Predicting the number of known and unknown species in European seas using rates of description

Mark J. Costello[1, 2, 3] and Simon P. Wilson[4]

[1] Leigh Marine Laboratory, University of Auckland, PO Box 349, Warkworth 0941, New Zealand. Email m.costello@auckland.ac.nz Tel. +64-9-425 7253

[2] Ecological Consultancy Services Ltd (EcoServe), Apex House, Greenmount Industrial Estate, Harold's Cross Road, Dublin 12, Ireland.

[3] The Huntsman Marine Science Centre, 1 Lower Campus Road, St Andrews, New Brunswick, E5B 2L7, Canada

[4] School of Computer Science and Statistics, Trinity College, University of Dublin, Dublin 2, Ireland.

KEYWORDS: biodiversity, species richness, discovery rates, taxonomy

Running title: Predicting species diversity

The authors declare they have no interests in conflict with this paper.

Biosketches

Mark J. Costello (BSc Galway, PhD Cork) is an ecologist with a particular interest in marine biogeography. He was the first chair of the Ocean Biogeographic Information System until 2008, and now chairs the World Register of Marine Species (www.marinespecies.org) and the Society for the Management of Electronic Biodiversity Data (www.smebd.eu). Related interests include the use of marine reserves as controls for human impacts on marine ecosystems, and the ecology of sea lice parasites on wild and farmed salmonids.

Simon P. Wilson (BA Oxon, PhD George Washington University) is a statistician with interests in statistical modelling and inference for environmental, engineering and telecommunications applications. He is an associate professor at Trinity College Dublin and the head of its statistics group. He is an elected member of the International Statistical Institute and a fellow of the Royal Statistical Society.

MJC conceived and wrote this paper, and conducted some of the analyses. SPW conducted the statistical modelling and co-wrote the paper.

ABSTRACT

Aim

In this paper, we compared species description rates to predict the numbers of undescribed species. These data were used to discuss the merits of various attempts to estimate species richness in the oceans.

Methods

Predictions of how many species may exist on earth have lacked an inventory of how many have been described except for a few small taxa. The ocean is a good place to start an inventory because it includes all but one of the phyla and most classes of life on Earth. The European Register of Marine Species (ERMS) was compiled by taxonomic experts, covered all marine taxa, and accounted for synonyms. Reflecting taxonomic history, Europe's species are the best described in the world.

Results

ERMS listed 29,713 species of animals, plants and protists, but excluded bacteria and viruses. An estimated 6,500 described species were not included. The best prediction of the number of species remaining to be described was 5,613. Plots of years when species were first described showed no decrease in the rate of description for any taxa except birds, mammals and krill. If taxonomic effort has increased, whether due to more resources globally, or greater efficiencies of productivity, then description rates per unit effort may be declining and the number of undescribed species may be lower than predicted. However, apart from reduced rates of description during the World Wars, there were no changes in description rates that could be easily attributed to such factors.

Conclusions

There are about 36,000 species described from European seas, and we predict that 40,000 to 48,000 may exist. This comprises 15% of the estimated 230,000 described marine species.

However, this area is well-known compared to other seas and the proportion of yet to be discovered species will be higher elsewhere.

INTRODUCTION

Why knowing species diversity is important

Understanding and conserving biodiversity are amongst the greatest challenges for mankind, as recognised in the world Convention on Biological Diversity. The most practical and commonly used indicator of biodiversity is the number of species within an area, habitat or sample (Costello 2001). Biogeographic and ecological theories, supported with experimental and field data, predict that habitat fragmentation and loss will result in species extinctions. Over harvesting, habitat loss, pollution and climate change are reducing species populations and leading to the current extinction crisis, such that half of all present species may be extinct within the next 100-300 years (Chapin et al. 2000, Jackson 2008). Yet many species, especially of smaller invertebrates, have never been described. Thus species are going extinct before they have been recognised or known. Some, if not many, of these species will be important to the functioning of ecosystems that provide services for society, and/or may become important as natural resources (e.g. food, biotechnology). The increased rate of extinction of species is occurring on land and in the oceans (Carlton et al. 1999, Dulvy et al. 2003). These extinctions are not only a permanently lost opportunity, but may have serious social and economic consequences. Considering these issues, it is not surprising that one of the most frequently asked questions about biodiversity is how many species are there. The number of species also provides the most practical measure of how much we know and do not know about biodiversity.

Curiously, most estimates of the number of undescribed species were based on informal expert opinions, rather than counts, of how many species have been described. It must seem astonishing to the public that after more than 250 years of using a standard approach to naming species, that biologists have not kept an inventory of what species have been described. This results in multiple descriptions of the same species, and continuous effort to unravel confusion regarding species names. Perhaps one fifth of all recently described species names are synonyms (Bouchet 2006). Thus analyses of species inventories, including those of importance for conservation, food and pests, must include quality control of names. A taxonomically authoritative inventory of species is not only important for quality control in science and resource management, but is of popular interest. For example, during the launch of the Census of Marine Life, public interest was captured by the phrase "how many fish are there in the sea"; and the popular media were still surprised at the discovery of new species (Bouchet 2006). A media release on the initiative to inventory all marine species was picked up by over 271 media outlets around the world (www.marinespecies.org). Even the wider scientific community needs to be reminded that new species are being described even in large animals; e.g. Guan (2007) reported in the journal *Nature* that three new species of Cetacea have been described since 2002. Efforts to inventory all known species are about half-complete for all (Bisby et al. 2009), and two-thirds of all marine species (Appeltans et al. 2009). However, this excludes viruses and bacteria for whom the species measure of biodiversity as applied to eukaryotes does not apply (e.g. O'Donnell et al. 1994). In Europe at least, such efforts have been compromised by there being insufficient taxonomic expertise in the taxa that are both the richest in species and contain the most species remaining to to be described (Costello et al. 2006).

Estimating species richness

The variety of approaches that have been taken to predict the number of species that may exist on earth have been reviewed by May (1992) and Hammond (1994), including: projecting from past trends in rates of discovery of species; educated guesses; polls of experts' opinions; extrapolation from the proportions of species in an intensively studied habitat that had not previously been described, to the total area of that habitat; multiplication of the proportions of species known to occur in certain habitats to estimate the likely numbers in other habitats (e.g. insects on other species of tree in the tropics, or ratios of plants to insects); using the ratios of numbers of species in different taxa in well-studied areas (e.g. Europe) to predict the total numbers in other parts of the world where only a limited proportion of the organisms have been well-described (typically birds and mammals); and relationships of body-sizes to species numbers. To these, Reaka-Kudla (1996) added extrapolation from species – area relationships using Island Biogeographic Theory. We classify these into those based on (1) expert estimates, (2) biogeographic theory and proportions of species by habitat or area, (3) body-size to richness relationships, and (4) past species description rates.

(1) Even experts may not have a sufficiently complete perspective of all organisms, parasitic and free-living, large and small, to provide reasonable estimates by opinion alone. Scientists initially estimated there were 20,000 to 25,000 marine species in Europe but later the same people produced an incomplete list of 15-33% more than this (Costello 2000). Similarly low estimates of known species numbers by experts have been found for terrestrial taxa (Hammond 1992) and fishes (Cohen 1970). In contrast, 25 scientists in the Fauna Europaea project estimated there would be 89,000 to 300,000 land and freshwater animals in Europe, but they inventoried 130,000, 16% less than the mean estimate (de Jong, pers. comm.). Thus expert estimates of known species richness can range from underestimates of 33% to overestimates of 16%, a range of nearly 50%.

(2) Most of the methods for estimating species richness by extrapolation suffer from weak or unsubstantiated relationships between the relative proportions of species in different habitats and geographic areas (Hammond 1992). May (1988) suggested that the factors influencing species richness might operate differently in different environments. Indeed, it is evident that the proportions of taxa vary significantly between habitats; take the marine, estuarine, freshwater and

land environments as extreme cases. Species vary greatly in their endemicity to a particular habitat, with some being obligate parasites or symbionts of particular host habitats. The well-established pattern where few species are abundant and most are rare needs to be accounted for in estimating how many species exist in any geographic area or habitat. Predictions of species richness using standardised 'sampling effort' methods do account for abundance in this way (Colwell and Coddington 1994).

