Welcome to GBIF Science Review No.10!

Inside this compendium, you’ll find new selection of 50 summaries of recent GBIF-enabled research. With almost five peer-reviewed papers published every day whose findings are based on data from the GBIF network, both you and our communications team has a wider range of examples to review and explore than ever.

Science Review No.10 also includes a feature section on research into invasive alien species (IAS), long one of the most frequent use cases for GBIF-mediated data. Watch this space, as they say: with the release of the latest IPBES thematic assessment on IAS this year, the GBIF community—including a new task group on invasives—will coordinate work to fill the data gaps it has identified.

For those of you who wish to stay up-to-date on the latest GBIF-enabled research, visit GBIF.org/data-use.
The careful reader may note that this issue has eliminated the year from its title. Let there be no doubt: we remain committed to producing a new collection every year that updates our visuals on the annual patterns, trends and statistics. However, the featured papers themselves will no longer be limited to those published during a single calendar year.

This change reflects the fact that each of these summaries appears first on GBIF.org, as well as our interest in profiling important research as it appears. No longer will we let a calendar-year delay on our design and print schedules keep the next Science Review from you.

As always, enjoy!
Uses and trends

These visualizations of the use of GBIF-mediated data reveal more than the growing number of peer-reviewed publications supported each year. The overlay on the graph below emphasizes authors’ increasing use of DOI-based citations, a crucial practice for maintaining data provenance and scientific transparency. At right, the charts reveal the expanded geographic distribution of institutions whose researchers apply the data in their research. Each trend highlights GBIF’s widening sphere of influence and impact on scientific research.

Annual number of peer-reviewed articles using GBIF-mediated data

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Citations</th>
<th>DOI Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>249</td>
<td>2</td>
</tr>
<tr>
<td>2014</td>
<td>350</td>
<td>7</td>
</tr>
<tr>
<td>2015</td>
<td>407</td>
<td>7</td>
</tr>
<tr>
<td>2016</td>
<td>438</td>
<td>41</td>
</tr>
<tr>
<td>2017</td>
<td>696</td>
<td>117</td>
</tr>
<tr>
<td>2018</td>
<td>676</td>
<td>99</td>
</tr>
<tr>
<td>2019</td>
<td>743</td>
<td>161</td>
</tr>
<tr>
<td>2020</td>
<td>987</td>
<td>309</td>
</tr>
<tr>
<td>2021</td>
<td>1300</td>
<td>553</td>
</tr>
<tr>
<td>2022</td>
<td>1428</td>
<td>674</td>
</tr>
</tbody>
</table>
Peer-reviewed articles by region

- **Europe**: 1102 (2022) vs 984 (2021)
- **Latina America**: 525 (2022) vs 445 (2021)
- **North America**: 371 (2022) vs 398 (2021)
- **Africa**: 175 (2022) vs 130 (2021)
- **Asia**: 577 (2022) vs 437 (2021)

Peer-reviewed uses by country

- **CN**: 331 (2022) vs 227 (2021)
- **US**: 309 (2022) vs 306 (2021)
- **BR**: 167 (2022) vs 142 (2021)
- **GB**: 149 (2022) vs 140 (2021)
- **MX**: 140 (2022) vs 131 (2021)
- **DE**: 126 (2022) vs 130 (2021)
- **ES**: 105 (2022) vs 102 (2021)
- **FR**: 78 (2022) vs 78 (2021)
- **AU**: 78 (2022) vs 70 (2021)
- **IT**: 66 (2022) vs 67 (2021)
Climate change expands future ranges of arbovirus vectors

Study of underappreciated carriers of disease-causing viruses reveals significant potential range expansions under future climatic conditions

Mosquitoes carrying arboviruses pose serious health threats to millions of people worldwide. The range of *Aedes* mosquitoes is expanding, but research has focused mainly on species *Ae. aegypti* and *Ae. albopictus*, known for their role in spreading dengue, chikungunya, and Zika viruses.

In this study, researchers from Morocco explored the global distribution under current and future climatic conditions of three understudied *Aedes* species (*Ae. japonicus*, *Ae. vexans* and *Ae. vittatus*)—all confirmed transmitters of several arboviruses.
Using mainly GBIF-mediated species occurrences, the authors modelled the current ecological niches of each species based on 19 bioclimatic variables at a 5x5 km resolution. Based on general circulation models with representative concentration pathways (RCPs) of varying emission scenarios, they projected the modelled distributions into future climatic conditions.

For *Ae. japonicus*, the models predicted high suitability across its native range of Japan, the Korean peninsula and southern China, as well as non-native ranges in the eastern United States and most of western Europe. *Ae. vexans* had a similar, but broader distribution, including high suitability in southern South America. The predicted distribution of *Ae. vittatus*, however, showed suitability in lower latitudes with occurrence probabilities highest in Central and South America, across Africa, India, Southeast Asia and northern Australia.

Future models predicted range changes to all three species under all RCP scenarios with broad increases observed for low and intermediate pathways and small decreases for high and worst-case pathways.

By 2050, *Ae. japonicus* was predicted to expand its range across much of Europe, the US, Russia and central Asia, while the range of *Ae. vexans* was likely to also include northern Africa and southern Australia. For *Ae. vittatus*, future projections included expansion in sub-Saharan Africa and the Middle East.
Study of more than 10,000 lichens show that the UV absorbance capability of fungal metabolites and their resistance to leaching in humid environments helped shape global distribution.

Seventeen per cent of the 110,000 described fungal species sustain mutualistic relationships with mainly green algal species in the lichen class *Trebouxiophyceae* or the order *Trentephiliales*. In their self-sustaining symbiotic relationship, the alga, or photobiont, supplies the fungal partner, or mycobiont, with carbon through photosynthesis, while receiving protection from ultraviolet (UV) radiation through chemical metabolites produced by the mycobiont.
In this study covering more than 10,000 lichenized fungal species, researchers explored the global distributions of trebouxiophycean and trentepohlialean photobionts, while assessing the role of lichen substances and their properties driving these patterns.

By combining GBIF-mediated occurrences of fungal species and their associated photobionts with data on lichen substances and derived properties, the authors modelled a network of relationships between the environment, occurrence probabilities and chemical properties of metabolites. The model revealed a clear dominance of lichenized Trebouxiophyceae at higher latitudes, while trentepohlialean lichens showed a higher occurrence probability in the tropics. Both the potential for environmental dissolution and UV-B radiation showed maximum values at lower latitudes.

The occurrence probability of Trebouxiophyceae photobionts was positively linked with the occurrence of lichen substances, which decreased with probability for leaching and thus dissolution. Further, Trebouxiophyceae substances had higher UV-B absorbance and lower solvability than those of Trentepohliales.

Together, the study’s findings suggest that the overall environmental drivers of lichen fungi and algae distributions, such as water availability and solar radiation, are mediated by chemical and UV-protective properties of lichen substances.
Downsizing protected areas intensifies landscape fragmentation and potential biodiversity loss

Study assesses impacts of draft law proposing a reduction of a large protected area in southern Brazil, finding greater fragmentation and increased risk to threatened species

The present study aimed to expose the damage that misguided policies may have on conservation actions. It assessed the impacts of a draft law that proposed to downsize the Devonian Escarpment Environmental Protection Area (DEEPA). This site in southern Brazil’s Paraná state features Atlantic Forest and Brazilian Savanna (Cerrado) biomes, two highly diverse landscapes rich in endemic species and under intense pressure from human-made impacts.
The authors showed that approval of the proposed legislation would reduce the area currently protected within DEEPA by more than 68 per cent. For remnants of natural vegetation within the area, only 36 per cent would remain under protection, while the average distance between protected fragments would increase by 27 per cent.

By leveraging GBIF-mediated occurrences in their analysis of the region, the authors found that the proposed downsizing would reduce species protection by 70 per cent. Among species classified as threatened, the proposed reduction would eliminate protection for 51 species currently classified as threatened under the IUCN Red List of Threatened Species™, of which one is Critically Endangered (CR), 19 Endangered (EN) and 31 Vulnerable (VU).

The authors conclude the case under study shows that when misaligned with national and global environmental goals, economic development may contradict sustainable development, leading to increased carbon emissions, greater loss of biodiversity and diminished ecosystem services.
Modelling Chagas disease in the Atlantic Forest of south-eastern Brazil

Study finds that humidity, temperature and mammalian richness influence vector occurrence and infection rates of disease-causing parasite.

Trypanosoma cruzi is a haemoflagellate protozoan with a complex life cycle and the causal agent of Chagas disease, affecting millions of people in Latin America, especially in rural areas. Little is known about the transmission patterns and the effect of the environment on this vector-borne disease.

This study based in the Atlantic Forest of Espírito Santo (ES) State, Brazil, explores the biotic and abiotic variables that modulated the occurrence of a highly infected insect vector, Triatoma vitticeps, and T. cruzi infections.
Using data collected via citizens capturing insects in their residences and submitting them for analysis at municipal public health agencies, the authors created models of vector occurrence and parasite infection based on variables for soil, vegetation, elevation, climate and mammalian species richness. Derived from GBIF-mediated species occurrences, the mammalian species richness was included as previous studies had identified a direct relationship between loss of mammal richness and increase of resilient *T. cruzi* reservoirs and thus, infection risk.

