



UvA-DARE (Digital Academic Repository)

Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale

Kissling, W.D.; Ahumada, J.A. ; Bowser, A.; Fernandez, M.; Fernández, N.; Garcia, E.A.; Guralnick, R.P.; Isaac, N.J.B.; Kelling, S.; Los, W.; McRae, L.; Mihoub, J.-B.; Obst, M.; Santamaria, M.; Skidmore, A.K.; Williams, K.J.; Agosti, D.; Amariles, D.; Arvanitidis, C.; Bastin, L.; De Leo, F.; Egloff, W.; Elith, J.; Hobern, D.; Martin, D.; Pereira, H.M.; Pesole, G.; Peterseil, J.; Saarenmaa, H.; Schigel, D.; Schmeller, D.S.; Segata, N.; Turak, E.; Uhlir, P.F.; Wee, B.; Hardisty, A.R.

Published in:
Biological Reviews of the Cambridge Philosophical Society

DOI:
[10.1111/brv.12359](https://doi.org/10.1111/brv.12359)

[Link to publication](#)

License
CC BY

Citation for published version (APA):

Kissling, W. D., Ahumada, J. A., Bowser, A., Fernandez, M., Fernández, N., Garcia, E. A., Guralnick, R. P., Isaac, N. J. B., Kelling, S., Los, W., McRae, L., Mihoub, J.-B., Obst, M., Santamaria, M., Skidmore, A. K., Williams, K. J., Agosti, D., Amariles, D., Arvanitidis, C., ... Hardisty, A. R. (2018). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological Reviews of the Cambridge Philosophical Society*, 93(1), 600–625. <https://doi.org/10.1111/brv.12359>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://www.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Appendix S1: Information on projects

Examples of projects with EBV-relevant data products

eBird

The eBird project (<http://ebird.org/content/ebird/>), launched in 2002 by the Cornell Lab of Ornithology, is one of the largest online citizen science programs aiming to collect basic information on bird species distributions, abundances and trends (Kelling *et al.*, 2015; Sullivan *et al.*, 2014, 2009; Wood *et al.*, 2011). The project gathers bird observations through checklists that record the number of individuals of each species observed at a specific date and location, data-collection variables (e.g. observer identity, start time, location, duration, travelled distance), and whether all species observed were reported (i.e. absence). The quantity of eBird data makes it one of the largest biodiversity data sets. However, the distribution of search effort across space and time is irregular, making it difficult accurately to infer directly from the data how patterns of occurrence or abundance are defined in certain areas during specific times of the year. To address this, continental-scale species distribution models (SDMs) estimate the relative abundance across a species' entire life history for every week of the year (Fink *et al.*, 2014) and can be used to identify species' land cover associations (Zuckerberg *et al.*, 2016). These models control for variation in species' detectability as a function of search effort. eBird has gathered sufficient data to model approximately 1,500 bird species from 2004 to the present. While models are currently generated only for the Western Hemisphere, data collection is global and eBird is making a considerable contribution to the collection and sharing of global bird records (Amano *et al.*, 2016).

TEAM wildlife monitoring surveys

The TEAM network (<http://www.teamnetwork.org/>) is a consortium of three core and more than 80 implementing partners. TEAM implements several standardized monitoring protocols, including a wildlife monitoring protocol for camera trap data (Jansen *et al.*, 2014). The camera-trap protocol is implemented in 23 tropical forest sites across the world, and monitors approximately 500 populations consisting of 250 species of ground-dwelling mammals and birds. At each site, the occupancy (not abundance) of each population is estimated using a hierarchical Bayesian model that estimates occupancy and detection probability (Ahumada *et al.*, 2013; Beaudrot *et al.*, 2016), which provides a detection history for each population since 2007. The data collected can be used to build an EBV data product for this group of mammals and birds in tropical forest protected areas, and the modelled occupancy data can be aggregated by region, taxonomic or functional groups. One current limitation for EBV use is that occupancy time series are calculated at the 'site level' (usually a protected area) and are not easily interpolated to other unsampled areas. However, as the network grows with more sites, this gap-filling could be eventually accomplished with modelling techniques. A recent analysis of the TEAM data showed that one-third of the investigated ground-dwelling vertebrates had stable occupancy dynamics, one-third increased, and one-third declined (Beaudrot *et al.*, 2016).