The tropics and Indo-West Pacific are centres of high species richness, and may be centres of origin and/or speciation for marine species (Marshall 2006). If the latter is true then there would have to be equal rates of dispersal, speciation and extinction across taxa for the relative richness of species to be the same with geographic distance. A recent analysis across 24 regions of the world, found that the relative number of marine species across phyla and classes was variable, although whether this reflected different evolution of biota or bias in taxonomic effort and knowledge was as yet unclear (Costello et al. submitted). Dolphin and Quicke (2001) attributed differences in the ratio of described braconid wasps to butterflies between geographic regions to reflect taxonomic effort, but also noted that two of the best known taxa globally differed in their relative richness between the tropics and temperate regions. The lack of clear correlations between proportions of taxa across habitats and areas compromises the use of biogeographic, habitat and area relationships for predicting global species richness.

Extrapolation from the richness of species in samples in one habitat or local area is compromised by insufficient knowledge of the relationships between sample, regional, and spatial variation in sample richness, i.e. alpha, gamma and beta diversity; in both marine (Lambshead and Boucher 2003) and terrestrial (Stork 2007) environments. Sample richness may be low in some habitats but high in the region if it has a large variety of habitats (i.e. little overlap in species between samples within the region and thus high beta-diversity), and thus high gamma diversity (Lambshead and Boucher 2003). Another region may have similar gamma diversity but high alpha and low beta diversity. Geographical variation in these relationships confound attempts to predict regional from local richness (e.g. Dolphin and Quicke 2001), as attempted for tropical rainforest insects (Erwin 1988) and deep-sea macrobenthos (Grassle and Maciolek 1992). Indeed, Hammond (1994) concluded that extrapolation from a few taxa in one habitat or location was the poorest method of estimating wider species richness.

(3) Within a taxon, species body sizes may form a pattern whereby gaps in the size distribution may suggest as yet undescribed species (e.g. Finlay et al. 1996, Hall and Greenstreet 1996, Gaston and Blackburn 1997, Cooper et al. 2006). In various marine and terrestrial taxa, larger sized species tend to be better known due to being more conspicuous (May 1988, Gaston 1991, Costello et al. 1996, Zapata and Robertson 2006). Holoplankton and fish species with a larger body-size, larger geographic range, and wider depth range, have been described earlier (Gibbons et al. 2005, Zapata and Robertson 2006). However, as it is not possible to record a species spatial distribution until it is described, the apparently narrower geographic and depth ranges for recently described species may be a consequence of when they were described rather than the reason for their later discovery. Orme et al. (2002) found no relationship between body size and the richness of marine metazoan species, and recognised that finding such patterns would require the proportion of described and undescribed species to be similar, which does not seem to be the case. A size frequency distribution model fitted to a tropical fish assemblage of 1,222 species predicted only 15% of the known number of undescribed species and 5% of all the estimated undescribed species (Zapata and Robertson 2006). Apart from being a poor predictor, there appears little basis on which to make predictions on species richness using body size relationships, if a prerequisite is that the richness of the taxon is well-known. Analysis of body-

size relationships to diversity need to account for variation during the growth of most marine species, and that the environment and habitat commonly vary between life-stages (e.g. planktonic larvae).

(4) Projecting from actual data should produce more accurate estimates of future discoveries. Using an authoritative inventory of species, Costello *et al.* (1996) studied the rate of description of species recorded from British and Irish seas, and showed that rates of description were still very high in the organisms with a smaller body size even in these well-studied areas. However, they did not predict how many species remained to be discovered. Paxton (1998) fitted a line to the post-1830 part of the description rate curve for 217 large (> 2m wide or long) marine animals to predict that 47 (20%) more species awaited description in the world. Also using description rates, Wittman (1999) predicted the global diversity of mysid crustaceans to be four times the number of described species. In contrast, euphausiidacean crustacean description rates decreased since the 1920's, suggesting that most species had been discovered (Mauchline and Murano 1977).

In a problem that is subject to so much uncertainty, it is not surprising that statistical methods have been used, which produce both a prediction for the number remaining and an estimate of the error in that prediction. It is important to distinguish between what we call 'sampling effort' and 'species description rate' curves. The former 'discovers' species recorded by a standard method in the same habitat and geographic location but most of which will have been previously recorded elsewhere. The later records species 'discovered' new to science, often by using different sampling methods in less explored locations and habitats. When the number of individuals sampled (per species) in the discovery process is known then there is a large statistics literature on how to compute estimates of the number remaining (reviewed by Bunge and Fitzpatrick 1993 and Colwell and Coddington 1994). Unfortunately, these data are usually not available and all that is known is the species description rate, which can be extrapolated and a statistical model used to estimate the error in the prediction. Solow and Smith (2005) used a Poisson process but did not account for variation between description rates between years. Bebber et al. (2007) also used the Poisson model to derive prediction error estimates, and they and other authors found the logistic equation to be the most appropriate for modelling description rates of a variety of taxa (Frank and Curtis 1979, Zapata and Robertson 2006, Woodley et al. 2009). Wilson and Costello (2005) described a new statistical model which has been used to predict how many additional species may yet be discovered from the year their description was published. We posit that this is the most statistically rigorous approach yet applied to predicting the numbers of species because it is based on verifiable data, and accounts for variation in description rates. The Wilson and Costello model is a generalization of the Poisson process that allows processes that have either more or less variability than it. This addresses one of the main disadvantages of the Poisson process, namely that the variance of the process is equal to its mean, hence differing amounts of variability from the mean are not permitted. For comparison, we computed the prediction from the special case of the Poisson process with the logistic function as the mean trend and it was shown that in general species discovery is more variable than the Poisson process, so that Poisson process models would tend to underestimate the amount of error in any prediction of numbers of remaining species. With the availability of additional species lists, such as for other oceans and terrestrial areas, the Wilson and Costello model can be rapidly applied to provide more realistic estimates of total species number than presently exist.

The rates of species description within smaller areas can be affected by sampling effort, and thus may not reflect the rate of description of species new to science. We consider the description rates of species in Europe to represent general species discovery rates because it is where taxonomy originated and developed. However, the use of description rates in smaller and adjacent geographic areas would include many species described elsewhere and so not represent real discoveries of new species. Because we use marine species, we covered almost all phyla and classes excluding bacteria and viruses, and thus a wide range of taxonomic specializations which should minimise any effects of sampling and description practices that may have arisen within specialisations.

METHODS

Costello et al. (2001) produced the largest all-taxon checklist of marine fauna and flora in the world in its time, the European Register of Marine Species (ERMS). They listed 29,713 valid animal and plant species in European seas. Of these, 10% of the taxa were not sufficiently rich in species, and/or did not have sufficient years of description available for the present analysis (Table 1). In all, 26,529 of the species included the year of description in an accessible format for analysis. Europe was defined as the area encompassed by the European Register of Marine Species; namely from the North Pole along the east coast of Greenland to Iceland, then along the Mid-Atlantic Ridge at 3,000m depth to the 26°N parallel which meets the African coast. It thus included the Baltic, Mediterranean and Black Seas, and islands of Madeira, Canaries and Azores.

The non-homogeneous renewal process model of Wilson and Costello (2005) was fitted to the data via a Bayesian inference method and a probability distribution of the total number of species remaining was then derived. The model assumes that the trend of the description curve follows a logistic type shape, so at *t* years after 1749 the number of species discovered follows a trend curve with form:

$$\frac{N}{1 + \exp(-v_1 (t - y_2))} .$$

where N is the expected total number to be discovered, y_2 is the year of greatest rate of description and v_1 is a scale parameter. Species description years were sampled as a random process that follows this trend but with random variation about it. The data were used to fit the logistic parameters and learn about the amount of random variability about the trend. These fitted values determined the prediction for the number of remaining species and the uncertainty in that prediction.

RESULTS

Rates of description

Description rates showed an initial ‘exploratory’ period of taxonomic expertise and knowledge development with few species described, followed by a period of publication of many species descriptions. In well-known taxa, this rate decreased as fewer species remained to be described, and the curve took a sigmoid shape. This complete pattern was evident in the better-known taxa, such as birds (Aves), mammals, and krill (Euphausacea, Crustacea) (Figure 1a). However, most taxa appeared to be increasing linearly (Figure 1, see Appendix for entire dataset).