Assessing correlations of vector occurrence and parasite infection with environmental variables, the authors found that the patterns were best explained by relative humidity, average temperature, soil type, altitude and mammalian richness. The central and southern regions of ES presented as transmission hotspots—with the highest distributions of *T. vitticeps* and *T. cruzi* infections—with optimal conditions for all variables present in these regions.
Determining the drivers of genetic variation in North American plants

Study coupling micro-satellite data with species distribution modelling reveals mechanisms that shape genetic diversity among 91 angiosperms and gymnosperms

In order to conserve biodiversity and predict responses to global change, understanding the drivers of genetic variation and their relative importance is essential. In this paper, researchers presented a phylogeographic framework for testing the contributions of historical climatic shifts versus the positions of populations within species ranges in shaping genetic diversity among North American plants.

By analysing genetic data from 1,400 populations of 91 native North American plants that flower (angiosperms)
and those that do not (gymnosperms), the authors derived summary statistics estimating levels of genetic diversity and population-specific differentiation.

They then used GBIF-mediated occurrences to produce species distribution models (SDM) representing the current climate and conditions at the Last Glacial Maximum (LGM, ~22,000 years ago). In addition to ecological suitability, the models were used to quantify distances to range edges and potential glacial refugia for each genetic population.

Combining the genetic statistics with the output of the SDMs, the authors found that genetic diversity significantly decreased as distance to potential glacial refugia increased. Likewise, as the suitability of a population’s LGM habitat increased, so did the corresponding genetic diversity.

The analysis also demonstrated a clear relationship between distance to range edge and suitability under current climatic conditions, accompanied by a significant increase in genetic diversity. While providing clear support for demographic effects related to the so-called “central marginal hypothesis,” the study’s overall results suggest that history and demography have roughly equivalent importance in shaping genetic variation.
Global decline in capacity of coral reefs to provide ecosystem services

Study of 50 years of reef surveys finds coral coverage halved, which coupled with steep declines in catch numbers, suggests severe losses in reef-associated species abundance

Millions of people worldwide rely on coral reefs and the ecosystem services they provide such as food, livelihoods and buffering against extreme climate events. For coastal Indigenous peoples in small-island developing states, seafood consumption is 15 times higher than non-Indigenous populations, constituting up to 90 per cent of dietary animal protein in some Pacific nations.

This study examined how coral coverage, reef-associated biodiversity and fisheries changed globally from 1957 to
2007. Analysing surveys of 3,500 reefs in 87 countries, researchers estimated a decline in coral coverage of ~5 per cent per decade and a total of 50 per cent decline from the 1957 baseline.

Using occurrences of more than 40,000 reef-associated macroalgae, macroinvertebrates and fish species mediated by OBIS, GBIF and other sources, the authors calculated species richness by exclusive economic zone (EEZ). These showed the greatest coral-associated biodiversity in Pacific countries and territories. Applying a regression model they further revealed a positive relationship between estimated reef area and the species richness among 94 EEZs.

The countries with the highest per-capita consumption of coral fish species among Indigenous peoples were Palau, Micronesia and Kiribati. Global reef-associated catches increased from 1950 until peaking in 2002 after which they declined. Global catch per unit effort (CPUE)—an indirect measure of species abundance—peaked in the 1970s and then dropped significantly. In the EEZs of many countries, the decline over 50 years was close to 100 per cent.
Old vs. young: testing hypotheses for the formation of the Amazon basin

Study uncovers biogeographic evidence supporting the Old Amazon hypothesis for the formation of the Amazon fluvial system

Covering eight countries and one overseas territory, the Amazon ecosystem represents the largest continuous tract of tropical rainforest in the world. The formation of the Amazon basin rivers is believed to be a main contributor to the high biodiversity of the ecosystem.

This study examines the fit of phylogenetic and biogeographic data to two different hypotheses around the timing of the formation of the Amazon fluvial system: the “Old Amazon” (up to 10 million years ago) vs. the “Young Amazon” (around 2.5 million years ago).
Researchers used genetic sequence data for six monophyletic taxa across four animal classes—amphibians, birds, insects and mammals—to reconstruct time-calibrated molecular phylogenies. Using GBIF-mediated occurrences they performed ancestral range inferences, testing different dispersal rates and potential isolation barriers.

Their results show evidence of initial dispersals to the basin ranging from around 16 to 10 million years ago with full taxa coverage between 12 and six million years ago. Not all taxa agreed with the temporal range, but in general the results of the molecular and biogeographical analyses provide support for the “Old Amazon” model.
Mapping the likely geography of SARS-CoV-2’s evolutionary origins

Study identifies novel coronaviruses closely related to COVID-19 pathogen in hotspots of high bat diversity

While millions of people have died as a result of Covid-19 since the beginning of the pandemic, the precise origin of the disease-causing pathogen, the SARS-CoV-2 virus, remains elusive. Massive scientific efforts have gone into sequencing viral genomes while sampling potential wildlife in order to determine hosts and/or reservoir species, as most human viral pathogens have zoonotic origins. So far, SARS-CoV-2 relatives have been identified in several species of Asian horseshoe bat (genus Rhinolophus) as well as the Malayan pangolin (Manis javanica).
In this study, researchers reviewed swabs and samples collected from bats in Yunnan province, China, between May 2019 and November 2020 and identified 24 novel coronavirus genomes from different bat species, among them four SARS-CoV-2-like viruses present in nine samples. One of the novel bat coronaviruses, RpYN06 (short for “Rhinolophus pusillus, Yunnan, sample no. 6”), exhibited 94.5 per cent overall sequence identity to SARS-CoV-2, making it the closest relative identified to date in certain genomic regions.

Strikingly, the new viruses described in the study as well as previously identified species with high genetic similarity to SARS-CoV-2 were all discovered within a relatively small area (11 km²) of Yunnan province.

In parallel, the authors also sought to narrow the geographic search for SARS-CoV-2’s origins by modelling ecological niches of 49 horseshoe bat species using GBIF-mediated occurrences. Their investigations showed suitable habitats for Rhinolophus in the tropical forests of Southeast Asia, with as many as 23 species co-occurring in hotspots extending from South Vietnam into Southern China.

While the virus’s immediate progenitor remains unknown, these lines of evidence clearly show the enormous phylogenetic and genomic diversity of coronaviruses in Southeast Asia. Increased surveillance of the region’s bats and other wild mammals may help track spillovers of pathogenic viruses from animals to humans.
Predicting the occurrence of mycetoma in Sudan

Study links probability of neglected tropical disease to aridity, proximity to water, sodium in soil and richness of thorny tree species

Mycetoma is a neglected tropical disease caused by bacterial and fungal infections leading to chronic inflammation and eventually deformities, if untreated. The disease has been reported in 102 countries, but Sudan currently has the world’s highest case burden.

In hope of helping to detect and manage cases in Sudan, this study sought to identify environmental predictors of the disease while mapping probability and identifying potential hotspots.
Based on a database of locations of patients at the onset of symptoms, researchers constructed an ensemble model using a wide range of candidate predictors, including data on climate, soil, livestock, waterways, vegetation and GBIF-mediated occurrences of thorny trees—believed to be a major cause of wounds through which infections can happen.

For eumycetoma (caused by fungal infections) cases, probability increased with proximity the nearest waterway and greater diversity of thorny trees. The latter, however, had no effect on probability of actinomycetoma (caused by bacterial infections), which in addition to proximity to waterways was affected by mean temperature, aridity and soil sodium concentration.

The maps produced for the study showed geographic variation in the risk of mycetoma linked to environmental factors, predicting the highest occurrence in the central and southeastern parts of Sudan and along the Nile and its tributaries.
Climate change and fungal pathogens threaten global tea production

Study modelling effects of climate change predicts loss of a third of tea growing areas and increased exposure to fungal pathogens

The most widely consumed beverage in the world after water, tea is a cultural staple playing a significant role in many countries' economies. Climate change may reduce suitable tea-growing habitat and increase pathogen infectivity and susceptibility.

In this study, researchers examined the potential effects of a warming climate on tea and important associated fungal pathogens. By combining GBIF-mediated occurrences with data from the literature, the authors modelled the potential distribution of the two main varieties of cultivated Camellia sinensis (L.) Kuntze observed in Hangzhou, China by rhinolin (CC BY-NC 4.0)
tea (*Camellia sinensis* var. *sinensis* and *C. sinensis* var. *assamica*) as well as three disease-causing fungal species—at baseline and three future climate scenarios.

The models showed significant reductions in suitable habitat for both varieties of up to third of the baseline area by 2050. Overlaps in the suitable area for tea and pathogens were predicted to range from 10 to 45 per cent—the worst case being the overlap of *Colletotrichum acutatum* on *C. sinensis* var. *sinensis* in southern China.