Living Planet Index

The Living Planet Index (LPI) data set (Collen *et al.*, 2009; Loh *et al.*, 2005) currently consists of over 18,500 time-series records, some starting in 1970, of more than 3,700 vertebrate species worldwide. The time-series data capture the change in population size, or a proxy, for a species monitored in a given location. Data primarily originate from peer-reviewed literature and online databases of long-term monitoring schemes. Data types are

heterogeneous and can reflect full counts of individuals, breeding pairs, density estimates, biomass or an index of abundance. Monitoring methods also vary, but are consistent within each time series. LPI data are from different spatial extents (local to global), with a bias towards Europe, North America, and Southern or Eastern Africa. The temporal resolution consists of annual measures of abundance, but not always for consecutive years. One limitation relating to the building of EBV data products is that the LPI data set does not represent a single systematic monitoring scheme that is updated periodically, but a collation of many individual monitoring studies that vary in length of time, spatial extent and resolution. In the production of the LPI as an indicator of abundance change, this is tackled by calculating relative annual changes in population sizes rather than absolute change. As such, the LPI does not represent the particular state of biodiversity in a given place or time. Rather, it is most commonly used as an aggregated indicator of changes and trends in species abundance. Most of the time series are monitored at small spatial scales. They could be aggregated in different ways geographically, taxonomically or by habitat, and could potentially be used to extrapolate information spatially, temporally and taxonomically.

Baltic Sea zooplankton monitoring

The Baltic Sea zooplankton monitoring (BALTIC) data set uses raw data from various national plankton monitoring programs in the Baltic region, as aggregated by the Swedish LifeWatch initiative (Gärdenfors *et al.*, 2014). The monitoring programs in the Baltic region are highly consistent with regard to the sampling protocols and currently include more than 60,000 abundance measurements collected in 26 stations during 1979–2013 (www.sharkdata.se). We used these data and analysed abundance trends across space and time. Details about data collection and data processing are given below.

Statistical analysis of mean abundances from the refined data set suggests that in principle it is possible to obtain EBV data products from aggregated marine data sets, but only for a few abundant and typical indicator species (Table S1). Neither sufficient nor consistent data exist to make EBV data products for rare species, species with taxonomic uncertainties, and species requiring specialists for identification. This example shows that basic taxonomic information is lacking, despite the high quality of the Baltic monitoring data and the long-term record at the regional scale. Aggregation of information at higher taxonomic levels, such as to the class/order designation, are therefore necessary, and this obstructs analysis of species-level trends. In addition, it is very common that marine organisms are identified to different taxonomic levels across programs, which likewise diminishes the comparability of abundance records across time and space. For these reasons it seems unlikely that EBV data products can be produced for a broad range of marine species from existing data sets. Although relevant biodiversity changes could be detected in some cases using EBVs data products prepared at a taxonomic resolution above the species level, this underscores the need to increase the amount of species-level measurements in marine biodiversity monitoring programs. High-throughput approaches, such as optical scanning (Gorsky *et al.*, 2010) and metabarcoding methods (Bourlat *et al.*, 2013), may provide technical solutions to overcome these problems in the future.

Details about data collection and data processing of the BALTIC data set

Data collection

After short exploration of various data sets available for parts of the European seas, covering longer time periods (e.g. benthos, fish, plankton, etc), we decided to analyse a zooplankton data set compiled by the Swedish LifeWatch initiative (www.svenskalifewatch.se) from national monitoring programs in the Baltic region across 26 stations between 1979 and 2013

(including the transitional regions Skagerrak and Kattegat). The overall data set included 60,511 observation records accessible at www.sharkdata.se. More information about the SHARKdata system is available at <http://sharkdata.se/documentation/>. Some of the data processing was conducted separately by the Emodnet initiative (for documentation see <http://rshiny.emodnet-biology.eu/SHARKshiny/>) and we elaborated on this analysis to answer specific questions on the validity of the Baltic long-term monitoring programs for the calculations of corresponding EBVs. The goal of our analysis was to identify generic problems in the structure and processing of marine data when calculating EBVs.

