A peer-reviewed open-access journal
NeoBiota
Advancing research on alien species and biological invasions

NeoBiota 109: 45–66 (2026)
DOI: 10.3897/neobiota.109.193861

Methods

A workflow for impact indicators of alien species: A demonstration with *Acacia* species

Mukhtar M. Yahaya[1], Sabrina Kumschick[2,3], Sandra MacFadyen[1,4], Pietro Landi[1,4], Cang Hui[1,4,5]

1 *Mathematical Biology Hub, Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa*
2 *Centre for Invasion Biology, Botany and Zoology Department, Stellenbosch University, Stellenbosch, South Africa*
3 *Kirstenbosch Research Centre, South African National Biodiversity Institute, Cape Town, South Africa*
4 *National Institute for Theoretical and Computational Sciences, Stellenbosch University, Stellenbosch, South Africa*
5 *Biodiversity Informatics Unit, African Institute for Mathematical Sciences, Cape Town, South Africa*
Corresponding author: Mukhtar M. Yahaya (yahayamukhtar.m@gmail.com)

Abstract

The negative impact of alien species is recognised as a major threat to biodiversity. An impact indicator, a measurable proxy of the status of alien species impacts over space and time, is essential for informing biological invasion policies and management strategies. We introduce an open-source workflow for computing impact indicators of alien species, combining occurrence data from the Global Biodiversity Information Facility (GBIF) with assessments of Environmental Impact Classification for Alien Taxa (EICAT). To implement the workflow, we developed an R package, impIndicator, which allows users to compute and visualise impact indicators for individual species and sites, as well as regional impact indicators. This tool can support ecological research and management by providing standardised insights into where alien species pose the greatest threats and whether current interventions are effectively reducing them. Such information is directly relevant to policy frameworks, including Target 6 of the Convention on Biological Diversity's Kunming–Montreal Global Biodiversity Framework and the Sustainable Development Goals. We demonstrate the workflow using *Acacia* species in the Iberian Peninsula, South Africa and California, showcasing the spatiotemporal dynamics of their impacts and highlighting sites with higher impact risk. The resulting impact indicators were sensitive to variation in sampling effort and the completeness of the underlying occurrence data, highlighting the importance of well-sampled and systematically curated biodiversity datasets for robust impact assessment.

Key words: Alien species, biodiversity, EICAT, GBIF, impact indicator, workflow

Academic editor: Joana Vicente
Received: 1 April 2026
Accepted: 6 August 2026
Published: 3 September 2026

Citation: Yahaya MM, Kumschick S, MacFadyen S, Landi P, Hui C (2026) A workflow for impact indicators of alien species: A demonstration with *Acacia* species. NeoBiota 109: 45–66. https://doi.org/10.3897/neobiota.109.193861

Introduction

The negative impact of alien species is one of the leading causes of biodiversity loss (Bacher et al. 2023). Effective management of these impacts is hindered by the lack of evidence-based impact indicators, defined as measurable proxies that quantify the impacts of alien species on native ecosystems across space and time (Vicente et al. 2022). At the same time, biodiversity data have improved substantially. Species occurrence records are becoming increasingly available through global databases (GBIF.org 2025), while assessments of alien species impacts are becoming more standardised through frameworks (Nentwig et al. 2010; Blackburn et al. 2014; Bacher et al. 2025). Together, these developments provide an opportunity to establish a standard approach for quantifying alien species impacts by integrating occurrence data with impact assessments. Such indicators can provide information for evidence-based policy and management by helping to prioritise species and sites for intervention, evaluate the effectiveness of control measures and forecast impacts under future scenarios (Kumschick et al. 2025).

This need is particularly pressing because alien species can severely affect ecosystem functioning and the ecosystem services that underpin human well-being. In response, governments and organisations worldwide are developing and implementing policies to prevent the introduction of new invasive alien species, control the spread of existing ones and restore affected ecosystems (IPBES 2023). These policies require standardised, evidence-based decision-support tools, including impact indicators, to monitor changes in impacts, assess spread and evaluate the effectiveness of control measures (Grêt-Regamey et al. 2017). Impact indicators can also complement spread forecasts to provide information for preventative and management actions.

For an indicator to be useful for monitoring and tracking progress, such as towards Target 6 of the Convention on Biological Diversity (CBD) Kunming-Montreal Global Biodiversity Framework which aims to eliminate, reduce and/or mitigate the negative impacts of alien species on biodiversity and ecosystem services, its underlying workflow should adhere, as far as possible, to the Findable, Accessible, Interoperable and Reusable (FAIR) Data Principles (Groom et al. 2017; Groom et al. 2019). To produce reliable and repeatable impact indicators, it is essential to employ a transparent, open-data workflow that consistently converts raw data into coherent, detailed and replicable indicators (Seebens et al. 2020; Boyd et al. 2023). Adherence to the FAIR Data Principles in both inputs and outputs promotes transparency, reusability and long-term viability. Furthermore, the workflow should be sufficiently flexible to be applicable across different spatial, temporal and taxonomic scales. Finally, impact indicators should be accompanied by estimates of uncertainty where possible, enabling users to interpret the results appropriately and avoid unwarranted confidence in the outputs (Vicente et al. 2022).

Several indicators have been proposed to monitor biological invasions over time (Rabitsch et al. 2016; Wilson et al. 2018; Henriksen et al. 2024). Indicators focusing on occurrence data for alien species include the cumulative number of established alien species (Henriksen et al. 2024) and the relative abundance of alien species (Wilson et al. 2018; Delavaux et al. 2023). While occurrence-based indicators provide a general measure of invasion trends, they do not necessarily reflect changes in ecological impacts. Furthermore, they do not account for variation in impact magnitude amongst alien species, such as when distinguishing between levels of severe impacts (e.g. Henriksen et al. (2024)). These limitations can be addressed by integrating occurrence data with information on impact magnitude.

The Red List Index (RLI) is an indicator that measures changes in average extinction risk of groups of species through time (Butchart et al. 2007, 2025). Changes in extinction risk are driven by multiple factors, including, but not limited to, invasive alien species. However, when analysed alongside the abundance, richness or temporal trends of co-occurring invasive alien species, the RLI can provide evidence of an association between biological invasions and changes in extinction risk, thereby offering insights into the potential ecological impacts of invasive alien species on ecosystems (Butchart 2008; Rabitsch et al. 2016; Philippe-Lesaffre et al. 2024). Although the RLI of alien species reflects variation in how alien species affect native species, it does not fully capture the magnitude of those impacts or the mechanisms by which alien species exert them on the broader ecosystem (i.e. it does not explicitly account for impact severity or the mechanisms through which impacts are exerted). Nevertheless, combining information on impact magnitude and mechanisms with species distributions can provide a more comprehensive representation of alien species impacts on recipient ecosystems.

The Environmental Impact Classification for Alien Taxa (EICAT), recently adopted by the IUCN, provides a standardised framework for classifying the magnitude and mechanisms of alien species impacts (IUCN 2020a). Alien species are categorised from Minimal Concern to Massive, depending on the reported magnitude of their impacts on native species (Blackburn et al. 2014; IUCN 2020b). Combining information on the impacts of alien species and their occurrence can provide a representation of their impacts in a particular region (Latombe et al. 2017). The combination of data on impacts and occurrences has been proposed as an indicator of invasion impacts (e.g. Wilson et al. (2018)), but, until recently, there was no standardised methodology for implementing this approach. Recent research has made important methodological advances to address this shortcoming (Boulesnane-Guengant et al. 2025; Oveisi et al. 2026). For example, Kumschick et al. (2025) and Boulesnane-Guengant et al. (2025) used EICAT assessments together with species distribution data to create spatial risk maps of the impact of *Acacia* species in South Africa. Their methods involve transforming EICAT categories documented across different impact mechanisms into numerical impact scores for each species using an aggregation metric (e.g. the mean). These impact scores are then aggregated across species within standardised grid cells, based on the species distributions derived from occurrence records or species distribution models (see the Materials and methods section for analogous aggregation procedures). Although the approach incorporates spatial variation in impacts, it does not explicitly quantify species-level, site-level and regional-level impacts through time, which is the focus of the workflow developed in this study.

To create scientifically robust, policy-relevant impact indicators, we derive species-level, site-level and regional impact indicators over time by combining occurrence data with impact magnitude. Building on the methods of Kumschick et al. (2025) and Boulesnane-Guengant et al. (2025), we integrate open-source EICAT assessments with species occurrence data to make the impact indicators taxonomically, spatially and temporally explicit. The workflow is embedded in an R package to compute and visualise these indicators, following standard software development guidelines (Huybrechts et al. 2024). The workflow produces three main outputs: (i) species-level impact indicators; (ii) site-level impact indicators and (iii) regional impact indicators.

Materials and methods

Occurrence data

The workflow uses occurrence data from the Global Biodiversity Information Facility (GBIF) (Fig. 1). GBIF is a large biodiversity data infrastructure that supports the FAIR Data Principles and is widely used for scientific purposes (Heberling et al. 2021; Ivanova and Shashkov 2021). The workflow uses an occurrence cube format, which aggregates occurrences along spatial (e.g. Extended Quarter