Overall, the models predicted that temperatures in current tea growing areas would reach optimum levels for the selected fungal species, favouring the growth and spread of the pathogens. While tea cultivation may be moved into more favourable climates, the wider climate tolerances of pathogens mean that innovative approaches will be needed to protect future tea production.
The evolution of fruit colour

Study finds support for the “disperser syndrome hypothesis” but also identifies roles of abiotic factors in the evolution and geography of fruit colour syndromes

Relying on animals for seed dispersal, many plants need to advertise their fleshy fruits and attract frugivorous birds and mammals—in accordance with the perceptual abilities of the dispersing clade. The disperser syndrome hypothesis (DSH) associates “contrastive colours” i.e. black, blue, red, with bird dispersal, while linking fruits of “cryptic” “colours” i.e. green, brown, yellow, with mammal dispersal.

In this study, researchers tested DSH, while attempting to flesh out a potential role of abiotic factors on fruit colour
evolution, as contrastive colouring of fruits by anthocyanins could also imply a means of simply protecting the photosynthetic apparatus against UV radiation—or the damaging effects of cold.

The authors set out with a dataset of fruit colouration for more than 3,100 fleshy-fruited plants. They used GBIF-mediated occurrences to determine the distributions of the plants and derive climatic factors. This was combined with bird and mammal breeding ranges and diet data to create a model to test the various hypotheses.

Their results found evidence for both biotic and abiotic variables contributing to the evolution of fruit colour. Cryptic colouration occurs largely in warm environments where the relative bird frugivore prevalence is low. When wet season temperatures are low and/or bird prevalence is high, however, most fruits display contrastive colours.
Impediments to understanding seagrasses response to global change

Review of seagrass knowledge reveals gaps and biases in sampling practices across spatial, temporal and taxonomic dimensions

A major clade of about 70 species belonging to the Alismatales order, seagrasses are important marine plants, playing key roles in productivity, nutrient cycling and carbon sequestration, while also providing food for marine herbivores and nursery grounds for many fishes and invertebrates.

Focused on exploring the response of seagrasses to global change, this paper assessed the current available knowledge of the group in order to identify potential gaps and biases that might impede successful predictions.
Exploring sampling biases using occurrence data available for the clade through GBIF.org, the study reveals spatial gaps in Western and Central Indo-Pacific and temporal bias towards spring and summer months—and very few records from before 1900. The authors also detected a weak phylogenetic signal in collection frequency, indicating some degree of taxonomic bias within the clade. The study also points to a lack of available phylogenetic and evolutionary knowledge for seagrasses.

With 31 per cent of seagrass species in global decline, conservation action requires addressing these shortfalls. The authors suggest increased funding for seagrass monitoring programmes and specimen digitization, support for technological advances, such as DNA sequencing, and improving analytical and computational tools available to researchers.
Synthesizing tree biodiversity data to understand global patterns of vegetation

Study assesses potential of efficient integration of tree biodiversity data—including distribution, abundance, traits, sequences and conservation—for producing a holistic view of global vegetation

In terrestrial ecosystems, trees constitute the majority of biomass, providing structure, habitats, and resources for surrounding lifeforms. Playing a major part in global climate through their role in carbon and water cycling, trees are among the best-studied plant groups.

With the aim of producing a more holistic understanding of vegetation, researchers evaluated the potential
for integration of all available data on trees, assessing representation and identifying gaps in data and steps required to fill them.

Based on a list of more than 58,000 tree species derived from GlobalTreeSearch and the Global Inventory of Flora and Traits (GIFT), the authors compiled all available data, including occurrence data mediated by GBIF, trait data from TRY, abundance data from sPlot, molecular sequences from GenBank, and conservation assessments from IUCN.

Only about 10 per cent of species had data available across all databases. Coverage of distribution data was very good with nearly 85 per cent of trees having at least one record in GBIF. Availability of multiple traits, molecular data and conservation assessments was far less—about 45 per cent.

The study highlights trees as good models for understanding vegetation patterns, but points to gaps in data, particularly in the tropics, suggesting technological advances such as DNA barcoding, near-infrared spectroscopy and LiDAR remote sensing as potential high-throughput solutions for filling gaps.
Separating the sedges using population genomics and niche modeling

Study using ecological niche modeling combined with DNA sequencing reveals species complex as distinct evolutionary entities

The science of delineating organisms within the proper taxonomic boundaries has improved significantly with the introduction of molecular based analysis, allowing for conclusions based on chromosome numbers, protein markers and other sequence information.

This study combined morphological analysis with population genomics and ecological niche modelling to assess taxon limits within the Carex jemtlandica–Carex lepidocarpa species complex, plants belonging
to the large Carex genus of more than 2,000 species of grass-like plants, commonly known as sedges.

By sequencing and analysing DNA from 192 individuals collected primarily in eastern and central Norway, the authors reveal two distinct genetic clusters, largely corresponding to C. jemtlandica and C. lepidocarpa.

Using GBIF-mediated occurrences of Carex spp. specimens combined with bioclimatic data, the authors employed a multiple ensemble modelling approach, effectively creating an average model of seven different algorithms. Comparing the resulting models indicated that the two species occupy completely different environmental niches.

The overall result suggest that C. jemtlandica and C. lepidocarpa are separate and evolutionarily significant entities.
Lagging behind climate change—debt for the North, bonus for the South

Study of European plants find climatic debts in colder Alpine and Boreal regions, but climatic bonus in warmer Mediterranean regions

When the climate changes, species can either track their climatic optimum by moving in space—or respond adaptively adjusting to new climatic conditions. Climate change, however, often happens faster than species are able to respond, leaving them with an unpaid “climatic debt”.

To verify the existence of such debt, French researchers studied temporal shifts in climatic conditions and occupancy trends for 4,000 of the most common vascular plant species in Europe over the last 65 years.
Using more than 100 million GBIF-mediated species occurrences combined with six bioclimatic variables on temperature and precipitation, the authors calculated annual species climatic indices (SCI) for each species and variable and estimated species occupancy trends.

While negative correlation between SCI and occupancy trends—indicating a climatic debt—was not significant at the continental level, the two coldest biogeographic regions (boreal and alpine) showed an increase in temperature associated with negative occupancy trends over time, suggesting insufficient range shifts, i.e., climatic dept.

Surprisingly, the opposite pattern was present in the warmest Mediterranean region, as the occupancy of species experiencing temperature increases throughout their range tended to increase, suggesting a climate “bonus”.

Duchenne F, Martin G and Porcher E (2021)

France

European plants lagging behind climate change pay a climatic debt in the North, but are favoured in the South

Ecology Letters

doi.org/10.1111/ele.13730
Exposure to plant biodiversity reduces risk of cancer in children

Large New Zealand birth-cohort study finds statistically significant and dose-dependent association between increased plant diversity and protection against childhood leukaemia

A diverse gut microbiota is essential for human health, as many diseases are associated with reduced microbial diversity in the intestines. Vascular plants are important habitats for microorganisms such as bacteria, with microbial diversity increasing with plant diversity.

Setting out to test the so-called biodiversity health hypothesis and demonstrate a direct link between exposure to plant diversity and specific health outcomes,
this study followed all children born in New Zealand between 1998 and 2013 (and their mothers) from the time of conception to age five, tracking plant exposure while identifying cases of acute lymphoblastic leukaemia (ALL), the most common childhood cancer.

The authors estimated the plant exposure of every child (and mother during pregnancy) based on where they lived and a composite metric of GBIF-mediated plant occurrences, land cover data, a satellite imagery-based vegetation index and census data classifying a location as urban or rural.

After five years, the study identified 264 cases of ALL among the almost 950,000 children tracked. Based on the calculated exposure metrics, the study found that children in the highest tertile of plant diversity had a 35 per cent reduced risk of developing ALL.
Hierarchical framework for mapping wild bee pollination at the local scale

Study combines species distribution modelling and ecosystem service mapping to predict relative pollination potential

Wild bees are important pollinators playing a key role in crop cultivation and wild plant reproduction. Like other insects, wild bees face declines due to pollution, land use change, agricultural intensification and climate change.

This study proposes a framework that combines advantages of species distribution modelling and ecosystem service mapping to develop a high-resolution model that predicts spatial patterns of pollination ecosystem services potential from wild bees at the local scale.
In the first step, researchers used GBIF-mediated occurrences combined with five bioclimatic variables and seven land use/cover classes to model the potential distribution of 25 important *Andrena* bee species at the global scale.

In the study area of Bornhöved Lake District in Schleswig-Holstein, Germany, the authors then applied habitat filters and foraging ranges of each species to refine the output of the species distribution models at a local scale, producing a map of relative pollination potential.

By deploying traps in rapeseed fields of the study area, the authors were able to test the predictive powers of their model, finding correlation between predicted pollination service potential and observed *Andrena* richness and abundance.
Assessing biases and broader applicability of expert range maps

Study of more than 50,000 animals shows bias of expert range maps at administrative borders and failure to capture all known occurrences

Having accurate maps of species distributions is fundamental to assessing conservation priorities and developing targeted conservation strategies, but also to understanding basic biodiversity patterns.

This study analyses expert range maps of 50,000 animal species to test the assumption that these provide consistent and standardized estimates of species’ ranges.

By rasterizing and stacking species range boundaries, the authors created boundary density maps grouped by
higher level taxa (mammals, dragonflies and damselflies, amphibians, birds and reptiles) and overlaid these with different features such as administrative boundaries.

