomb–Benford law to detect species misreporting in mixed pelagic catches

Eros Quesada *, Massimiliano Cardinale, Nuno Prista, Valerio Bartolino, Patrik Börjesson, Mikaela Bergenius Nord, Niklas Larson, Annelie Hilvarsson, Katja Ringdahl

Institute of Marine Research, Department of Aquatic Resources (SLU Aqua), Swedish University of Agricultural Sciences, Turistgatan 5, 453 30 Lysekil, Sweden

*Corresponding author. Institute of Marine Research, Department of Aquatic Resources (SLU Aqua), Swedish University of Agricultural Sciences, Turistgatan 5, 453 30 Lysekil, Sweden. E-mail: eros.quesada@slu.se

Abstract

Modern stock assessment models used to provide management advice on sustainable catches rely on unbiased catch data. Distortion of this data, intentional or not, may increase the uncertainty in the stock perception, jeopardize the assessment of marine resources, and compromise their sustainable management with negative ecological and socio-economic effects. In this study, we apply an analysis of anomalous numbers based on the Newcomb–Benford law (NBL) to test for fisheries catch misreporting. We focus on the Swedish small pelagic fisheries targeting herring and sprat in the Baltic Sea, which are known to be highly problematic due to the pronounced mixing of the two species in their catches and the existence of potential incentives for misreporting. The analyses also include fishery-independent data from international scientific surveys, which are used as standards for the interpretation of the anomalies in the commercial catch data. We demonstrate that data from two Baltic fishery independent surveys conformed to the NBL, while Swedish commercial catch data recorded at sea (logbooks) and onshore (landing declarations) did not, indicating inaccurate reporting of commercial catches. While non-conformity to the NBL may not be considered as proof of misreporting, and to determine the intentionality of misreporting, if any, goes beyond the scope of the paper, we discuss the possible reasons for the observed deviations from the model and recommend the application of this method for quality control of fishery data. Further research (i.e. testing new tools both for detection and estimation of misreporting) should be carried on this fishery with the aim of improving the accuracy of the reported catches. Furthermore, we open the discussion to whether the management should rely on less accurate but more spatially resolved or more accurate but spatially unresolved commercial data. The application of the NBL presented in this study can be readily implemented to other stocks and fishery as a supporting tool to investigate potential misreporting and contribute to improve our understanding of self-reported fisheries data.

Keywords: misreporting; Newcomb–Benford law; small pelagic; baltic; stock assessment

Introduction

Fisheries management worldwide is generally based on stock assessment models that rely on information from fisheries catches (Hilborn and Walters 2013). Catch data of good quality are important for the evaluation of stock status, and incomplete or distorted catch information likely bias the estimates of stock assessment models and might mislead management decisions (Clarke et al. 2006, Pauly and Zeller 2016). Most modern stock assessment methods assume that the catch is a random variable, therefore subject to at least observation error, but require the catches to be nearly unbiased (Methot et al. 2020). For processes highly dispersed in space and time, such as commercial fishing, control or scientific observation of all fishing activities is financially unsustainable (Blank and Gavin 2009). One approach to complement commercial fishing data involves combining a mosaic of a few expensive, high-quality samples collected by scientists in specific areas and times (also called fishery-independent data) with a large amount of data obtained directly from the fishing industry (also called fishery-dependent data). Fishery-dependent data are generally characterized as having a good spatio-temporal coverage (Dennis et al. 2015), but of questionable quality (Watson and Pauly 2001, Pitcher et al. 2002, Tesfamichael and Pitcher 2007) and can be distorted by misreporting (i.e. the inaccurate transmission, intentional or not, of catch information). Misreporting may occur in a fishery because commercial fishing is an economic activity subjected to numerous regulations and reporting requirements, in which the eventual manipulation of reported data could lead both to higher revenues or the avoidance of high fines (Watson and Pauly 2001, Le Gallic and Cox 2006, Blank and Gavin 2009). There are also intrinsic challenges in characterizing fisheries catches such as in situations where catches are large and the weight of various species in a mixed fishery needs to be individually reported (Thiessen and Blasius 2012).

When validation of data is not possible due to the lack of independently verified information, one has to face the apparent paradox of searching for distortion in the data, using the potentially distorted data. Unsupervised anomaly detection techniques have been developed to analyze these situations. These techniques track a common behavior shared by most of the observations and use it as a reference to identify those records that deviate from it as outliers, and thus can be considered potentially anomalous (Bolton and Hand 2002). A similar approach is to use theoretical models to which intrinsic features of the observations should conform to highlight possible anomalies in the data. The Newcomb–Benford law (NBL) is a mathematical model that serves this scope. Against the popular belief that numbers extracted randomly from nature are equally likely to have as leading digit any number from 1 to 9, the NBL predicts that they are more likely to have a small number as the leading digit (Raimi 1969, Graham et al. 2009). NBL predicts that leading digits from 1 to 9 will occur in empirical data with expected proportions and that these proportions can be calculated using logarithmic functions. The patterns described by the NBL were first observed by the American astronomer and mathematician Simon Newcomb (1835–1909, Newcomb 1881) and then documented by the physicist Dr Frank Benford (1938) in a surprising amount of different natural phenomena (e.g. the specific heat of chemical compounds, the area of rivers, etc.).

Many authors from different backgrounds attempted to explain the ubiquity of NBL in natural systems (Berger et al. 2009). Natural populations' consistency to the NBL has been explained through exponential growth, implying that populations spend more time in regions of their exponential increase characterized by a low leading digit (Ross 2011). Scott and Fasli (2011), Rodriguez (2004), and later Fang and Chen (2020) showed that log-normally distributed data have a good fit to the NBL when the underlying distribution has a sufficiently high shape parameter (i.e. $\sigma \geq 1.2$). Data on the individuals caught from natural populations (i.e. catch data, as the one treated in this work) are expected to be conform to the NBL as they typically distribute according to a log-normal distribution (Hoyle et al. 2024).

In the modern literature, this model has been popularized as an auditing tool (e.g. Nigrini 2012). The expected frequencies of records starting with digits from 1 to 9 according to NBL are used as a null model to test empirical data and search for deviations that could be related to external sources of alteration such as, in the case of fisheries, catch misreporting.

The NBL has been used in 2002 studies (gathered on the website https://www.benfordonline.net/, Berger et al. 2009) from different disciplines. In fishery science, it has not been used extensively as an auditing tool, but it has been used in a small number of applications (i.e. Graham et al. 2009, Tsagbey et al. 2015, 2017, Noleto-Filho et al. 2022). For example, Graham et al. (2009) used NBL to investigate the outputs from a Canadian lobster fishery in two areas subjected to different conservation measures. The analysis showed that in the standard fishing areas, data did not conform to the NBL, while conformity was found in the highly regulated area. Tsagbey et al. (2017) found departures from the NBL in a Ghanaian fishery for tunas. Noleto-Filho et al. (2022) estimated that 20% of fish and invertebrate landings in the Brazilian small-scale fisheries were potentially biased.

Misreporting is expected to occur in this fishery for biological, methodological, and economic reasons. First, the two species are often caught together and, as they are morphologically similar, it is difficult to separate them for non-experts. Second, they are often caught in large quantities, which makes both the estimation of the species composition by the fishers and controls of the declared species proportions by authorities on board difficult (Patterson 1998, Bray 2001). During fishing operations, the catch is pumped on board and into tanks. The species composition is estimated by the fishers visually or through the sorting of several catch samples. The total amount caught of each species is sometimes derived using the volume in the tank and the (approximate) density of the catch. Landing is an opportunity for verifying the quantities and the species composition, but catches from different fish events may be stored together and mixed to varying degrees, making an accurate estimation challenging and leaving room for misreporting (Hentati Sundberg et al. 2014). Third, the two species are subject to Total Allowable Catches (a factor known to constitute a strong incentive for fishers to misreport, Petterson 1998), which may vary for the two species and for fleets belonging to different countries fishing on them in the same area. Finally, the catch comprising both species may be destined to industrial purposes (e.g. production of fish products) leading to less incentive for fishers to have an accurate species identification.

Suspicions of misreporting in this mixed fishery have been raised and discussed for the last 20 years, both in the scientific (Sjöstrand 2000, ICES 2015, 2016, 2017, 2018, 2019, 2020, 2021b), the divulgative (*sensu* scientific, but not peer-reviewed, FishSec 2019, Berkow 2021) and the fisheries community (Svensson 2019). Hentati Sundberg et al. (2014) showed possible discrepancies both in the total quantity reported and in the species composition in the historic catch data from the Swedish pelagic fishery (1996–2009) for both species. They estimated that in some time period (2002–2007) and fleet segments about half of the catch (57% of sprat and 48% for herring) was possibly misreported. ICES has progressively recognized the importance of possible misreporting in the mixed fishery and declared in 2021 that

> Species misreporting of herring has occurred in the past (Hentati Sundberg et al. 2014) and there are again indications of sprat being misreported as herring. This has not been quantified but may affect the quality of the assessment. (ICES 2019, 2020, 2021a).