All samples were collected by hauling plankton nets vertically through the water. The sampled volume was estimated based on either the flowmeter reading on the net, or from the landing net opening area and the length of the haul. The animals in the net were preserved for later determination, counted and quantified. As the number of animals can be very large sometimes only a sub-sample was counted which was then extrapolated to the amount in the whole sample. For each species (and sometimes developmental stage and gender) present in the samples, the number of animals counted (# counted), number of animals per volume of water (abundance, density), and sometimes weight (wet mass/area, wet mass/volume) and the mean or median length (length) were recorded. We used this data set to calculate trend lines for abundance and distribution of Baltic plankton species.

Data processing

Exploration of variance. The data set was explored at three different spatial scales, including (i) the entire Baltic Sea, (ii) four sub-regions with different hydrographical and biogeographic features: the Kattegat–Skagerrak, the Baltic Proper, as well as the Upper and Lower Bothnian Bay and, (iii) 26 individual stations. Explored variables included year, season, sampling effort, depth, sampling time, and species abundance.

Annual variance. Most stations started with reasonably consistent annual sampling programs from year 2000 onwards. The Upper Bothnian Bay had only two stations from 2000 onwards and three stations from 2005. Lower Bothnian Bay had four stations from 2000 and five stations from 2005. The Baltic Proper had very little continuous annual sampling, with only two stations sampling from 2000 onwards and six stations from 2007. The Kattegat–Skagerrak had only one station from 2000 and four stations from 2006. In order to be able to compare trends among the sub-regions, we chose to include only data from 2006 onwards.

Seasonal variance. Only a few stations had an even sampling program across the seasons. Most stations sampled between April and September. Upper Bothnian Bay had three stations with good summer coverage. Lower Bothnian Bay had five stations with good summer coverage. Kattegat–Skagerrak had four stations with both summer and winter coverage. Baltic Proper had six stations with good summer coverage and a few samplings in the winter. Because of the large seasonal variation of plankton abundance and the lower sampling number during winter, we decided to compare only measurements from the early summer. We therefore lumped all abundance records from April to July for further analysis.

Diurnal variance and depth slices. Samples were often split into several depth slices. Since plankton migrates in the water column during day and night, depth slices obtained at different times of the day are not comparable. The data set contained reference to time of sampling for 65.4% of the records (38,538 records) ranging from 22:00 to 10:00, with most intense sampling during the morning hours (6:00–9:00). In order to avoid sampling bias due to diurnal migration of plankton, we lumped all depth slices to obtain depth-integrated abundance of taxa per cubic meter.

Replicates and depth slices. Field codes in the original data set suggested that at some sites more than one replicate was taken. We standardized all measurements by dividing the abundance counts by the number of replicates.

Taxonomic variance. Taxonomic inconsistencies were removed manually after inspecting the species abundances across stations. We used the World Register of Marine Species (WoRMS) taxonomic name service to synonymise all species names.

Life-history variation. The data set included both measurements of adult and non-adult stages (eggs, juveniles, larvae). We excluded all records of non-adult life stages from the analysis.

Visualization. We plotted boxplots from all abundance values (number of individuals/m³) during April–July at the stations in the sub-region for all years. The mean abundance values were connected in the graph with a line over the years 2007–2013. The same approach was used for creating boxplots for the entire region.

Data storage. All input files, output files, R code, and geographic coordinates of the stations are stored at <https://ecds.se/> under file identifier: ccc84507-49c1-43df-9887-97d2232bcb89 (<https://ecds.se/dataset/calculating-essential-biodiversity-variables-for-species-population-abundance-in-the-baltic-sea>).