This exercise demonstrated an average of 20—30 per cent of non-coastal species range boundaries coinciding with country borders, many with no clear geophysical boundaries. When considering species richness, 60 per cent of areas with the highest spatial turnover in species occurred at political boundaries.

Finally, the authors compared the expert maps with GBIF mediated data, finding 80 per cent of taxa having more than 30 per cent of their occurrences outside the corresponding range map.

Taken together, these results reveal high bias of expert range maps at administrative borders, suggesting a need for alternative approaches to reconstructing patterns of distribution.
Automated species identification using 19th-century zoological illustrations

Study explores large-scale zero-shot learning for automated classification of scientific illustrations to reduce time needed for digitization

Scientific illustrations have historically served as perhaps the most important medium for conveying a species’ characteristic traits, offering a means of highlighting and delineating miniscule details not always well-suited to photography. Such illustrations are often stored in museum repositories and archives in undigitized form, where they remain unavailable for generic use.

The use of automated methods for reducing the time and effort required for digitization—including
interpretation of historical names and other metadata—may help facilitate access to such important sources of historical knowledge.

In this study, a team of Dutch researchers explore “zero-shot learning” to address the problem. In brief, this approach allows for recognition of objects for which no direct examples of the object class are observed during training. Instead, the method relies on embedded class information from other sources.

The authors trained their model on a dataset of 14,502 illustrations of 7,973 animal species from *Iconographia Zoologica*, embedding a class hierarchy based on the GBIF backbone taxonomy, literature from the Biodiversity Heritage Library (BHL) and dimensional features from photographs from iNaturalist.

Evaluating the model on an unannotated dataset of digitized illustrations of historical fauna of Indonesia, the researchers achieved an overall classification accuracy of around 35 per cent. While this may seem low, illustrations from 80 classes—with zero examples for training—were categorized correctly, showing the potential of computational methods for embedded models of species classification.
Climate change shifts European cephalopod distributions north-east

Study of cephalopods in regional waters suggests shifts in environmental suitability drive populations increases north—and decreases south—of the English Channel

Squid, octopuses and cuttlefish are important species in fisheries worldwide. This short-lived group is sensitive to changes in environmental conditions, so climate change may negatively influence their sustainable exploitation.

In this study, researchers aggregated species occurrences of Octopus vulgaris (common octopus), Sepia officinalis (common cuttlefish) and Loligo vulgaris
(common squid) from OBIS, GBIF and SeaLifeBase to model their ecological niches in European waters.

The authors considered various environmental parameters such as salinity and temperature, applied in an ensemble modelling framework. The approach combined several different algorithms that integrated machine learning and regression-based methods, to produce predictions about the current and future environmental suitability and distribution for each species.

Overall, the models predicted increases in environmental suitability in all areas north of the English Channel, shifting the centres of the species’ distributions north-eastward. The highest increases were observed in the North Sea for cuttlefish and squid and in the Baltic Sea for octopus. General decreases, however, were observed for all three species in the Bay of Biscay and across the Mediterranean.
Global assessment of the effects of climate change on freshwater fish

Study of more than 16,000 species suggests the extinction of almost half of current freshwater fish by mid-century

Several studies of freshwater fishes have examined the effects of climate change on individual species or in specific countries, many suggesting that the group may be especially vulnerable considering changes in hydrological regimes.

Taking a holistic approach, this study provides a global assessment of predicted climate-induced changes to the diversity and richness of watersheds around the world, modelling the current and future distributions of more than 16,000 freshwater fish species based on GBIF-mediated occurrences.
Under two different climate scenarios, the study predicts the complete disappearance of distributional areas of about half of all freshwater fish species by the year 2070, a decline more pronounced in tropical river basins and particularly high in southeast Asia, mainly affecting species with smaller body size.

On average, across all watersheds, approximately 20 per cent of species are predicted to be lost by 2050. The most important climatic predictors observed were annual precipitation and isothermality (the relationship between mean daily temperature and annual temperature range), such that species richness declined at higher rates in watersheds with high annual rainfall and temperature seasonality. In addition to a reduction in number of species, the predicted richness decline will also lead to a significant reduction in phylogenetic and functional diversity.
Current and future coverage of protected areas in the Appalachian Mountains

Climate change may shift species ranges to improve coverage of protected areas for some species while worsening for others

Protected areas (PA) are an important tool to help conserve global biodiversity. As the climate changes, the coverage of PAs for some species and habitats may change— for better or worse.

In this study, researchers evaluated the current and future PA coverage in the US Appalachian Mountains for the species ranges of 258 amphibians, birds, mammals and reptiles of conservation policy concern. Using GBIF-mediated occurrences, the authors modelled the current bioclimatic niche of each species, forecasting
future range shifts by integrating climate projections from circulation models and greenhouse gas concentration pathways.

The current PA coverage was found to be rather poor with more than 90 per cent of species having less than 15 per cent of their estimated ranges covered. Assuming species are able to disperse freely, future climatic conditions would aggravate coverage for amphibians and mammals, but improve coverage for birds and reptiles.

The authors suggest the Blue Ridge region as a candidate for increasing protection with both current and future environments suitable for the greatest number of species.
Global decline in wild bee diversity

Study finds that the number of bee species observed globally has been dwindling since the late 20th century

Wild bees are the most important group of insect pollinators, essential to the reproduction of hundreds of thousands of wild plant species and 85 per cent of all cultivated plants. There is increasing evidence of declines in wild bee populations, but most studies to date have been limited to smaller taxonomic groups and local, national or regional contexts.

In this study, researchers reasoned that a global decline among bees would be observable as a decrease in population sizes and ranges, resulting in a diminished
chance of observing bees and thus a reduced number of species recorded in GBIF.

By analysing all occurrences of six bee families (the “true bees”), the authors first demonstrated a clear increase in the number of occurrences per year from 1900 until 2015. While the number of species with records available through GBIF per year also increased steadily throughout the first half of the century, it plateaued and then declined—particularly after 1990.

When exploring the geographic distribution of bees by continent, the authors found marked differences in contributions to global diversity and the periods of decline. Europe and North America have the best temporal and spatial coverage in GBIF and exhibited the steepest declines. However, all continents except Oceania appeared to be contributing to the worldwide decline in bee diversity observed.

While these results might be affected by heterogeneous data collection, the authors found it more likely that these trends represent real scenarios of declining bee diversity. In the best scenario, thousands of species have become too rare to be observed, while in the worst scenario, they have gone extinct.
Prioritizing conservation of Indonesian medicinal plants

Study analyses priority medicinal plant diversity in Indonesia, providing recommendations on short and long-term conservation action.

Medicinal plants—defined as higher plants used for medicinal purposes—number more than 50,000 species worldwide, and global export value is estimated to exceed $3.3 billion.

Encompassing two biodiversity hotspots, Indonesia is a megadiverse country with as many as 7,500 medicinal plant species. In this study, researchers used GBIF-mediated occurrences of 233 priority species in a gap analysis to identify sites ideal for in situ reserves and ex situ collection.
The study identifies the western part of Java as the richest area with as many as 82 unique species. Forty-one 50 x 50 km cells were identified as potential candidates for *in situ* conservation, of which 33 overlapped with protected areas.

For *ex situ* conservation, the study pointed to Western Java and Maluku as most in need of further collection. Thirty-eight species had fewer than five known occurrences and, of these, only six were present in national *ex situ* conservation.

The authors of the study propose establishing active conservation of priority medicinal species inside current protected areas. They also suggest intense domestication and propagation of the six priority species already present in living collections to support their re-introduction in natural habitats.
Gaps in North American butterfly inventory knowledge

Study assessing inventory completeness of butterflies in Canada, the US and Mexico reveals narrowing data gaps, but increasing bias

Gaps in openly-accessible biodiversity data limit attempts to address important ecological and evolutionary questions. Knowing where and when gaps occur is critical not only for choosing scales for analyses, but also for prioritizing efforts to fill such gaps.

In this study, researchers assessed the inventory completeness of North American butterflies, i.e., how many species have been recorded versus how many species would be expected. With range maps derived from field guides as a baseline for expected richness,
the author used species occurrences from GBIF, iDigBio and eButterfly as the comparison for recorded richness.

Overall, more than 90 per cent of occurrences fell within delineations from range maps. Inventory completeness was spatially heterogeneous, with noticeable gaps in the far north of Canada, midwestern United States and northern Mexico. A comparison between occurrence records based on specimens with those based on community observations revealed a slightly higher average completeness ratio for the latter. Incomplete sampling was highest in regions facing the most severe threats from a rapid changing climate—particularly in areas surrounding Hudson Bay.

The study proposes initiatives of community partnerships to increase both opportunistic sampling, as well as structured surveys. As human footprint is found to be the main driver behind inventory completeness, the authors argue that such efforts must be directed far from densely-populated regions, if they are to be effective in filling data gaps.
Indicators to support assessment and decision-making around trends in biodiversity information

Framework introduced to assess global and national biodiversity data coverage suggests need to shift incentives from numbers of records to value, provide support for more targeted data mobilization and capacity building.