The aim of this work is to detect possible species misreporting in the small pelagic fishery targeting sprat and herring in the central Baltic. Estimates of misreporting in the pelagic fisheries from the Baltic Sea would be highly relevant for a correct evaluation of the herring and sprat stocks (ICES 2021a). However, a correction is not trivial as misreporting is expected to have been changing in time and space and among different fleet components. In this study, we use the NBL for the first time for a fishery in the Northeast Atlantic. Our work is also particularly novel as it is based on a comparative analysis of scientific and commercial fisheries data. First, the NBL model is validated on scientific survey catch data, which are collected with the highest possible standards. Once demonstrated that the scientific catches conform to the NBL, the model is used to test the fishery catch data.

Materials and methods

Study area and subject

The Baltic Sea is a semi-enclosed sea (Ducrotoy and Elliott 2008). Semi-enclosed basins like the Baltic Sea have been shown to be particularly susceptible to anthropogenic pressures (e.g. the Black Sea, Tokarev and Shulman 2007). Of the ~100 fish species recorded in the Baltic (McKenzie et al. 2007), three (i.e. herring, sprat, and cod) constitute 90% of the overall catches and feed and/or contribute to the economy of the 16 million people, by nine countries, living along Baltic coastal waters (Ojaveer et al. 2010, Burns and Stöhr 2011). Among the Baltic species, herring and sprat historically dominate the biomass available in the basin (Ojaveer et al. 2010). The small pelagic fishery targeting these two species

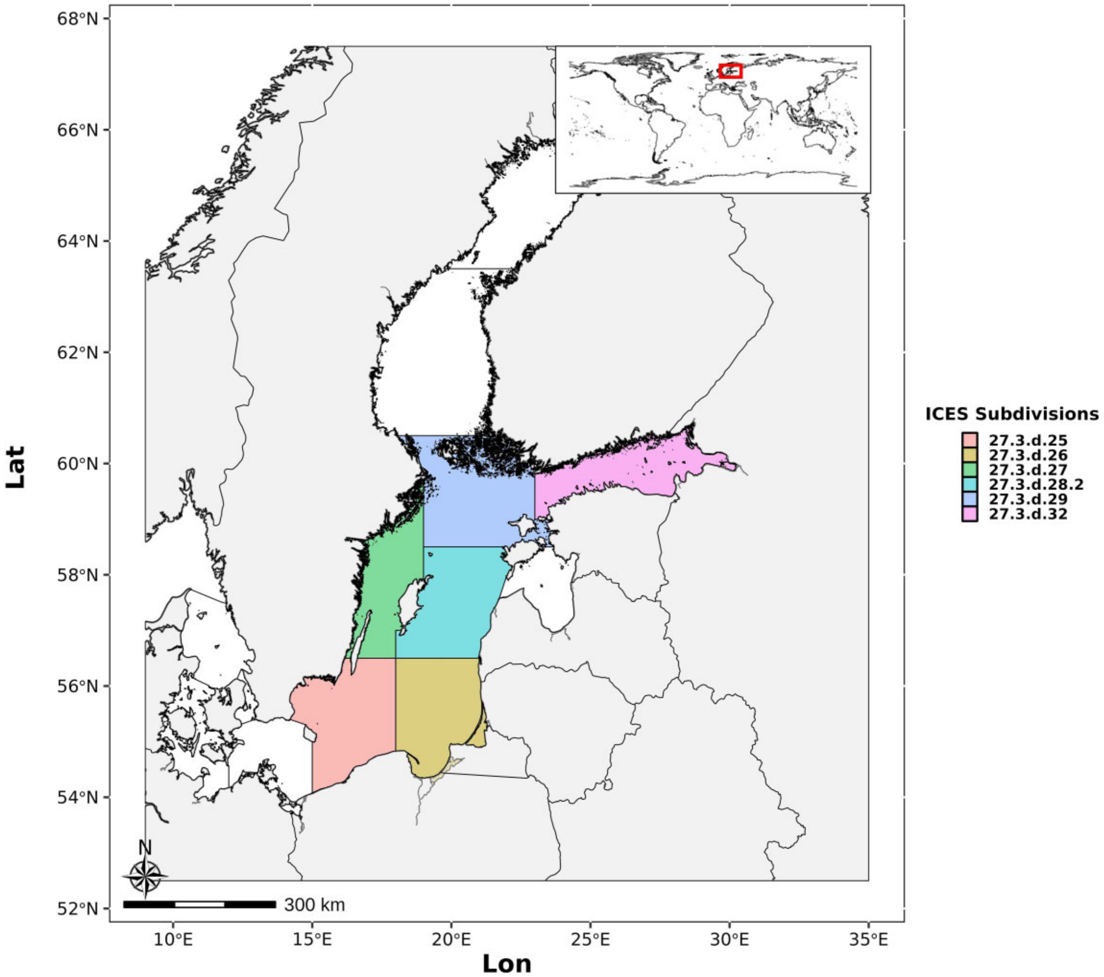

Figure 1. A map illustrating the Baltic Sea and the area analyzed in this study. The latter is filled with colors in order to distinguish the ICES Subdivisions: Southern Central Baltic—West (Subdivision 27.3.d.25), Southern Central Baltic—East (Subdivision 27.3.d.26), West of Gotland (Subdivision 27.3.d.27), East of Gotland (Open Sea) (27.3.d.28.2), Archipelago Sea (Subdivision 27.3.d.29), and Gulf of Finland (Subdivision 27.3.d.32).

represents the largest and most economically important commercial fishery in the Baltic Sea (Arrhenius and Hansson 1993, ICES 2021a).

This study concentrates on Central Baltic herring and sprat stock inhabiting the Baltic in FAO Subdivision Southern Central Baltic—West (Subdivision 27.3.d.25), Southern Central Baltic—East (Subdivision 27.3.d.26), West of Gotland (Subdivision 27.3.d.27), East of Gotland (Open Sea) (27.3.d.28.2), Archipelago Sea (Subdivision 27.3.d.29), and Gulf of Finland (Subdivision 27.3.d.32) (Fig. 1).

Data

The dataset used in this study covers a 23-year timespan (1999–2021) for both commercial and survey data (Table S3).

Data on scientific surveys

Fisheries-dependent data of abundance of sprat and herring were derived from two surveys: the Baltic International Trawl Surveys (BITS, ICES 2014) and the Baltic International Acoustic Survey (BIAS).

BITS is a demersal survey carried out during the first and fourth quarters in ICES subdivisions 22–28 (ICES 2014). For BITS, we used data collected by all available countries between 1999 and 2020. To compare the output of the application of the NBL to commercial and survey data, BITS hauls made using International Standard Trawl for Baltic Demersal Survey TV-3#930TVL gear (TVL, ICES 2014) were used. TVL gear was selected since this gear closely mimics the commercially used Otter and Paired Bottom Trawls, and thus can be regarded as the best available control for conformity of small pelagic trawl catches with the NBL.

BIAS is an international pelagic survey carried out during the third quarter (October) and across ICES subdivisions 25–29 (Larson 2020). In this survey, fish biomass is estimated using hydroacoustic methods. Trawling and biological sampling are carried out to divide the acoustic data into species, lengths, and ages. This analysis focuses on the part of the BIAS data resulting from these hauls and that were collected by Sweden between 2001 and 2021. In the case of BIAS, we used data belonging to pelagic gears (midwater otter trawl, indicated with code "321" hereafter).

Data on Swedish commercial fishery catches

Commercial fishery data both in the forms of logbooks and landing declarations from Swedish vessels were used. These data result from reports by the fishers, at sea and at landing site, respectively, carried out under the legal framework of the EU control regulation. The data contain references in time and space, information on the effort (vessel and gear features), and the species caught (quantities, taxonomy, contribution of the species to the catch, among the others). In Sweden, there are two sources of official catch data: logbook data, which are the estimated catches recorded on a haul-by-haul basis, and landing declarations, which represent the landed weight per species, area, and gear, measured at the end of the fishing trip. Being reported at sea, logbooks have a higher spatio-temporal resolution with expected lower accuracy, while the landing declaration, which is based on weights measured at the landing site, has a lower spatio-temporal resolution, but it is expected to have a higher accuracy. Note that while we will refer to this latter data as "landing declaration," the data reported by the fishermen at the landing site (i.e. the landing declaration) were not available to our study in the original format they were collected. Rather, a transformed dataset, with landed weights corrected for onboard processing and disaggregated by area (based on logbooks) by the Swedish authorities, was available. The procedure used by the Swedish authorities includes the assignation of the weight declared by the fishermen to the areas reported in the logbook and proportionally to the catch the fishermen declared in the latter document. Furthermore, species-specific conversion factors taking account of the presentation of the catch (e.g. gutted or not) are used to complete the conversion. The presence of conversion factors makes it difficult to rebuild the original declared catch via aggregation, except for those landings that are landed whole, with no correction factor applied. Although not representing the original landing dataset as reported directly by the fishermen, this dataset can be considered as an example of data used in the stock assessment context to input commercial catches. For this reason, it can be used as an example of commercial catch data set used as input in the stock assessment and hence suitable for the application of the methodology presented here.

