



Advancing our understanding of biological invasions with long-term biomonitoring data

Phillip J. Haubrock · Laís Carneiro · Rafael L. Macêdo · Paride Balzani · Ismael Soto · Jes Jessen Rasmussen · Peter Wiberg-Larsen · Zoltan Csabai · Gábor Várбірó · John Francis Murphy · J. Iwan Jones · Ralf C. M. Verdonschot · Piet Verdonschot · Gea van der Lee · Danish A. Ahmed

Received: 13 April 2023 / Accepted: 8 July 2023 / Published online: 5 August 2023
© The Author(s) 2023

Abstract The use of long-term datasets is crucial in ecology because it provides a comprehensive understanding of natural fluctuations, changes in ecosystems over extended periods of time, and robust comparisons across geographical scales. This information is critical in detecting and analysing trends and patterns in species populations, community dynamics, and ecosystem functioning, which in turn helps in predicting future changes and impacts of human

activities. Additionally, long-term data sets allow for the evaluation of the effectiveness of conservation efforts and management strategies, enabling scientists and decision makers to make evidence-based decisions for biodiversity conservation. Although the use of long-term data is recognized as highly important in several scientific disciplines, its usage remains undervalued regarding questions in invasion science. Here, we used four regional subsets (i.e. England, Hungary, Denmark and the Dutch-German-Luxembourg) of a recently collated long-term time series database to investigate the abundance and dynamics of occurring non-native species over space and time in Europe. While we found differences in the numbers of non-native species across the studied regions (Dutch-German-Luxembourg region = 37; England = 17, Hungary = 34; Denmark = 3), non-native species detection

Phillip J. Haubrock, Laís Carneiro and Rafael L. Macedo have equally contributed to this work.

Danish A. Ahmed is the senior author contributed.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10530-023-03141-0>.

P. J. Haubrock (✉)
Department of River Ecology and Conservation,
Senckenberg Research Institute and Natural History
Museum Frankfurt, Gelnhausen, Germany
e-mail: Phillip.haubrock@senckenberg.de

P. J. Haubrock · P. Balzani · I. Soto
Faculty of Fisheries and Protection of Waters,
South Bohemian Research Center of Aquaculture
and Biodiversity of Hydrocenoses, University of South
Bohemia in České Budějovice, Vodňany, Czech Republic

P. J. Haubrock · D. A. Ahmed
CAMB, Center for Applied Mathematics
and Bioinformatics, Gulf University for Science
and Technology, Hawally, Kuwait

L. Carneiro
Laboratory of Ecology and Conservation, Department
of Environmental Engineering, Federal University
of Parana - UFPR, Curitiba, Brazil

R. L. Macêdo
Graduate Program in Ecology and Natural Resources,
Federal University of São Carlos, UFSCAR, São Carlos,
Brazil

R. L. Macêdo
Centre National de Recherche Scientifique, AgroParisTech,
Ecologie Systématique Evolution, Université Paris-Saclay,
Orsay, France

rates were continuous over time. Our results further show that long-term monitoring efforts at large spatial scales can substantially increase the accuracy and rate at which non-native species are detected. This information can inform management endeavours dealing with non-native species, underlining the need for invasion scientists and authorities-stakeholders to make more effort in collecting, analysing and making available long-term datasets at broader geographic ranges.

Keywords Biological invasion · Monitoring efforts · Ecology · Time series · Non-native species

Introduction

Globalisation has led to an accompanying increase in propagule and colonisation pressure (Briski et al. 2012) and, thereby, to an increase in the number of established non-native species reaching unprecedented levels (Seebens et al. 2017; Daly et al. 2023). In particular, while the European continent has been a historical epicentre for species translocations and introductions worldwide (Pyšek et al. 2008), it has experienced a substantial increase in the number of

non-native and invasive species in recent decades (Seebens et al. 2017, 2021). This notable trend is a growing concern due to the harmful consequences they pose to biodiversity and human well-being. These include, but are not limited to, the impairment of crops and infrastructure (Lavery et al. 2015), and biotic and functional homogenisation (Olden et al. 2004). Also, the lack of understanding of which non-native species will ultimately become invasive limits our ability to undertake preventive and control measures (Jarić et al. 2019; Pyšek et al. 2020). Therefore, it is vital to understand the invasion process, providing insights into the dynamics of ecosystem functioning, the response of native species to the invader, and the impact of human activities on the environment (Arim et al. 2006). At local and regional scales, increasing rates of introduction have been associated with a loss of ecosystem resilience and the facilitation of subsequent non-native species (Haubrock et al. 2021; Le Hen et al. 2023). As the different drivers of invasions are context dependent and can act synergistically, more integrative monitoring efforts are needed to improve management investments from being focused solely on the few invasive species that are better known and considered to be problematic (Watkins et al. 2021). Despite the well-known risk that biological invasions pose worldwide, only a few countries have fulfilled the requirements to achieve the “Aichi Target number 9” of the Convention on Biological Diversity’s (CBD) relating to biological invasions (McGeoch et al. 2016), underlining the need for future improvements. In this sense, historical data can be an important source to understand how variations in the introduction of non-native species and their establishment in ecosystems occur, and the possible implications (Clavero & Villero 2014).

Long-term biomonitoring data are immeasurably valuable to ecological sciences, as they can provide a fine-scale record of changes in the composition and abundance of species and thus, community compositions over time (Haubrock et al. 2023a, b; Haubrock and Soto 2023; Fig. 1). The information within long-term biomonitoring data could be used to understand the dynamics of non-native species introductions and to track their spread across a range of spatio-temporal scales, modulated by the biological resistance of native species and the availability of local resources, but also, possibly, to gain a better understanding of the influence of environmental conditions on their success or failure

J. J. Rasmussen
Norwegian Institute for Water Research (NIVA Denmark),
Copenhagen, Denmark

P. Wiberg-Larsen
Department of Ecoscience, Aarhus University, Aarhus,
Denmark

Z. Csabai
Department of Hydrobiology, Faculty of Sciences,
University of Pécs, Pécs, Hungary

Z. Csabai
Balaton Limnological Research Institute, Tihany, Hungary

G. Várbró
Centre for Ecological Research, Institute of Aquatic
Ecology, Debrecen, Hungary

J. F. Murphy · J. I. Jones
School of Biological and Behavioural Sciences, Queen
Mary University of London, London, UK

R. C. M. Verdonschot · P. Verdonschot · G. van der Lee
Wageningen Environmental Research, Wageningen UR,
Wageningen, The Netherlands