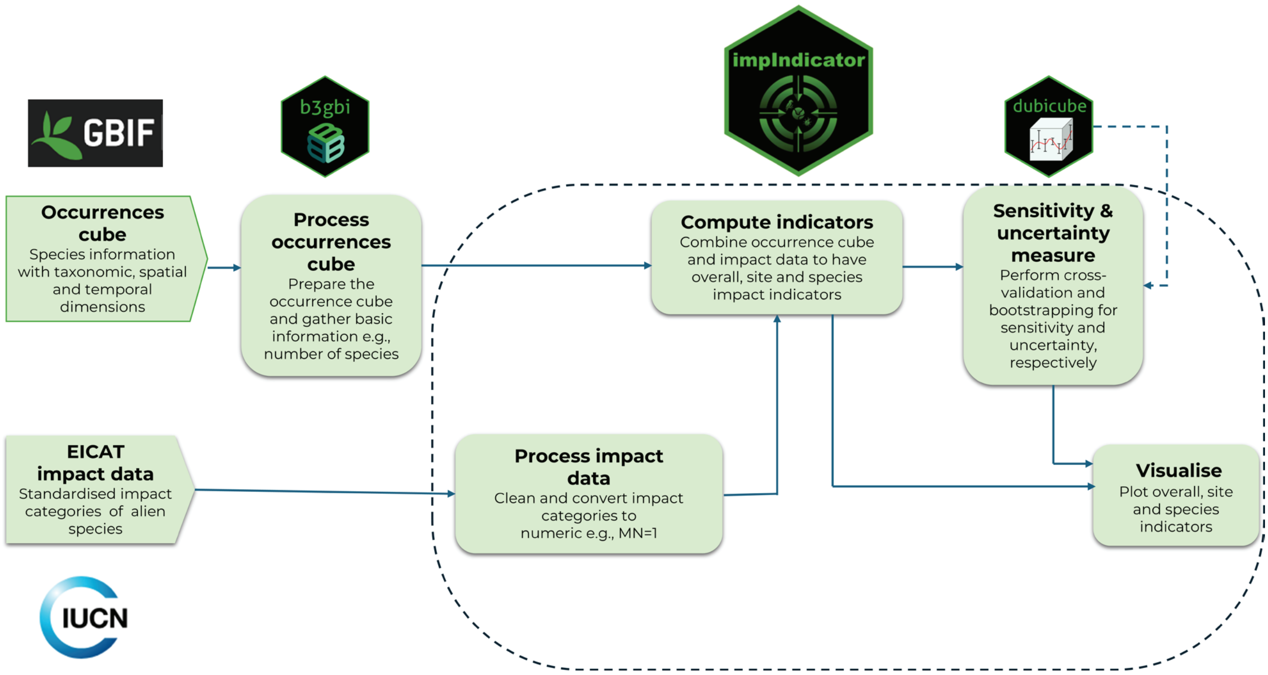

Figure 1. Diagrammatic representation of the workflow embedded in impIndicator to compute and visualise impact indicators for alien species (https://github.com/b-cubed-eu/impIndicator). The workflow integrates GBIF occurrence data and EICAT impact assessments.

Degree Grid Cell (EQDGC)), temporal (e.g. yearly) and taxonomic (e.g. species) dimensions (Desmet et al. 2023), making the data suitable for modelling and computational analysis (Oldoni et al. 2020). The cube is processed using the b3gbi R package (Dove 2026). This ensures standardised input data for further analysis and verifies that the data format is correct. To minimise the influence of sampling bias, multiple records of a species from the same site (grid cell, i.e. a single spatial unit on the Earth's surface, for example, E033S02BB of EQDGC) within a given year are converted to binary presence (1) or absence (0) records for all computations (Boria et al. 2014).

The status of species (alien or native) constitutes an Essential Biodiversity Variable for monitoring biological invasions (Latombe et al. 2017). It is necessary to confirm that the species included in the analysis are alien within the region of interest before computing impact indicators. Several published datasets now facilitate the verification of species status. For example, Gómez-Suárez et al. (2025) developed a harmonised dataset of species status using the standardised SinAS workflow (Seebens et al. 2020), which integrates and harmonises global sources such as the Global Register of Introduced and Invasive Species (GRIIS) and the World Checklist of Vascular Plants (WCVP). In addition, reliable sources such as peer-reviewed literature and technical reports can be used to determine alien status.

Impact data

The workflow uses EICAT data as evidence of the impact of alien species. Impacts of alien species, based on EICAT assessments are classified into Minimal Concern (MC), Minor (MN), Moderate (MO), Major (MR) or Massive (MV), according to the severity of their negative effects on native species in the recipient community. Specifically, a species is assigned MC if there is no reduction in the performance of individuals of a native species; MN if the performance of individuals is reduced, but there is no decrease in native species population size; MO if the native species' population is reduced; MR if the impact has led to local extinction, but is reversible; and MV if the impact has led to naturally irreversible extirpation or extinction. The impacts are also classified according to the 12 mechanisms described by the IUCN, which are competition, predation, hybridisation, transmission of disease, parasitism, poisoning/toxicity, biofouling or other direct physical disturbance, grazing/herbivory/browsing, chemical impacts on ecosystems, physical impacts on ecosystems, structural impacts on ecosystems and indirect impact through interactions with other species. Additional information, such as the location of the impact, is also reported (IUCN 2020a). EICAT data are available through the Global Invasive Species Database (GISD; https://www.iucngisd.org/gisd/). However, the EICAT assessments, hosted in GISD, are not yet fully open or FAIR, as they currently do not provide persistent identifiers and are not fully accessible for bulk or machine-readable extraction. For species without published assessments, users can conduct and add their own assessments within this workflow.

For computational purposes, the semi-quantitative (ordinal) EICAT categories must first be transformed into numerical scores. The transformation depends on the assumed encoding scale for the impact categories, such as a linear or exponential relationship and on how the MC category is interpreted. Specifically, MC may be treated either as zero impact, because no impact on native species has been recorded or as a non-zero category reflecting some degree of concern. We considered three encoding schemes, which can be used depending on the assumptions of the study. The first, MC = 0, MN = 1, MO = 2, MR = 3 and MV = 4, assumes a linear relationship and interprets MC as zero impact (Hagen and Kumschick 2018). This scheme is consistent with the interpretation of MC as the absence of detectable effects on native species. The second, MC = 1, MN = 2, MO = 3, MR = 4 and MV = 5, also assumes a linear relationship, but treats MC as a non-zero category to represent some degree of concern, despite no documented impact on native species (e.g. Jansen and Kumschick (2022)). The third, MC = 1, MN = 10, MO = 100, MR = 1000 and MV = 10,000, assumes an exponential relationship between categories and scores, while likewise treating MC as non-zero. This scheme gives disproportionate weight to higher-impact categories (Suppl. material 1: appendix 1, table S1). There are many other possible schemes not considered in this framework, but these may be included in further versions of the software after careful methodological evaluation.

Impact indicators

To produce impact indicators, we used a hierarchical aggregation framework with three levels: species, site and region (a whole study area, for example, South Africa). First, because species are often assigned multiple impact categories across study locations and impact mechanisms, we aggregated these records to obtain a single species-level impact score intended to represent the overall documented impacts of each alien species outside its native range. For example, *Acacia dealbata* has been classified as MR in the Drakensberg, South Africa, through structural impact, but as MN in the Biobío Region, Chile, through chemical impact (see Jansen and Kumschick (2022)). To reconcile such variation, we converted impact categories to scores and aggregated them within species using one of three approaches: Maximum, Mean or Sum of Maximum across Mechanisms (Table 1, Fig. 2).

Table 1. The three types of impact indicators, their definitions, functions implemented in impIndicator, computation methods, assumptions, limitations and intended applications. The regional impact indicator is computed as the spatial sum of the corresponding site-level impact indicators for each year. Therefore, the site-level and regional impact indicators share the same underlying computation method (aggregation of impact scores within and across species), assumptions and limitations.

	Name	Definition	Methods	Assumptions	Limitations	Intended applications
1	Species-level Impact Indicator; compute_species_indicator (https://b-cubed-eu.github.io/impIndicator/reference/compute_species_indicator.html)	The impact per individual alien species over time.	**Maximum**: Assigns a species the highest recorded impact across all available scores (Blackburn et al. 2014; IUCN 2020a; Kumschick et al. 2024).	(i) The highest observed impact best reflects management concern; (ii) A species can potentially achieve its highest impact wherever it occurs; (iii) Rare, but severe impacts are more important than average impacts.	(i) Can substantially overestimate species impacts; (ii) Ignores the frequency and consistency of impact records; (iii) Sensitive to exceptional cases and local conditions; (iv) Does not distinguish species with one severe record from species consistently causing severe impacts.	Precautionary risk assessment, prioritisation for regulation, early detection and rapid response.
			Mean: Computes the average impact across all scores for a species (D'hondt et al. 2015). This method provides an expected impact value and is most appropriate when numerous impact records exist for the species.	(i) All impact records are equally informative; (ii) The average impact reflects the most likely ecological outcome; (iii) The mean adequately represents variation amongst locations and mechanisms.	(i) Can underestimate species capable of causing severe impacts; (ii) Rare high-impact events are diluted by numerous low-impact records; (iii) Sensitive to uneven sampling effort amongst species.	Comparing species objectively, estimating realised impacts.
			Sum of Maximum across Mechanisms: Calculates a species' impact score by summing the maximum impact for each mechanism through which the species exerts its impacts (Nentwig et al. 2010).	(i) Different mechanisms contribute independently to the total impact; (ii) Species affecting ecosystems through many mechanisms are more harmful than species affecting them through a single mechanism; (iii) Impacts from different mechanisms are additive.	(i) Can favour species with many documented mechanisms rather than truly greater impacts; (ii) Depends heavily on completeness of EICAT assessments; (iii) May double-count overlapping ecological effects; (iv) Species with more research effort may receive higher scores.	Comprehensive risk assessment, prioritising species with multiple impact pathways, understanding ecosystem-wide invasion threats.
2	Site-level Impact Indicator; compute_site_indicator (https://b-cubed-eu.github.io/impIndicator/reference/compute_site_indicator.html)	The impact of alien species per site in the study region over time.	**Precautionary**: This method uses the Maximum method to aggregate each species' impact and then computes the maximum impact across species in each site.	(i) The highest recorded impact is the most relevant for management; (ii) Species can potentially achieve their maximum impact wherever they occur; (iii) The most damaging species adequately represents site risk.	(i) Overestimates impacts; (ii) Ignores all other species once the highest-impact species is identified; (iii) Cannot distinguish between sites with one versus many invaders; (iv) Sensitive to rare or exceptional impact records.	Early warning, conservation planning, protected areas, risk-averse decision making, identifying sites requiring immediate attention.
			Precautionary Cumulative: Uses the Maximum method to aggregate each species' impact and then computes the summation of all impacts in each site. The Precautionary Cumulative method provides the highest impact score possible for each species, but considers the number of co-occurring species in each site.	(i) Each species can achieve its maximum impact; (ii) Impacts of different species are additive; (iii) Worst-case impacts can occur simultaneously.	(i) Tendency to overestimate impacts; (ii) Assumes all species express their maximum impacts at the same site and time; (iii) Does not account for synergistic or antagonistic interactions amongst species; (iv) Sensitive to rare impact observations.	Identifying invasion hotspots, prioritising management across sites, estimating cumulative invasion risk under a precautionary framework.
			Mean of Means: Uses the Mean method to aggregate each species' impact and then computes the mean of all species in each site. The Mean of Means provides the expected impact within individual species and across all species in each site.	(i) Mean impact reflects the most likely impact of a species; (ii) All impact records are equally informative; (iii) Average impact across species adequately represents site conditions.	(i) Can underestimates high-risk species; (ii) Rare but severe impacts are diluted; (iii) Sensitive to uneven numbers of impact records amongst species.	Estimating expected impacts, comparisons amongst sites, long-term monitoring.
			Mean Cumulative: Uses the Mean method to aggregate each species' impact and then computes the sum of all impact scores in each site. The Mean Cumulative provides the expected impact score within individual species, but adds co-occurring species' impact scores in each site.	(i) Mean impact reflects expected species impact; (ii) Species impacts are additive; (iii) Multiple species collectively contribute to site impact.	(i) Underestimates rare severe impacts; (ii) Assumes additive impacts.	Monitoring cumulative invasion pressure, comparing invasion burden amongst sites, assessing management outcomes.
			Cumulative: Uses the Sum of Maximum across Mechanisms method to aggregate each species' impact and then computes the sum of all species' impacts per site. The Cumulative method provides a comprehensive view of the site-level impact while considering the impact and mechanisms of multiple species.	(i) Species affecting ecosystems through multiple mechanisms have greater impact potential; (ii) Different mechanisms contribute independently and additively to the total impact; (iii) Mechanism diversity reflects ecological severity.	(i) Can strongly favour species with many documented mechanisms; (ii) Depends heavily on completeness of EICAT assessments; (iii) May inflate scores when mechanisms overlap ecologically.	Comprehensive risk assessments, prioritising sites affected with multiple impact pathways, understanding ecosystem-wide invasion threats.