The commercial database in the form of a logbook is considered particularly relevant as this is the one currently representing the fishery-dependent data provided by Sweden to the stock assessment of the two stocks given its higher resolution (i.e. haul level). On the other hand, the commercial database in the form of a landing declaration was used in this analysis as it is considered a more accurate indication of the quantities caught. For this study, only data from pelagic and bottom trawlers that landed their catch in Sweden or Denmark in the time frame 1999–2021 were considered. The study does not consider differences in landings below or above minimum landing size.

Anomaly detection with the Newcomb–Benford law

The general significant digit law [Equation 1, Nigrini (2012), adapted from Hill (1995)] describes the expected relative frequency of any number in the interval (1–9) at different digit (or combination of digits) position in a number, according to the NBL:

$$Pr\left(D_1 = d_1, \ldots, D_k = d_k\right) = log_{10}\left[1 + \left(\frac{1}{\sum_{i=1}^{k} d_i \cdot 10^{k-i}}\right)\right]. \quad (1)$$

For all positive integers k, all $d_1 \in \{1, 2, \ldots, 9\}$ and all $d_j \in \{1, 2, \ldots, 9\}$, with $j = 2, \ldots, k$. Equation (1) is a generalization of the following more specific formulas, predicting the probability of numbers from 1 to 9, at first (Equation 2), second (Equation 3), and first two digits (Equation 4) (with other possible configurations) (Nigrini 2012):

$$Pr\left(D_1 = d_1\right) = \log_{10}\left(1 + \frac{1}{d_1}\right) \text{ with } d_1 \in \{1, 2, \ldots, 9\}, \quad (2)$$

$$Pr\left(D_2 = d_2\right) = \sum_{d_1=1}^{9} \log_{10}\left(1 + \frac{1}{d_1 \cdot d_2}\right) \text{ with } d_2 \in \{0, 1, \ldots, 9\}, \quad (3)$$

$$Pr\left(D_1 D_2 = d_1 d_2\right) = \log_{10}\left(1 + \frac{1}{d_1 \cdot d_2}\right) \text{ with } d_1 d_2 \in \{10, 11, \ldots, 99\}. \quad (4)$$

The Equation (2) predicts the behavior of numbers at the first digit under the NBL model and serves as reference for the first-digit test (F1T). The Equation (3) is the reference for the second-digit test (F2T), while the Equation (4) is known as the first-two-digit test (F12T). F12T will be considered the main test for this work. We report the F1T results as well for historical reasons but highlight to the reader that this test may suffer from false negatives. The presence of false negatives is due to the fact that at F1T a set of numbers may resemble NBL like even if the mantissas (i.e. the part of a logarithm after the decimal point) are not evenly distributed, and an even distribution of mantissas is a mathematical basis for the NBL model (see Nigrini 2012 for a formal explanation). Data on commercial Swedish catches were tested against the NBL model to detect the possible presence of anomalies.

In this study, we used the mean absolute deviation (MAD) test to assess if the data conform to the NBL (Nigrini 2012). We also report the χ^2 goodness of fit test results for historical reasons (for details on the test, see Sheskin 2003), but this test has been found to exhibit an excess of power problem (i.e. may lead to false positives when a high number of records, usually ~5000, is provided and requires equal sample sizes to compare their results across different databases, see Nigrini 2012).

The MAD statistic corresponds to the mean discrepancy between the values observed for each digit and those expected under the NBL model, across all categories (da Silva Azevedo et al. 2021, Equation 5 re-adapted from Nigrini 2012):

$$\text{MAD} = \frac{\sum_{i=1}^{k} \left|O_i - E_i\right|}{k}, \quad (5)$$

where O_i represents the observed frequency at the ith-category, E_i represents the expected (or theoretical) frequency at the ith-category, while k is the number of categories. Similar to the χ^2 goodness of fit test, the k parameter is given in the case of the NBL by the test that is carried out on ($k_{F1T} = 9$, $k_{F12T} = 90$) (Nigrini 2012, da Silva Azevedo et al. 2021). MAD test does not consider the number of records and thus does not exhibit excess of power problem. MAD is therefore a suitable tool for testing the consistency to NBL for large datasets and comparing datasets of different magnitude (Nigrini 2012). The advantage of a sensitivity independent from the number of records comes with the cost that the critical threshold to judge the conformity of a dataset to NBL is expert-based (da Silva Azevedo et al. 2021). Nigrini (2012) reports cut-off scores that can be used as a reference for this purpose. Following Nigrini (2012), we used the threshold levels presented in Table 1.

Table 1. Critical cut-off scores (readapted from Nigrini 2012).

F1D	F2D	Conformity
0.000–0.006	0.0000–0.0012	Close conformity
0.006–0.012	0.0012–0.0018	Acceptable conformity
0.012–0.015	0.0018–0.0022	Marginally acceptable conformity
>0.015	>0.0022	Non-conformity

Data handling and analyses were performed in R (R version 4.2.1) and using the *benford.analysis* R package (version 0.1.5, Cinelli 2018). The analysis was carried out on the commercial and survey datasets. For the commercial data, we proceeded further with the analysis of all species caught by all trawlers. Afterward, we considered the small pelagics caught by bottom trawlers (merging catches from paired bottom trawlers and otter bottom trawlers) and those caught by mid-water trawlers (paired mid-water trawlers and otter mid-water trawlers). In this part of the analysis, we used records involving herring and sprat catches to maximize the number of records available to the test with the NBL. Similarly, for survey data, we based the analysis first on all the species and secondly on small pelagic only (i.e. sprat and herring only). To produce a basis for comparison for commercial data, we restricted the analysis on survey data collected with comparable gears (i.e. catches of commercial bottom trawlers can be compared with surveys using "TVL" gear and catches from commercial pelagic trawlers can be compared with those of scientific gear used in BIAS). The number of orders of magnitude the dataset is covering may have a role in its consistency to the NBL and hence should be taken into account for this type of comparison. Fewster (2009) indicates that "*data from any distribution will tend to be 'Benford', as long as the distribution spans several integers on the log10 scale—several orders of magnitude on the original scale—and as long as the distribution is reasonably smooth*." Survey data may be regarded as datasets spanning over a higher number of orders of magnitude, as they do not target aggregation of fishes. However, in our case, for the (unfiltered) scientific BITS and BIAS data and the commercial logbook and landing declaration data, the orders of magnitude observed are, respectively, 8, 7, 7, and 7 (Fig. S1). When filtered for herring and sprat, the orders of magnitude become 8, 7, 6, and 7 (Fig. S2). The order of magnitude of the datasets relative to small pelagic fish and filtered by specific gear is 6 for logbook pelagic trawlers and 6 for logbook bottom trawlers, 7 for landing declaration pelagic trawlers and 5 for landing declaration bottom trawlers, and 8 for BITS selected bottom trawlers and 7 for BIAS pelagic trawlers (Table S4, Fig. S3 for bottom gears, Fig. S4 for pelagic gears). Hence, we consider the datasets comparable and the differences in the orders of magnitude negligible.

We used the "*benford.analysis:: get.Suspects*" function to retrieve the observations belonging to the digit groups ($n = 5$) most inconsistent with the NBL (Noleto-Filho et al. 2022). This was done on both data types with the exclusion of those strata having <1000 records. In this analysis with "record," we indicated the information relative to the quantity reported for one species in one haul (in the case of commercial logbook or survey data) or in one landing event split by fishing date, ICES Rectangle (according to the logbook proportions, in the case of the landing declaration, see the "Data on Swedish commercial fishery catches" section for details).

Results

Scientific data performance against the NBL

The scientific BITS data

The survey BITS data ($n = 34\,344$) showed "close conformity" to the NBL at the first digit and "close conformity" at the first two digits. The survey BITS data, including herring and sprat species ($n = 9885$ records), showed "close conformity" to the NBL at the first digit (F1T, MAD = 0.00204) and "close conformity" at the first two digits (F12T, MAD = 0.0009). The results were consistent with the analyses using only the data of small pelagics caught with the TVL (Figs 2 and 3, Table 2 and Table S5 for details).