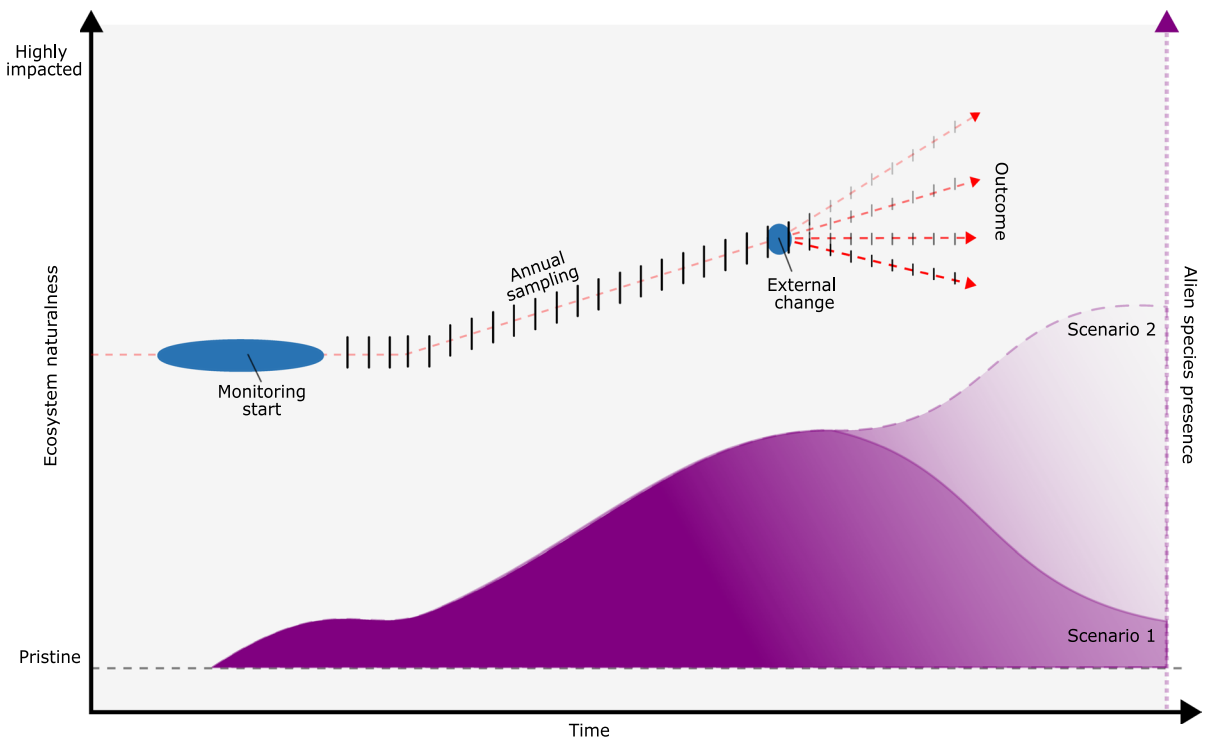


Fig. 1 Two possible scenarios of non-native species presence (i.e., abundance or occurrence) over time (right axis) following an external change (i.e. climatic, anthropogenic disturbance, etc.) and how long-term biomonitoring can capture it. In scenario 1, an external change could lead to a decrease in non-native species presence, and in scenario 2 the changes

can result in increases. External changes can also vary, altering ecosystem naturalness (left axis) from pristine to highly impacted through time. The long term biomonitoring is assumed to capture these changes in community composition and ecosystem alterations

(Guareschi et al. 2021). This could help to develop strategies to prevent their introduction, control their spread, or eradicate established populations. Yet, the ability of long-term biomonitoring to assess introduction dynamics and potentially introduction rates has not been tested (but see Haubrock et al. 2022; Soto et al. 2023). Long-term biomonitoring data, therefore, present enormous potential, e.g. covering large areas across Europe, but also a major challenge because they have not been tested to the extent of benefits they may provide to invasion biologists.

Hence, we used long-term biomonitoring data collected across four European regions—Denmark, Hungary, England, and the Dutch-German-Luxembourg region—to evaluate their capability to assess non-native species introductions in regard to regional differences in temporal dynamics and introduction rates. We hypothesised that (i) introduction rates of

non-native species have been consistent over time, not reflecting any distinguishable temporal patterns. We consequently hypothesised that (ii) long-term biomonitoring data can be indicative of non-native species introductions at large scale, albeit being taxonomically limited, whereas (iii) regional differences will be prominent, but depend on sampling effort. Therefore, the information can be used to identify trends and patterns in the spread of introduced species and to develop effective strategies for their management and control. Ultimately, our goal is to better understand the dynamics of introduced non-native species in ecosystems and promote international cooperation to develop more effective approaches for mitigating their potentially harmful effects.

Methods

To investigate the rate of non-native species introduction and their occurrence in long-term biomonitoring data, we utilised a recently collated database of aquatic macroinvertebrate abundances (Haase et al., in review, Haubrock et al. 2022; Fig. 2) and considered four regional clusters located within Europe, with sufficiently large spatial coverage over Denmark, Hungary, England, and the Dutch-German-Luxembourg regions (henceforth referred to as DGL). These time series contained abundance data of macroinvertebrate groups identified at the species level, exclusively collected from streams and rivers and covered a minimum of eight, not necessarily consecutive sampling years. Each time series was surveyed at the

same geographic location throughout the sampling period (Table 1). Comparable sampling techniques and protocols [e.g. RIVPACS (in England) or DIN 38410 (in Germany)] were used across the sampling sites within each region and were uniform across each site's sampling period.