The years that description curves began to increase, often linearly, were related to how well groups were known. Figure 1 provides examples of contrasting description rates for the more species rich taxa. From around the time of publication of Linne’s descriptions of 9,000 species in 1758, the rate of description of birds, mammals, echinoderms, bryozoans, ctenophores and free-living barnacles (Cirripedia) increased. For European seas, the description rates for molluscs, decapods, insects, scyphozoans, and boney (Osteichthyes) and cartilaginous (Chondrichtyes) fish have been increasing at a relatively constant rate since ca. 1750; the siphonophores, rotifers, acanthocephalans, ascidians and thalacians since 1775; digeneans, parasitic nematodes, polychaete worms, cestodes and sipunculans since 1800; and monogeneans, nemerteans, isopods, and siphonostomatoids since 1825. From about 1800 to 1840, description rates for anemones (Actinaria), hydroids (Hydrozoa), leeches (Hirudinea), stomatopods, parasitic barnacles (Rhizocephala, Cirripedia), and euphausiidaceans, increased. The description of some taxa began more recently. Pogonophores have only been described since 1960, gnathostomulids and entoprocts since 1940, gastrotrichs since 1920, tardigrades since 1900, oligochaetes, halacarid mites (Acarina), and myxozoans since 1875, and turbellaria since 1850. For all taxa combined, the rate of description was still high and linear (Figure 2).

The number of discoveries per year showed a decline during the two World Wars (Figure 3). The duration of this decline was about a decade in taxa such as Nematoda and Foraminifera, but descriptions of Mollusca were relatively low for 70 years (1900 to 1970's) (Figure 1 and Appendix). In plotting trend lines, we excluded the earliest species descriptions (largely by Linnaeus) and recent years to avoid any effects of delays in new species descriptions being recorded in the database, as did Dolphin and Quicke (2001) and Bebber et al. (2007). This time-delay was also found for crustaceans by Martin and Davis (2006). Apart from the world wars, there was no evidence of particular trends in the rate of description over time, such as may affect small numbers of species with the sampling of new habitats, new diagnostic tools, or greater taxonomic resources or efficiencies.

There was a strong positive correlation between the numbers of species described since 1900 with the number of species known per taxon ($r^2 = 0.878$). If the most species rich taxon (Mollusca) was omitted the correlation was still high ($r^2 = 0.6292$). Thus the more species in a taxon the more likely it was that that more new species would be described in that taxon. However, there was no significant relationship between the rate of description of new species and the number of species in a taxon ($r^2 = 0.0236$), nor between the number of species described before 1900 and the number known today ($r^2 = 0.1231$). Thus, while more species would be discovered in species rich taxa, richness did not necessarily indicate how well a taxon was known.

Estimating the unknown

To estimate the total number of marine species in Europe, described and undescribed, we first added the present list of 29,713 species, to the taxa not covered within Costello et al. (2001). The latter included about 300 species of non-halacarid mites (P.J.A. Pugh, pers. comm.), and perhaps 200 lichens (H. Fox, pers. comm.). The protists excluded from ERMS comprised about 4000 diatoms (D. Mann, pers. comm.), 1000 ciliates and 1000 other protist species including cyanobacteria (S. Brandt, pers. comm.). In addition the lists of Brachiopoda and Rotifera were not complete for the Mediterranean, although these were not species rich taxa. Adding the present list to the omitted taxa indicated there were about 36,213 marine species described from European seas. Wilson and Costello (2005) estimated that 4,230 species remained to be discovered for a selection of 12,763 of the more species rich taxa (Table 1). This was 33% of the total number of species for those taxa. However, Isopoda contributed disproportionally to this figure; for 656 isopod species the predicted number to be discovered had a median of 2,916, mean 4,353, and 95% confidence limits from 594 to 16,237. If the Isopoda were excluded the proportion remaining to be discovered was 11%. If this proportion was applied to the entire known list, it predicted an additional 3,930 species remained to be described, and that there could be about 40,000 marine species in Europe.

Because the standard deviation (SD) in the predicted number of species was dependent on the mean (r^2=0.6881), the coefficient of variation (CV = (mean/SD)) was used as an index of variation in the rate of description. The CV ranged from 1.15 to 2.04, and all but two of the 32 taxa for which the CV was computed had a value that was significantly more than 100% (CV = 1.0) (Table 1). This indicates that the rate of description is highly variable for most taxa, and more than would be accurately modelled by a Poisson process model.

When our model was applied to all of the 26,529 species (with year of description in ERMS) it predicted between 7,000 and 11,000 species remained to be discovered in European seas, with a median at around 8,500 (Table 2) (i.e. 32% more species, but this included Isopoda). The predictions for numbers to be discovered from between the end of the data in 1999 and 2010, and then 2020, ranged from 1,200 to 2,300, assuming the past rates of description continued. The Poisson model was broadly in agreement with the renewal model, although the latter predicted a slightly slower rate of description in the next 20 years but a slightly higher overall number. Our model indicated that the data show a higher variability about the mean trend than can be explained by the Poisson process of about 20%. For example, the standard deviation in the process was about 20% larger than would be allowed in a Poisson process (where the standard deviation must be the square root of the mean). The renewal model was a good fit to the data trend in descriptions and correctly modelled the amount of variation about the trend (Figure 2).

DISCUSSION

Socio-economic factors can affect the rate of description of taxa, such as the effect of the World Wars noted here and in other studies (Gaston et al. 1995, Finlay et al. 1996, Barnes 1989, Costello et al. 1990, 1996, Kelly and Costello 1996, Martin and Davis 2006). However, apart from the wars, and variation between years (Figure 3), the rates of description of species in the more species-rich taxa were consistently linear over time when plotted cumulatively (Figures 1, 2). Most description curves increased most rapidly in the late 19th century as found for global taxa in general (May 1992). This is despite changes in authors, exploration of new localities and habitats, application of new sampling methods, technologies, modern work practices; having more authors per species described, and notable impacts of a few authors who published many species in a few years (Martin and Davis 2006, Zapata and Robertson 2006, Eschemeyer et al. submitted). Even though more authors published more papers describing new tropical fish species in recent decades, the description rate has not changed (Zapata and Robertson 2006). An analysis of beetle species description rates similarly found no indication of changes in taxonomic effort since the mid-19th century, and suggested any increase in effort may be offset by the greater time needed to account for the past literature and rationalise synonymies (Frank and Curtis 1979). Similarly, the number of new crustaceans described per year since 1860 (Martin and Davis 2006), and new mollusc species listed in the Zoological Record has "remained remarkably stable" over the years 1960-1993 (Bouchet 1997). The numbers of new animal species in the Zoological Record (ZR) from 1979 to 1988 varied minimally between years (Hammond 1992). These findings increase confidence that past description trends are a reasonable predictor of future discoveries.

A difficulty in using some inventories of species names is that an unknown proportion would be synonyms (Frank and Curtis 1979, Solow et al. 1995, Bouchet 1997). This can be especially a problem for the taxa that receive more attention from amateurs, such as some insects, molluscs and fish. For example, there were an average of 1.62 names per species for world molluscs (Bouchet 1997), 5.4 names per freshwater fish species in Europe (Kottelat 1997), and 2 names per species of marine fish worldwide (Eschemeyer et al. submitted). The rate of synonymy of Thysanoptera and other insects has been 20% or more (Gaston et al. 1995, Solow et al. 1995); and on average, the rate of synonyms may be about 20% globally (Bouchet 2006). The proportion of names that are known to be synonyms decreases with time (Gaston et al. 1995, Solow et al. 1995). This is not necessarily because of improved practices, but because it takes time to recognise synonyms. Thus, analyses using species names need to account for the variation in synonymy rates over time as well as between taxa. These problems were minimised in ERMS because regional experts compiled the lists and reconciled synonymies. However, detection of further recent synonyms would decrease apparent recent description rates.