The scientific BIAS data

When all the species caught with the available gear are considered, the performance of BIAS data ($n = 2974$) was weaker (resulting in "acceptable conformity" at F1T and "acceptable conformity" at F12T). The BIAS data consist of a relatively small dataset, when compared to the others, and the observed frequency of digits is less clearly matching the frequencies expected for those digits under the NBL model. Nevertheless, this dataset falls within the MAD conformity limits with the NBL model. Comparable results to the former cases were observed when only small pelagic catches were analyzed for BIAS (Figs 2 and 3, Tables 2 and S5 for details).

Commercial data performance against the NBL

The commercial logbook data

When commercial logbook catches for all the species having more than 1000 records are considered, the data ($n = 291\,672$) have an "acceptable conformity" at F1T, but were found to have non-conformity to the model at F12T (Fig. 4, Tables 2 and S5). Catch data on herring and sprat ($n = 119\,525$) exhibited "non-conformity" to the NBL at the first digit (F1T, MAD = 0.017) and "non-conformity" at the first two digits (F12T, MAD = 0.01203) (Figs 3 and 4, Tables 2 and S5).

Around 45.12% of the records ($n = 53\,928$, herring and sprat species only considered) were found to be potentially biased according to F12T. The number of suspect cases varied between ICES areas (range: 40.95%–51.42% of records) and ICES Rectangles (range: 30.41%–56.29% of records), but also with differences by destination country (range: 42.85%–50.19% of records), year (range: 31.5%–55.04% of records), gear type (range: 41.28%–48.96% of records), and vessel (range: 12.7%–77.96% of records) (Table S6). Comparable results are observed for the commercial logbook data using the catches for small pelagics performed by bottom trawlers and pelagic trawlers (Table 2 and Table S5 for details).

The commercial landing declaration

The commercial landing declaration dataset collecting catches of all species ($n = 225\,841$) was found to be in "close conformity" to the NBL at the first digit but showed "non-conformity" at the first two digits (Fig. 4, Table S5 for details). When only the catches of herring and sprat species ($n = 92\,868$) were tested, the data exhibited "acceptable conformity" to the NBL at the first digit (F1T,

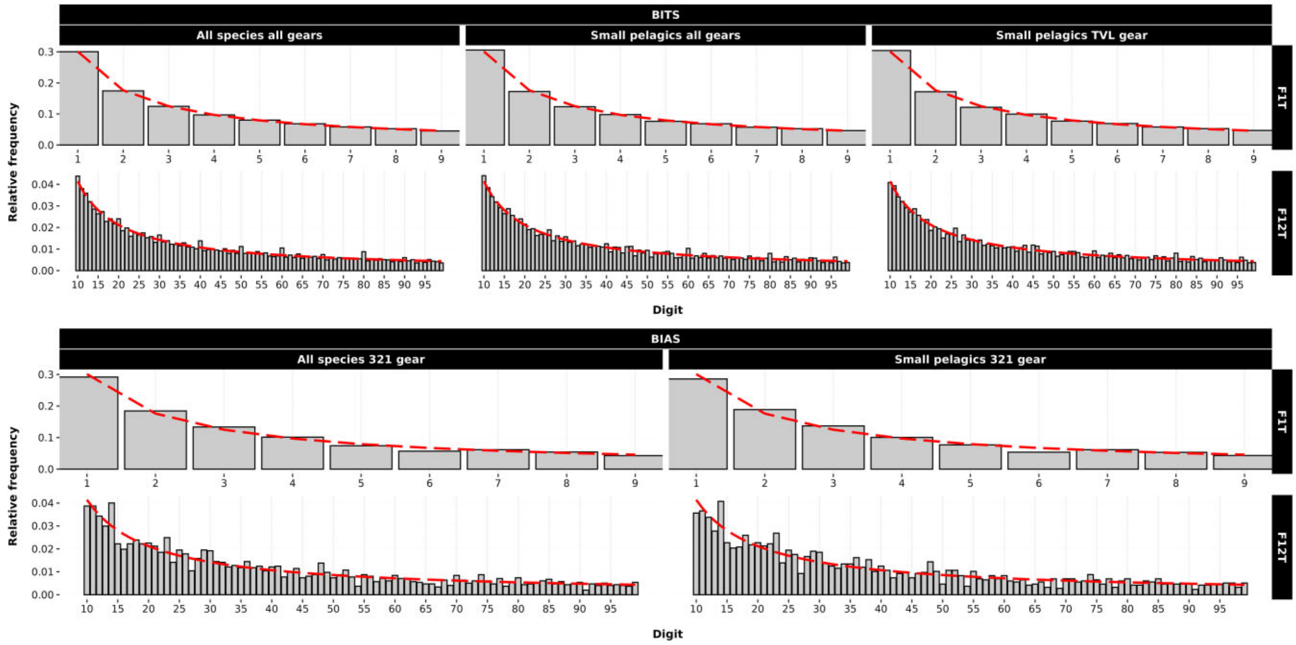

Figure 2. Survey data against the NBL. The top panels (first six plots) show the performance of BITS data when tested against the NBL model. The bottom panels (last four plots) refer to the BIAS data. The first and third rows refer to the distribution of the first digit, while the second and fourth rows refer to the distribution of the first two digits. The survey data tested includes all species (in the first column, for BITS and BIAS), the small pelagic for all gears (in the second column for BITS, in BIAS only one gear is available) and the small pelagic caught by specific gears (in the third column for BITS with TVL, in the second column for BIAS with the gear "321").

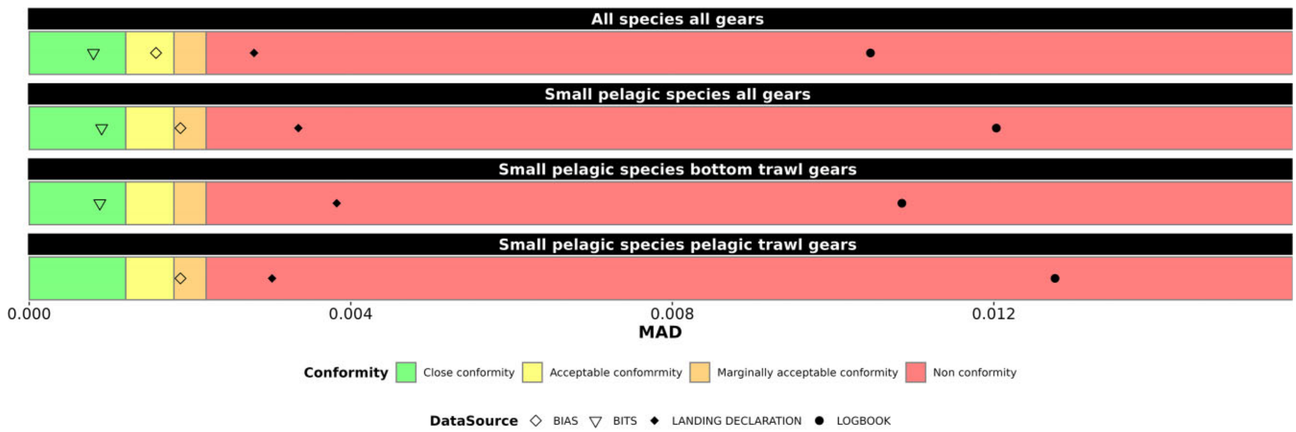

Figure 3. Comparison between commercial and scientific MAD values when tested against the NBL. The vertical lines are MAD value thresholds identifying areas of consistency with the NBL. The panels describe the performance of the data in different instances. From top to bottom: (i) all species with all gears available, (ii) small pelagic species (herring and sprat) with all gears available, (iii) small pelagic species with bottom trawl gears, and (iv) small pelagic species with pelagic trawl gears.

MAD = 0.00863) and "non-conformity" at the first two digits (F12T, MAD = 0.00335) (Figs 3 and 4, Tables 2 and S5). About 19.22% of the records ($n = 17\,852$, herring and sprat species considered) were found to be potentially biased according to F12T. The number of suspect cases varied between ICES areas (range: 14.3%–21.38% of records) and ICES Rectangles (range: 13.35%–34.04% of records), but also with differences by year (range: 14.38%–31.26% of records), gear type (range: 16.74%–20.04% of records), and vessel (range: 8.71%–37.59% of records) (Table S6). Comparable results to the previous case were observed when the data were filtered for the catches performed by pelagic or bottom trawlers (Tables 2 and S5).