Identification of non-native species

Non-native species were identified and verified by consulting the *Centre for Agriculture and Bioscience International* (CABI 2023); Google Scholar (<https://scholar.google.com/>), the *Global Biodiversity Information Facility* (GBIF; <https://www.gbif.org/>), and the *Global non-native Species First Record Database* (sTWIST; Seebens et al. 2017). A species was only

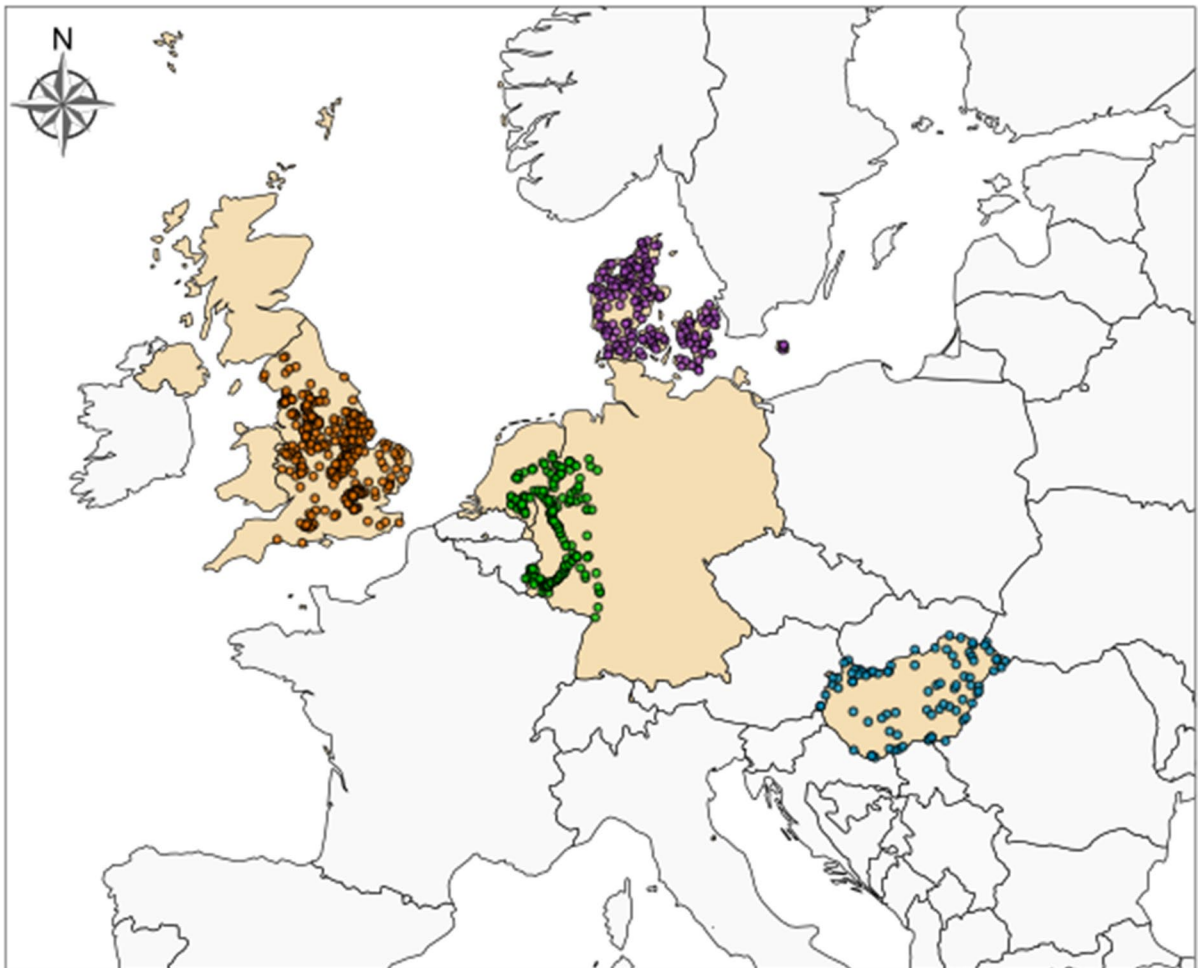


Fig. 2 Location of time series from Denmark (n=174, purple), Hungary (n=84, blue), England (n=296, orange), and Dutch-German-Luxembourg region (n=178, green) containing non-native species

Table 1 Description of the study time-series, by region, number of time series, the range sampling duration, the average duration of time series, and the average of annual samplings

Region	No. of times series (n)	Range of sampling duration	Average duration of time series in years (mean ± SD)	Average number of annual samplings in time series in years (mean ± SD)
Denmark	174	1992–2019	20.71 ± 3.60	22.65 ± 3.42
Hungary	84	2005–2019	11.13 ± 1.38	11.08 ± 1.57
England	296	1994–2019	14.37 ± 4.16	11.81 ± 2.84
DGL	178	1968–2019	23.95 ± 9.53	14.41 ± 5.90

considered as non-native when confirmed by three of the four sources.

Unique species occurrences over time

To investigate the temporal occurrence of unique species, we plotted the Empirical Cumulative Distribution Functions (ECDF) for both non-native and native species over time within each region (Fig. 4). The ECDF is a step function that increases by $1/n$ for each additional species occurrence, where n is the total number of non-native or native species identified in a given region.

We considered a Weibull probability distribution $p(t)$ of unique species occurrences over time t , given by

$$p(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1} e^{-\left(\frac{t}{\lambda}\right)^k} \tag{1}$$

where t is measured in years after the first species was recorded in a given region. $\lambda > 0$ is a scale parameter for the probability distribution, and $k > 0$ is a shape parameter that indicates whether the rate of unique species occurrences decreases ($k < 1$), is constant ($k = 1$) or increases ($k > 1$) over time. The mean time to new species detection (mean; $E(t)$) and variance $\text{var}(t)$ is given by

$$E(t) = \lambda\Gamma(1 + 1/k), \quad \text{var}(t) = \lambda^2[\Gamma(1 + 2/k) - (\Gamma(1 + 1/k))^2] \tag{2}$$

where Γ is the Gamma function defined by $\Gamma(a) = \int_0^\infty z^{a-1} e^{-z} dz$.

We fitted the cumulative probability distribution function $P(t)$ (CDF) for the Weibull distribution given by

$$P(t) = 1 - e^{-\left(t/\lambda\right)^k} \tag{3}$$

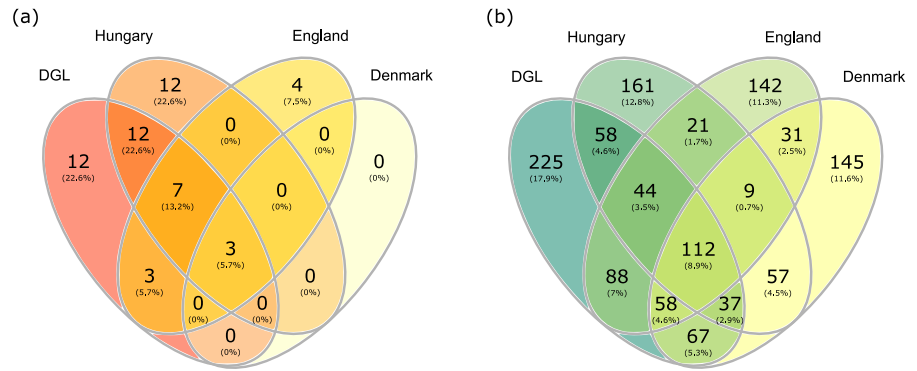
to the ECDF empirical data points. The non-linear regression curve fitting tool *fitlm* from Matlab was used for the distribution fitting, and the strength of the fitting was quantified by the coefficient of determination (R^2) and the root mean square error (RMSE). We reported the estimated model parameters k and λ for non-native and native species, across each region. In the context of multiple species occurrences, the CDF can be used to determine the time to saturation in recorded species. Also, the number of non-native species per region was compared with the number of non-native species identified in the four regions according to the *Global non-native Species First Record Database* (Seebens et al. 2017).