	Name	Definition	Methods	Assumptions	Limitations	Intended applications
3	Regional Impact Indicator; compute_regional_indicator (https://b-cubed-eu.github.io/impIndicator/reference/compute_regional_indicator.html)	The cumulative impact across all alien species and sites in the study region over time.	Same as above.	Same as above.	Same as above.	Risk assessments of a region, comparing impacts between regions.

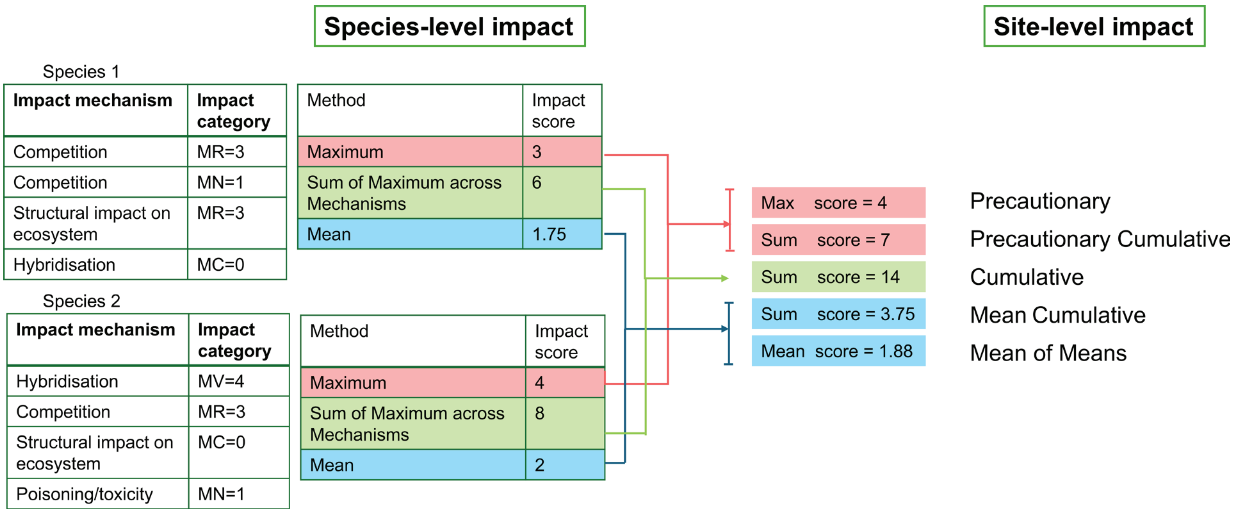

Figure 2. Methods for computing impact indicators. Given two hypothetical species occurring simultaneously at a site, each with recorded EICAT impact categories and impact mechanisms, their numerical impact scores are first aggregated within species to obtain species-level impact indicators and then aggregated across species within sites to compute site-level impact indicators. Regional impact indicators are subsequently calculated as the spatial sum of the site-level indicators. Adapted from Boulesnane-Guengant et al. (2025).

We then used these species-level scores to estimate site-level impact. As multiple alien species can co-occur at a site, the impacts of all species recorded at a site were aggregated using one of five methods proposed by Boulesnane-Guengant et al. (2025): Precautionary, Precautionary Cumulative, Mean of Means, Mean Cumulative and Cumulative. These methods differ in how impacts are aggregated within and across species (Table 1, Fig. 2).

Finally, to estimate the impact of alien species at the scale of the study region, we aggregated site-level impact values across all sites within each year. The regional impact was, therefore, computed as the spatial sum of site-level impacts using the same five aggregation approaches as those used for the site-level indicator: Precautionary, Precautionary Cumulative, Mean of Means, Mean Cumulative and Cumulative (Table 1, Fig. 2).

Uncertainty and sensitivity

To estimate the uncertainty associated with the impact indicators, we used the bootstrapping method. For each year in the occurrence cube, the procedure re-samples individual records (rows) from the cube (including impact scores) with replacement and then re-computes the indicator for each re-sampled dataset. The original cube is treated as the population and re-computing the indicator across multiple bootstrap samples yields a distribution of possible indicator values. The standard error and confidence interval are derived from this distribution (Suppl. material 1: appendix 2). We implemented the bootstrapping and calculated the confidence interval using the dubicube R package (Langeraert and Van Daele 2026). The uncertainty measure is currently available only for the regional impact indicator.

To estimate how sensitive the regional impact indicator is to individual species and to identify species that disproportionately influence the regional impact indicator, we used a leave-one-species-out cross-validation (LOSO-CV) technique. This technique evaluates the extent to which the regional impact indicator is driven by individual species by systematically recalculating the indicator after excluding one species at a time and comparing the resulting values to those obtained from the full occurrence cube. The percentage deviation between the indicator calculated after omitting a species and the full-data indicator provides a measure of the contribution of the left-out species to the regional impact indicator (Suppl. material 1: appendix 3). We implemented the LOSO-CV using the dubicube R package (Langeraert and Van Daele 2026).

Software description

R has become the primary programming language for ecological data analysis over the past decade (Lai et al. 2019; Gao et al. 2025). Therefore, we developed an R package, impIndicator (Yahaya et al. 2026), to facilitate the use of the workflow by ecologists and other stakeholders. The package implements the workflow described in sections: Occurrence data, Impact data, Impact indicators and Uncertainty and sensitivity (Fig. 1) and is fully open source. Its source code is available through the Biodiversity Building Blocks for Policy (B3) GitHub repository (https://github.com/b-cubed-eu/impIndicator). B3 is a European Union Horizon-funded project that aims to standardise access to biodiversity data for proactive policy-making (https://b-cubed.eu/). Each version release is automatically archived on Zenodo (https://doi.org/10.5281/zenodo.15052675). The package can be installed directly from the B3 R-universe repository (https://b-cubed-eu.r-universe.dev/builds):

```r
# install impIndicator in R
install.packages("impIndicator", repos = "https://b-cubed-eu.r-universe.dev")

# Load packages
library(impIndicator)
```

Workflow demonstration

We demonstrated the workflow using *Acacia* species in the continental regions of the Iberian Peninsula (Spain and Portugal combined), South Africa and California. We produced species-level, site-level and regional impact indicators for *Acacia* in these three regions, as well as uncertainty and sensitivity analyses for the regional impact indicators, using the impIndicator R package. We downloaded the GBIF occurrence cube for *Acacia* in Spain, Portugal, South Africa and California via the GBIF website (GBIF.org 2026). The downloaded cubes have an EQDGC Level 2 (0.25° by 0.25° grid cell) spatial resolution and yearly temporal resolution. We used the EQDGC because it has been widely used in African atlases and is well-suited to global analyses (Harrison et al. 1997; Larsen et al. 2009). In contrast, the European Environment Agency (EEA) reference grid is only available for Europe. The yearly interval is the basic format of the cube, which can be changed depending on the study. To test the sensitivity of the impact indicators to spatial and temporal resolution, we performed analyses at finer and broader spatial resolutions (0.125° by 0.125° and 0.5° by 0.5° grid cells, respectively) and finer and broader temporal resolutions (1/2-year and 2-year intervals) (Suppl. material 1: appendix 4). We used the EICAT assessment data from Jansen and Kumschick (2022), because these are currently more detailed than those available through the GISD. The data provide the *Acacia* species name, categories and the mechanism of the impact reported in each source. For all the analyses in this study, we used the default encoding scheme in impIndicator for the transformation of impact category to impact score; MC = 0, MN = 1, MO = 2, MR = 3 and MV = 4 because it is consistent with the original EICAT categories.

Species-level impact indicator of *Acacia* species

As an example, we first prepare acacia_cube_IBR, eicat_acacia and iberian_sf, representing the occurrence cube for *Acacia* species, EICAT data and the Iberian Peninsula's shapefile, respectively (see Suppl. material 1: appendix 5). As the spatial object iberian_sf represents only the mainland of the Iberian Peninsula, the impact indicator is computed exclusively for the mainland. Consequently, only species occurrences located within the Iberian mainland area are included in the computation.