Similar patterns in the species description rates have been found for terrestrial species in Europe where about 600-700 new species have been described annually (Fontaine 2005), compared to 150 marine species in Europe per year since 1950. Mammals and birds, but also dragonflies and molluscs, had reached an asymptote (since 1850), but other taxa showed linear rates of description since 1750 (Lepidoptera, Amphibia, Reptilia), 1820's (Diptera, Nematoda, Trichoptera, Hymeoptera, Coleoptera, Annelida, Caridea, Cnidaria), or 1900 (Siphonaptera, Tardigrada, Acari, Myriapoda, Collembola, Nematomorpha, Gastrotricha) (Fontaine 2005). Globally, butterflies, birds and marine mammals have reached an asymptote (Robbins and Opler 1997, Bebber et al. 2007, Woodley et al. 2009).

About 11% or 33% of European marine species have been predicted to be undescribed, depending on whether the Isopoda are included or not. In either case, outside of European seas it appears that the proportion of undescribed species is much higher (e.g. Poore et al. 1994, 2008, Bouchet 1997, Koslow et al. 2001, Brandt et al. 2007). Lambshead and Boucher (2003) reported that 30% to 40% of free-living Nematoda in European seas found in field surveys were new to science, but that they were likely to have a similar number of species to Polychaeta. In contrast, the present study suggested only 3% of free-living Nematoda in Europe were undescribed, but 11% of Polychaeta. Polychaeta have 223 (12%) more species than free-living Nematoda and there was a positive relationship between number of species and number of species predicted to be discovered.

A comparison with the European list suggested only half of the Western Indian Ocean and South African species were described (Griffiths 2005). A survey of taxonomists opinions for 12 taxa in eight marine areas indicated that about 80% of species on the USA coast were described but only half of the species in the Eastern Pacific and Great Barrier Reef (Winston 1988). The deep-sea also contains many undescribed species. For example, Bouchet (1997) found that 20% of the NE Atlantic deep-water molluscs had been described only in the previous 20 years. A review of global mysid crustacean biodiversity found more species per unit area at lower than higher latitudes, and suggested the greater number described in the northern than southern hemispheres reflected sampling effort rather than reality (Wittman 1999). It has been estimated that there are 230,000 described marine species (Bouchet 2006). Thus the relatively well-studied European seas contained only about 15% of the world’s described marine fauna, and future discoveries in other oceans will reduce this proportion further. A recent analysis found similar numbers of species per unit seabed area and volume for the Mediterranean, Caribbean and Gulf of Mexico; but about half that number for Atlantic Europe (Costello et al. submitted). However, the same study found no significant relationship between marine species richness and area or volume across 24 regions of the world but noted that this may also reflect inadequate sampling and taxonomic effort, as well as possible biodiversity – area relationships in the oceans.

Considering the small contribution of the European marine species to the world total, and that the relative proportions of taxa appear to vary between regions and habitats, we do not attempt to extrapolate from the present findings to other oceans. A better approach would be to collate species checklists from other areas and conduct the same predictive analyses on them, and new online databases of species distribution data should enable this (Costello and Vanden Berghe 2006). These inventories should be stratified by sampling method (and thus habitat, life-style, and body size) and include species abundance, so as to use ‘sampling effort curves’ (Hammond 1994, Colwell and Coddington 1994). The availability of distribution data for more marine species may aid prediction of global richness by identifying patterns in turnover and richness ratios between taxa (e.g. parasites per host, copepods per fish species) (Hammond 1992). However, if the rates of description of taxa in southern hemisphere oceans are still in the exploratory phase, as the high proportion of undescribed species suggests, then description rate curves will have large confidence limits.

What this study has illustrated is that even in the best studied, and one of the more species poor oceans in the world (Poore and Wilson 1993), that significant numbers of species remain to be described in certain groups. Indeed, more species remain to be described in the most species rich taxa, indicating that we know least about the most diverse taxa. This has clear implications for marine research funding, in that priority should be given to taxonomic research on these groups. A review of about 150 years of research publications in Ireland found that the geographic areas and species studied most, continue to be studied most (Kelly and Costello 1995, 1996). A proactive approach to fill taxonomic gaps is thus necessary. Taxonomic effort needs to be directed to the most species rich invertebrate taxa in the less well-studied areas (Costello et al. 2006), notably the deep sea, tropics and southern hemisphere. A collaborative approach between the northern hemisphere countries which generally have the largest specimen collections and greatest resources for sampling in the deep sea and remote areas, and developing countries where taxonomic expertise is required to understand ecosystems, ecology and harvest natural resources, may be the most cost-effective way forward.

ACKNOWLEDGEMENTS

The foundation of this study was part-funded by the European Commission Marine Science and Technology research programme as part of the project “European Register of Marine Species” (MAS3-CT97-0146). We thank the many participants in that project for their cooperation that created this database (Costello 2000); Yde de Jong, Jacob van der Land, Howard fox, David Mann, sue Brandt, and Philip Pugh for helpful personal communications; and Gary Poore for his very helpful comments that improved the paper. Simon Wilson’s work was partly supported by the STATICA project, funded by the Principal Investigator Programme of Science Foundation Ireland, grant number 08/IN.1/I1879. This paper was finalised as a contribution to the Ocean Biogeographic Information System, a sub-programme of the Census of Marine Life; the European research network on Marine Biodiversity and Ecosystem Function (MarBEF); and the Pan European Species-directories Infrastructure (PESI), a project arising from the European Distributed Institute of Taxonomy (EDIT). This is MarBEF publication number MPS-09043.

REFERENCES

Appeltans W., Bouchet P., Boxshall G.A., Fauchald K., Gordon D.P., Hoeksema B.W., Poore G.C.B., van Soest R.W.M., Stöhr S., Walter T.C. & Costello M.J. (eds) (2009). World Register of Marine Species. www.marinespecies.org, Accessed 20 November (2009).

Barnes, R.D. (1989). Diversity of organisms: How much do we know? *American Zoologist* **29**, 1075-1084.

Bebber D.P., Marriott F.H.C, Gaston K.J. Harris S.A. & Scotland R.W. (2007). Predicting unknown species numbers using discovery curves. *Proceedings of the Royal Society* B **274**, 1651–1658.

Bisby F.A., Roskov Y.R., Orrell T.M., Nicolson D., Paglinawan L.E., Bailly N., Kirk P.M., Bourgoin T. & Baillargeon G. (eds) (2009). Species (2000) & ITIS Catalogue of Life: (2009) Annual Checklist. CD-ROM; Species (2000): Reading, UK.

Bouchet, P. (1997). Inventorying the molluscan diversity of the world. What is our rate of progress? *The Veliger* **40**, 1-11.

Bouchet P. (2006) The magnitude of marine biodiversity. In: Duarte CM (ed.), *The exploration of marine biodiversity: scientific and technological challenges.* Fundación BBVA, Madrid, 33-64.

Bouchet P., Lozouet P. & Sysoev A. (2009). An inordinate fondness for turrids. *Deep-Sea Research II* **56**, 1724-1731.

Brandt A., Gooday A. J., Brandão S.N., Brix S., Brökeland W., Cedhagen T., Choudhury M., Cornelius N., Danis B., De Mesel I., Diaz R.J., Gillan D.C., Ebbe B., Howe J.A, Janussen D., Kaiser S., Linse K., Malyutina M., Pawlowski J., Raupach M. & Vanreusel A. (2007). First insights into the biodiversity and biogeography of the Southern Ocean deep sea. *Nature* **447**, 307-311.

Bunge J. & Fitzpatrick M. (1993). Estimating the number of species - a review. *Journal of the American Statistical Association* 88, 364-373.

Carlton, J.T., Geller J.B., Reaka-Kudla M. & Norse E.A. (1999). Historical extinctions in the sea. *Annual Reviews in Ecology and Systematics* **30**, 515-538.

Chapin III, F.S., Zavaleta E.S., Eviner V.T., Naylor R.L., Vitousek P.M., Reynolds H.L., Hooper D.U., Lavorel S., Sala O.E., Hobble S.E., Mack M.C. & Diaz S. (2000). Consequences of changing biodiversity. *Nature* **405**, 234-242

Cohen, D.M. (1970). How many recent fishes are there? *Proceedings of the California Academy of Sciences Forth Series* **38**,341-346.

Colwell R.K. & Coddington J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. R. Soc. Lond. B* **345**, 101-118.