Table S1. Principal steps of the information supply chain to build Essential Biodiversity Variable (EBV) data sets as applied to the Baltic Sea zooplankton monitoring (BALTIC) data set with indication of abundance records and taxa available for trend analysis.

Generic steps	Specific data-processing steps applied to BALTIC	No. of abundance records	No. of taxa
Sampling (raw observations)	Assembly and integration of raw data from different plankton monitoring programs in the Baltic region	60,511	217
EBV-useable data set	Taxonomic refinement and basic cleaning, e.g. making location and species names consistent	59,374	181
EBV-ready data set	Cropping the data set to relevant seasons and years	18,856	135
Derived & modelled EBV data	Cropping the data set to comparable species observations with >100 measurements (no modelling)	7,067	33

Table S2. Detailed description of workflow steps used in the eBird project.

	Workflow step	Detailed description of workflow steps
1	Identify and import raw data and associated metadata	Data collection in the form of checklists, lists of presence and/or counts of bird species observed and identified during a search. Each eBird checklist contains seven data-collection variables that identify the observer, the location the observations were made, the duration spent searching for birds, the distance travelled during a search (which could be zero if the observer stood at one location), the number of individuals of each species observed, the number of people in the search party, and whether they were submitting a complete checklist of all the birds that they observed or only a partial list of the species they observed (Sullivan <i>et al.</i> , 2014).
2	Check data-sharing agreements and licenses	eBird offers free access to its data <i>via</i> pre-packaged and customizable file downloads and Application Programming Interfaces (APIs). For all downloads and APIs, it is first required that a request with a short abstract is submitted to learn more about how eBird data are being used. All non-commercial requests will be granted as soon as possible, but it may take a few days to process requests. All parties requesting data must adhere to eBird's data use agreement, which is included in every downloaded data bundle. The data sets mentioned below contain only validated records, unless noted.
3	Check data completeness and consistency (e.g. dates, units, spatial information, etc.)	The eBird project has developed methods to ensure that data conform to basic quality standards during data submission. Prior to a checklist being accepted in the eBird database the observer must enter all data on the observation process during submission. Additionally, all species reported on the checklist are passed through data quality filters that rely on historic eBird data and expert opinion, and any unusual records are flagged for further review (Kelling <i>et al.</i> , 2011).
4	Combine and join data sets from different sources	All eBird records are gathered as checklists of the presence or (in the vast majority of cases) counts of birds of all reported species, and all information on the observation process. Since data are gathered in the same web portal, all data have already been combined.
5	Match taxonomy	eBird species and subspecies taxonomy follows the Clements Checklist, which is a global bird taxonomy that follows regional authorities. For example, in the New World, the Clements Checklist largely defers to the two American Ornithologists Union committees – the North American Classification Committee (NACC) and the South American Classification Committee (SACC) – with the goal of near-complete compliance. Clements updates occur once a year in the late summer/autumn, and are documented in full, and can be downloaded directly. eBird taxonomic updates coincide with the Clements updates in August.
6	Check data quality and clean data (errors, outliers, duplicates, etc.)	At data submission, the observer must enter all data on the observation process. All species reported on the eBird checklist must pass data quality filters, and any unusual records are flagged for further review. Feedback is provided immediately to the submitter to ensure data were entered correctly. If the submitter believes their flagged record to be accurate, it is sent to one of more than 1700 regional experts for review. In any given year, approximately 5% of all observations submitted are flagged for review.
7	Check data coverage and fit for purpose, create input files	While the volume of eBird data is high, the search effort across space and time is irregular. This makes it difficult to infer abundance patterns from the raw data. To overcome these challenges in raw eBird data, species distribution models are used to 'fill-in' geographical gaps between locations of data collection and to 'standardize' search effort to correct for variation in detection rates from checklist to checklist.
8	Identify analysis tool (and covariates if needed)	A two-stage species distribution model called the Spatio-Temporal Exploratory Model (STEM) has been developed to model eBird data (Fink <i>et al.</i> , 2010, 2014). Specifically, the STEM uses zero-filled eBird checklists data together with elevation data and environmental covariates. Elevation