The conservation and management of biodiversity relies on comprehensive evidence to establish baselines and track changes. This dynamic setting highlights both the need for and importance of reliable indicators whose metrics are integrated within flexible analytical frameworks that are readily updated.

In this paper, researchers propose two novel indicators—the Species Status Information Index (SSII) and the
Species Sampling Effectiveness Index (SSEI)—in support of global assessment, monitoring and decision-making around trends in biodiversity information. The authors also provide a platform for the continuous tracking of the indices in the future.

To demonstrate the indicator framework, the authors used nearly half a billion GBIF-mediated species occurrences to assess more than 31,000 terrestrial vertebrates. Their results found that 42 per cent of nations showed increasing trends in coverage over the previous decade, driven largely by a rapid increase in bird data. Fewer than 20 per cent of all countries had increased trends for non-bird taxa. While citizen science initiatives provide invaluable contributions to biodiversity data, these findings indicate that they have not yet been optimized to close gaps in knowledge. To ensure improvements in data coverage, the authors recommend shifting incentives from numbers of records and species collected to value of records contributed.

Some regions that previously showed comprehensive sampling, e.g. Western Europe and Australia, appeared to have slowed in their progress, likely a result of nations operating at maximum capacity. For these, shifting investments to support data mobilization and capacity building in nations lagging behind targets, may help contribute to goals defined through the UN Convention on Biological Diversity (CBD). As examples of such projects, the authors highlighted GBIF’s BID and BIFA programmes.
Prioritizing conservation targets for Indo-West Pacific mangrove forests

Study maps habitat suitability of ten mangroves, predicts future richness decreases and identifies priority areas and species

Mangrove forests provide a range of essential ecosystem services with direct impact on coastal community livelihoods. With climate change and human activity in general threatening mangroves, identifying conservation targets is critical, especially in areas with rich mangrove diversity but limited response capacity.

In this study, authors from Sun Yat-sen University in China selected 10 representative mangrove species from different families, with genus-level distributions limited to the Indo-West Pacific (IWP) region, home to more than 50 species occurrences.
species. Seven species were true mangroves, while three species were mangrove-associates.

Using GBIF-mediated species occurrence records rarified to 5 km grid cells and 30 environmental variables, the authors developed ensemble models based on eight distinct algorithms, mapping the potential distribution of the 10 IWP species in past, present and future climates. The produced maps revealed that about 10 per cent of the IWP area was currently suitable for mangroves and associates. The overall suitable area was predicted to increase by 2050, however, increases were driven by low species richness, and the models predicted a sharp decrease in cells with moderate-to-high species richness.

The study identified priority sites for conservation action in several countries, including areas identified as past refugia, while ranking five species as the highest priority. The authors evaluated the current coverage of sites within protected areas, and highlighted the need for specific mangrove management plans, especially in developing countries.

A data-driven geospatial workflow to map species distributions for conservation assessments

Workflow makes it easier to map species for conservation action, produces estimates with higher accuracy than expert-derived range maps

Assessing species’ risk of extinction for conservation actions usually involves a range of metrics based on the distribution of the species. Species distribution modelling is a common approach used to predict the extent of suitable habitat, however, such models may rely on idiosyncratic decisions and be challenging to implement for multiple species.

In this study, authors developed a simple five-step geospatial workflow relying on a user-friendly and deterministic method called inverse distance weighting (IDW) to map
the area of habitat (AOH) within a species’ range—suitable for conservation planning and decision-making.

Using GBIF-mediated occurrences of forest birds from the Americas, the authors chose 723 species as their case study to evaluate the performance of the IDW workflow. They used the occurrence data to draw and define the extent of occurrence (EOO) of each species and derived absences from eBird hotspots in which a species was never recorded.

When compared to expert-drawn maps from BirdLife, the IDW workflow produced maps of 40 per cent higher accuracy with substantially fewer errors of both omission and commission. While the authors found no significant difference in the size of ranges, AOH values from expert-drawn maps were consistently smaller than IDW-derived estimates.

With a lower technical threshold than species distribution modelling and an overall improved accuracy over expert-drawn maps, the authors argued that the reproducible IDW workflow may be used reliably for identifying critical areas for conservation actions and establishing protected areas.
The colonial legacy of herbaria

Study suggests that colonial exploitation of plant specimens has inverted the relationship between where plant diversity naturally occurs and where key scientific evidence is catalogued and housed.

Four hundred million specimens residing in herbaria worldwide form the basis of our understanding of the planet’s flora. Originally collected to describe species and identify species of potential economic value, herbarium specimens have served repeatedly as the basis of further innovation, but these collections carry persistent colonial legacies that arose from the framework of exploitation often surrounding early botanical exploration.
In a study published in *Nature Human Behaviour*, Park and colleagues from 31 countries on every continent used data on more than 85 million plant specimens to analyse geographic collection and storage patterns for specimens. The authors drew on more than 50 million GBIF-mediated records along with data from another 92 herbaria in 39 countries.

The compilation of country-by-country matrices for the periods before and after the second World War (WWII), revealed colonialism’s strong imprints in collections. Most of the world’s flora was stored in temperate regions, and institutions in the United States and several European nations housed more than twice the number of species that occur in these nations. The historical trend of specimen movement to Europe and North America from the rest of the world has remained largely constant after WWII. The proportion of specimens collected from other continents has increased after WWII, with the US emerging as the largest collector of overseas specimens.

The authors suggested that, to help ensure inclusiveness in future herbarium curation and use, acknowledgements of their colonial legacy should appear in collections alongside specimen interpretations and role in scientific research. They also highlighted targeted funding initiatives, such as Biodiversity Information for Development (BID), as apotential means of addressing disparities. Most importantly, the authors stressed that the needs and wishes of the people who lived under colonial rule must guide efforts to address the colonial history of these collections.
Extinction risk of Mesoamerican crop wild relatives

Expert group assessed more than 200 species of Mesoamerican relatives of important crops, finding a high proportion severely threatened.

Crop wild relatives (CWR) are the wild cousins of domesticated crops and an important source of genetic diversity that can help strengthen important food sources in the face of climate change and other threats. Mesoamerica is a significant centre of domestication, wild relative diversity and origin of many important crops, such as maize, squash, avocado, vanilla and cotton.

In this study, a stakeholder workshop selected 224 priority CWR taxa related to crops cultivated or domesticated in Mesoamerica and assessed their extinction risk according
to IUCN Red List categories and criteria. Using occurrence data from GBIF and other sources, experts derived species distribution and richness maps, calculating the extent of occurrence (EOO) and area of occupancy (AOO) for each species of interest.

Among the assessed CWR taxa, 35 per cent were considered threatened including seven Critically Endangered (CR), 48 Endangered (EN) and 16 Vulnerable (VU). The crop with the highest proportion of threatened relatives was Vanilla at 100 per cent (eight taxa, all CR or EN), followed by cotton (Gossypium) with 92 per cent, avocado (Persea) with 60 per cent and maize with 44 and 33 per cent for Zea and Tripsacum, respectively.

The highest diversity of CWR taxa was located in the Mexican states of Jalisco (31 taxa) and Oaxaca (28 taxa), while most threatened taxa were present in Veracruz. The dominant threat processes included land use change introduced by agriculture and farming, pests and invasive alien species.

With climate change affecting existing crops and agricultural practices threatening CWR and the genetic potential they represent, the study highlights a need for promoting more sustainable agriculture and implementing policies supporting CWR conservation in Mesoamerica to help ensure future food security.
Global warming reshuffling marine plankton communities

Study featuring ensemble models of global plankton distributions suggests major restructuring of assemblages towards the end of the century with potential threat to plankton-mediated ecosystem services

Forming the basis of the ocean’s food web, marine plankton are essential to a number of ecosystem services including carbon sequestration. Increasing sea temperatures may promote speciation and diversity, but also force species to shift their ranges to more suitable climates.

In this paper, researchers built an ensemble of species distribution models for more than 800 species of phyto- and
zooplankton based on occurrence records mediated by GBIF and other sources, coupled with environmental variables such as surface temperature, oxygen concentration, solar irradiance and macronutrient levels.

For the present time, the constructed models predicted strong latitudinal richness gradients with overall annual means decreasing from the equator towards the poles. Phytoplankton displayed richness peaks in tropical upwellings, whereas zooplankton peaked in the subtropics.

Projecting the models into a future climatic scenario of high greenhouse gas emissions, the researchers found an overall global increase in mean species richness by the end of the century, driven primarily by changes in ocean surface temperature. The models, however, also predicted poleward shifts of species’ distributions at a velocity of 35 km/decade, leading to net increases in species diversity in temperate latitudes and decreases in the tropics.

As the ranges of species shift, the composition of plankton communities would be restructured with novel species interactions. The authors suggest that the combined changes could weaken the biological carbon pump effect of plankton carbon sequestration with the risk of functional ecosystem collapse in particularly sensitive regions.
Study finds surprising positive role of human disturbance on macrofungal diversity and distribution, and identifies potential biodiversity hotspots in southern Scandinavia

Playing a key role in nutrient cycling and carbon storage, fungi form a large and diverse group about which our knowledge remains limited compared to animals and plants, hampering conservation and environmental management efforts.