The impIndicator R package provides the function compute_species_indicator to compute the impact indicator per species using the Mean method (Table 1). The function returns a dataset containing the aggregation method used, the number of species, species names and the impact values for each species. The results demonstrate variability in the impacts of different species across regions. Overall, eleven unique *Acacia* species with impact evidence occur across the three regions: *Acacia baileyana*, *Acacia cyclops*, *Acacia dealbata*, *Acacia decurrens*, *Acacia elata*, *Acacia longifolia*, *Acacia mearnsii*, *Acacia melanoxylon*, *Acacia pycnantha*, *Acacia salicina* and *Acacia saligna*. Specifically, *Acacia dealbata*, *Acacia mearnsii* and *Acacia baileyana* have the highest impact values in the Iberian Peninsula, South Africa and California, respectively (Fig. 3). The output can be displayed in the console using the print function. The following shows an example of the code and output for the species-level impact indicator of *Acacia* in the Iberian Peninsula.

```r
# Compute species-level impact
# Iberian
IBR_species<-compute_species_indicator(
  cube = acacia_cube_IBR,
  impact_data = eicat_acacia,
  method = "mean",
  region = iberian_sf
)


print(IBR_species)
```

```
## Species impact indicator
##
## Method: mean
##
## Number of species: 10
##
##
## First 10 rows of data (use n = to show more):
##
## # A tibble: 15 × 11
##     year `Acacia baileyana` `Acacia cyclops` `Acacia dealbata`
##    <dbl>              <dbl>            <dbl>             <dbl>
##  1  2010                  3             1.33              40
##  2  2011                 NA            NA                 53.9
##  3  2012                 NA            NA                 47.0
##  4  2013                  3            NA                137.
##  5  2014                 NA            NA                113.
##  6  2015                 NA            NA                176.
##  7  2016                  3            NA                170.
##  8  2017                 NA             1.33             170.
##  9  2018                  3             2.67             207.
## 10  2019                  3             2.67             303.
## # ℹ 5 more rows
## # ℹ 7 more variables: `Acacia decurrens` <dbl>, `Acacia longifolia` <dbl>,
## #   `Acacia mearnsii` <dbl>, `Acacia melanoxylon` <dbl>,
## #   `Acacia pycnantha` <dbl>, `Acacia salicina` <dbl>, `Acacia saligna` <dbl>
```

Site-level impact indicator of *Acacia* species

We used the compute_site_indicator function to compute the site-level impact indicator using the Mean Cumulative method. The function returns a dataset containing the aggregation method used, the number of sites, the number of species and the impact value for each site. We visualised the site-level impact indicator for 2014, 2019 and 2024 in each region, highlighting their spatial patterns of impact (Fig. 4). The following shows an example of the code and output for computing the site-level impact indicator for *Acacia* in the Iberian Peninsula.

```
# Compute site-level impact
IBR_site<-compute_site_indicator(
  cube = acacia_cube_IBR,
  impact_data = eicat_acacia,
  method = "mean_cum",
  region = iberian_sf
)

print(IBR_site)
```

```
## Site impact indicator
##
## Method: mean_cum
##
## Number of cells: 468
##
## Number of species: 10
##
##
## First 10 rows of data (use n = to show more):
##
## # A tibble: 468 × 18
##    cellCode  xcoord ycoord `2010` `2011` `2012` `2013` `2014` `2015` `2016`
##    <chr>      <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
##  1 E002N41BC  2.62    41.6   1.74 NA     NA         NA     NA  NA     NA
##  2 W000N38BB -0.125   38.9   1.95   1.95 NA         NA     NA  NA     NA
##  3 W000N39BC -0.375   39.6   1.95   1.95 NA         NA     NA  NA     NA
##  4 W000N39CA -0.875   39.4   1.95 NA     NA         NA     NA  NA     NA
##  5 W000N39DA -0.375   39.4   1.95 NA      1.95      NA     NA   1.14   1.95
##  6 W004N36AD -4.62    36.6   3.10   1.95 NA         NA     NA  NA      1.95
##  7 W004N38DA -4.38    38.4   1.74   1.74 NA         NA     NA  NA     NA
##  8 W004N40CB -4.62    40.4   1.74 NA     NA         NA     NA  NA     NA
##  9 W005N36DA -5.38    36.4   1.95 NA     NA         NA     NA  NA     NA
## 10 W005N39CA -5.88    39.4   1.74 NA     NA         NA     NA  NA     NA
## # ℹ 458 more rows
## # ℹ 8 more variables: `2017` <dbl>, `2018` <dbl>, `2019` <dbl>, `2020` <dbl>,
## #   `2021` <dbl>, `2022` <dbl>, `2023` <dbl>, `2024` <dbl>
```

Regional impact indicator of *Acacia* species

We used the compute_regional_indicator function to compute the regional impact indicator using the Mean Cumulative method. Confidence intervals were estimated from the bootstrap replicates using the percentile method. The function returns a dataset containing the method used to compute the indicator, the number of sites, the number of species and the regional impact value together with its confidence interval. The resulting plot shows the regional impact indicator of *Acacia* for each region, indicating that regional impact values are higher in the Iberian Peninsula and South Africa than in California (Fig. 5). The following shows an example of the code and output for computing the regional impact indicator for *Acacia* in the Iberian Peninsula.

```
# Compute regional impact
IBR_regional<-compute_regional_indicator(
  cube = acacia_cube_IBR,
  impact_data = eicat_acacia,
  method = "mean_cum",
  region = iberian_sf,
  ci_type = "perc"
)
```

```
## [1] "Performing whole-cube bootstrap."
```

```
print(IBR_regional)
```

```
## Regional impact indicator
##
## Method: mean_cum
##
## Number of cells: 468
##
## Number of species: 10
##
##
## First 10 rows of data (use n = to show more):
##
## # A tibble: 15 × 7
##      year diversity_val est_boot se_boot bias_boot    ll    ul
##     <dbl>         <dbl>    <dbl>   <dbl>     <dbl> <dbl> <dbl>
##  1  2010          74.5     48.8    5.23     -25.7  64.3  84.5
##  2  2011         100.      66.6    6.22     -33.6  87.5 112.
##  3  2012          82.1     55.9    5.30     -26.3  71.3  92.5
##  4  2013         253.     202.     8.02     -50.7 237.  268.
##  5  2014         208.     159.     7.44     -49.4 194.  223.
##  6  2015         316.     242.     8.91     -74.0 299.  333.
##  7  2016         379.     288.    10.4      -91.0 359.  399.
##  8  2017         322.     235.    10.1      -87.4 302.  342.
##  9  2018         372.     280.    10.0      -92.4 353.  392.
## 10  2019         541.     410.    11.9     -131.  518.  564.
## #   5 more rows
```

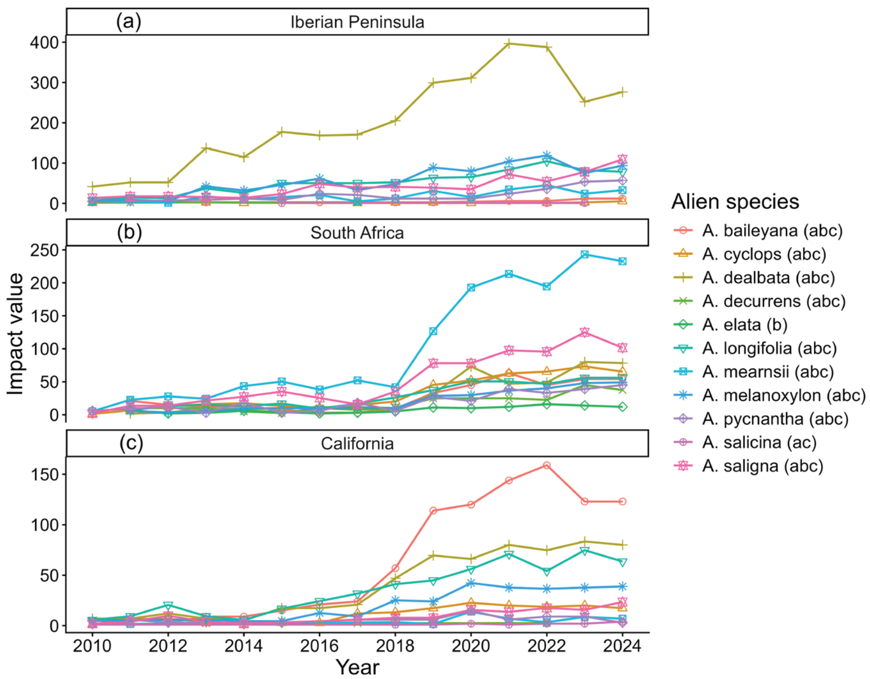

Figure 3. Species-level impact indicator of *Acacia* species in: **a.** The Iberian Peninsula; **b.** South Africa, and **c.** California. Each species' impact value is calculated as the mean impact score across its recorded assessments multiplied by the total number of sites occupied by the species. An increase in the impact indicator reflect a growing number of occupied grid cells, which may result from genuine species spread, increased sampling effort or both. The letters (a), (b) and (c) preceding each species name indicate that the species is present in the corresponding study region included in the analysis.