Cooper R.A., Maxwell P.A., Crampton J.S., Beu A.G., Jones C.M. & Marshall B.A. (2006). Completeness of the fossil record: estimating losses due to small body size. *Geology* **34**, 241-244.

Costello, M.J. (2000). Developing species information systems: the European Register of Marine Species. *Oceanography* 13, 48-55.

Costello, M. J. (2001). To know, research, manage, and conserve marine biodiversity. *Océanis* **24** (4), 25-49.

Costello M.J., Holmes J.M.C., McGrath D. & Myers A.A. (1990). A review and catalogue of amphipod Crustacea in Ireland. *Irish Fisheries Investigation Series* B, No. 33 (1989), 70 pp.

Costello M.J., Emblow C.S. & Picton B.E. (1996). Long term trends in the discovery of marine species new to science which occur in Britain and Ireland. *Journal of the marine biological Association of the United Kingdom* **76**, 255-257.

Costello, M.J., Emblow, C. & White R. (editors) (2001). European Register of Marine Species. A check-list of marine species in Europe and a bibliography of guides to their identification. *Patrimoines naturels* **50,** 1-463.

Costello, M.J. & Vanden Berghe E. (2006). "Ocean Biodiversity Informatics" enabling a new era in marine biology research and management. *Marine Ecology Progress Series* **316**, 203-214.

Costello, M.J., Emblow C.S., Bouchet P. & Legakis A. (2006). European marine biodiversity inventory and taxonomic resources: state of the art and gaps in knowledge. *Marine Ecology Progress Series* **316**, 257-268.

Costello M.J., Coll M., Danovaro R., Halpin P., Ojaveer H., & Miloslavich P. A. census of marine biodiversity knowledge, resources and future challenges . PLoS One , submitted.

Dolphin K. & Quicke D.L.J. (2001). Estimating the global species richness of an incompletely described taxon: an example using parasitoid wasps (Hymenoptera: Braconidae). *Biological Journal of the Linnean Society* **73**, 279-286.

Dulvy, N.K., Sadovy, Y. & Reynolds, J.D. (2003). Extinction vulnerability in marine populations. *Fish and Fisheries* **4,** 25-64

Embley T.M., Hirt R.P. & Williams D.M. (1994). Biodiversity at the molecular level: the domains, kingdoms and phyla of life. *Philosophical Transactions of the Royal Society of London* B **345**, 21-33.

Erwin, T.L (1988). The tropical rainforest canopy: the heart of biotic diversity. In: *Biodiversity,* E.O. Wilson (ed), National Academy Press, Washington D.C., 123-129 pp.

Eschemeyer W N, Fricke R, Fong JD, & Polack D Marine fish biodiversity: a history of knowledge and discovery (Pisces). Submitted for publication.

Finlay B.J., Corliss J.O., Esteban G. & Fenchel, T. (1996). Biodiversity at the microbial level: the number of free living ciliates in the biosphere. *Quarterly Journal of Biology* **71**, 221–237.

Finlay B.J., Thomas J.A., McGavin G.C., Fenchel T. & Clarke R.T. (2006). Self-similar patterns of nature: insect diversity at local to global scales. *Proceedings of the Royal Society* B **273**, 1935-1941.

Fontaine, B. (2005). Fauna Europaea gap analysis. Final Report, Fauna Europaea project. Unpublished report, National Museum d'Histoire Naturelle, Paris.

Frank J.H. & Curtis G.A. (1979). Trend lines and the number of species of Staphylinidae. *The Coleopterists Bulletin* **33**, 133-149.

Gaston K.J. (1991). Body size and probability of description: the beetle fauna of Britain. *Ecological Entomology* **16,** 505-508.

Gaston K.J. & Blackburn T.M.. (1997). A critical assessment of the form of the interspecific relationship between abundance and body size in animals. *Journal of Animal Ecology* **66,** 233-

Gaston K.J., Scoble M.J. & Crook A. (1995). Patterns in species description: a case study using the Geometridae (Lepidoptera). *Biological Journal of the Linnean Society* **55,** 225-237.

Gibbons M.J., Richardson A.J., Angel M.V., Buecher E., Esnal G., Fernandez Almon M.A., Gibson R., Itoh H., Pugh P., Boettger-Schnack R., & Thuesen E. (2005). What determines the likelihood of species discovery in marine holozooplankton: is size, range or depth important? *Oikos* **109**, 567-576.

Grassle J.F. & Maciolek N.J. (1992). Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. *The American Naturalist* **139**, 313-341.

Griffiths C.L. (2005). Coastal marine biodiversity in East Africa. *Indian Journal of Marine Sciences* **33**, 35-41.

Guan S. (2007). New species show how little we know of the sea. *Nature* **446**, 492.

Hagström Å, Pommier T, Rohwer F, Simu K, Stolte W, Svensson D, & Zweifel UL. (2002). Use of the 16S ribosomal DNA for delineation of marine bacterioplankton species. *Applied and environmental microbiology* **68**, 3628-3633.

Hall S.J. & Greenstreet S.P. (1996). Global diversity and body size. *Nature* **383**, 133.

Hammond P.M. (1992). Species inventory. In: Groombridge B. (Ed), *Global biodiversity: status of the Earth's living resources.* Chapman and Hall, London, 17-39.

Hammond P.M. (1994). Practical approaches to the estimation of the extent of biodiversity in speciose groups. *Philosophical Transactions of the Royal Society of London* B **345**, 119-136.

Jackson J.B.C. (2008). Ecological extinction and evolution in the brave new ocean. *Proceedings of the National Academy of Sciences* **105**, 11458-11465.

Kelly, K.S., & Costello, M.J. (1995). Marine related papers published in the Irish Naturalists' Journal, 1925-1993. *Irish naturalists' Journal* **25**, 89-98.

Kelly, K.S., & Costello, M.J. (1996). Temporal trends and gaps in marine publications in Irish periodicals. In: Keegan, B.F. and O'Connor, R. Editors, *Irish Marine Science (1995)*, Galway University Press, Galway, 37-48.

Koslow, J. A., Gowlett-Holmes K., Lowry J.K., O'Hara T., Poore G.C.B. & Williams A. (2001). Seamount benthic macrofauna off southern Tasmania: community structure and impacts of trawling. *Marine Ecology Progress Series* **213**, 111-125.

Kottelat, M. (1997) (1997). European freshwater fishes. *Biologia, Bratislava.* 52/Supplement 5, 1-23

Lambshead, P.J.D. & Boucher G. (2003). Marine nematode deep-sea biodiversity – hyperdiverse or hype? *Journal of Biogeography* **30,** 475-485.

Marshall C.R. (2006). Fossil record reveals tropics as cradle and museum. *Science* **314**, 66-67.

Martin J.W. & Davis G.E. (2006). Historical trends in crustacean systematics. *Crustaceana* **79**, 1347-1368.

Mauchline J. & Marano M. (1977). World list of the Mysidacea Crustacea. *Journal of the Tokyo University of Fisheries* **64**, 39-88

May, R.M. (1988). How many species are there on Earth? *Science* **241**, 1441-1449.

May, R.M. (1992). Past efforts and future prospects towards understanding how many species there are. In: Solbrig, O. T., van Emden, H. M. and van Oordt, P.G.W. (editors), *Biodiversity and global change*. International Union for Biological Sciences Monograph No. 8, Paris, 71 –81.

May R.M. (1994). Conceptual aspects of the quantification of the extent of biological diversity. *Philosophical Transactions of the Royal Society of London* B **345**, 13-20.

O'Donnell A.G., Goodfellow M., & Hawksworth D.L. (1994). Theoretical and practical aspects of the quantification of biodiversity among microorganisms. *Philosophical Transactions of the Royal Society of London* B **345**, 65-73.

Orme C.D.L., Quicke D.L.J., Cook J.M., & Purvis A. (2002). Body size does not predict species richness among the metazoan phyla. *Journal of Evolutionary Biology* **15**, 235-247.

Paxton, C.G.M. (1998). A cumulative species description curve for large open water marine animals. *Journal of the marine biological Association of the United Kingdom* **78**, 1389-1391.

Poore, G.C.B. & Wilson, G.D.F. (1993). Marine species richness. *Nature* **361**, 597-598.