		data are derived from the ASTER instrument on-board NASA's Terra Satellite, and annual land cover from the MODIS global land cover product (MCD12Q1). The University of Maryland classification scheme is used to classify each 500 m × 500 m MODIS pixel into one of 14 land cover classes, which are then summarized as proportions within a 3 km × 3 km (900 hectare) set of pixels centred at each eBird observation location.
9	Apply statistical analysis	<p>The two-stage STEM is applied. In the first stage, the study extent is partitioned into a very large number of 'local' regions and time periods. These local regions are small enough to assume stationarity and a species distribution model is fitted within each local region. In the second stage the local models are combined into a single STEM 'ensemble' model that covers the entire large region and calendar year. Because the regions and time periods overlap across the local models, the set of local models can be stitched together into a seamless STEM model that describes distribution or abundance at any location and at any time of year.</p> <p>The STEM models individual species one at a time and produces one daily occurrence and abundance estimate per week for all 52 weeks of the year. For each week, estimates are made across a regular grid of ~3 million locations that span terrestrial North America at a resolution of ~ 3 km × 3 km. Variation in detectability associated with the search effort is controlled by standardizing the estimates to be the expected count of the species on a search conducted from 7–8 a.m. while traveling 1 km on the given day at the given location by a typical eBirder.</p>
10	Calculate uncertainty	Uncertainty estimates are produced as part of the STEM workflow by randomly resampling training data within local regions. Thus, each occurrence and abundance estimate is computed as an ensemble average of up to 100 local models producing a smoothed bootstrap or bagged estimator. For each occurrence and abundance estimate the upper and lower limits are also computed as the 10 th and 90 th percentiles across the ensemble. These limits are conservative estimates of the sampling variation of the smoothed bootstrap estimator.
11	Visualize results	For each species, geospatial layers are generated for occurrence and abundance predictions, maps and animations of weekly predictions for the year, predictive performance metrics for both occurrence and abundance at a weekly resolution, and predictor importance and partial dependence information for the environmental covariates used to train the models. Additionally, a variety of eBird data products are made available (Sullivan <i>et al.</i> , 2014). eBird data are published annually in the Ecological Metadata Language (EML) and Darwin Core (DwC), as well as published in a well-described CSV file on the eBird website.

Table S3. Detailed description of workflow steps used in the Tropical Ecology Assessment and Monitoring (TEAM) project.

	Workflow step	Detailed description of workflow steps
1	Identify and import raw data and associated metadata	Camera trap images are imported directly from the memory card of the camera trap into a computer using software developed by TEAM (Wild.ID). Software grabs all metadata automatically, and the user completes the species information. Once images are identified, records and images are compressed into a special package that is then uploaded to a server for staging (Fegraus <i>et al.</i> , 2011).
2	Check data-sharing agreements and licenses	Since primary data are used, this step is not part of this workflow. All TEAM data providers agree to share the data publicly and give Conservation International a license to use the data and images for non-commercial purposes. Data providers agree to share the data for non-commercial use and data users must contact data providers if they intend to use the data for publication. Data users are encouraged to offer data