Evaluating the usefulness of long term data for invasion dynamics

To gather insights into regional differences in the presence of non-native species, we investigated the the full number of non-native species records (*i*), the number of time series with at least one non-native record and, (*ii*) the average relative abundance of non-native species (*iii*), across time. For this, we only used the respective response variable (*i-iii*) as a function of time and the number of time series per year to account for temporally varying data availability. As a means of evaluation, the slopes of (*i-iii*) for each of the four regions as well as their respective 95% confidence intervals were compared using linear mixed-effect models using the *lmer* function of the nlme R package (Pinheiro et al. 2013).

To determine whether sample size (i.e. number of time series over time) was sufficient to describe the presence of non-native species in the four regions, the cumulative number of identified non-native and native taxa were plotted separately against the cumulative number of time series investigated per year

Fig. 3 Venn-diagram showing the relative overlap of non-native (a) and native species (b) across Denmark, Hungary, England, and the Dutch-German-Luxembourg region (DGL) according to the available long-term data



(Haubrock et al. 2023). For this, we used the *specaccum* function of the vegan R package (Oksanen et al. 2013), randomising each time series ten times (Ferry & Cailliet 1996; Ferry et al. 1997). Cumulative curves were considered to be asymptotic if ten previous values of the total number of prey taxa were within ± 0.5 of the range of the asymptotic number of taxa, indicating the minimum monitored time series years required to describe the diversity of non-native and native taxa (Huvencers et al. 2007).

Results

All 732 time series analysed from all four regions contained records of non-native species (Fig. 3). The number of native species in Denmark (n=517), Hungary (n=508), England (n=511) and the DGL (n=716) outweighed non-native species in each region (n=3 in Denmark; 34 in Hungary; 17 in England; 37 in the DGL; Supplement 1) in a ratio between 172:1 in Denmark and 15:1 in Hungary. The species

records within our time series showed that the overlap of non-native species present (%) was higher than for native species for DGL, Hungary, DGL+Hungary, and DGL+Hungary+England, (Fig. 3).

Unique species occurrences over time

The ECDF plots highlight that in each region, all non-native species reported over time were identified more rapidly than native species were collected over the same period. While not accounting for the possibility of not having found or correctly identified all non-native and native species, it shows that the effective observations of non-native and native species initiated at the same point in time, occurring in short tandem repeats (Fig. 4). Also, the ECDF plots show that with consistent sampling efforts non-native species are identified in approximately the same proportion over time as native species (Fig. 4e–h).

We found that the rate of unique species occurrences decreased over time ($k < 1$) in Hungary and England for both non-native and native species, and

Table 2 Estimated distribution parameter values λ , k from fitting the CDF (Eq. 3) to the empirical proportion of non-native and native species occurrences across different regions. The fit-

ting is quantified by the coefficient of determination (R^2) and the root mean squared error (RMSE). The mean and variance of time to detection was computed using Eq. (2)

Region	Group	λ	k	R^2	RMSE	Mean	Variance
Denmark	Non-natives	7.78	3.68	0.52	0.33	7.02	4.50
	Natives	5.57	0.72	0.86	0.08	6.90	96.46
Hungary	Non-natives	1.86	0.68	0.12	0.19	2.40	13.03
	Natives	3.30	0.86	0.90	0.08	3.56	17.21
England	Non-natives	10.45	0.93	0.91	0.10	10.79	133.76
	Natives	9.70	0.97	0.97	0.04	9.83	102.68
DGL	Non-natives	24.03	1.37	0.96	0.06	21.99	265.50
	Natives	31.82	2.05	0.97	0.05	28.19	207.64

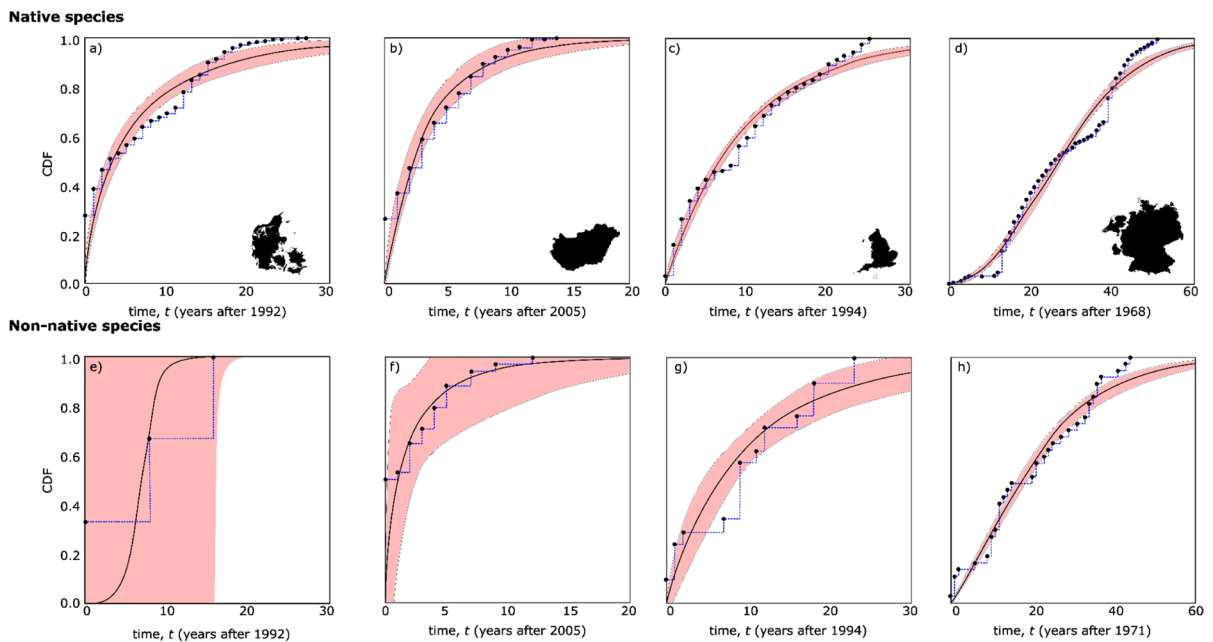


Fig. 4 Empirical Cumulative Distribution Functions (ECDF; blue step functions) for the proportion of non-native and native species from Denmark (a, b), Hungary (c, d), England (e, f), and Dutch-German-Luxembourg region (DGL) (g, h). The Cumulative Distribution Function (CDF) given in Eq. (1)