Poore, G.C.B., Just, J., & Cohen, B.F. (1994). Composition and diversity of Crustacea Isopoda of the southeastern Australian continental slope. *Deep-Sea Research* **41**, 677-693.

Poore, G.C.B., McCallum, A.W. & Taylor, J. (2008). Decapod Crustacea of the continental margin of southwestern and central Western Australia: preliminary identifications of 524 species from FRV Southern Surveyor voyage SS10-2005. *Museum Victoria Science Reports* 11, 1-106.

Reaka-Kudla, M. (1996). The global biodiversity of coral reefs: a comparison with rain forests. In: Reaka-Kudla, M.L., Wilson, D.E. and Wilson E.O. (editors), *Biodiversity II, Understanding and protecting our biological resources.* Joseph Henry Press, Washington D.C., 83-108.

Robbins R.K. & Opler P.A. (1997). Butterfly diversity and a preliminary comparison with bird and mammal diversity. In: Wilson DE, Reaka-Kudla ML, Wilson EO (Eds.), *Biodiversity II, understanding and protecting our biological resources.* Joseph Henry Press, Washington DC, 69-82.

Solow A.R., Mound L.A., & Gaston K.J. (1995). Estimating the rate of synonymy. *Systematic Biology* **44**, 93-96.

Solow A.R. & Smith W.K. (2005). On estimating the number of species from the discovery record. *Proceedings of the Royal Society* B **272**, 285-287.

Stork N.E. (2007). World of insects. *Nature* **448/9**, 657-658.

Wittman, K.J. (1999). Global biodiversity in Mysidacea, with notes on the effects of human impact. In: Schram F.R. and Vaupel Klein J.C. von (editors), *Crustaceans and the biodiversity crisis. Proceedings of the fourth international crustacean congress, Amsterdam, The Netherlands, July 20-24, (1998) (1998), vol. 1.* Brill, Leiden, 511-525.

Wilson, S. P. & Costello, M.J. (2005). Predicting future discoveries of European marine species by using a non-homogeneous renewal process. *Applied Statistics* **54** (5), 897-918.

Winston, J.E. (1988). The systematists' perspective. *Memoirs of the California Academy of Sciences* **13**, 1-6.

Woodley M.A., Naish D., & Shanahan H.P. (2009). How many extant pinniped species remain to be discovered? *Historical Biology* **1,** 1-11.

Zapata F.A. & Robertson D.R. (2007). How many species of shore fishes are there in the Tropical Eastern Pacific? *Journal of Biogeography* **34,** 38-51.

Table 1. The numbers of species per taxa from the European Register of Marine Species compiled in 1999 (Costello *et al.*2001), and percent described by 1900. Abbreviations: nd = no data because the year a species was described was not provided for species in the list; CL = confidence limits. Cephalorhyncha included Loricifera, Priapulida, Kinorhyncha, and Nematomorpha. Macroalgae included the phyla Rhodophycota, Phaeophycota, Chlorophycota, and two genera of Xanthophycota.

	Species group name		1999	% by 1900	Predicted number of species		Coefficient of variation
					Mean	95% CL	Mean (95% CL)
	Cnidaria	Cubozoa	1	100	-	-	
	Minor phylum	Cycliophora	1	100	-	-	
	Crustacea	Remipedia	1	100	-	-	
	Protozoa (protists)	Kathablepharids	2	0	-	-	
	Minor phylum	Placozoa	2	100	-	-	
	Crustacea	Branchiura	2	0	-	-	
	Crustacea	Pentastomida	2	50	-	-	
	Crustacea	Mystacocarida	2	0	-	-	
	Crustacea	Copepoda – Mormonilloida	2	100	-	-	
	Crustacea	Thermosbaenacea	2	100	-	-	
	Arthopoda	Diplopoda	2	100	-	-	
	Minor phylum	Cephalochordata	2	100	-	-	
	Protozoa (protists)	Apicomplexa (free-living species)	3	100	-	-	
	Protozoa (protists)	Apusomonads	3	0	-	-	
	Crustacea	Copepoda – Platycopioida	3	0	-	-	
	Protozoa (protists)	Stramenopiles incertae sedis	4	25	-	-	
	Crustacea	Aspidogastrea	4	100	-	-	
	Flowering plants	Seagrass	5	nd	-	-	
	Vertebrata	Pisces – Agnatha	5	80	-	-	
	Vertebrata	Tetrapoda – Reptilia	5	100	-	-	
	Arthopoda	Chilopoda	6	50	-	-	
	Crustacea	Cladocera – Branchiopoda	9	100	-	-	
	Minor phylum	Phoronida	9	44	-	-	
	Protozoa (protists)	Labyrinthulids	10	30	-	-	
	Crustacea	Cirripedia – parasitic Ascothoracida	10	30	-	-	
	Algae	Euglenids – kinetoplastids	13	38	-	-	
	Crustacea	Tantulocarida	13	0	-	-	
	Algae	Crytophytes	14	-	-	-	
	Protozoa (protists)	Thaustrochytrids	15	0	-	-	

	Arthopoda	Insecta – Chironomidae	15	40	-	-	
	Crustacea	Copepoda – Misophrioida	16	6	-	-	
	Protozoa (protists)	Bicosoecids	17	18	-	-	
	Protozoa (protists)	Thaumatomonads	17	0	-	-	
	Minor phylum	Hemichordata	17	59	-	-	
	Minor phylum	Brachiopoda	*18	100	-	-	
	Minor phylum	Echiura	19	74	2.2	0-19	1.15 (0.85-1.57)
	Arthopoda	Insecta	19	100	-	-	
	Protozoa (protists)	Xenophyophora	20	30	-	-	
	Crustacea	Stomatopoda	22	50	44.8	2-210	1.52 (1.12-2.00)
	Polychaeta (formerly a Minor phylum)	Pogonophora	23	0	57.1	0-269	1.22 (0.85-1.69)
	Algae	Prasinophytes	24	nd	-	-	
	Minor phylum	Gnathostomulida	25	0	37	0-280	1.25 (0.92-1.96)
	Protozoa (protists)	Euglenids – heterotrophic	26	27	-	-	
	Cnidaria	Antipatharia	28	61	0.2	0-1	1.41 (1.15-1.78)
	Crustacea	Cirripedia – parasitic Rhizocephala	28	46	13	0-68	1.28 (1.00-1.64)
	Protozoa (protists)	Ciliates – folliculinids	30	3	-	-	
	Crustacea	Copepoda – Monstrilloida	33	36	266.4	3-1346	1.77 (1.06-2.45)
	Tunicate	Thaliacea	35	71	-	-	
	Algae	Haptophytes	36	3	-	-	
	Minor phylum	Mesozoa	36	50	-	-	
	Annelida	Hirudinea	36	58	17.5	0-2	1.46 (1.16-1.91)
	Protozoa (protists)	Ciliates – Chonotricha	37	14	-	-	
	Cnidaria	Octocorallia - Pennatulacea	37	78	14.8	0-96	1.43 (1.14-1.77)
	Cnidaria	Ctenophora	38	76	30.9	0-211	1.52 (1.20-1.91)
	Protozoa (protists)	Protista incertae sedis (heterotrophic species)	40	13	-	-	
	Crustacea	Euphausiacea	41	61	0.08	0-1	1.51 (1.27-1.82)
	Protozoa (protists)	Ciliates – Rhynchodida	42	7	-	-	
	Minor phylum	Chaetognatha	42	38	-	-	
	Minor phylum	Sipuncula	44	52	-	-	
	Minor phylum	Entoprocta	45	27	85	4-468	1.95 (1.45-2.43)
	Vertebrata	Tetrapoda – Mammalia	50	98	0.3	0-2	1.42 (1.20-1.70)
	Minor phylum	Cephalorhyncha	52	17	-	-	
	Cnidaria	Scyphozoa	53	64	-	-	
	Tunicata	Appendicularia	53	60	39.4	1-178	1.94 (1.59-2.34)
	Crustacea	Isopoda, Epicaridea, Bopyridae	54	52	-	-	
	Minor phylum	Acanthocephala	67	36	-	-	
	Protozoa (protists)	Amoebae – naked	74	8	-	-	
	Vertebrata	Tetrapoda – Aves	74	97	0.3	0-2	1.61 (1.39-1.87)