		providers co-authorship in papers that might result from further analyses of the data. For details see: http://www.teamnetwork.org/data/use
3	Check data completeness and consistency (e.g. dates, units, spatial information, etc.)	At staging, scripts check each data set (site) for spatial and temporal consistency and continuity, and whether species are within IUCN estimated geographic ranges.
4	Combine and join data sets from different sources	Since primary data are used, this step is not applied in this workflow.
5	Match taxonomy	This is done at the moment of data entry (step 1). TEAM follows standard taxonomic authorities for birds and mammals (IUCN Red List) which are already loaded on Wild.ID at the point of data processing.
6	Check data quality and clean data (errors, outliers, duplicates, etc.)	Images from species that are difficult to identify are shared with specialist groups (e.g. duikers, new world deer). This is not automated at present, but the goal is to do it through ZOOUNIVERSE soon.
7	Check data coverage and fit for purpose, create input files	A cube is created that is a trinary matrix (0,1,NA) for each population/species at each site with camera trap points (rows), time in days (columns) and year as the third dimension in the cube. Time is then compressed into 15 periods (7-days each). These are the data-input files. For covariates a similar cube is created for each covariate (Ahumada <i>et al.</i> , 2013; Beaudrot <i>et al.</i> , 2016).
8	Identify analysis tool (and covariates if needed)	A Bayesian-formulated dynamic occupancy model is used. This uses extinction and colonization probabilities to fit occupancy data through time and corrects for detection probability (Ahumada <i>et al.</i> , 2013; Beaudrot <i>et al.</i> , 2016). Data sets are imported for covariates: elevation, temperature and rainfall. Distance to border from each camera trap point is calculated using GIS scripts. For covariates a similar cube as described in step 7 is created.
9	Apply statistical analysis	Since the model does not have a closed solution, it is fitted using Monte Carlo Markov Chain (MCMC) algorithms using JAGS (Just Another Gibbs Sampler) in R (Beaudrot <i>et al.</i> , 2016).
10	Calculate uncertainty	Output of model includes the distribution of all parameters estimated, so uncertainty is a by-product of model fit.
11	Visualize results	The result of the above steps is the occupancy trend for each population at each site (~ 500 overall). TEAM combines occupancy trends using the Wildlife Picture Index (WPI) (O'Brien <i>et al.</i> , 2010). The WPI is the geometric mean of occupancies relative to the first year of measurement. It is presented graphically at a global level, continental level, protected area, and by species, functional groups, etc. (wpi.teamnetwork.org).

Table S4. Detailed description of workflow steps used in the Living Planet Index (LPI) project.

	Workflow step	Detailed description of workflow steps
1	Identify and import raw data and associated metadata	Systematic and targeted data searches for vertebrate population time series from the scientific literature, government reports, NGO reports and websites, and online databases. Once potential data sources are found, the details are saved in a reference library for processing along with a copy of the data source. Metadata are not extracted at this step. A first check of the data for quality is carried out, e.g. identifying a potential table or figure with data that has population trend data for single species. In-depth quality checking is done in step 3.
2	Check data-sharing agreements and licenses	Most raw data are publically available. If not, then a data-sharing agreement is signed for data entering the Living Planet Database (LPD), based on a template from the data provider if they have one. The agreement states the data use, storage and sharing but does not usually specify licenses. The data can remain confidential at the data provider's request whereby the data are held in the LPD but not made publicly available. There is a data-use policy