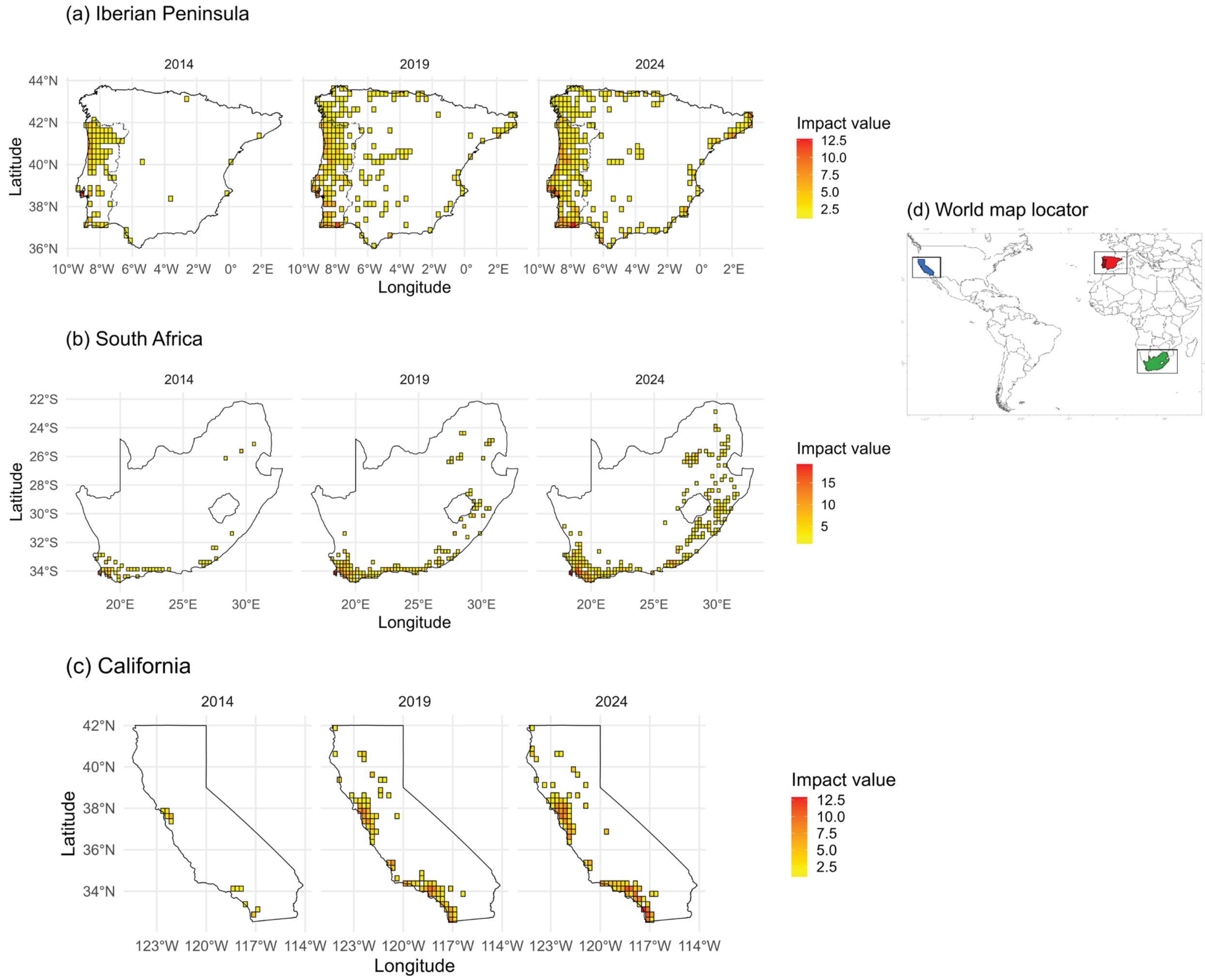

Figure 4. Site-level impact indicator of *Acacia* species in: **a.** The Iberian Peninsula; **b.** South Africa, and **c.** California for 2014, 2019 and 2024. The site-level impact indicator represents the Mean Cumulative impact score (see Materials and methods) aggregated across all *Acacia* species at each site. It is calculated by assigning each species its mean impact score of all recorded impact assessments and summing these values within each site (Fig. 2), thereby providing a spatial representation of Fig. 3, with average species impact scores summed within each grid cell. Changes in the distribution of coloured grid cells should be interpreted with caution. The disappearance of a coloured grid cell does not necessarily indicate local disappearance of a species, but may instead reflect incomplete or uneven sampling, including the absence of observations in subsequent surveys. The world map locator (d) indicates the Iberian Peninsula (red), South Africa (green) and California (blue).

Sensitivity of the regional impact indicator to *Acacia* species

We applied the LOSO-CV Percentage Error analysis to the regional impact indicator for *Acacia* in the three regions using the Mean Cumulative method. The LOSO-CV Percentage Error quantifies how individual species contribute to the impact indicators across the three regions. For example, *Acacia dealbata* contributed approximately 55% to the regional impact indicator in the Iberian Peninsula in 2010 (Fig. 6). Specifically, *Acacia dealbata*, *Acacia mearnsii* and *Acacia baileyana* made the largest contributions to the regional impact indicator in the Iberian Peninsula, South Africa and California, respectively.

Sensitivity to spatial and temporal resolution

The impact values varied consistently with spatial and temporal resolution. Specifically, the range of site-level impact values across the three regions decreased as spatial resolution

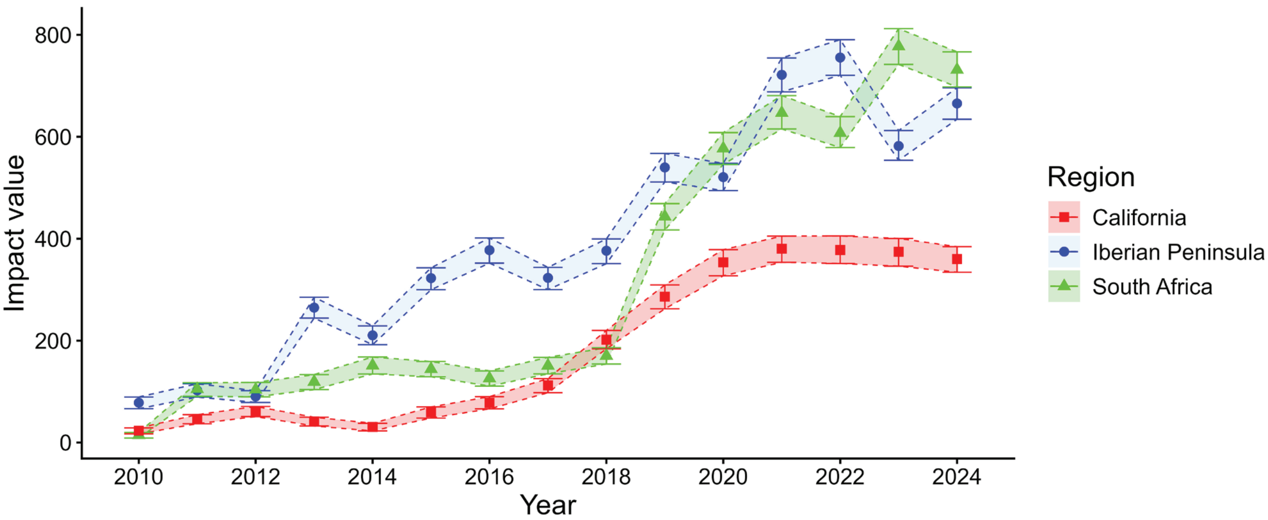

Figure 5. Regional impact indicator of *Acacia* species in the Iberian Peninsula (blue circle), South Africa (green triangle) and California (red square), illustrating the cumulative impact value for each year. The impact indicator is calculated using the Mean Cumulative method, where each species is assigned its mean impact score across all recorded impact assessments and these scores are summed across all sites for each year (Fig. 2). It therefore represents the spatial sum of the site-level impact indicators derived from the species-level indicators shown in Fig. 3. The error bars represent 95% confidence intervals, calculated using the percentile bootstrap method. Wider confidence intervals indicate greater uncertainty in the estimated impact indicator for that year. The number of bootstrap replicates was 1000. Changes in the regional impact indicator should be interpreted with caution. Increases may reflect genuine species spread, increased sampling effort or both, whereas decreases may be an artefact resulting from reduced detection or reporting rather than true declines in species distributions or impacts.

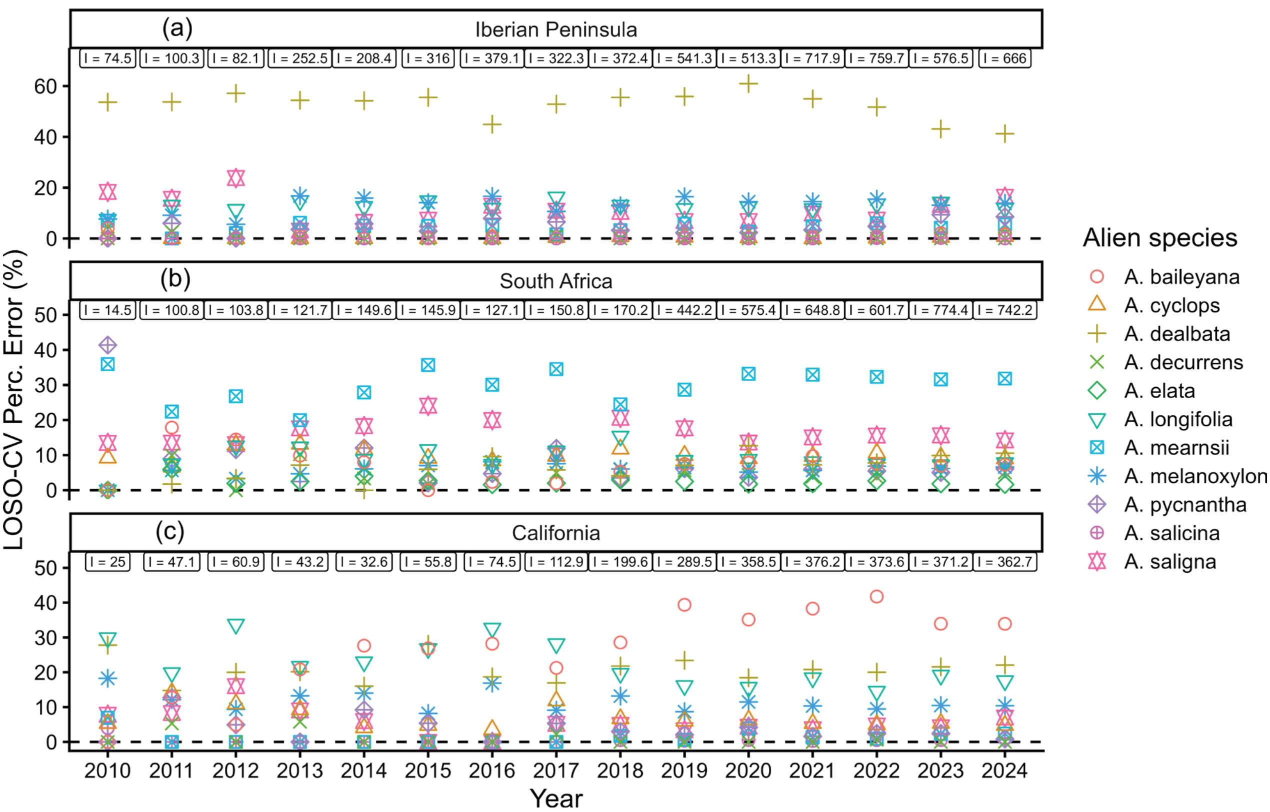

Figure 6. Sensitivity analysis of the regional impact indicator using the LOSO-CV Perc. (Percentage) Error for each *Acacia* species in Fig. 5. The LOSO-CV Percentage Error indicates the percentage contribution of each species to the regional impact indicator (denoted as *I* = at the top of each year) for each year shown on the x-axis.

increased (Fig. 4, Suppl. material 1: appendix 4, figs S2, S5). In contrast, species-level and regional impact values increased consistently with increasing spatial resolution (Figs 3, 5, Suppl. material 1: appendix 1, figs S1, S3, S4, S6). This pattern arises because finer spatial resolution aggregates fewer species within individual sites (grid cells), resulting in lower site-level impact values. At the same time, finer resolution increases the number of occupied sites, leading to higher species-level and regional impact values.