To determine the macrofungal diversity and distribution at a continental scale, researchers combined GBIF-mediated occurrences of 1,845 species with

Netherlands, Australia, Germany

The critical role of tree species and human disturbance in determining the macrofungal diversity in Europe

Global Ecology and Biogeography
doi.org/10.1111/geb.13372

data on 25 environmental variables to create a predictive model at 5 km resolution across eight European countries. The study distinguished between the two main functional groups: ectomycorrhizal (mutualistic symbionts) and saprotrophic (decomposers) fungi.

Analysing the model output, the authors found the highest distribution probability and species richness in eastern Denmark and southern Sweden, identifying new fungal biodiversity hotspots for both functional groups. On the opposite end of the probability scale, France and Spain had relatively low species richness.

The patterns revealed by the model were best explained by the dominant tree species, tree cover and human footprint. Saprotrophic diversity increased with number of beech (Fagus spp.) trees while ectomycorrhizal species were most likely to occur in areas dominated by spruces (Picea spp.).

While generally found to have a negative impact on biodiversity, human activity in this study had a positive correlation with macrofungal distribution probability. The authors speculated that anthropogenic alterations may create novel, otherwise rare habitats benefitting certain fungi.
Climate change threatening medicinal plants of Indonesia

Species distributions models predict significant future losses of suitable climate for two-thirds of Indonesian medicinal plant species

Encompassing two biodiversity hotspots, Indonesia is one of the megadiverse countries of the world, home to many plants relevant to medicine with immense potential value both locally and worldwide. Changes to climate, however, are affecting and threatening the distribution of these species.

Using GBIF-mediated occurrences of 139 priority medicinal plants in Indonesia, the authors of this paper created models of the species’ distributions under current climatic conditions and future conditions.
based on projections of greenhouse gas emissions for the years 2050 and 2080.

The models predicted that two-thirds of Indonesian medicinal plants will lose rather than gain areas of suitable climate. Over half of the plant species examined stood to lose up to 80 per cent of their suitable area of distribution—with the largest losses seen in Papua, Java and Sulawesi.

By assessing the future threat level of medicinal plants according to IUCN Red List criteria, the authors identified twenty species as the highest priorities for recommended long-term conservation efforts in the country.
Using machine learning to predict plant trait distributions from citizen science photos

Combining trait data with observational photographs, study creates neural network model for automated predictions of global distributions of plant traits

How ecosystems respond to global change can best be assessed by examining the functional traits available among species within the ecosystem. Rapid ecosystem monitoring could be facilitated by effective community trait measurements aided by machine learning techniques.

In this paper, authors designed a Convolutional Neural Network (CNN) aimed at predicting plant traits from photographs, trained using the TRY database of known species traits, e.g. leaf area, seed mass and...
stem density, combined with GBIF-mediated iNaturalist observations with photographs.

Adding data on trait plasticity and local climate at the location of an observation, the authors created an ensemble model showing high replicability and low error rates, generalizing well across growth forms, taxa and biomes.

To demonstrate the potential of the model, the authors used a separate set of iNaturalist observations to derive a global map of plant trait distributions. The produced map expressed a unimodal latitudinal distribution peaking around the equator for traits such as leaf area, height and seed mass, while leaf nitrogen concentration also showed peaks in northern temperate and polar zones.

In North America, a longitudinal gradient of high (in the east) to low (in the west) leaf area was apparent, while the opposite trend was observed for nitrogen concentration and stem density. These predictions combined correlated significantly with existing global maps of plant traits.
More than half the species in the IUCN Red List’s Data Deficient category face risk of extinction

New GBIF-enabled study analysing machine learning-derived probabilities of extinction suggests that up to 85 per cent of amphibians categorized as Data Deficient could disappear

In this study, researchers from the Norwegian University of Science and Technology (NTNU) present results from a novel machine learning classifier suggesting that more than half the species categorized as Data Deficient (DD) in the IUCN Red List of Threatened Species™ (Red List) face increased probability of extinction.

The threats are even more shocking for some groups, with the classifier predicting that 85 per cent of DD-listed amphibians and more than half of the DD

Oedipina capitalina Solis, Espinal, Valle, O’Reilly, Itgen & Townsend, 2016 observed in Comayagüela, Honduras by Josue Ramos Galdamez (CC BY-NC 4.0).
marine invertebrates, insects, mammals and reptiles are threatened by extinction. The Red List categorizes nearly 15 per cent of the nearly 150,000 assessed species as Data Deficient, which often gives an impression of lower risks and leads to their exclusion from studies of biodiversity impact and change. In line with some previous studies, the present analysis suggests that a larger portion of DD species may actually be more threatened than data-sufficient species.

The classifier relied on a range of more than 400 predictors, human pressures and environmental stressors including taxonomy, habitat preferences, expert range maps and GBIF-mediated occurrences, based on which variables on climate, land cover, human footprint, pesticide uses and several other factors were derived. The authors used a dataset of 28,363 DS species for training and testing the classifier, obtaining an overall accuracy of 85 per cent. Applied to a dataset of 7,699 DD species, probabilities of being threatened by extinction were higher on average than for DS species.

The paper’s authors have created a interactive website on which users can test the algorithm of the classifier and explore the full dataset of DD species and their extinction risk estimates.

Earlier this year, GBIF and IUCN achieved a milestone in their collaboration with the release of a new feature that allows users to filter occurrences by global IUCN Red List Category, including Data Deficient.

Borgelt J, Dorber M, Høiberg MA and Verones F (2022) Norway

More than half of data deficient species predicted to be threatened by extinction

Communications Biology
doi.org/10.1038/s42003-022-03638-9
Study: More than 60,000 people in Southeast Asia may be infected with SARS-related coronaviruses every year

Research estimates the risk of human infections from SARS-related coronaviruses originating in 23 bat host species across Southeast Asia

In this study, researchers estimate that more than 66,000 people in Southeast Asia are directly infected each year with SARS-related coronaviruses (SARSr-CoV) originating in 23 bat species.

These same viruses have caused massive disruptions worldwide since the emergence of SARS in 2002, MERS in 2012 and COVID-19 in 2019. SARSr-CoV spillovers from bats are thought to have occurred via intermediate hosts, such as domestic food animals or wild animals in close contact with humans, but recent evidence also
points to direct human-bat transmission. Many infections that produce mild or no symptoms with little human-to-human transmission may go undetected, but they still represent an opportunity for viral adaptation and a risk of improved human-to-human spread. To estimate the risk of human infection with SARSr-CoVs, the authors of this study used a probabilistic assessment that defines risk as a combination of number of people living in an area and the probabilities of human/bat contact and detecting infections.

The authors derived maps of areas of habitat for 26 known bat SARSr-CoV host species occurring in Southeast Asia from IUCN range maps that were refined by habitat suitability and verified with GBIF-mediated occurrence data. The results comprise a consensus area of SARSr-CoV-carrying bats that spans more than 5 million km² and is home to nearly 500 million people. Incorporating data from literature on human-bat contacts, virus-positive blood samples from humans reporting bat encounters, and human SARS antibody duration, the authors estimate that within this consensus area, a median of 66,280 people are infected with SARSr-CoVs every year.

The findings and accompanying maps may improve public health measures, pandemic preparedness, and prevention for future outbreaks by identifying the regions most at risk and encouraging targeted surveillance and early detection of spillover events and identification of novel bat coronaviruses.
Farming seaweed to feed, power and save the planet

Study explores the potential for large-scale seaweed cultivation to advance sustainable development goals and provide environmental and socio-economic benefits through a carbon-sequestering blue economy

Is it possible that large-scale seaweed farming could simultaneously reduce agricultural land conversion, mitigate greenhouse gas emissions and provide an economically valuable alternative source of food, animal feed and biofuels?

In this study, researchers first estimated the global cultivation potential by modelling the distribution of 34 seaweed species based on occurrence data from GBIF, the
Atlas of Living Australia (ALA) and the Ocean Biodiversity Information System (OBIS) combined with a suite of spatially explicit oceanographic data layers. Using existing literature, the authors then reviewed the potential and sustainability for this suite of species to be produced for human food, livestock feed and biofuel conversion.

Informed by these findings, they modelled different scenarios of impact of seaweed substitution on terrestrial agriculture and global sustainability, focussing on food, feed or biofuel, both separately and combined. They also considered a special scenario of supplementing livestock diets with a specific seaweed taxon (Asparagopsis), which reduces methane production.

The distribution models found ~650 million hectares (Mha) of ocean suitable for seaweed cultivation—roughly three times the size needed to satisfy the demands of all main scenarios combined. In the food scenario alone, in which seaweed constituted 10 per cent of global diets, as many as 110 Mha of natural land could be spared, also reducing emissions by more than a billion tonnes of CO₂ per year.