(black solid curves) were fitted against the proportion of species occurrences over time, with the reported R^2 and RMSE values that quantify the fitting given in Table 2. The red shaded areas represent 95% confidence regions for the range of predicted CDF values

in Denmark only for native species: in all other cases, this rate increased ($k > 1$). The mean time to detection was shortest for non-native species in Hungary, and longest for native species in the DGL region (Table 2).

Among the four investigated regions, no temporal pattern in the occurrence of non-natives could be identified as these non-native species were identified

sporadically. Solely in Hungary, the majority of species ($n = 15$) were already present in 2005, in the first years of monitoring campaigns (Supplement 2). New native species were identified periodically in bursts, possibly reflecting the inclusion of new time series. Comparing non-native species within the time series from the four investigated regions and freshwater macroinvertebrates reported in sTWIST (Seebens

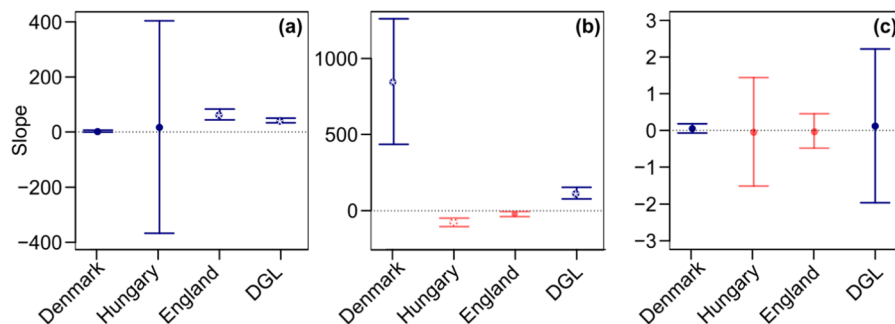
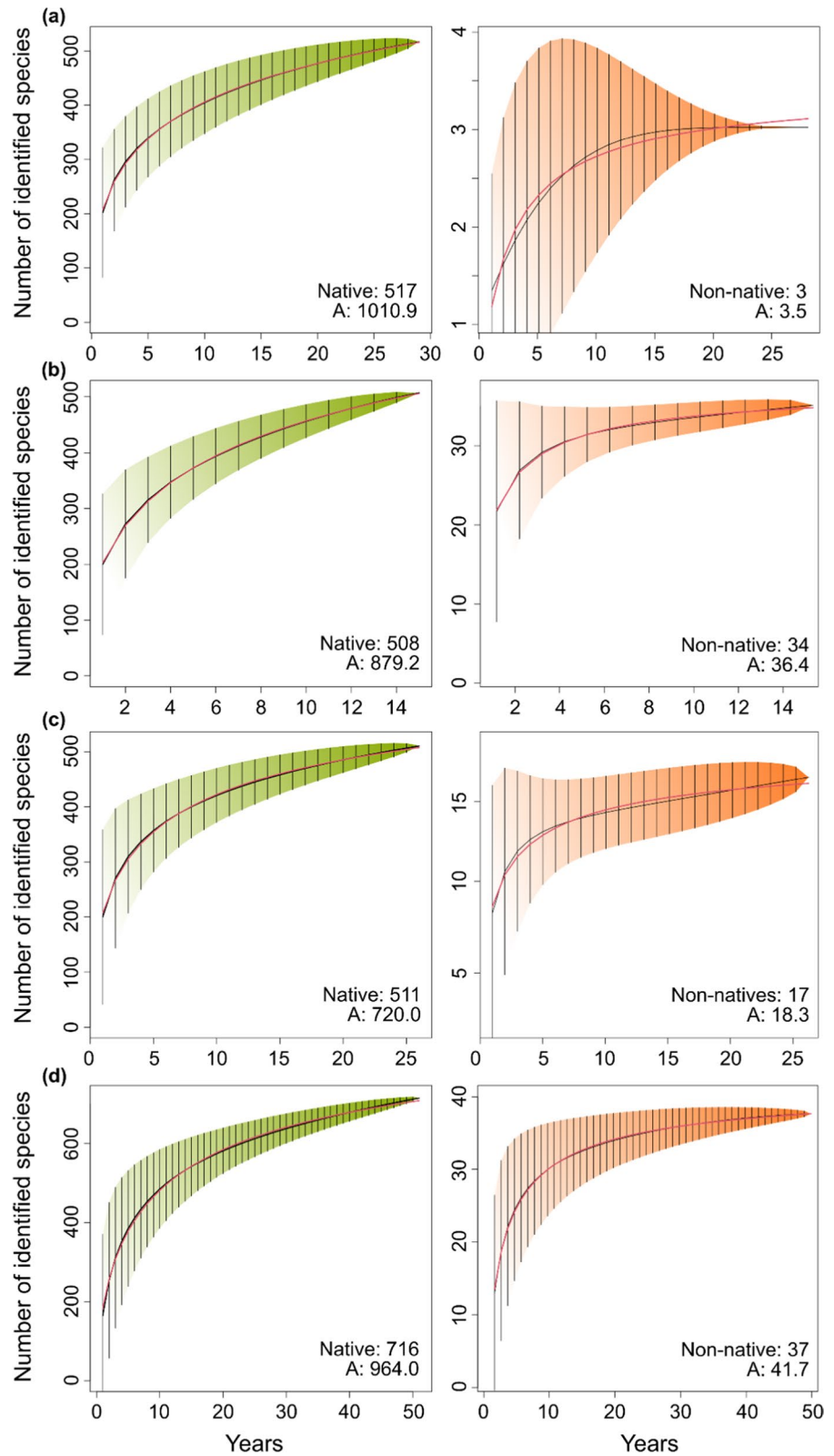