The species-level and regional impact values decreased with coarser temporal resolution (Figs 3, 5, Suppl. material 1: appendix 1, figs S7, S9, S10, S11), whereas site-level impact values showed little variation across the three temporal resolutions (Fig. 4, Suppl. material 1: appendix 1, figs S8, S12). This is because finer temporal resolution partitions occurrence records into shorter time intervals, reducing the number of species and sites aggregated within each interval and consequently lowering species-level and regional impact values.

Discussion

Workflow

Rapid and reliable monitoring information is required to enable effective management of threats to biodiversity (UNEP 2022). In this study, we developed a reproducible workflow and an accompanying R package, impIndicator, to quantify species-level, site-level and regional impact indicators that track temporal and spatial changes in the impacts of alien species. Together, the workflow and software provide a practical framework for translating occurrence and impact data into indicators that provide information for evidence-based management and policy decisions related to biological invasions. By packaging the workflow into an accessible tool, the impIndicator R package facilitates the application of these indicators by researchers, conservation practitioners and policy stakeholders.

The values of the impact indicators generated by the workflow are driven by changes in the distribution of alien species within the study region. For example, the regional impact indicator increases when an alien species spreads to additional sites or when a new alien species is recorded in the study region (Fig. 5). Consequently, the impact indicators reflect both spatial expansion and the accumulation of impacts through time. As the indicators are sensitive to the spatial and temporal resolution of the underlying data, their absolute values should not be interpreted in isolation. Instead, they are most informative when compared across species, sites, regions or time using the same methodological settings. For example, the regional impact value for 2022 can only be interpreted as indicating a higher or lower impact relative to other years. Accordingly, the primary value of these indicators lies in supporting comparative assessments rather than providing absolute measures of impact. Nevertheless, additional factors, such as sampling completeness and species detectability, may influence the estimated indicator values. We discuss the implications of sampling completeness for the *Acacia* case study later in this section.

Comparable impact indicators

Species can be assigned their maximum recorded impact score, consistent with the approach adopted by the IUCN to support a precautionary framework for biodiversity conservation (Kumschick et al. 2024). However, this practice has been questioned by

Cassini (2023), who argued that the maximum impact approach may not be appropriate for all purposes and departs from classical measures of central tendency, such as the mean, commonly used in ecological analyses. Moreover, the maximum method ignores the variation in magnitude, thereby obscuring the ranking of species for prioritisation (Volery et al. 2021). By offering flexible aggregation options, including the mean impact score and the sum of maximum impact across mechanisms, our workflow addresses these concerns by enabling users to select aggregation methods aligned with their specific research or management objectives. This flexibility also facilitates comparison with and integration of existing impact frameworks. For example, the impact risk maps described by Kumschick et al. (2025) are equivalent to our site-level impact indicators, while our workflow further extends this approach by allowing multiple options for aggregating species impact scores within sites (Boulesnane-Guengant et al. 2025).

As each indicator method employs a different aggregation strategy, the resulting indicators can produce substantially different estimates of impact. The choice of method should be guided by the objectives of the analysis and the assumptions underlying each approach. For example, the precautionary site-level indicator assigns the maximum impact score amongst species present at a site. As a result, when all species considered have been classified as causing high impacts (e.g. Major or Massive) at least once, the indicator may show little or no variation amongst sites because each site is assigned the same maximum score. In contrast, methods such as the Mean Cumulative indicator incorporate the contributions of all species and their average impacts, allowing greater variation in site-level impact values and providing a representation of cumulative invasion pressure (Boulesnane-Guengant et al. 2025).

Furthermore, the site-level impact indicator serves as a visual and analytical tool to represent the intensity of biological invasions across different parts of an area (e.g. Fig. 4). By enabling spatial comparisons, such as between provinces, states or conservation areas, it highlights hotspots and areas at risk of impact caused by alien species. These spatial data are useful for prioritising management actions, coordinating restoration projects and fostering cross-regional collaboration to address alien species impacts effectively (Potgieter et al. 2022).

The regional impact indicator offers a nuanced representation of temporal changes in alien species impacts at local, regional or global scales (e.g. Fig. 5). By tracking the increase and decrease of ecological threats over time, the indicator can provide insights into the dynamics of alien species impacts, helping assess whether current management practices are effective or need adjustment. This temporal perspective can also support more targeted resource allocation and proactive interventions to mitigate biodiversity loss and ecosystem degradation.

The LOSO-CV Percentage Error evaluates the sensitivity of the regional impact indicator to the removal of individual species, highlighting how strongly the regional impact depends on particular species and providing insight into the potential consequences of species control or removal. In contrast, the species-level impact indicator is intended for direct comparison amongst species, enabling the ranking and prioritisation of species, providing information for control strategies and advancing research on alien species' ecological roles and adaptation patterns.

The three regions

The Iberian Peninsula is one of the European hotspots for invasions by *Acacia* species (Botella et al. 2023). This prominence has spurred research on their management in recent years (Marchante et al. 2024), reflecting both ecological impacts and socioeconomic costs associated with these invasions. In Portugal, the significant environmental damage caused by *Acacia* species has led to strict regulatory measures: all *Acacia* species are currently subject to legal restrictions (Decreto-Lei no. 92/2019). Although *Acacia* species are widely distributed across mainland Portugal and the north-western and coastal regions of Spain, our results revealed substantial spatial variability in their impact across the Peninsula. Thus, a widespread distribution does not necessarily translate into uniformly high impact. This heterogeneity is especially relevant for management planning. Comprehensive control of all *Acacia* species across the entire Iberian Peninsula would require considerable financial and logistical resources. Therefore, species-level and site-level impact indicators provide a more strategic approach, enabling managers to prioritise the most impactful species and the most affected sites. Such targeted prioritisation can improve cost-effectiveness and enhance the overall efficiency of invasion management efforts in the region.

The California Invasive Plant Council (Cal-IPC) is a non-governmental organisation that works to protect California's environment from invasive plants through science and policy. The Cal-IPC classifies species according to their impacts into five categories (High, Moderate, Limited, Alert and Watch), based on state-wide spread as a proxy for their ecological impact. Ten *Acacia* species have been listed in their inventory as having a significant current impact and/or a risk of becoming invasive (The Cal-IPC Inventory – California Invasive Plant Council 2026). Most of these species are classified as Watch due to their potential future impact, leading to the conclusion that *Acacia* species currently have relatively minor ecological impacts in California compared with other Mediterranean-climate regions (Rundel 2023). This aligns with our results, which showed lower indicator values in California than in the Iberian Peninsula and South Africa (Fig. 5).

Since the EICAT impact scores assigned to each species remain constant through time, the increase in impact values, for instance between 2018 and 2021 in South Africa (Figs 3b, 5), is mainly due to additional occurrence records, which may reflect either genuine spread to new sites or increased recording effort. However, the apparent rate of spread inferred from our indicators differs from that reported by Kotzé et al. (2025), who surveyed South African sites during 2007–2008 and 2016–2023 using observers in low-flying aircraft and estimated an increase of only 10.6% between the two survey periods. This discrepancy is most likely attributable to differences in sampling effort and data sources. In particular, the dataset used by Kotzé et al. (2025) is not included in the GBIF occurrence records analysed here and the increase in records between 2018 and 2021 is likely associated with the incorporation of iNaturalist observations into GBIF. Similar effects may also influence the impact indicators in the other study regions.

Challenges and limitations

Apart from the sensitivity of the impact indicators to spatial and temporal resolution discussed above, we highlight that the indicators should be interpreted with caution, because observed trends may reflect changes in sampling effort in addition to genuine changes in species distributions. This is a common limitation of biodiversity occurrence datasets, arising from factors such as (i) delays between field observations and their publication; (ii) the exclusion of datasets that are not openly available; (iii) temporal changes in data availability as large datasets become publicly accessible (e.g. the incorporation of iNaturalist records) and (iv) non-systematic sampling. One way to reduce these limitations is to incorporate locally collected occurrence data, particularly those based on systematic or high-intensity sampling, even when they are not yet available through GBIF. Ideally, this should be complemented by more rapid and comprehensive integration of such datasets into GBIF.

Another limitation of this workflow is the current lack of EICAT impact assessments for most alien taxa, although ongoing initiatives are expanding the taxonomic coverage through coordinated assessment project, and few assessments are available in the scientific literature (e.g. Evans et al. (2016); Canavan et al. (2019); Volery et al. (2021); Jansen and Kumschick (2022)). At present, these assessments are neither fully open nor fully FAIR, requiring users to provide their own assessment data for many taxa and regions. The workflow also assumes that impact magnitudes recorded in one location are transferable to other locations (Hayes and Barry 2008), although this assumption may not hold in all ecological contexts, such as island ecosystems. Since most available impact assessments are not specific to the study region, the impact indicators presented here represent potential impacts, which may differ from the realised effects of alien species within the study region or at individual sites. As additional impact assessments become available, it will become possible to filter for impacts relevant to the region of interest or other ecological criteria. However, such geographically explicit impact data remain unavailable for most invaded regions.

Further developments

Feedback will be collected via the issue tracker on the impIndicator GitHub repository (https://github.com/b-cubed-eu/impIndicator/issues). Future versions of the package will extend uncertainty calculation and visualisation to species-level and site-level impact indicators. Recent advances in deep learning species distribution models (SDMs) (Ryckewaert et al. 2025) provide an opportunity to extend the workflow to project the potential impact of alien species across space and time. Such models combine species occurrence records with environmental predictors to estimate species distributions, including in areas with limited sampling or undetected occurrences (Pearson et al. 2007; Algar et al. 2009). This approach could help reduce biases associated with uneven or unknown sampling effort, including delays between field observations and their subsequent publication through GBIF, thereby improving the robustness of the resulting impact indicators.

Conclusions

This study presents a workflow and accompanying software for computing impact indicators for alien species using open-source data. As alien species continue to spread and increase in abundance, they can have substantial negative impacts on native species and ecosystem functioning. By integrating species occurrences with assessments of ecological impacts, we developed a workflow and an R package that computes impact indicators for alien species. These indicators represent both the spatial extent of species distributions and the severity of their potential impacts. We demonstrate the application of this workflow using *Acacia* species in three regions: the Iberian Peninsula, South Africa and California.