		on the LPD website which outlines the terms of use of the data for non-commercial purposes (http://livingplanetindex.org/documents/data_agreement.pdf). However, there is no requirement to sign an agreement.
3	Check data completeness and consistency (e.g. dates, units, spatial information, etc.)	This is a quality check. It ensures the data passes the protocol for the LPD e.g. the method and survey effort are consistent throughout the time period, the results are robust.
4	Combine and join data sets from different sources	Data are entered into the LPD and annotated with a core set of metadata. Data source information is retained in the database so that each record is traceable.
5	Match taxonomy	The taxonomy used in the database is in accordance with the latest authority. The taxonomy used in the data source is not automatically followed.
6	Check data quality and clean data (errors, outliers, duplicates, etc.)	This step is done during step 3, before data entry. Data are checked that they pass the protocol and that there are no duplicate records.
7	Check data coverage and fit for purpose, create input files	The first check is not carried out as data are assumed to be fit for purpose before being entered into the database. Data are exported from the LPD in a format ready for analysis. The code used for analysis processes the data according to a set of variables and subsets are created for the data input files.
8	Identify analysis tool (and covariates if needed)	Generally, a Generalized Additive Modelling (GAM) framework is used to process the data each time as this has been selected as the most suitable for this type of data (Collen <i>et al.</i> , 2009). The GAM is always the first step in the analysis. Subsequent statistical models can be used once the data have been processed and the rates of change calculated for each time series. Covariates are generally used with the annual rates of change for each time series using a mixed model. We use a range of geographic, ecological and life-history traits as covariates in this sort of model. This step is performed after step 9 as the data must be processed first using the GAM framework.
9	Apply statistical analysis	As part of the GAM framework, time series are modelled according to the number of data points whereby we use a GAM to process those with 6 or more points and the Chain method (a log-linear model) to process those with less than six points, or those that fail the GAM model test. A mixed model with covariates can then be applied to calculate the annual rates of change for each time series.
10	Calculate uncertainty	For calculating indices of abundance, we use bootstrapping to produce 95% confidence intervals. These show uncertainty surrounding the index values and not uncertainty in the abundance estimates of the individual populations. The uncertainty is propagated through the time series and so at any year, they show the cumulative uncertainty in the index value since 1970.
11	Visualize results	Data are publicly available <i>via</i> a website (http://www.livingplanetindex.org/data_portal) where users can search the online database and download time series data. Results are published formally every two years in WWF's <i>Living Planet Report</i> and in policy publications such as <i>Global Biodiversity Outlook</i> . Thematic reports and papers are also published on a more <i>ad hoc</i> basis (e.g. McRae <i>et al.</i> , 2017).

Table S5. Detailed description of workflow steps used in the Baltic Sea zooplankton monitoring (BALTIC) project.

	Workflow step	Detailed description of workflow steps
1	Identify and import raw data and associated metadata	The Swedish Meteorological and Hydrographical Institute (SMHI, www.smhi.se) currently aggregates marine monitoring data from national monitoring programs as a part of the Swedish LifeWatch initiative. The data

		are (i) identified, (ii) stored in a single place (e.g. the national archive for climate and environmental data), (iii) consistently formatted, and (iv) annotated with basic metadata. The data are made available in various ways, e.g. through REST API (www.sharkdata.se) as described in the <i>Biodiversity Catalogue</i> (https://www.biodiversitycatalogue.org/services/72) and in the sharkweb data portal (http://sharkweb.smhi.se/).
2	Check data-sharing agreements and licenses	All data are generated by national monitoring programs (not scientific projects). Specific data-sharing agreements for re-use were signed (http://kodu.ut.ee/~riina82/policy.html) and the data are made freely available under the license Creative Commons Attribution 4.0 SE (http://sharkdata.se/datapolicy/). In some cases there may be practical problems in obtaining assistance (i.e. work hours) from the original data provider to help in collecting and annotating the data. Such assistance can be provided by the Swedish LifeWatch program (http://www.svenskalifewatch.se) and the Nordic-Baltic Collaboration on e-Infrastructures for Biodiversity Informatics (https://wiki.neic.no/wiki/Biodiversity).
3	Check data completeness and consistency (e.g. dates, units, spatial information, etc.)	The basic cleaning and formatting was done already by SMHI as a part of the aggregation (described under step 1), e.g. regarding latitude/longitude, time format, tabular structure, metadata templates, etc.
4	Combine and join data sets from different sources	As described above, SMHI provides a consistent format for all aggregated data, so they can be downloaded as a single package. Here, data for all zooplankton were downloaded for the time period 1979–2013 from all available stations. After downloading, some inconsistencies had to be removed. This included spelling errors in the location names (e.g. ‘Å17’, ‘A17’, and ‘A_17’ were transformed to ‘A17’). In addition there were some mismatches in the spatial references. This is because when a station is re-sampled the boat will never be exactly in the same position but rather in the same area. For example, some records of station A17 were with latitude-longitude 58.278-10.5133, while others had 58.274-10.5123. These small mismatches were removed by calculating the mean of all A17 latitudes and longitudes, and re-assigning this mean to all A17 records.
5	Match taxonomy	In principle, there are two taxonomic matching steps. The first one is fully automated at the level of the data provider and generates some basic taxonomic consistency (e.g. deletes synonymies). Some basic checks are performed automatically on all biological data hosted by SMHI. Here, all taxon names in the database are matched according to the Swedish taxonomic backbone (http://www.slu.se/en/Collaborative-Centres-and-Projects/dyntaxa/). However, the database is not entirely up to date with regard to the European Marine Taxonomic backbone (based on the World Register of Marine Species) and this may lead to inconsistencies that have to be treated manually. The second step needs to be done by the scientist after downloading the data. This step is ambiguous and the decisions for how to transform the data depend on the questions under investigation. Here, all species of the genus <i>Acartia</i> were lumped into one group because some monitoring programs identified <i>Acartia</i> spp. to the species level while others identified them as <i>Acartia</i> sp. Such manual taxonomic processing was performed using R scripts.
6	Check data quality and clean data (errors, outliers, duplicates, etc.)	Most of the errors and duplicates were already removed by the data provider (SMHI). No further data cleaning was applied.
7	Check data coverage and fit for purpose, create input files	Processing was done to filter comparable records using R scripts. This included: <i>Annual variance.</i> Most stations started with reasonably consistent annual sampling programs from year 2000 onwards. The Upper Bothnian Bay had only two stations from 2000 onwards and three stations from 2005. Lower