Fig. 5 Slopes ($\pm 95\%$ CI) of the relationships between the number of non-native species records (a), the number of invaded time series (b), and the average relative abundance of non-native species (c) and time for Denmark, Hungary, Eng-

land, and the Dutch-German-Luxembourg region (DGL). Positive trajectories are displayed in blue, negative trajectories in red. Slopes with asterisk were significantly different from zero ($p < 0.05$)

Fig. 6 Saturation curves indicating the years it takes until the completeness of species inventories (A) is reached, showing all native (left side; with the number of species indicated in the bottom right) and all non-native species (right; with the number of species indicated in the bottom right) detected in Denmark (a), Hungary (b), England (c), and the Dutch-German-Luxembourg region (d) based on the available long-term data



et al. 2017), we found considerable inconsistencies in the number of non-native species reported in our data compared with sTWIST. This was especially evident in England, where sTWIST reported 1.8-times more non-native freshwater macroinvertebrate species than reported in the available time series, followed by the DGL- region (1.7:1). In Denmark, our time series reported 3 non-native freshwater macroinvertebrate species versus 5 reported in sTWIST, while for Hungary, our time series reported 8-times higher numbers of non-native freshwater macroinvertebrate species than reported in sTWIST (Supplement 3).

Evaluating the usefulness of long-term data for invasion dynamics

Model assumptions were met in all cases, except for the number of invaded time series in the DGL and changes in the relative abundance of alien species identified in Hungary. Our models hence indicated that over time, the number of non-native species recorded in long-term data from the four investigated regions was growing, albeit significantly increasing only in England and the DGL (Fig. 5a; Supplement 4). The number of invaded time series, however, increased significantly in Denmark and the DGL, while decreased significantly in Hungary (Fig. 5b; Supplement 5). The average relative abundance of non-native species remained constant over time across all four regions (Fig. 5c; Supplement 6). Species accumulation curves further revealed that non-native species observations saturated significantly earlier than observations for native species, reaching their asymptote after 3.5 to 41.7 monitoring years compared to 720.2 to 1010.9 monitoring years (Fig. 6; Supplement 7).

Discussion

Recognition of the importance of using long-term data for a more in depth understanding of impacts (e.g. costs and magnitude) is not new (Blossey 1999; Gill et al. 2021; Le Hen et al. 2023). However, demands have not been aligned with realistic efforts in using wider datasets to allow more effective resource allocation in combating introductions and mitigating negative outcomes from non-native species. In this study, we investigated the occurrences of

non-native and native species within long-term data from four European regions with particularly high spatio-temporal coverage of time series. We found staggering differences not only between non-native and native species, but also across regions. We further demonstrated that new non-native species occurrences were rare but continuous over time, detached from temporal patterns in native species occurrences. Continuous monitoring efforts recording occurrences of non-native species over time, likely represent only a subset of the non-native species established in the respective region, suggesting that time series provide a powerful tool for the detection and effective management and control of invasive populations.

Unique species occurrences over time

Empirical Cumulative Distribution Functions (ECDF) plots are a useful tool to highlight features of an investigated dataset. Compared with a histogram or density plot, they have the advantage of visualising each observation directly, meaning that there are no binning or smoothing parameters that need to be adjusted (Magurran 2013; Gatti 2014; Langrené & Warin 2021). The ECDF plots indicated that with prolonged sampling effort, non-native species are identified at the same pace as native species (Fig. 3c, d). This suggests that despite the known lag time in identifying and reporting of non-native species (McGeoch et al. 2012), established monitoring efforts will be able to identify all non-native species within studies sites, possibly also new arrivals. However, it should be noted that detecting non-native species with similar characteristics to native species can often be difficult, often prolonging the lag time. Interestingly, ECDF plots also revealed a decrease in first identifications of new native species, whereas first non-native species occurrences remained constant, indicating that more non-native species will possibly appear over time. In practical applications, the mean time can be used to make predictions about the expected rate of species detection in a population over time. It can help to guide conservation efforts aimed at preserving biodiversity by providing estimates of the time remaining before a certain proportion or number of species are lost.

However, ECDF plots have to be taken with care if saturation in terms of species identifications has not been reached. Here, non-native species identifications

were continuous over the period—likely in the process of saturating—potentially reaching saturation in the future. The comparison with sTWIST further indicated that, albeit a considerable disparity in terms of total species recorded emphasised by the example of Hungary (1 non-native species in sTWIST vs 34 in our long term data), many more non-native species are likely to be identified in the respective regions. Yet, it should be noted that recent analyses based on sTWIST present established species cumulatively, ignoring that not all species reported eventually established, thereby ignoring the possibility of non-native species disappearing again post establishment, and possibly inadequately representing the true regional presence of non-native species nowadays (Seebens et al. 2018, 2021). In our dataset, non-native species occurred over multiple years and, therefore, were considered as established, although some likely failed to establish and to form populations that created new propagules (Briski et al. 2012). This discrepancy originated from the underlying assumption made by the sTWIST database, recording first records and assuming non-native species to remain in the country as established.

Evaluating the usefulness of long-term data

Globally, occurrence records of new non-native species are growing (Seebens et al. 2017). The resulting data are usually compiled and consequently extracted again from large databases such as GBIF. The uncertainty associated with the occurrence data from these databases can lead to misrepresentations of current distributions of non-native species and do not provide data on their permanence in the environment. Hence, the ecology of invasions still lacks specific long-term monitoring to detect species compared with other long-term ecological studies (Sukhotin & Berger 2013; Harvey et al. 2020). Local and regional biomonitoring is, therefore, an important temporal record of both the first records of non-native species and the local variation in establishment and dispersal, particularly for underrepresented but highly impactful taxa such as invasive vertebrates (Haubrock et al., 2023; Le Hen et al. 2023). Additionally, gaps in historical monitoring can be detrimental for accurately determining the environmental space where species can establish populations. As anthropogenic disturbances, i.e. landscape and climate changes, cause

uncertain impacts on biodiversity, tracking long-term data on species distributions will become more important as it will provide information on species tolerances and survivorship outside optimal ranges (Hellmann et al. 2008). Thus, long-term information can avoid false positives (locations where species are no longer able to sustain populations) and false negatives (accounting for robust data on distributions of rare species, or those with detection issues “sleeper populations”; Spear et al. 2021; Bracken et al. 2022). In doing so, we will be able to use statistical and computational techniques to accurately predict potential distributions and map areas at risk of invasions.