Supplementing just 0.5 per cent of ruminant diets with Asparagopsis would have the potential of mitigating CO₂-equivalent emissions by around 2.6 billion tonnes per year, requiring only a tiny fraction of sea area compared to all other scenarios. This represents 7.5 per cent of global emissions—equivalent to all carbon emissions from fossil fuels in India—in 2021.
The role community ecology in zoonotic disease outbreaks

Reduced richness of mammals in deforested landscapes in India’s Western Ghats linked to risks of outbreaks of Kyasanur Forest disease virus (KFDV)

Outbreaks of zoonotic viral diseases are becoming more prevalent as humans put more pressure on wildlife habitats—which makes the understanding of the influence of community ecology on the processes of pathogen circulation and spillover events ever more critical.

Kyasanur Forest disease virus (KFDV) is an emerging tick-borne pathogen in the Western Ghats region of India, and in this study, researchers focused on testing the role of species richness and forest loss in shaping
outbreaks. As a measure of species richness, the authors first modelled the combined distributions of 24 mammal species using GBIF-mediated occurrences along with variables of climate, forest cover and human footprint.

When juxtaposed with the distribution of KFDV outbreaks, the models uncovered significant and substantial associations between disease, species richness and forest loss. As such, increased mammalian richness was associated with increased KFDV outbreak occurrence, however, only in the context of minimal forest loss. As forest loss increased, the association was reversed, as lower species richness meant increased disease risk.

The study identified specific primates—the bonnet macaque (*Macaca radiata*), the black-footed gray langur (*Semnopithecus hypoleucos*) and the tufted gray langur (*Semnopithecus priam*)—whose relative abundance was associated with increased KFDV outbreaks. The authors suggested the incorporation of monitoring for these species into surveillance programmes as sentinels of KFDV circulation.

Overall, the study enforced previous findings indicating pathways from habitat fragmentation and biodiversity loss to increased spillover. While unable to pinpoint specific mechanisms of viral maintenance and infection ecology, the results suggested that species resilient to landscape change might be most relevant to KFDV spillover events.
Integrating multiple indices of biodiversity to identify novel areas of significance for conservation

Study uses measures of weighted endemism to reveal a series of regional species-rich areas worthy of conservation in the Qinghai-Tibet Plateau

Species richness is not the only measure of biodiversity, and while the number of distinct species in a given area is important, factors such as endemity, phylogenetic diversity and evolutionary distinctiveness are also relevant when identifying priority areas for biodiversity conservation.

In this study of the Qinghai-Tibet Plateau, researchers used GBIF-mediated data of 12,000 seed plant species.
to assess several measures of biodiversity including species richness, weighted endemism, phylogenetic diversity, phylogenetic endemism and biogeographically weighted evolutionary distinctiveness at a resolution of 50 × 50 km.

Their overall results showed a consistent gradual decrease of diversity from the southeast to the northwestern part of the region. To be able to identify areas of high endemism which do not necessarily contain a particularly high number of species overall, the authors used the corrected weighted endemism (CWE) measure and discovered several unique, previously unrecognized diverse areas. These were in contrast with other diversity metrics, which overlapped with already known hotspots.

Nine of the newly identified areas of high endemicity were located where protection is absent or inadequate, suggesting a need for re-evaluating regional conservation strategies.
Sampling biases shape our view of the natural world

Massive study describes extensive taxonomic and geographic biases in available species occurrence data in terrestrial and marine realms, highlighting need for increased sampling in less developed, distant and inaccessible regions of the world.

In a comprehensive global analysis of more than 740 million species occurrences from GBIF and OBIS, researchers explored spatial and taxonomic coverage of a representative selection of vertebrate and invertebrate groups to identify biases and their potential drivers.

The results showed that at a five km resolution, less than seven per cent of the Earth’s surface—five percent of oceans and 11 per cent of land—has been sampled.
disregarding birds, the coverage dropped to just four per cent of oceans and seven per cent of land.

In terms of taxonomic bias, less than half a per cent of genera in OBIS data and a mere 100 bird species in GBIF, accounted for more than 50 per cent of all records, respectively. Geographically, nearly 80 per cent of GBIF data came from just 10 countries, and coverage was strongly correlated with GDP per capita.

The study also found that eighty per cent of all records were within 2.5 km of roads. Between 22 (for mammals) and 47 (for arachnids) per cent of records were found within 1 km of a city. Despite representing only nine per cent of land, temperate broadleaf forests contained around 50 per cent of terrestrial records.

In the marine realm, 18 per cent of records and 41 per cent of species coincided with the busiest shipping routes covering only two per cent of the oceans. Representing 65 per cent of the total area, open oceans had 32 per cent of all marine records.

The study calls for increased support for museum specimen digitization, strategic standardized surveys and data sharing—highlighting GBIF initiatives such as BID and BIFA as examples—to improve coverage in less-accessible areas, such as high mountains and deep seas, where citizen science approaches are insufficient.
If corals go extinct, half the world’s tropical fish may follow

Statistical models show that coral extinctions due to reef bleaching will impact far more species than expected and may reduce global reef fish richness by more than 40 per cent.

Coral reefs contribute food and structural complexity for habitats of fish and other reef organisms. While many reef fish may survive increasing ocean temperature, corals are more sensitive, and effects of extinctions caused by bleaching may extend far beyond the species that strictly depend on corals.

In this study, researchers mapped the global distribution of coral and tropical fish richness using
data from GBIF, OBIS and IUCN. They then estimated fish richness dependency on corals using limited natural history information and a statistical model of indirect environmental effects of coral richness on fishes.

Simulating a world without corals, the model predicted that 41 per cent of tropical reef fish species depend on corals and would be lost as well. For a more comprehensive view of the effects of coral loss on fish communities, the authors also considered phylogenetic and functional diversity, for which the model predicted a decline of 32 and 23 per cent, respectively.

While total coral reef extinction may sound pessimistic, climate-based projections predict that 100 per cent of the world’s coral reefs will have been exposed to at least one bleaching event by the year 2060 in the most severe carbon emissions scenario.
Using soil hydromechanics to predict global earthworm distribution

Researchers develop a mechanistic biomechanical model that highlights strong correlation between habitat suitability and observed distributions for key subterranean “ecological engineers”

Earthworms are considered important “ecosystem engineers” as the soil biopores they form by burrowing can serve as pathways for water flow, aeration and reuse by growing roots. Previous research has suggested, for example, that earthworm activity may enhance certain crop yields by up to 25 per cent.

To study how earthworms disturb soils, a Swiss team of researchers started by calculating that the maximum...
pressure sustained by a worm hydroskeleton was 200 kilopascals (kPa). Applying this measure to global data on soil moisture and texture, they then modeled 0.1 degree grids around the world to identify where earthworms could burrow based on mean expansion limit pressures. After removing regions with permafrost, acidic pH, sand content and other limiting factors, they fine-tuned their mechanistic biophysical model to predict hospitable soil zones delineating the potential for earthworm migration in changing climates.

Using GBIF-mediated earthworm data they tested the predicted distributions against observed occurrences, finding that 86 per cent of reported earthworm presences agreed with the model classification. Upon examining the remaining false negatives, however, the authors found that these records often corresponded to local geographical features, such as river banks and anomalous precipitation zones.
INVASIVES
About this special section

Invasive alien species remain a dominant topic in GBIF-enabled research, and this small subsample of eight papers mirrors the larger corpus in its range of interests and concerns. Together, they demonstrate the broad applicability of location-specific biodiversity data from the GBIF network in understanding invasion dynamics and biosecurity threats viewed from the perspective of the past, present or future.

The research can sometimes focus and zero in on the role of single species. Elsewhere it pulls back to scan centuries’ worth of invaders on our planet. Meanwhile, the ability to analyse and reveal the evolving interplay of traits and interactions in new or changing environments has taken on greater urgency as climate-change impacts alter or strengthen the many costs of invasive species. In every case, GBIF remains a trusted resource by providing easy access to information essential for supporting action to prevent or reduce threats from biological invasions.
Anthropogenic habitat disturbances associated with human development determine botanical naturalization

Study finds that both climate and socioeconomic factors help explain differences in the spread of European plants around the world

Half of all widely naturalized species in the world are native to Europe, and more than 20 per cent of the continent’s native vascular plants have naturalized outside the continent. This vast collection of introductions is three times more than would be expected by chance, making European flora an obvious candidate for studying naturalization patterns.

In this study, researchers used GBIF-mediated vascular plant occurrences to model the potential distribution of 1,485 European endemics outside Europe. Comparing
these results with the realized ranges of the same species present on the Global Naturalized Alien Flora (GloNAF) database, they calculated a “naturalization debt”—that is, a proportion of species that do not yet occupy an otherwise suitable region.

The authors found that European endemic plants were only naturalized in less than five per cent of their suitable regions outside Europe, meaning that in terms of area, more than 95 per cent of expected naturalizations have not yet been observed.

Exploring socio-economic drivers, the study also found that naturalization debt decreased with the level of anthropogenic disturbances associated with human development, meaning that more naturalized species are observed in more developed regions.

Counter-intuitively, the average proportion of unsuccessful naturalizations in unsuitable environments was lower in developed regions with more international treaties relevant to invasive alien species. This, however, is more likely a result of regions with many invasive species joining treaties—rather than treaties promoting invasions.
Factors influencing successful naturalization and invasion of alien plants in India

Case study explores variables of biogeography, introduction pathways, use, traits and climate for more than 700 plants to determine why some alien species become naturalized, and some naturalized species become invasive.