Discussion and conclusion

Misreporting is a long-standing issue both in statistical survey studies (e.g. Sloan et al. 2004, Maurer et al. 2006, Rosaz and Villeval 2012, Selb and Munzert 2013, Amuedo-Dorantes and Arenas-Arroyo 2022) and in fishery science (Watson and Pauly 2001, Ainsworth and Pitcher 2005, Clarke et al. 2006, Hentati Sundberg et al. 2014, Pauly and Zeller 2016, Rudd and Branch 2017, Van Beveren et al. 2017). Many countermeasures to misreporting have been discussed in fishery science, but no "one fits all" solution has been offered to date.

In this work, we have shown how the analysis of anomalous numbers via the application of the NBL to fishery data, and the comparison of performance with a scientific dataset can

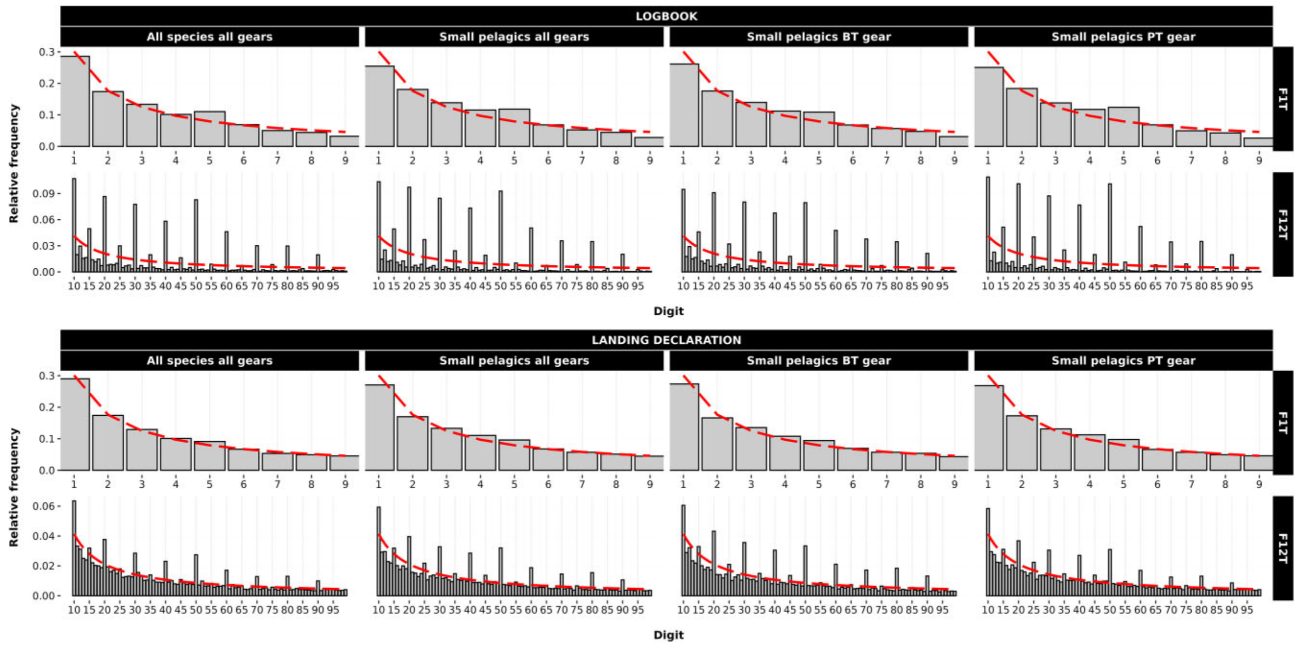

Figure 4. Commercial data against the NBL. The top panels (first eight plots) show the performance of logbook data when tested against the NBL model. The bottom panels (last eight plots) refer to the landing declaration data. The first and third rows refer to the distribution of the first digit, while the second and fourth rows refer to the distribution of the first two digits. The commercial data tested includes all species in the first column, the small pelagic herring and sprat in the second column, and the small pelagic caught by the bottom and pelagic trawlers, respectively, in the third and fourth columns.

Table 2. Main results from analysis of first ("Test" = "F1T") and first two ("Test" = "F12T") digits of different databases ("Dataset"): logbook data ("LB"), landing declaration data ("LD"), Baltic International Trawl Survey ("BITS"), Baltic International Acoustic Survey ("BIAS").

Dataset	Instance	*N*	Test	MAD	Conformity (MAD)	SU (%)	SU VE (range %)
BITS	AS-AAG	34 344	F1T	0.00 083	CC	-	-
BITS	SP-AAG	9885	F1T	0.00 204	CC	-	-
BITS	SP-TVL	8542	F1T	0.00 232	CC	-	-
BITS	AS-AAG	34 344	F12T	0.00 080	CC	6.82%	6%–12.07%
BITS	SP-AAG	9885	F12T	0.00 090	CC	11%	9.77%–11.02%
BITS	SP-TVL	8542	F12T	0.00 088	CC	9.01%	8.59%–9.21%
BIAS	AS-AAG	2974	F1T	0.00 612	AC	-	-
BIAS	SP-AAG	2162	F1T	0.00 747	AC	-	-
BIAS	AS-AAG	2974	F12T	0.00 158	AC	11.74%	-
BIAS	SP-AAG	2162	F12T	0.00 188	MAC	13.27%	-
LB	AS-AAG	291 672	F1T	0.01 022	AC	-	-
LB	SP-AAG	119 525	F1T	0.01 700	NC	-	-
LB	SP-PTG	73 750	F1T	0.01 934	NC	-	-
LB	SP-BTG	45 775	F1T	0.01 330	MAC	-	-
LB	AS-AAG	291 672	F12T	0.01 046	NC	41.21%	6.86%–77.89%
LB	SP-AAG	119 525	F12T	0.01 203	NC	45.12%	12.7%–77.96%
LB	SP-PTG	73 750	F12T	0.01 276	NC	47.5%	16.22%–66.78%
LB	SP-BTG	45 775	F12T	0.01 086	NC	41.29%	12.58%–77.96%
LD	AS-AAG	225 841	F1T	0.00 448	CC	-	-
LD	SP-AAG	92 868	F1T	0.00 863	AC	-	-
LD	SP-PTG	52 177	F1T	0.00 892	AC	-	-
LD	SP-BTG	40 691	F1T	0.00 908	AC	-	-
LD	AS-AAG	225 841	F12T	0.00 280	NC	18.02%	8.45%–36.9%
LD	SP-AAG	92 868	F12T	0.00 335	NC	19.22%	8.71%–37.59%
LD	SP-PTG	52 177	F12T	0.00 302	NC	18.36%	11.88%–24.7%
LD	SP-BTG	40 691	F12T	0.00 383	NC	20.32%	8.71%–37.59%

These datasets were analyzed in instances ("Instances") whose acronym refers to all species ("AS"), small pelagics ("SP") caught by all available gears ("AAG"), pelagic trawl ("PTG"), or bottom trawl gear ("BTM" for commercial and "TVL" for survey). Column *N* refers to the number of records. Mean absolute deviation ("MAD") values are reported. The conformity to the NBL model, assessed through MAD values, is one of "close conformity" (CC), "acceptable conformity" (AC), "marginally acceptable conformity" (MAC), and non-conformity (NC). The total amount of suspects ("SU (%)") is the number of records belonging to the first five groups of digits "not-conform" to the NBL in the F12T, divided by the total number of records of the dataset and multiplied per 100. The "SU VE (range %)" is the same percentage but stratified by vessel, meaning that the range indicates the extremes of the percentage of suspect records belonging to the vessels having the fewest and most suspect records. In the Supplementary Material, a version of this table (Table S5), including additional statistics (χ^2, Table S5), and a version of it deepening the suspects stratification (i.e. by year and ICES Subdivision, Table S6) are available.

inform about potential misreporting. This study represents the first application of the NBL not only in a Baltic fishery and survey but in the Northeast Atlantic geographical region.

Overall, research data conformed to the NBL frequency distribution, while the commercial data departed from it. As highlighted in the literature, non-conformity with the NBL cannot be considered as proof of misreporting (Graham et al. 2009, Nigrini 2012). In the context of the analysis presented here, several reasons can explain the lack of conformity between the NBL and the commercial fishery data. The non-conformity of these data is likely the result of inaccurate reporting of the quantities caught due to the introduction of non-random error in the figures communicated to authorities and used by the scientific community, but this does not necessarily imply that the inaccuracy is driven by the human intention to misreport. Discrepancies from the NBL model may be due to inaccuracy in the measurement tools or procedures used by the fishers. The large amount caught in this mixed fishery and the requirements to report information at a haul level may have resulted in estimates of the species ratio and total amount caught that are associated with various degrees of measurement errors. When the distribution of digits from the commercial data is considered, the observed frequencies seem to retain the shape expected under the model but on various levels are characterized by high maxima and low minima exceeding the values expected for frequencies that conform to the NBL.