Here, the applied model found comparable patterns in the presence of non-native species across regions, even significant increases in England and the DGL. Due to the consistent sampling methodology over time, obtained trends are more reliable than if based on large occurrence databases, as the origin of the data and methods used are traceable, as well as the temporal variation, allowing analysis of uncertainties (Hughes et al. 2021). The positive patterns found in the regions of England and the DGL may be related to a greater number of time series, showing that with more extensive monitoring networks, the detection of new species, both native and non-native, is improved. Interestingly, the number of time series with non-native species increased significantly in the DGL and Denmark, in the latter case artificially increased by *P. antipodarum* spreading passively (Haubrock et al. 2022). This suggests progressing invasions, as these regions are also those with time series collected over a longer time frame. However, the number of time series with non-native species was decreasing in Hungary. While management interventions could be a reason, it is also possible that local conditions resulted in non-native species failing to establish or that time series were not continuously monitored, underlining that monitoring has to be consequent and consistent (Haase et al. 2018). Hungary was also the country with the smallest number of time series spanning the shortest time interval, but the country with the second highest number of non-native species, likely due to the arrival of Ponto-Caspian species (Soto et al. 2023b).

Increasing introduction rates at regional and local scales can indicate that ecosystems are becoming less resilient to new introductions based on decreasing ability to support native species, thus being at

risk of being functionally overwhelmed (Haubrock et al. 2021; Le Hen et al. 2023). On the other hand, the ongoing introduction of non-native species may also be the result of the loss of biodiversity due to the large-scale anthropisation of ecosystems, which creates new opportunities for invasions and replacement of ecosystem functions that otherwise have become extinct (Lundgren et al. 2020). This can exacerbate biodiversity loss (Bellard et al. 2022), loss of ecosystem functions (Vilà et al. 2010), and ultimately, affect human well-being (Bacher et al. 2018). On the other hand, decreasing introduction rates can indicate that efforts to prevent, control, or eradicate non-native and potential invasive species are successfully hindering their spread, dispersal is limited due to new (i.e. geographic) barriers, or that a form of ‘saturation’ has been reached, i.e. enabling additional management possibilities to bolster native biodiversity (Watkins et al. 2021). Such management interventions are crucial, yet difficult to implement as non-native aquatic macroinvertebrate species are particularly hard to detect due to the hidden nature of life underwater. In the case of Denmark, due to the number of time series on a small area and the widespread species *P. antipodarum*, we found that only a low number of monitoring years is needed to identify all non-natives. In both England and the DGL, which are both more species-rich, more monitoring years were needed. Thus, our results show that high monitoring effort, i.e. many monitored sites, decreases the time needed to identify all non-native species in a specific region. High spatial monitoring coverage—evenly distributed regionally but also globally (Hughes et al. 2021)—will result in non-native species and their spread more rapidly being identified. While our analysis only focused on riverine ecosystems, we believe that monitored lakes could on the one hand result in more non-native species being identified within each region, but also lead to a decrease in monitoring years required due to more sites being monitored over time.

Conclusion

Long-term data across wide geographic scales play a critical role in ecology and conservation biology as

it shows its usefulness in deciphering a realistic picture of the status and trends of native and non-native species accumulation across geographical scales. Overall, long-term macroinvertebrate monitoring data could become critical in enabling invasion biologists and authorities-stakeholders to detect invasions early and apply rapid responses. Here, we showed a consistently increasing trend across macroinvertebrate groups, which was supported by the observed increase in new records and number of invaded time series. Of particular interest is the number of monitoring sites needed to collect sufficient monitoring years before all non-native species are identified. As more introductions are expected worldwide, continuous and widespread monitoring efforts will be needed to quantify the effectiveness of management interventions, and to propose priority areas for control and mitigation of plausible impacts aiming at the protection of biodiversity.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors have not disclosed any funding.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arim M, Abades SR, Neill PE, Lima M, Marquet PA (2006) Spread dynamics of invasive species. *Proc Natl Acad Sci* 103:374–378
- Bacher S, Blackburn TM, Essl F, Genovesi P, Heikkilä J, Jeschke JM et al (2018) Socio-economic impact classification of alien taxa (SEICAT). *Meth Ecol Evol* 9:159–168