	Minor phylum	Tardigrada	76	1	-	-	
	Protozoa (protists)	Ciliates – aloricate oligotrichs	82	20	19.9	6-43	1.66 (1.52-1.83)
	Cnidaria	Scleractinia	86	70	160.1	3-789	2.02 (1.69-2.36)
	Cnidaria	Octocorallia (excl. Pennatulacea)	92	65	183.2	11-880	1.85 (1.59-2.15)
	Protozoa (protists)	Amoebae – testate	97	16	-	-	
	Protozoa (protists)	Choanoflagellates	98	15	-	-	
	Cnidaria	Siphonophora	105	49	-	-	
	Crustacea	Cirripedia – non-parasitic Thoracica	107	81	303.9	15-1280	2.04 (1.18-2.45)
	Minor phylum	Rotifera	139	58	-	-	
	Vertebrata	Pisces – Chondrichthyes	145	72	-	-	
	Arthopoda	Pycnogonida	146	55	-	-	
	Crustacea	Copepoda – Cyclopoida	177	32	-	-	
	Crustacea	Cumacea	188	52	-	-	
	Annelida	Oligochaeta	190	16	-	-	
	Crustacea	Mysidacea	198	44	-	-	
	Nematoda	Nematoda – parasitic	212	49	-	-	
	Arthopoda	Acarina – Halacaridae	214	27	-	-	
	Cnidaria	Myxozoa	230	15	-	-	
	Minor phylum	Gastrotrichia	240	3	-	-	
	Cnidaria	Actiniaria	243	57	6.4	1-14	1.39 (1.28-1.51)
	Crustacea	Tanaidacea	280	30	-	-	
	Platyhelminthes	Cestoda	312	56	-	-	
	Fungi	Fungi	318	30	-	-	
	Platyhelminthes	Monogenea	353	35	91.2	48-176	1.39 (1.28-1.50)
	Crustacea	Copepoda – Poecilostomatoida	353	48	-	-	
	Crustacea	Copepoda – Siphonostomatoida	354	60	14.6	5-29	1.40 (1.29-1.54)
	Tunicata	Ascidiacea & Sorberacea	393	46	-	-	
	Minor phylum	Nemertea (Nemertini)	478	40	28.5	13-50	1.53 (1.43-1.65)
	Platyhelminthes	Digenea	592	34	-	-	
	Crustacea	Isopoda – excluding Epicaridea	605	39	4553	594-16237	2.26 (2.08-2.42)
	Echinodermata	Echinodermata	648	67	-	4-21	1.31 (1.22-1.40)
	Crustacea	Copepoda – Calanoida	649	34	-	-	
	Crustacea	Decapoda	672	70	-	-	
	Cnidaria	Hydrozoa	684	64	39.4	21-65	1.36 (1.26-1.47)
	Algae	Dinoflagellates	718	nd	-	-	
	Minor phylum	Bryozoa	724	61	-	-	
	Crustacea	Ostracoda	769	53	-	-	
	Platyhelminthes	Turbellaria	1137	18	-	-	
	Protozoa (protists)	Foraminifera	1167	56	40.5	26-58	1.52 (1.46-1.57)
	Crustacea	Amphipoda	1183	57	75.7	51-106	1.34 (1.28-1.39)
	Vertebrata	Pisces – Osteichthyes	1199	69	132.2	69-209	1.63 (1.54-1.73)
	Crustacea	Copepoda –	1357	64	207.4	155-271	1.27 (1.22-1.32)

		Harpacticoida					
	Nematoda	Nematoda, free living	1625	12	48.2	31-68	1.38 (1.32-1.44)
	Porifera	Porifera	1640	55	30.6	18-45	1.35 (1.30-1.41)
	Algae	Macroalgae	1702	nd	-	-	
	Annelida	Polychaeta	1848	54	-	-	
	Mollusca	Mollusca	3353	69	-	-	
		TOTAL	**29713**	**45**	**1935.4**	**0-13206**	
		AVERAGE	258	47	62.4		

Table 2. Summary of the predictions for the number of marine species to be discovered, assuming a logistic mean trend, according to the renewal process model of Wilson and Costello (2005) and the Poisson process.

	Wilson and Costello (2005)			**Poisson process**		
Number species to be discovered	Mean	Median	95% Probability	Mean	Median	95% Probability
Total	8670	8580	(7000, 10840)	9040	8980	(7110, 11530)
By 1999–2010	1280	1280	(1120, 1450)	1260	1260	(860, 1430)
By 1999–2020	2320	2310	(2020, 2630)	2290	2300	(1990, 2600)

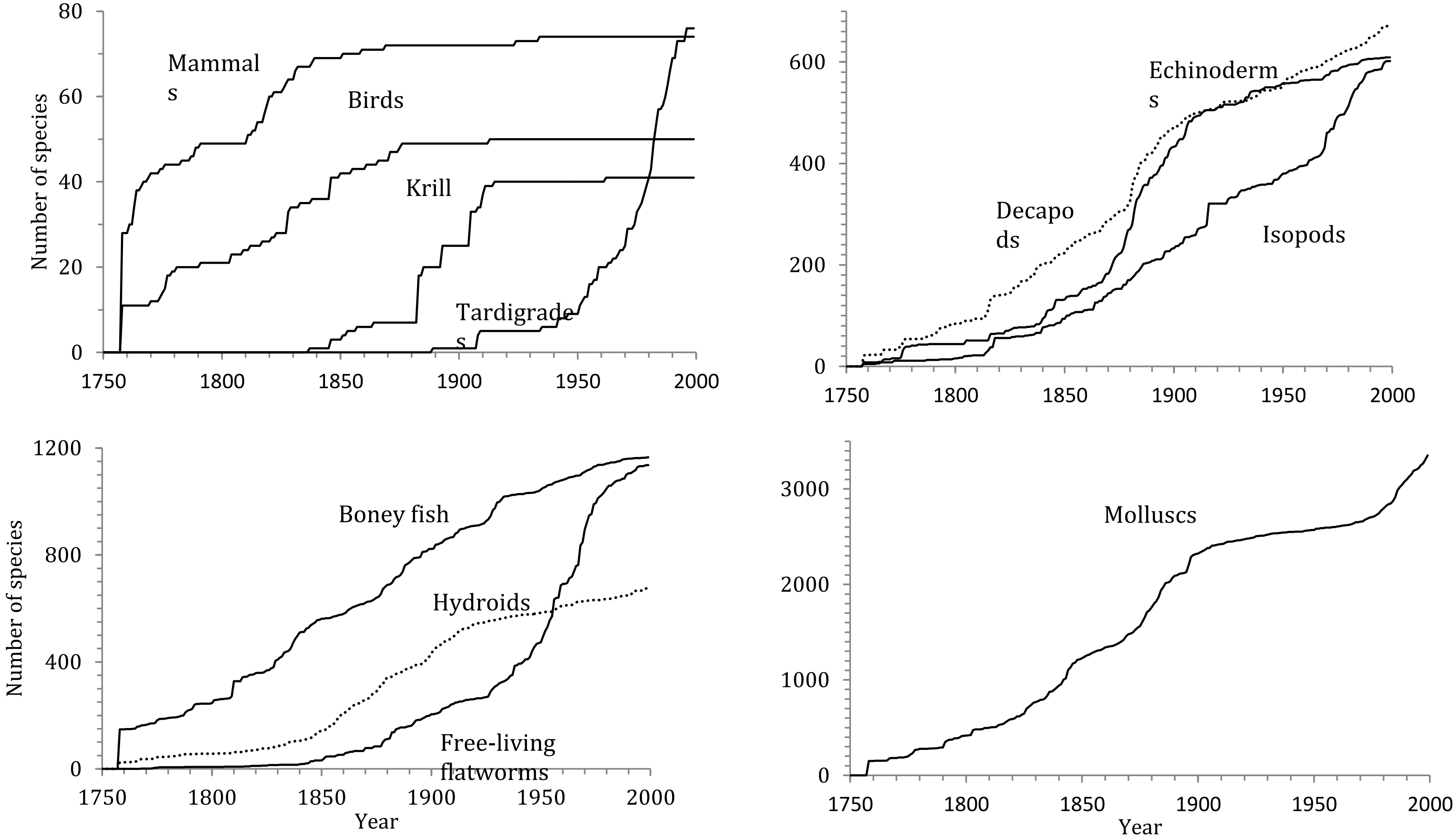

Figure 1. Cumulative number of species described for a selection of taxa contrasting in their pattern over time and number of species (a) mammals, birds, krill, tardigrades, (b) echinoderms, decapod and isopod crustaceans, (c) boney fish (Osteichthyes), hydroids and free-living flatworms (Turbellaria), and (d) molluscs. Note scale varies on vertical axis.