When a plant is introduced to a new region, it only becomes naturalized, and later, invasive, if able to overcome certain barriers. Studies indicate that only about 10 per cent of introduced plant species naturalize, and of these, only a tenth become invasive.

In order to identify parameters influencing successful transitions on the introduced—invasive continuum, this
study focussed on India, home for 11 per cent of the world’s plant diversity and three major biodiversity hotspots, collecting data for 715 alien plant species in the country.

For each species, the authors determined 13 variables including level of establishment, the size of native and naturalized ranges, introduction pathways, and relevant functional traits, relying on GBIF-mediated occurrence data to infer climatic preferences, land use classes and the time since introduction.

Overall, the study found that invasive alien plants had larger naturalized ranges, a greater number of uses and higher specific leaf area than their naturalized and casual counterparts. In addition to range sizes, path analyses showed that growth form directly influenced naturalization success, whereas time since introduction directly influenced invasion success, further affected by number of native congener.

The authors emphasized inclusion of trading regulations on alien plants while promoting native species in national policy frameworks. With time having a direct impact on invasion success, alien plants already naturalized should be prioritized for early detection and removal.
The role of temperature in shaping the impacts of marine invasions

Ecological impact of marine invasions is most severe within the thermal niche of the invader

When marine species are redistributed by human transport systems, such as ballast water, the success of establishment and magnitude of ecological impact depend on a variety of factors. Temperature is a fundamental driver of physiological performance of species, however, its role in shaping impacts of invasive marine species remains largely unexplored.

In this study, researchers set out to quantify the impact of invasive species in marine ecosystems, exploring how these may relate to differences in temperature. The authors identified 50 species with recorded ecological
impact, and then used GBIF-mediated data to assess the ranges of the species and characterize their realized thermal niches.

Their analysis showed that most invasions occurred poleward of the cool edge of a species’ range with an overall mean shift of 6.8 degrees latitude. Thermal mid-points at these locations were on average 1.2 degrees cooler than those of the native ranges.

The ecological effect of invasions on abundance of native taxa was most severe at sites 2.2 degrees cooler than the origin of the invader, with decreasing effect toward higher and lower differences in temperature.

The study concludes that seemingly harmless species established outside their native ranges need to be monitored as they may develop larger impacts with future warming oceans.

Plant and insect invasions follow waves of globalization

Study analyses invasion dynamics of plants and insects across three centuries, finding links to world trade openness and clear impact of major global events with reduced trade such as wars and financial crises

Increased global trade and human movement have facilitated the spread of alien and invasive species worldwide. Since the Industrial Revolution, the world has experienced several waves of globalization that may have accelerated biological invasions, but the influence of these waves on invasion dynamics remains unclear.

In this study, Swiss researchers performed a global temporal analysis of rates of alien species first records between...
1750 and 2000 using data of more than 10,000 insects and plants from the Alien Species First Records database. To account for variation in sampling effort, the authors extracted first records for more than 500,000 native species from GBIF.org for the same period.

The analysis revealed that increased alien first record rates were strongly correlated with sampling effort. For insects, an invasion wave was observed between 1820 and 1870 with first record rates 143 per cent higher than expected. Plant invasions started earlier, with first record rates 39 per cent higher than expected from 1750 to 1870.

Both groups experienced reduced invasion rates from around 1900 to 1960 with observed first records ~20 per cent fewer than model expectations, coinciding with major global events (World War I, the Great Depression and World War II).

To test the impact of large-scale variations in globalization on overall invasion dynamics, the authors used a global trade openness index, calculated based on the relationship between total import/export and gross domestic product (GDP). For both groups, they found strong correlations between trade openness and invasions, explaining both increases and decreases over the past centuries, making globalization a major socio-economic predictor of biological invasions.
Native to China, the spotted lanternfly planthopper (*Lycorma delicatula*; SLF) has spread, established and become invasive in the eastern United States, feeding on vineyard grapes, threatening wine production through vine deaths and near-total yield losses.

Recognizing SLF as an emerging paninvasive species and potential threat to the $300B global wine industry, researchers from Pennsylvania—where the first US introduction took place ca. 2014—developed a paninvasion severity assessment framework, presented in this study.
Based on US Centers for Disease Control and Prevention pandemic assessment methods adapted to invasion process theory, the framework relied on the estimation of three invasion potentials in uninvaded US states and countries worldwide: transport (metric tonnage of goods imported from invaded states), establishment (ensemble of species distribution models using GBIF-mediated occurrences) and impact (tonnage of grapes and wine produced).

Illinois, Texas and California had the heaviest trade with invaded states and also the highest establishment potential, and, for California, the highest impact potential. This pattern of aligned invasion potentials carried through on a global scale where the major grape-producing regions (i.e. highest potential for impact) also had the highest risk of transport and establishment.

Concluding with an overall estimated risk of SLF disrupting the global wine market of 80 per cent, the authors strongly recommend immediate coordinated efforts to reduce transport, establishment and impact potentials globally. They suggest regular updates of SLF transport potential as more territories become invaded, but also research into long-term control methods, such as SLF-specific RNA-based insecticides.
Predicting marine invaders under climate change using functional and ecological traits

A new trait-based index for assessing the Invasive Potential (IP) of a species correlates well with observed spreading of Mediterranean invaders.

Invasive alien species represent one of the main threats to marine biodiversity. Human activities are impacting patterns of invasions by providing new routes of introduction, while at the same time climate change is altering environmental suitability.

In the paper studying potential fish invaders in the Mediterranean, researchers developed an empirical index for the Invasive Potential (IP) of a species assessing a series of traits related to reproduction, nutrition, morphology and means of introduction.
Applying the framework to 13 candidate species with origins in the Red Sea or Atlantic Ocean, the authors demonstrated a significant correlation between IP and the current observed spreading of the species. Traits related to reproduction, especially, determined high invasion potential.

Using occurrence data from the Ocean Biodiversity Information System (OBIS), GBIF and others combined with data on seafloor topography, distance to coast, surface salinity and temperature variables, the authors modelled the current and future distributions of the species with the highest IP scores.

The models found the highest overall contemporary suitability for invaders in the South-Eastern part of the Mediterranean, while the northern part and the Adriatic Sea had relatively low suitability. In these waters, however, projections under future climate scenarios predicted significant increases in suitability, highlighting the need for revised management strategies for protected areas and fisheries in the region.

Improving predictions of invasive fish ranges combining functional and ecological traits with environmental suitability under climate change scenarios

Global Change Biology
doi.org/10.1111/gcb.15896
Predicting hotspots for invasive species introduction in Europe

Study uses machine-learning approach to produce risk hotspot maps for pest introductions with high predictive accuracy

Invasive plant pests have a massive social impact in Europe, costing billions of euros every year. Accurate predictions of invasive hotspots—areas with the highest risk of incurring invasions—could help focus efforts on prevention and control.

For this study, researchers developed a machine-learning method to create risk maps of pest introductions. Rather than focussing on a single species, the authors trained their model on a dataset covering 243 invasive species.
Following cleaning and thinning the final dataset feeding the authors’ “elastic-net” algorithm consisted of 170,460 GBIF-mediated presence records enhanced with a rich feature set covering climate, soil, erosion risk, land cover and water indicators, as well as population and road densities, anthropogenic pressure, distance to cities and ports, and nightlight radiance as a proxy for GDP.

The analysis revealed the highest risk of introductions in the BeNeLux states, northern Italy, the northern Balkans and the United Kingdom, with areas around container ports like Antwerp, London, Rijeka and Saint Petersburg meriting particular attention as potential invasion hotspots.
Human activity influences genetic dynamics of widespread invader

Study of invasive grass species in two closely located sub-Antarctic islands shows clear effect of human activity on genetic diversity and population structure

Human activities can affect the spread and invasion of alien plant species, but distinguishing external factors from intrinsic population dynamics can be challenging and require comparisons of invasions in areas both with and without human presence.

This study identified an ideal model system in the sub-Antarctic region: two South African islands, Marion Island and Prince Edward Island, that share similar
biotic and abiotic conditions as well as a common invader (*Poa annua* L.). While Marion is home to a research station with year-round operations, Prince Edward remains largely unvisited.

To better understand the degree of invasion, the authors modelled the ecological niches of *P. annua* on both islands using GBIF-mediated occurrences. After gathering samples from different populations on both islands, they analyzed variations in genomic size, genetic diversity and structure.

The niche models for both islands showed that nearly their entire landmasses were climatically suited for the turfgrass, even though current distributions were limited to the coastlines. On occupied Marion Island *P. annua* populations had high levels of genetic diversity and low levels of structure, while the opposite was the case for unoccupied Prince Edward.

These results combined suggest that continued human activity on Marion facilitated ongoing gene flow and diversification. To mitigate increased spread and abundance on Prince Edward, the study recommends minimizing human activity to limit further introductions and thus avoid increasing the genetic diversity of the island’s *P. annua* population.


---

GBIF Secretariat (2023) GBIF Science Review N° 10

doi.org/10.35035/37pp-rt84

CC BY 4.0