The workflow is sensitive to the spatial and temporal resolution of the underlying occurrence data and to the completeness of those data. Caution should be exercised when interpreting the absolute values of the impact indicator because they depend on the methodological settings and the characteristics of the underlying data. Nevertheless, comparisons of indicator values across regions or through time provide a robust framework for tracking changes in the potential impacts of alien species. Where available, users can improve the robustness of the indicators by incorporating more complete or region-specific occurrence datasets. Like all impact indicators, those developed here are proxies for realised ecological impacts. By explicitly incorporating standardised information on impact magnitude in addition to species occurrence data, our workflow provides a more informative representation of potential impacts than occurrence-based indicators alone. The workflow can support researchers, conservation practitioners and policy-makers in identifying alien species and regions with the greatest potential for significant harm. By highlighting species and regions where impacts may be greatest, this approach can help guide management actions and prioritise efforts to mitigate biological invasions, contributing to the protection of biodiversity and the maintenance of healthy ecosystems.

References

Algar AC, Kharouba HM, Young ER, Kerr JT (2009) Predicting the future of species diversity: Macroecological theory, climate change, and direct tests of alternative forecasting methods. Ecography 32(1): 22–33. https://doi.org/10.1111/j.1600-0587.2009.05832.x

Bacher S, Galil B, Nuñez MA, Ansong M, Cassey P, Dehnen-Schmutz K, Fayvush G, Hiremath A, Ikegami M, Martinou A, McDermott S, Preda C, Vilà M, Weyl OLF, Fernandez R, Ryan-Colton E (2023) Impacts of invasive alien species on nature, nature's contributions to people, and good quality of life. IPBES Invasive Alien Species Assessment, 1–222. https://doi.org/10.5281/zenodo.7430731

Bacher S, Ryan-Colton E, Coiro M, Cassey P, Galil BS, Nuñez MA, Ansong M, Dehnen-Schmutz K, Fayvush G, Fernandez RD, Hiremath AJ, Ikegami M, Martinou AF, McDermott SM, Preda C, Vilà M, Weyl OLF, Aravind NA, Angelidou I, Athanasiou K, Atkore V, Barney JN, Blackburn TM, Brockerhoff EG, Carbutt C, Carisio L, Castro-Díez P, Céspedes V, Christopoulou A, Cisneros-Heredia DF, Cooling M, de Groot M, Demetriou J, Dickey JWE, Duboscq-Carra VG, Early R, Evans TG, Flores-Males PT, Gallardo B, Gruber M, Hui C, Jeschke JM, Joelson NZ, Khan MA, Kumschick S, Lach L, Lapin K, Lioy S, Liu C, MacMullen ZJ, Mazzitelli MA, Measey J, Mrugała-Koese AA, Musseau CL, Nahrung HF, Pepori A, Pertierra LR, Pienaar EF, Pyšek P, Rivas Torres G, Rojas Martinez HA, Rojas-Sandoval J, Ryan-Schofield NL, Sánchez RM, Santini A, Santoro D, Scalera R, Schmidt L, Shivambu TC, Sohrabi S, Tricarico E, Trillo A, van't Hof P, Volery L, Zengeya TA (2025) Global Impacts Dataset of Invasive Alien Species (GIDIAS). Scientific Data 12: 832. https://doi.org/10.1038/s41597-025-05184-5

Blackburn TM, Essl F, Evans T, Hulme PE, Jeschke JM, Kühn I, Kumschick S, Marková Z, Mrugała A, Nentwig W, Pergl J, Pyšek P, Rabitsch W, Ricciardi A, Richardson DM, Sendek A, Vilà M, Wilson JRU, Winter M, Genovesi P, Bacher S (2014) A unified classification of alien species based on the magnitude of their environmental impacts. PLoS Biology 12(5): e1001850. https://doi.org/10.1371/journal.pbio.1001850

Boria R, Olson L, Goodman S, Anderson R (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecological Modelling 275: 73–77. https://doi.org/10.1016/j.ecolmodel.2013.12.012

Botella C, Marchante H, Celesti-Grapow L, Brundu G, Geerts S, Ramirez-Albores JE, González-Moreno P, Ritter M, Richardson DM (2023) The Global Distribution of *Acacia*. In: Richardson DM, Le Roux JJ, Marchante E (Eds) Wattles: Australian Acacia Species Around the World. CABI, Wallingford, 131–147. https://doi.org/10.1079/9781800622197.0009

Boulesnane-Guengant O, Rouget M, Becker-Scarpitta A, Botella C, Kumschick S (2025) Spatialising the ecological impacts of alien species into risk maps. Global Ecology and Conservation 61: e03660. https://doi.org/10.1016/j.gecco.2025.e03660

Boyd RJ, August TA, Cooke R, Logie M, Mancini F, Powney GD, Roy DB, Turvey K, Isaac NJB (2023) An operational workflow for producing periodic estimates of species occupancy at national scales. Biological Reviews 98(5): 1492–1508. https://doi.org/10.1111/brv.12961

Butchart SHM (2008) Red List Indices to measure the sustainability of species use and impacts of invasive alien species. Bird Conservation International 18(S1): S245–S262. https://doi.org/10.1017/S095927090800035X

Butchart SHM, Akçakaya HR, Chanson J, Baillie JEM, Collen B, Quader S, Turner WR, Amin R, Stuart SN, Hilton-Taylor C (2007) Improvements to the Red List Index. PLoS ONE 2(1): e140. https://doi.org/10.1371/journal.pone.0000140

Butchart SHM, Akçakaya HR, Berryman AJ, Brooks TM, Burfield IJ, Chanson J, Dias MP, Donaldson JS, Hermes C, Hilton-Taylor C, Hoffmann M, Luedtke JA, Martin R, McDougall A, Neam K, Polidoro B, Raimondo D, Rodrigues ASL, Rondinini C, Rutherford C, Scott T, Simkins AT, Stuart SN, Vine J (2025) Measuring trends in extinction risk: A review of two decades of development and application of the Red List Index. Philosophical Transactions of the Royal Society B: Biological Sciences 380(1917): 20230206. https://doi.org/10.1098/rstb.2023.0206

Canavan S, Kumschick S, Le Roux JJ, Richardson DM, Wilson JR (2019) Does origin determine environmental impacts? Not for bamboos. Plants, People, Planet 1(2): 119–128.

Cassini MH (2023) A critical review of the precautionary approach of the IUCN impact classification for non-native taxa. Conservation Biology 37(2): e14037. https://doi.org/10.1111/cobi.14037

Delavaux CS, Crowther TW, Zohner CM, Robmann NM, Lauber T, Van den Hoogen J, Kuebbing S, Liang J, De-Miguel S, Nabuurs G (2023) Native diversity buffers against severity of non-native tree invasions. Nature 621(7980): 773–781. https://doi.org/10.1038/s41586-023-06440-7

Desmet P, Oldoni D, Blissett T, Robertson T (2023) Specification for species occurrence cubes and their production. https://docs.b-cubed.eu/guides/occurrence-cube/

D’hondt B, Vanderhoeven S, Roelandt S, Mayer F, Versteirt V, Adriaens T, Ducheyne E, San Martin G, Grégoir J, Stiers I, Quoilin S, Cigar J, Heughebaert A, Branquart E (2015) Harmonia+ and Pandora+: Risk screening tools for potentially invasive plants, animals and their pathogens. Biological Invasions 17(6): 1869–1883. https://doi.org/10.1007/s10530-015-0843-1

Dove S (2026) b3gbi: General Biodiversity Indicators for Biodiversity Data Cubes (Version 0.9.1) [R]. https://github.com/b-cubed-eu/b3gbi

Evans T, Kumschick S, Blackburn TM (2016) Application of the Environmental Impact Classification for Alien Taxa (EICAT) to a global assessment of alien bird impacts. Diversity and Distributions 22(9): 919–931.

Gao M, Ye Y, Zheng Y, Lai J (2025) A comprehensive analysis of R’s application in ecological research from 2008 to 2023. Journal of Plant Ecology 18(1): rtaf010. https://doi.org/10.1093/jpe/rtaf010

GBIF.org (2025) What is GBIF? https://www.gbif.org/what-is-gbif

GBIF.org (2026) [22 June 2026] GBIF Occurrence Download. https://doi.org/10.15468/dl.474trp [Data set]

Gómez-Suárez M, Laeseke P, Seebens H (2025) A global dataset of native and alien distributions of alien species. Scientific Data 12(1): 1914. https://doi.org/10.1038/s41597-025-06379-6

Grêt-Regamey A, Sirén E, Brunner SH, Weibel B (2017) Review of decision support tools to operationalize the ecosystem services concept. Ecosystem Services 26: 306–315. https://doi.org/10.1016/j.ecoser.2016.10.012

Groom Q, Strubbe D, Adriaens T, Davis AJ, Desmet P, Oldoni D, Reyserhove L, Roy HE, Vanderhoeven S (2019) Empowering citizens to inform decision-making as a way forward to support invasive alien species policy. Citizen Science: Theory and Practice 4(1): 33. https://doi.org/10.5334/cstp.238

Groom QJ, Adriaens T, Desmet P, Simpson A, De Wever A, Bazos I, Cardoso AC, Charles L, Christopoulou A, Gazda A, Helmisaari H, Hobern D, Josefsson M, Lucy F, Marisavljevic D, Oszako T, Pergl J, Petrovic-Obradovic O, Prévot C, Ravn HP, Richards G, Roques A, Roy HE, Rozenberg M-AA, Scalera R, Tricarico E, Trichkova T, Vercayie D, Zenetos A, Vanderhoeven S (2017) Seven recommendations to make your invasive alien species data more useful. Frontiers in Applied Mathematics and Statistics 3: 13. https://doi.org/10.3389/fams.2017.00013

Hagen BL, Kumschick S (2018) The relevance of using various scoring schemes revealed by an impact assessment of feral mammals. NeoBiota 38: 37–75. https://doi.org/10.3897/neobiota.38.23509

Harrison JA, Allan DG, Underhill LG, Brown CJ, Herremans M, Parker V, Tree AJ (1997) The atlas of southern African birds (Vol. 1). BirdLife South Africa, Johannesburg.