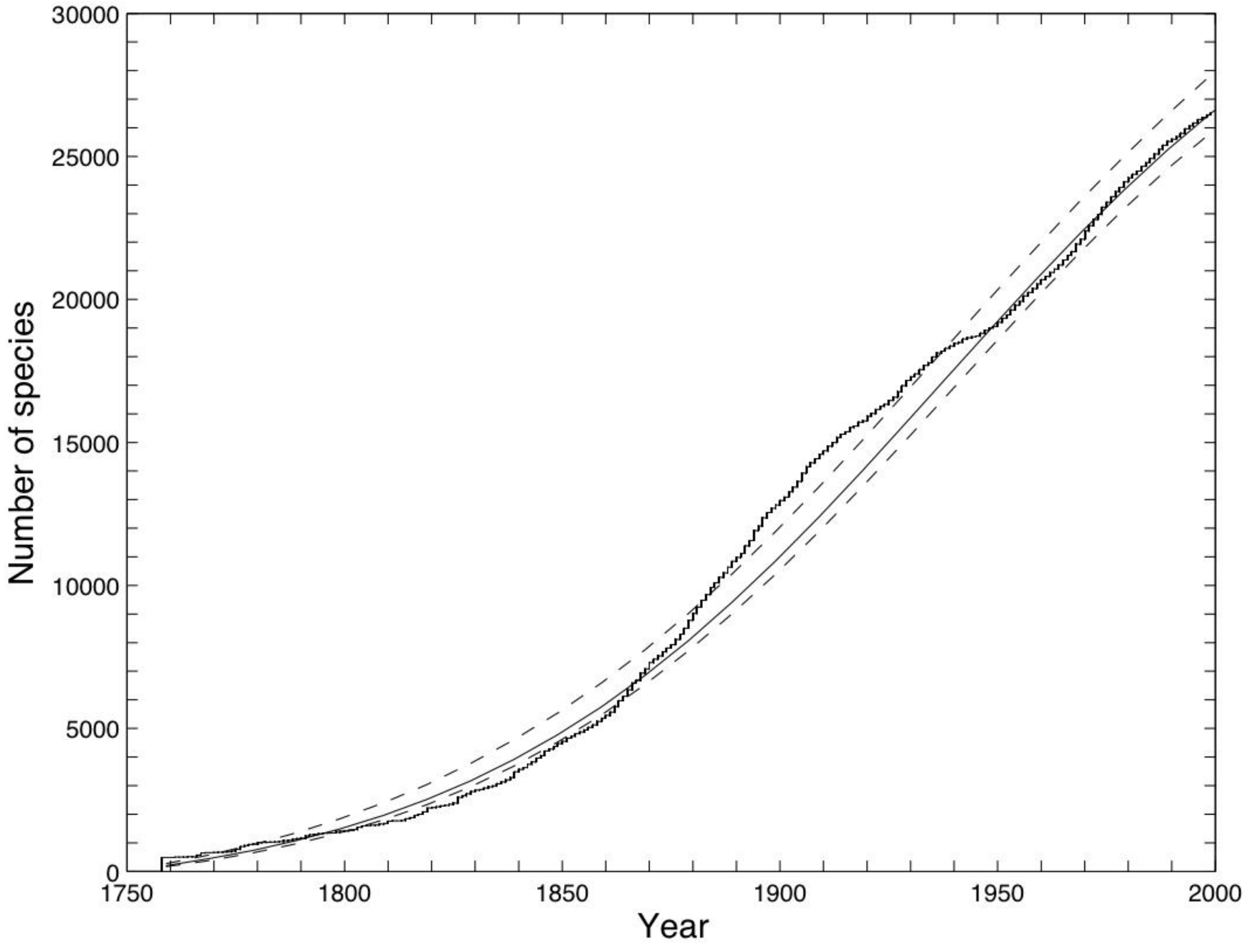

Figure 2: Description curve for 26,529 species (solid line) along with renewal process model fit (dashed lines); upper and lower dashed lines are 2.5 and 97.5 percentiles of fitted model distribution for number of species described, and central line is the median.

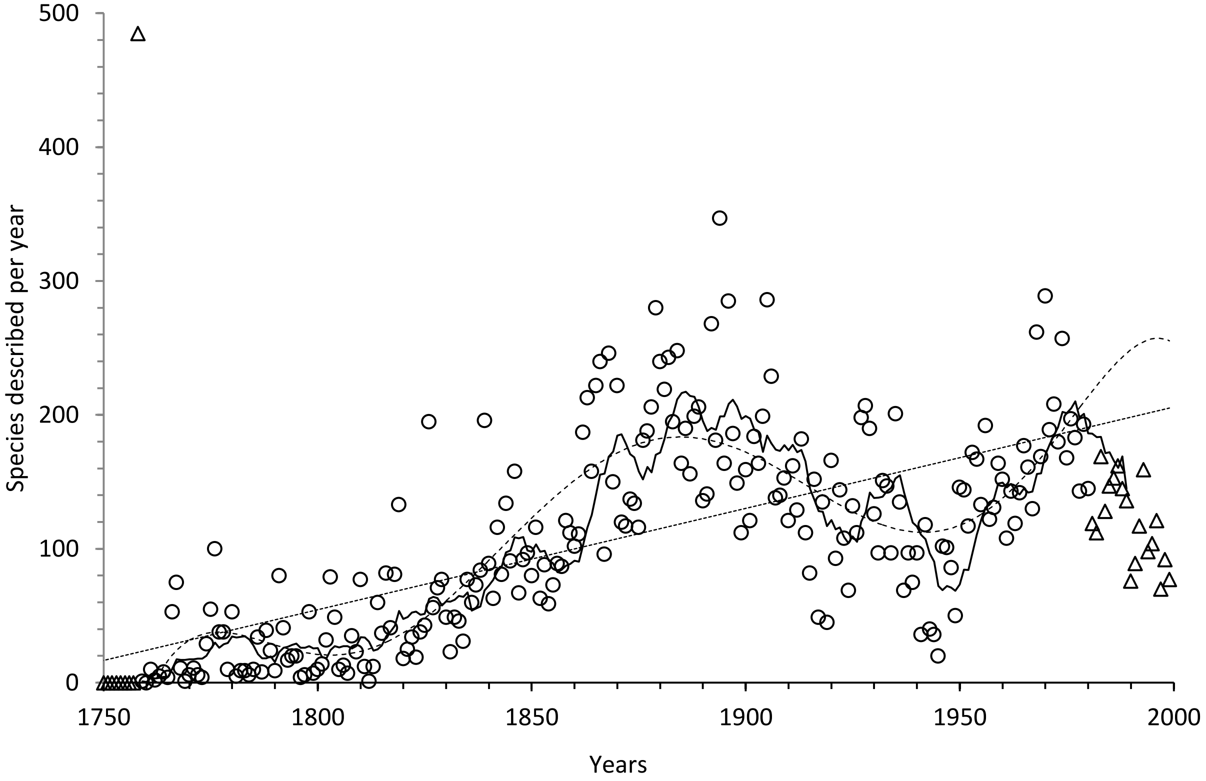

Figure 3. The number of new species described each year from 1747 to 1999 and reported from European seas. The lines plotted are a 10 year moving average (solid line), linear (r^2=0.320; dotted line) and 6th order polynomial (r^2 = 0.662; dashed line). The trend lines exclude species described prior to 1759 and after 1980 to avoid the effects of Linnaeus initially describing a lot of species at once, and the delayed entry of recently described species into the inventories; these data are shown as triangles. The polynomial line decreased if 1980's data was included.