The peaks in the frequency in multipliers of five seem to highlight a tendency to round results, a behavior that is likely to occur, especially considering the intensity of fishing in the high seas. This may explain the decrease in the number of anomalies of the landing declaration data that are collected in a better controlled condition environment. However, the improvement in performance does not remove doubts about the possible presence of misreporting because these latter data also do not conform to the NBL model. Alteration of data, if any, may occur in several ways; therefore, any reasoning around its occurrence, extent, and direction remains speculative (Nigrini 2012). In the commercial case presented here, a possible explanation for the observed distribution is that fishers are rounding at the second digit to the nearest multiplier of five. The peaks at first and second digits at 5 and 50, respectively, may derive from a coarse estimation of the proportion of the two species coupled with the fact that the catch data follow the NBL. For instance, if a given quantity is extracted from the sea, as demonstrated by our analysis of survey data, its leading digit is more likely to be 1 than any other digit. In this example, any coarse splitting of this quantity in 50% herring and 50% sprat will increase the frequency of 5 and 50, respectively, at F1T and F12T.

In the Baltic, herring and sprat are quota-limited species, and concerns have routinely been raised in the context of their assessment about the possible misreporting of their catches and its impacts (ICES 2021a). Which species may provide the incentive to misreporting in which period depends on the quota set. In any case, modifications to the numbers within each year will most likely not be random, but directional. Impacts on stock assessment are likely considering the large quantities being involved.

Assessing the intentionality of providing inaccurate results is out of the scope of this paper, and determining the modalities by which the reported data may diverge from the real one (i.e. the truth) is out of the capabilities of the tool used in our analysis. Instead, we would like to stress three important points related to the accuracy of the data highlighted by our research.

First, the numbers analyzed span across several orders of magnitude (e.g. for commercial logbook data up to 6, with a range of 0.1–500 000.0 kg, Table S4), and inconsistencies were found at the first two digits. This means that the inaccuracy (e.g. due to rounding), if any, may involve a large amount of fish. The risk of the presence of large approximations in the catch data collected at sea should motivate further research on the impact of possible inaccurate reporting (in each form, including rounding) on the stock perception (Watson and Pauly 2001).

Second, logbooks are generally used in fishery science, including their use for compilation of catch data for stock assessment (Sampson 2011). They provide unique information on the spatial and temporal distribution of catches (Cotter and Pilling 2007). This resolution is unavailable in the original landing declaration, which pools together the time and space resolution of the different hauls of the trips into unique points in time (i.e. landing dates) and space (i.e. landing sites). A spatially oriented stock assessment requires both high spatio-temporal resolution and accuracy of the data (Cotter and Pilling 2007, Cadrin 2020). If the logbook data are used for assessing the status of the resources, as is the case for the stocks analyzed in this study, they need to be accurate to avoid bias in the stock assessment model estimates. If the landing declaration data are used, they need to be spatially and temporally informative. The version of the landing declaration we used in this study disaggregates in space and time (according to the logbook information) the weights declared by the fishers at the landing site and given its improved conformity with the NBL may be considered, using a case-by-case approach, as a data input source for the assessment.

Third, while we appreciate the difficulties of estimating accurately the amounts of fishes caught by the large trawls and their separation by species, we believe that both industrial and research efforts are needed to assess the accuracy of catch data. Estimation of species composition on board is a complex task, which is carried on by fishermen using traditional methods that may be improved with *ad hoc* experiments. State-of-the-art, science-based methods to estimate the species composition are developing and include, among the others, applying computer vision and machine learning to videos or photos of the fish caught in order to classify them (e.g. Allken et al. 2019) as well as environmental DNA processing (e.g. Urban et al. 2024) in order to detect the species and possibly determine the percentage with which each of the species caught is contributing to the catch. Providing the fishers with more accurate and efficient ways of estimating the catch at the haul level would be beneficial for the science, the management, the resources, and the fishers as it would improve the quality of the commercial data at the highest resolution possible (Cotter and Pilling 2007), reducing the risk of bias in stock status and thus increasing the likelihood of viability for the stocks and the fishery economy, decreasing the rate of unintentional misreporting (e.g. because of rounding) and thus decreasing the risk of fines for fishers.

The conformity of survey data to NBL highlights that this model is adequate to describe the distribution of catches from pelagic and demersal species in fisheries based on trawling operations in the Baltic Sea. To identify the causes of misreporting goes beyond the aim of this study. It is important, however, that future work is undertaken to investigate the possible drivers of potential non-compliances as well as other possible sources of inaccuracy in the reports from this fishery, so that misreporting can be mitigated or alternatively quantified.

The approach presented here might be applied to many other contexts as misreporting affects or is suspected to affect many stocks and fisheries. Documented examples are numerous in the literature, including the eastern Baltic cod (*Gadus morhua*) fishery (Bastardie et al. 2008), the anglerfish (*Lophius piscatorius* and *Lophius budegassa*) caught in the Scottish fishery in the North Sea (Dobby et al. 2008), the herring and mackerel (*Scomber scombrus*) in the Dutch freezer-trawlers pelagic fishery occurring in European waters (Borges et al. 2008), many of the rockfishes species, including pacific ocean perch (*Sebastes alutus*) and northern rockfish (*Sebastes polyspinis*), caught in the Gulf of Alaska (Faunce 2011), the groundfish stocks in New England (e.g. Gulf of Maine haddock, *Melanogrammus aeglefinus*; Kerr et al. 2022), Northeast Atlantic spurdog (*Squalus acanthias*, De Oliveira et al. 2013), the Northeast Atlantic stocks of deep-water red crab (*Chaceon affinis*; Robinson 2008), the European common megrim (*Lepidorhombus whiffiagonis*), and the four-spotted megrim (*L. boscii*) in the Spanish mixed megrim fishery (Crego-Prieto et al. 2012), the North American off-shore (*Merluccius albidus*) and silver (*M. bilinearis*) hakes (Garcia-Vazquez et al. 2012), the red snapper (*Lutjanus campechanus*) in the fisheries Gulf of Mexico fishery (Marko et al. 2004), the Southern bluefin tuna (*Thunnus maccoyii*) caught in the Japanese longline tuna fishery (Polacheck 2012), and potentially many others (see Pauly and Zeller 2016 for a review concerning the accuracy of global fishery catch statistics). Fishery-dependent data are self-reported by the fishermen (e.g. fishery logbooks, sale notes, etc.) and constitute the main source of information on species catch composition and amounts extracted from the sea and therefore are a primary input to fish stock assessments. While advancements in fishery regulation (e.g. linked to the introduction of electronic logbook, the implementation of observer programs) have contributed to reduce some of the uncertainties, there is a general lack of supporting tools to assess the quality of fishery self-reported data. We argue that the NBL can contribute to filling this gap, and the methodology presented in this study can be readily implemented to other stocks and fishery as a supporting tool to investigate potential misreporting and contribute to improve our understanding of self-reported fisheries data. While the ignorance about the existence of this tool by the fishers is seen as an aspect that safeguards its applicability, we think that the realization by the fishers that post hoc control measures exist for the reported data may act as a deterrent to avoid intentional misreporting, if any.

Based on the results from this study, we recommend using the NBL model for quality control of commercial fishery information before their use as input data in stock assessment. While we recognize that the tool is not able to quantify the misreporting, we argue that it may assist in identifying where, and potentially when, further efforts are required to make such quantifications. Here, only commercial catch information from Swedish vessels was analyzed, and it is important to understand if similar non-conformities would also appear in catch data from other countries fishing in the Baltic.

To develop and implement tools able to quantify species misreporting in fisheries that take place in areas shared by multiple country fleets, an internationally coordinated effort against this phenomenon is needed. Future research based on these foundations shall focus on the production or implementation of these tools to produce alternative time series of the catches for the different species and evaluate the impact of the relative different scenarios on the stock perception. For example, alternative time series of catches inspected by our methodology could represent equally plausible hypotheses on the amount of fish removed and therefore be used as an additional dimension of the uncertainty in the current ensemble of stock assessment models used to provide management advice for the stocks (e.g. Stock Synthesis, Methot and Wetzel 2013). This may help the understanding of the consequences of the different catch time series that include misreporting on the status of the stock and reference points used to produce the scientific advice for the two stocks analyzed here and potentially many others, ecologically and economically important stocks in the Northeast Atlantic.

Acknowledgments

The authors are grateful to Sofia Carlshamre (Institute of Marine Research, Department of Aquatic Swedish University of Agricultural Sciences, SLU Aqua hereafter) and Barbara Bland (SLU Aqua) for the support in data retrieving, to Johan Lövgren (SLU Aqua) and Annica De Groote (SLU) for providing the opportunity of receiving feedback on the analysis from various colleagues in the context of the data preparation meeting for the benchmark of central Baltic herring and sprat and in the statistical community, and to Christopher A. Griffiths (SLU Aqua) for the support provided during the submission of this manuscript for the review. The authors would like to thank all the personnel involved in the BITS and BIAS surveys for collecting data on the field and thus creating a term of comparison for the commercial fishery data as presented in this work. The authors are grateful to the three anonymous reviewers and the editor for providing valuable notes on the initial draft of the manuscript and thus allowing to improve further our scientific work.