		<p>Bothnian Bay had four stations from 2000 and five stations from 2005. The Baltic Proper had very little continuous annual sampling, with only two stations sampling from 2000 onwards and six stations from 2007. The Kattegat–Skagerrak had only one station from 2000 and four stations from 2006. In order to be able to compare trends among these sub-regions, only data from 2006 onwards were included.</p> <p><i>Seasonal variance.</i> Only a few stations had an even sampling program across the seasons. Most stations sampled between April and September. Upper Bothnian Bay had three stations with good summer coverage. Lower Bothnian Bay had five stations with good summer coverage. Kattegat–Skagerrak had four stations with both summer and winter coverage. Baltic Proper had six stations with good summer coverage and a few samplings in the winter. Because of large seasonal variation of plankton abundance and the lower sampling number during winter, only measurements from the early summer were compared. Hence, all abundance records from April to July were lumped for further analysis.</p> <p><i>Diurnal variance.</i> Plankton abundance shows a high diurnal variance due to vertical migration during day and night. The data set contained reference to time of sampling for 65.4% of the records (38,538 records). The sampling times ranged from 22:00 to 10:00, with most intense sampling during the morning hours (6:00–9:00). Most stations with time information had their most intense sampling during the morning hours, with few exceptions in the Kattegat–Skagerrak and the Upper Bothnian Bay. In order to reduce potential sampling bias due to diurnal variance in plankton abundance it may be necessary to include only observations from one particular period, e.g. from 6:00 to 8:00. This was not applied here because it would have halved the number of observations. Instead, the variance between midnight and midday was included in the abundance calculations, making sure that there was not systematic bias towards a specific year or station. Since the sampling times were very similar it is expected that there is little likelihood of systematically over/under sampling a specific part of the day at any particular year, station, or sub-region.</p> <p><i>Replicates and depth slices.</i> Field codes in the original data set suggested that at some sites more than one replicate was taken. In addition, the samples were often split into several depth slices. All depth slices were lumped to obtain depth-integrated abundance of taxa per cubic meter, while the suggested replicates were averaged.</p> <p><i>Exclude larval stages.</i> We excluded all species records of non-adult stages.</p> <p><i>Select most frequent species.</i> We excluded species that were recorded less than 100 times.</p>
8	Identify analysis tool (and covariates if needed)	Only summary statistics were calculated. No co-variates were used.
9	Apply statistical analysis	After initial exploration of the abundance trends, the mean, minimum, maximum and range of abundance was calculated using number of individuals/m ³ as the measurement unit.
10	Calculate uncertainty	No uncertainty was calculated as no statistical test was performed.
11	Visualize results	Boxplots were plotted for all abundance values. The mean abundance values were connected with a line in the graph over the years 2007–2013.