- Bellard C, Marino C, Courchamp F (2022) Ranking threats to biodiversity and why it doesn't matter. *Nat Commun* 13:2616
- Blossey B (1999) Before, during and after: the need for long-term monitoring in invasive plant species management. *Biol Invasions* 1:301–311
- Bracken JT, Davis AY, O'Donnell KM, Barichivich WJ, Walls SC, Jezkova T (2022) Maximizing species distribution model performance when using historical occurrences and variables of varying persistency. *Ecosphere* 13:e3951
- Briski E, Bailey SA, Casas-Monroy O, DiBacco C, Kaczmarzka I, Levings C et al (2012) Relationship between propagule pressure and colonization pressure in invasion ecology: a test with ships' ballast. *Proceed Royal Soci B: Biol Sci* 279:2990–2997
- CABI (2023) CABI Compendium. Wallingford, UK: CAB International. www.cabidigitallibrary.org
- Clavero M, Villero D (2014) Historical ecology and invasion biology: long-term distribution changes of introduced freshwater species. *Bioscience* 64(2):145–153
- Daly EZ, Chabrierie O, Massol F, Facon B, Hess MC, Tasiemski A et al (2023) A synthesis of biological invasion hypotheses associated with the introduction–naturalisation–invasion continuum. *Oikos*. <https://doi.org/10.1111/oik.09645>
- Ferry LA, Cailliet GM (1996) Sample size and data analysis: are we characterizing and comparing diet properly? In: MacKinlay D, Shearer K (eds) "GUTSHIO'96" Feeding ecology and nutrition in fish: international congress on the biology of fishes. American Fisheries Society, San Francisco, California, pp 71–80
- Ferry LA, Clark SL, Cailliet GM (1997) Food habits of spotted sand bass (*Paralabrax maculatofasciatus*, Serranidae) from Bahia de Los Angeles, Baja, California. *Bull South Calif Acad Sci* 96:1–21
- Gatti RC (2014) Biodiversità. In teoria e in pratica. Libreria universitaria. it Edizioni
- Gill NS, Mahood AL, Meier CL, Muthukrishnan R, Nagy RC, Stricker E et al (2021) Six central questions about biological invasions to which NEON data science is poised to contribute. *Ecosphere* 12:e03728
- Guareschi S, Laini A, England J, Barrett J, Wood PJ (2021) Multiple co-occurrent alien invaders constrain aquatic biodiversity in rivers. *Ecol Appl* 31:e02385
- Haase P, Tonkin JD, Stoll S, Burkhard B, Frenzel M, Geijzen-dorffer IR et al (2018) The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity. *Sci Total Environ* 613:1376–1384
- Harvey CJ, Fisher JL, Samhouri JF, Williams GD, Francis TB, Jacobson KC et al (2020) The importance of long-term ecological time series for integrated ecosystem assessment and ecosystem-based management. *Progr Oceanogr* 188:102418
- Haubrock PJ, Pilotto F, Innocenti G, Cianfanelli S, Haase P (2021) Two centuries for an almost complete community turnover from native to non-native species in a riverine ecosystem. *Glob Change Biol* 27:606–623
- Haubrock PJ, Ahmed DA, Cuthbert RN, Stubbington R, Domisch S, Marquez JR et al (2022) Invasion impacts and dynamics of a European-wide introduced species. *Glob Change Biol* 28:4620–4632
- Haubrock PJ, Soto I (2023) Valuing the information hidden in true long-term data for invasion science. *Biol Invasions* 25:2385–2394. <https://doi.org/10.1007/s10530-023-03091-7>
- Haubrock PJ, Cuthbert RN, Haase P (2023) Long-term trends and drivers of biological invasion in Central European streams. *Sci Total Environ* 876:162817
- Haubrock PJ, Pilotto F, Soto I, Kühn I, Verreycken H, Seebens H et al (2023) Long-term trends in abundances of non-native species across biomes, realms, and taxonomic groups in Europe. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2023.163808>
- Hughes AC, Orr MC, Ma K, Costello MJ, Waller J, Provoost P et al (2021) Sampling biases shape our view of the natural world. *Ecography* 44:1259–1269
- Huveneers C, Otway NM, Gibbs SE, Harcourt RG (2007) Quantitative diet assessment of wobbegong sharks (genus *Orectolobus*) in New South Wales, Australia. *ICES J Mar Sci* 64:1272–1281
- Jarić I, Heger T, Monzon FC, Jeschke JM, Kowarik I, McConkey KR et al (2019) Crypticity in biological invasions. *Trends Ecol Evol* 34:291–302
- Langrené N, Warin X (2021) Fast multivariate empirical cumulative distribution function with connection to kernel density estimation. *Comput Stat Data Anal* 162:107267
- Laverty C, Nentwig W, Dick JT, Lucy FE (2015) Alien aquatics in Europe: assessing the relative environmental and socioeconomic impacts of invasive aquatic macroinvertebrates and other taxa. *Manag Biol Invasions*. <https://doi.org/10.3391/mbi.2015.6.4.03>
- Le Hen G, Balzani P, Haase P, Kouba A, Liu C, Nagelkerke LA et al (2023) Alien species and climate change drive shifts in a riverine fish community and trait compositions over 35 years. *Sci Total Environ* 867:161486
- Magurran AE (2013) Measuring biological diversity. Wiley-Blackwell, Oxford
- McGeoch MA, Spear D, Kleynhans EJ, Marais E (2012) Uncertainty in invasive alien species listing. *Ecol Appl* 22:959–971
- McGeoch MA, Genovesi P, Bellingham PJ, Costello MJ, McGrannachan C, Sheppard A (2016) Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biol Invasions* 18:299–314
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, O'hara RB, Simpson GL, Solyomos P, Henry M, Stevens H, Szoecs E, Maintainer HW (2013) Package "vegan" Title Community Ecology Package Version 2, pp. 5–7
- Olden JD, Poff NL, Douglas MR, Douglas ME, Fausch KD (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends Ecol Evol* 19:18–24
- Pinheiro J, Bates D, DebRoy S, Sarkar D, R Development Core Team (2013) Nlme: Linear and nonlinear mixed effects models
- Pyšek P, Richardson DM, Pergl J, Jarošík V, Sixtová Z, Weber E (2008) Geographical and taxonomic biases in invasion ecology. *Trends Ecol Evol* 23:237–244

- Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, Jeschke JM, Kühn I, Liebhold AM, Mandrak NE, Meyerson LA, Pauchard A, Pergl J, Roy HE, Seebens H et al (2020) Scientists' warning on invasive alien species. *Biol Rev* 95:1511–1534
- Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM et al (2017) No saturation in the accumulation of alien species worldwide. *Nat Commun* 8:14435
- Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM et al (2018) Global rise in emerging alien species results from increased accessibility of new source pools. *Proceed Nat Acad Sci* 115(10):E2264–E2273
- Seebens H, Bacher S, Blackburn TM, Capinha C, Dawson W, Dullinger S et al (2021) Projecting the continental accumulation of alien species through to 2050. *Glob Change Biol* 27(5):970–982
- Soto I, Ahmed DA, Balzani P, Cuthbert RN, Haubrock PJ (2023) Sigmoidal curves reflect impacts and dynamics of aquatic invasive species. *Sci Total Environ* 872:161818
- Soto I, Cuthbert RN, Ricciardi A, Ahmed DA, Altermatt F, Archambaud-Suard G et al (2023) The faunal ponto-caspi-ization of European waterways. *Biol Invasions*. <https://doi.org/10.1007/s10530-023-03060-0>
- Spear MJ, Walsh JR, Ricciardi A, Zanden MJV (2021) The invasion ecology of sleeper populations: prevalence, persistence, and abrupt shifts. *Bioscience* 71:357–369
- Sukhotin A, Berger V (2013) Long-term monitoring studies as a powerful tool in marine ecosystem research. *Hydrobiologia* 706:1–9
- Vilà M, Basnou C, Pyšek P, Josefsson M, Genovesi P, Gollasch S et al (2010) How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Front Ecol Environ* 8:135–144. <https://doi.org/10.1890/080083>
- Watkins HV, Yan HF, Dunic JC, Côté IM (2021) Research biases create overrepresented “poster children” of marine invasion ecology. *Conserv Lett* 14:e12802
- Wickham H, Chang W, Wickham MH (2016) Package ‘ggplot2’. Create elegant data visualisations using the grammar of graphics. Version 2(1):1–189

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.