Author contributions

The authors participated in the manuscript with the following contributions:

- [E.Q.]—Concept, analysis, writing of the first draft of the manuscript.
- [M.C.]—Concept and writing of the second draft of the manuscript.
- [N.P.]—Conceptualization, Writing—review & editing.
- [V.B.]—Concept, review and editing.
- [P.B.]—Scientific data provision. Review and editing.
- [M.B.N.]—Concept, review and editing.
- [N.L.]—Scientific data provision. Review and editing.
- [A.H.]—Description of fishing operations, review and editing.
- [K.R.]—Concept, review and editing.

Supplementary data

Supplementary data is available at *ICESJMS* online.

Conflict of interest: The authors declare that there is no conflict of interest.

Funding

No funding was provided for this work.

Data availability

This work used four databases, with different level of availability:

- BITS Data: The data underlying this article are available in DATRAS repository (ICES, 2022; https://gis.ices.dk/geonetwork/srv/metadata/f2f6f003-4dec-41c2-95d1-29dd48876662).
- BIAS Data: The Swedish contribution to the BIAS database is hosted at the Institute of Marine Research, Swedish University of Agricultural Sciences (SLU Aqua, Lysekil, Sweden). The relative data underlying this article will be shared on reasonable request to the corresponding author.
- Logbook Data: The commercial database is hosted at the Swedish Agency for Marine and Water Management (Göteborg, Sweden). The data underlying this article were provided by the Swedish Agency for Marine and Water Management under licence/by permission. Data will be shared on request to the corresponding author with permission of the Swedish Agency for Marine and Water Management.
- Landing declaration Data: The commercial database is hosted at the Swedish Agency for Marine and Water Management (Göteborg, Sweden). The data underlying this article were provided by the Swedish Agency for Marine and Water Management under licence/by permission. Data will be shared on request to the corresponding author with permission of the Swedish Agency for Marine and Water Management.

References

Ainsworth CH, Pitcher TJ. Estimating illegal, unreported and unregulated catch in British Columbia's marine fisheries. *Fish Res* 2005;**75**:40–55. https://doi.org/10.1016/j.fishres.2005.05.003

Allken V, Handegard NO, Rosen S *et al*. Fish species identification using a convolutional neural network trained on synthetic data. *ICES J Mar Sci* 2019;**76**:342–9. https://doi.org/10.1093/icesjms/fsy147

Amuedo-Dorantes C, Arenas-Arroyo E. Police trust and domestic violence among immigrants: evidence from VAWA self-petitions. *J Econ Geogr* 2022;**22**:395–422. https://doi.org/10.1093/jeg/lbab007

Arrhenius F, Hansson S. Food consumption of larval, young and adult herring and sprat in the Baltic Sea. *Mar Ecol Prog Ser* 1993;**96**:125. https://www.jstor.org/stable/24833538

Bastardie F, Nielsen JR, Kraus G The eastern Baltic cod fishery: a fleet-based management strategy evaluation framework to assess the cod recovery plan of 2008. *ICES J Mar Sci*. 2010;**67**:71–86. https://doi.org/10.1093/icesjms/fsp228

Benford F. The law of anomalous numbers. *Proc Am Philos Soc* 1938;**78**:551–72. https://www.jstor.org/stable/984802

Berger A, Hill TP, Rogers E. *Benford Online Bibliography*. 2009. http://www.benfordonline.net, (17 February 2024, datelast accessed).

Berkow C. *Where Have All the Herring Gone?* Stockholm University Baltic Sea Centre. 2021. https://www.su.se/stockholm-university-baltic-sea-centre/web-magazine-baltic-eye/fisheries/where-have-all-the-herring-gone-1.609372, (17 February 2024, datelast accessed).

Blank SG, Gavin MC. The randomized response technique as a tool for estimating non-compliance rates in fisheries: a case study of illegal red abalone (*Haliotis rufescens*) fishing in Northern California. *Environ Conserv* 2009;**36**:112–9. https://doi.org/10.1017/S037689290999004X

Bolton RJ, Hand DJ. Statistical fraud detection: a review. *Stat Sci* 2002;**17**:235–55. https://doi.org/10.1214/ss/1042727940

Borges L, Van Keeken OA, van Helmond AT *et al*. What do pelagic freezer-trawlers discard?. *ICES J Mar Sci* 2008;**65**:605–11. https://doi.org/10.1093/icesjms/fsn050

Bray K. A global review of illegal, unreported and unregulated (IUU) fishing. *FAO Fisheries Reports* 2001;88–134.

Burns TR, Stöhr C. Power, knowledge, and conflict in the shaping of commons governance. The case of EU Baltic fisheries. *Int J Commons* 2011;**5**:233–58. https://doi.org/10.18352/ijc.260

Cadrin SX. Defining spatial structure for fishery stock assessment. *Fish Res* 2020;**221**:105397. https://doi.org/10.1016/j.fishres.2019.105397

Cinelli C. *benford.analysis: Benford Analysis for Data Validation and Forensic Analytics. R Package Version 0.1.5*. 2018. https://CRAN.R-project.org/package=benford.analysis, (15 February 2025, datelast accessed).

Clarke SC, McAllister MK, Milner-Gulland EJ. Global estimates of shark catches using trade records from commercial markets. *Ecol Lett* 2006;**9**:1115–26. https://doi.org/10.1111/j.1461-0248.2006.00968.x

Cotter AJ, Pilling GM. Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. *Fish Fish* 2007;**8**:123–52. https://doi.org/10.1111/j.1467-2679.2007.00241.x

Crego-Prieto V, Campo D, Perez J *et al*. Inaccurate labelling detected at landings and markets: the case of European megrims. *Fish Res* 2012;**129**:106–9.

da Silva Azevedo C, Gonçalves RF, Gava VL *et al*. A Benford's law based method for fraud detection using R Library. *MethodsX* 2021;**8**:101575. https://doi.org/10.1016/j.mex.2021.101575

Dennis D, Plagányi É, Van Putten I *et al*. Cost benefit of fishery-independent surveys: are they worth the money?. *Mar Policy* 2015;**58**:108–15.

De Oliveira JA, Ellis JR, Dobby H. Incorporating density dependence in pup production in a stock assessment of NE Atlantic spurdog *Squalus acanthias*. *ICES J Mar Sci* 2013;**70**:1341–53.

Dobby H, Allan L, Harding M *et al*. Improving the quality of information on Scottish anglerfish fisheries: making use of fishers' data. *ICES J Mar Sci* 2008;**65**:1334–45. https://doi.org/10.1093/icesjms/fsn144

Ducrotoy JP, Elliott M. The science and management of the North Sea and the Baltic Sea: natural history, present threats and future challenges. *Mar Pollut Bull* 2008;**57**:8–21. https://doi.org/10.1016/j.marpolbul.2008.04.030

Fang G, Chen Q. Several common probability distributions obey Benford's law. *Physica A* 2020;**540**:123129. https://doi.org/10.1016/j.physa.2019.123129

Faunce CH. A comparison between industry and observer catch compositions within the Gulf of Alaska rockfish fishery. *ICES J Mar Sci* 2011;**68**:1769–77. https://doi.org/10.1093/icesjms/fsr093

FishSec. *The 20 Year Baltic Sea Scandal*. 2019. https://www.fishsec.org/2019/11/05/the-20-year-baltic-sea-scandal/, (17 February 2024, datelast accessed).

Garcia-Vazquez E, Machado-Schiaffino G, Campo D *et al*. Species misidentification in mixed hake fisheries may lead to overexploitation and population bottlenecks. *Fish Res* 2012;**114**:52–5. https://doi.org/10.1016/j.fishres.2011.05.012

Graham SD, Hasseldine J, Paton D. Statistical fraud detection in a commercial lobster fishery. *NZ J Mar Freshwater Res* 2009;**43**:457–63. https://doi.org/10.1080/00288330909510014

Hilborn R, Walters CJ (eds), *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Berlin Germany: Springer Science & Business Media, 2013.