Hayes KR, Barry SC (2008) Are there any consistent predictors of invasion success? Biological Invasions 10(4): 483–506. https://doi.org/10.1007/s10530-007-9146-5

Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D (2021) Data integration enables global biodiversity synthesis. Proceedings of the National Academy of Sciences 118(6): e2018093118. https://doi.org/10.1073/pnas.2018093118

Henriksen MV, Arlé E, Pili A, Clarke DA, García-Berthou E, Groom Q, Lenzner B, Meyer C, Seebens H, Tingley R, Winter M, McGeoch MA (2024) Global indicators of the environmental impacts of invasive alien species and their information adequacy. Philosophical Transactions of the Royal Society B: Biological Sciences 379(1902): 20230323. https://doi.org/10.1098/rstb.2023.0323

Huybrechts P, Trekels M, Abraham L, Desmet P (2024) B-Cubed software development guide. B-Cubed Documentation. https://docs.b-cubed.eu/guides/software-development/

IPBES (2023) Summary for Policymakers of the Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. In: Roy HE, Pauchard A, Stoett P, Renard Truong T, Bacher S, Galil BS, Hulme PE, Ikeda T, Sankaran KV, McGeoch MA, Meyerson LA, Nuñez MA, Ordonez A, Rahlao SJ, Schwindt E, Seebens H, Sheppard AW, Vandvik V (Eds) IPBES Secretariat, Bonn, Germany, 1–56. https://doi.org/10.5281/zenodo.7430692

IUCN (2020a) IUCN EICAT Categories and Criteria: The Environmental Impact Classification for Alien Taxa. IUCN Gland, Switzerland.

IUCN (2020b) Guidelines for using the IUCN Environmental Impact Classification for Alien Taxa (EICAT) Categories and Criteria. Version 1.1. IUCN Gland Switzerland, Cambridge, UK.

Ivanova NV, Shashkov MP (2021) The Possibilities of GBIF Data Use in Ecological Research. Russian Journal of Ecology 52(1): 1–8. https://doi.org/10.1134/S1067413621010069

Jansen C, Kumschick S (2022) A global impact assessment of *Acacia* species introduced to South Africa. Biological Invasions 24(1): 175–187. https://doi.org/10.1007/s10530-021-02642-0

Kotzé JDF, Wannenburgh A, van Wilgen BW (2025) Changes in the cover of selected invasive alien plant taxa between 2008 and 2023 in South Africa. Biological Invasions 27(3): 98. https://doi.org/10.1007/s10530-025-03558-9

Kumschick S, Bertolino S, Blackburn TM, Brundu G, Costello KE, de Groot M, Evans T, Gallardo B, Genovesi P, Govender T, Jeschke JM, Lapin K, Measey J, Novoa A, Nunes AL, Probert AF, Pyšek P, Preda C, Rabitsch W, Roy HE, Smith KG, Tricarico E, Vilà M, Vimercati G, Bacher S (2024) Using the IUCN environmental impact classification for alien taxa to inform decision-making. Conservation Biology 38(2): e14214. https://doi.org/10.1111/cobi.14214

Kumschick S, Journiac L, Boulesnane-Guengant O, Botella C, Pouteau R, Rouget M (2025) Mapping potential environmental impacts of alien species in the face of climate change. Biological Invasions 27(1): 43. https://doi.org/10.1007/s10530-024-03490-4

Lai J, Lortie CJ, Muenchen RA, Yang J, Ma K (2019) Evaluating the popularity of R in ecology. Ecosphere 10(1): e02567. https://doi.org/10.1002/ecs2.2567

Langeraert W, Van Daele T (2026) dubicube: Calculation and Interpretation of Data Cube Indicator Uncertainty (Version 0.11.0) [R]. https://b-cubed-eu.github.io/dubicube/

Larsen R, Holmern T, Prager SD, Maliti H, Røskaft E (2009) Using the extended quarter degree grid cell system to unify mapping and sharing of biodiversity data. African Journal of Ecology 47(3): 382–392. https://doi.org/10.1111/j.1365-2028.2008.00997.x

Latombe G, Pyšek P, Jeschke JM, Blackburn TM, Bacher S, Capinha C, Costello MJ, Fernández M, Gregory RD, Hobern D, Hui C, Jetz W, Kumschick S, McGrannachan C, Pergl J, Roy HE, Scalera R, Squires ZE, Wilson JR, Winter M, Genovesi P, McGeoch MA (2017) A vision for global monitoring of biological invasions. Biological Conservation 213: 295–308. https://doi.org/10.1016/j.biocon.2016.06.013

Marchante E, Colaço MC, Skulska I, Ulm F, González L, Neto Duarte L, Neves S, Gonçalves C, Maggiolli S, Dias J, Máguas C, Fernandes N, Fernandes L, Marchante H (2024) Management of Invasive Australian *Acacia* Species in the Iberian Peninsula. In: Richardson DM, Le Roux JJ, Marchante E (Eds) Wattles: Australian Acacia Species Around the World. CABI, Wallingford, 438–454. https://doi.org/10.1079/9781800622197.0027

Nentwig W, Kühnel E, Bacher S (2010) A Generic Impact-Scoring System Applied to Alien Mammals in Europe. Conservation Biology 24(1): 302–311. https://doi.org/10.1111/j.1523-1739.2009.01289.x

Oldoni D, Groom Q, Adriaens T, Davis AJS, Reyserhove L, Strubbe D, Vanderhoeven S, Desmet P (2020) Occurrence cubes: A new paradigm for aggregating species occurrence data (p. 2020.03.23.983601). bioRxiv. https://doi.org/10.1101/2020.03.23.983601

Oveisi M, Hazrati Z, Sun Y, Hejazi M, Bacher S, Schaerer HM (2026) Regional impact risk classification (RIRC): A predictive framework for plant invasion impacts. NeoBiota 107: 243–268. https://doi.org/10.3897/neobiota.107.182605

Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34(1): 102–117. https://doi.org/10.1111/j.1365-2699.2006.01594.x

Philippe-Lesaffre M, Arbieu U, Bang A, Camacho M, Cuthbert RN, Genovesi P, Kumschick S, Pili A, Seebens H, Wang S, Latombe G (2024) Extinction potential from invasive alien species (p. 2024.09.01.610685). bioRxiv. https://doi.org/10.1101/2024.09.01.610685

Potgieter LJ, Aronson MFJ, Brandt AJ, Cook CN, Gaertner M, Mandrak NE, Richardson DM, Shrestha N, Cadotte MW (2022) Prioritization and thresholds for managing biological invasions in urban ecosystems. Urban Ecosystems 25(1): 253–271. https://doi.org/10.1007/s11252-021-01144-0

Rabitsch W, Genovesi P, Scalera R, Biała K, Josefsson M, Essl F (2016) Developing and testing alien species indicators for Europe. Journal for Nature Conservation 29: 89–96. https://doi.org/10.1016/j.jnc.2015.12.001

Rundel PW (2023) Australian *Acacia* Species in California: Introduction Pathways, Invasion and Impact. In: Richardson DM, Le Roux JJ, Marchante E (Eds) Wattles: Australian Acacia Species Around the World. CABI, Wallingford, 167–180. https://doi.org/10.1079/9781800622197.0011

Ryckewaert M, Marcos D, Joly A (2025) Report on deep learning development. B3 project deliverable D4.2.

Seebens H, Clarke DA, Groom Q, Wilson JR, García-Berthou E, Kühn I, Roigé M, Pagad S, Essl F, Vicente J (2020) A workflow for standardising and integrating alien species distribution data. NeoBiota 59: 39–59. https://doi.org/10.3897/neobiota.59.53578

The Cal-IPC Inventory – California Invasive Plant Council (2026) The Cal-IPC Inventory – California Invasive Plant Council. https://www.cal-ipc.org/plants/inventory/

UNEP (2022) Decision adopted by the Conference of the Parties to the Convention on Biological Diversity. https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf

Vicente JR, Vaz AS, Roige M, Winter M, Lenzner B, Clarke DA, McGeoch MA (2022) Existing indicators do not adequately monitor progress toward meeting invasive alien species targets. Conservation Letters 15(5): e12918. https://doi.org/10.1111/conl.12918

Volery L, Jatavallabhula D, Scillitani L, Bertolino S, Bacher S (2021) Ranking alien species based on their risks of causing environmental impacts: A global assessment of alien ungulates. Global Change Biology 27(5): 1003–1016. https://doi.org/10.1111/gcb.15467

Wilson JRU, Faulkner KT, Rahlao SJ, Richardson DM, Zengeya TA, van Wilgen BW (2018) Indicators for monitoring biological invasions at a national level. Journal of Applied Ecology 55(6): 2612–2620. https://doi.org/10.1111/1365-2664.13251

Yahaya MM, Kumschick S, MacFadyen S, Landi P, Hui C (2026) impIndicator: Impact Indicators of Alien Taxa (Version 0.6.1) [R]. https://github.com/b-cubed-eu/impIndicator

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Artificial Intelligence (AI) use

The first author used Grammarly and ChatGPT to improve the grammar of this work during its preparation. After using the tools, all authors revised the content and take full responsibility for the work.

Funding

The authors are part of the team of B3 (Biodiversity Building Blocks for policy), funded by the European Union's Horizon Europe Research and Innovation Programme (ID No. 101059592). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them. We also acknowledge additional support from the National Research Foundation of South Africa (NRF 89967).

Author contributions

MMY coded the workflow in R. Supervision was performed by SK, SM, PL and CH. All authors contributed to the conceptualisation and writing of the manuscript.

Author ORCIDs

M.M. Yahaya https://orcid.org/0009-0008-9200-0863
S. Kumschick https://orcid.org/0000-0001-8034-5831
S. MacFadyen https://orcid.org/0000-0002-5316-440X
P. Landi https://orcid.org/0000-0002-6373-2992
C. Hui https://orcid.org/0000-0002-3660-8160

Data availability

The package implementing the workflow described in this study is fully open source. Its source code is hosted on the GitHub repository of the European Union Horizon B3 project (https://github.com/b-cubed-eu/impIndicator), with each released version automatically archived on Zenodo (https://doi.org/10.5281/zenodo.15052675). The package can be installed directly from the B3 R-universe repository (https://b-cubed-eu.r-universe.dev/builds).

Supplementary material 1

Supplementary information

Authors: Mukhtar M. Yahaya, Sabrina Kumschick, Sandra MacFadyen, Pietro Landi, Cang Hui

Data type: docx

Link: https://doi.org/10.3897/neobiota.109.193861.suppl1