Hill TP. A statistical derivation of the significant-digit law. *Stat Sci* 1995;**10**:354–63. https://www.jstor.org/stable/2246134

Hentati Sundberg J, Hjelm J, Österblom H. Does fisheries management incentivize non-compliance? Estimated misreporting in the Swedish

Baltic Sea pelagic fishery based on commercial fishing effort. *ICES J Mar Sci* 2014;**71**:1846–53. https://doi.org/10.1093/icesjms/fsu036

Hoyle SD, Campbell RA, Ducharme-Barth ND *et al.* Catch per unit effort modelling for stock assessment: a summary of good practices. *Fish Res* 2024;**269**:106860. https://doi.org/10.1016/j.fishres.2023.106860

ICES. Herring in subdivisions 25–29 and 32 (excluding Gulf of Riga herring). *ICES Advice: Recurrent Advice. Report*. 2014. https://doi.org/10.17895/ices.advice.18669569.v1

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2015. https://doi.org/10.17895/ices.advice.18668819.v1

ICES. Herring (Clupea harengus) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2016. https://doi.org/10.17895/ices.advice.18668195.v1

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2017. https://doi.org/10.17895/ices.pub.3127

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2018. https://doi.org/10.17895/ices.pub.4384

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2019. https://doi.org/10.17895/ices.advice.4748

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2020. https://doi.org/10.17895/ices.advice.5828

ICES. Baltic Sea ecoregion—fisheries overview. *ICES Advice: Fisheries Overviews. Report*. 2021a. https://doi.org/10.17895/ices.advice.9139

ICES. Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). *ICES Advice: Recurrent Advice. Report*. 2021b. https://doi.org/10.17895/ices.advice.7767

ICES. *Database on Trawl Surveys (DATRAS)*. Copenhagen: ICES. 2022. https://datras.ices.dk, (15 February 2025, datelast accessed).

Kerr L, Barajas M, Wiedenmann J. Coherence and potential drivers of stock assessment uncertainty in Northeast US groundfish stocks. *ICES J Mar Sci* 2022;**79**:2217–30. https://doi.org/10.1093/icesjms/fsac140

Larson N. Baltic International Acoustic Survey report, October 2019. *Aqua Reports*. Swedish University of Agricultural Sciences, 2020.**22**.

Le Gallic B, Cox A. An economic analysis of illegal, unreported and unregulated (IUU) fishing: key drivers and possible solutions. *Mar Policy* 2006;**30**:689–95. https://doi.org/10.1016/j.marpol.2005.09.008

MacKenzie BR, Gislason H, Möllmann C *et al.* Impact of 21st century climate change on the Baltic Sea fish community and fisheries. *Global Change Biol* 2007;**13**:1348–67. https://doi.org/10.1111/j.1365-2486.2007.01369.x

Marko PB, Lee SC, Rice AM *et al.* Mislabelling of a depleted reef fish. *Nature* 2004;**430**:309–10. https://doi.org/10.1038/430309b

Maurer J, Taren DL, Teixeira PJ *et al.* The psychosocial and behavioral characteristics related to energy misreporting. *Nutr Rev* 2006;**64**:53–66. https://doi.org/10.1111/j.1753-4887.2006.tb00188.x

Methot RD , Wetzel CR. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fish Res* 2013;**142**:86–99. https://doi.org/10.1016/j.fishres.2012.10.012

Methot RD, Wetzel CR, Taylor IG *et al. Stock Synthesis User Manual: Version 3.30. 15*. 2020.

Newcomb S. Note on the frequency of use of the different digits in natural numbers. *Am J Math* 1881;**4**:39–40. https://doi.org/10.2307/2369148

Nigrini MJ. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. New Jersey, U.S.: Wiley, 2012.

Noleto-Filho EM, Carvalho AR, Thome-Souza MJ *et al.* Reporting the accuracy of small-scale fishing data by simply applying Benford's law. *Front Mar Sci* 2022;**9**:947503. https://doi.org/10.3389/fmars.2022.947503

Ojaveer H, Jaanus A, MacKenzie BR. Status of biodiversity in the Baltic Sea. *PLoS One*. 2010;**5**:e12467. https://doi.org/10.1371/journal.pone.0012467

Patterson KR. Assessing fish stocks when catches are misreported: model, simulation tests, and application to cod, haddock, and whiting in the ICES area. *ICES J Mar Sci* 1998;**55**:878–91. https://doi.org/10.1006/jmsc.1998.0351

Pauly D, Zeller D. Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nat Commun* 2016;**7**:10244. https://doi.org/10.1038/ncomms10244

Pitcher TJ, Watson R, Forrest R. Estimating illegal and unreported catches from marine ecosystems: a basis for change. *Fish Fish* 2002;**3**:317–39. https://doi.org/10.1046/j.1467-2979.2002.00093.x

Polacheck T. Assessment of IUU fishing for southern bluefin tuna. *Mar Policy* 2012;**36**:1150–65. https://doi.org/10.1016/j.marpol.2012.02.019

Raimi RA. The peculiar distribution of first digits. *Sci Am* 1969;**221**:109–20. https://doi.org/10.1038/scientificamerican1269-109

Robinson M. (2008). Minimum landing size for Northeast Atlantic stocks of deep-water red crab, *Chaceon affinis* (Milne Edwards and Bouvier, 1894). *ICES J Mar Sci* 2008;**65**:148–54. https://doi.org/10.1093/icesjms/fsm189

Rodriguez RJ. Reducing false alarms in the detection of human influence on data. *J Account Audit Finance* 2004;**19**:141–58.

Rosaz J, Villeval MC. Lies and biased evaluation: a real-effort experiment. *J Econ Behav Organ* 2012;**84**:537–49. https://10.1016/j.jebo.2012.09.002

Ross KA. Benford's law, a growth industry. *Am Math Mon* 2011;**118**:571–83. https://doi.org/10.4169/amer.math.monthly.118.07.571

Rudd MB, Branch TA. Does unreported catch lead to overfishing?. *Fish Fish* 2017;**18**:313–23. https://doi.org/10.1111/faf.12181

Sampson DB. The accuracy of self-reported fisheries data: oregon trawl logbook fishing locations and retained catches. *Fish Res* 2011;**112**:59–76. https://doi.org/10.1016/j.fishres.2011.08.012

Scott PD, Fasli M. Benford's law: an empirical investigation and a novel explanation. Unpublished manuscript. 2011

Selb P, Munzert S. Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys. *Electoral Studies* 2013;**32**:186–96. https://10.1016/j.electstud.2012.11.004

Sheskin DJ. *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL, USA: Chapman and Hall/CRC. 2003. https://doi.org/10.1201/9780429186196

Sjöstrand B. Strömming—hur lite finns kvar?. *Östersjö*. 2000. 17–9.

Sloan JJ, Bodapati MR , Tucker TA. Respondent misreporting of drug use in self-reports: social desirability and other correlates. *J Drug Issues* 2004;**34**:269–92. https://10.1177/002204260403400202

Svensson A. Felrapportering av sill och skarpsill från Östersjön. *Svenssons Nyheter—Njord*. 2019. https://fiske.zaramis.se/2019/09/16/felrapportering-av-sill-och-skarpsill-fran-ostersjon/, (17 February 2024, datelast accessed).

Tesfamichael D, Pitcher TJ. Estimating the unreported catch of Eritrean Red Sea fisheries. *Afr J Mar Sci* 2007;**29**:55–63. https://doi.org/10.2989/AJMS.2007.29.1.5.70

Thiessen V, Blasius J. *Assessing the Quality of Survey Data*. Thousand Oaks, California: SAGE Publications Ltd. 2012. https://doi.org/10.4135/9781446251874

Tokarev Y, Shulman G. Biodiversity in the Black Sea: effects of climate and anthropogenic factors. In *Biodiversity in Enclosed Seas and Artificial Marine Habitats: Proceedings of the 39th European Marine Biology Symposium, Held in Genoa, Italy, 21–24 July 2004*. Springer Netherlands. 2007;23–33. https://10.1007/978-1-4020-6156-1_2

Tsagbey S, de Carvalho M, Page G. What's wrong with official fishery statistics? Evidence from tuna catch data from the Gulf of Guinea. *ISI Proceedings*. 2015.

Tsagbey S, de Carvalho M, Page GL. All data are wrong, but some are useful? Advocating the need for data auditing. *Am Stat* 2017;**71**:231–5. https://doi.org/10.1080/00031305.2017.1311282

Urban P, Jacobsen MW, Bekkevold D *et al*. eDNA based bycatch assessment in pelagic fish catches. *Sci Rep* 2024;**14**:2976. https://doi.org/10.1038/s41598-024-52543-0

Van Beveren E, Duplisea D, Castonguay M *et al*. How catch underreporting can bias stock assessment of and advice for northwest Atlantic mackerel and a possible resolution using censored catch. *Fish Res* 2017;**194**:146–54. https://doi.org/10.1016/j.fishres.2017.05.015

Watson R, Pauly D. Systematic distortions in world fisheries catch trends. *Nature* 2001;**414**:534–6. https://doi.org/10.1038/35107050

Handling Editor: Ruben Roa-Ureta