Supporting information

Supplementary methods

1. Preventing future biological invasions

For a set of well-studied invasive species that are known to have had significant negative impacts we determined where in regions with contiguous countries future biological invasions could occur and whether these invasions could cause impacts. We classified the invasions according to the six invasion scenarios (see the main paper) and assessed whether the invasions are likely to be prevented. The procedure followed is set out in Figure 2 in the main paper: 1) collate data for study species and determine where each species is likely to 2) be introduced ('introduction threat'), 3) become invasive ('invasion threat'), and 4) have an impact; 5) classify the invasions according to the relevant invasion scenarios; and 6) for each invasion use information on biosecurity to determine whether the invasion is likely to be prevented.

1.1. Species data

For this study our aim was to select a sample of well-studied invasive species from a variety of environments and taxonomic groups that have had serious impacts in places where they have been introduced. Therefore, we selected the species on the Global Invasive Species Database's (GISD) list of 100 of the world's worst invasive species (Lowe, Browne, Boudjelas, & De Poorter, 2000; Luque et al., 2014). Although this list is based on expert opinion and standard methodologies now exist to evaluate the impacts of alien species (e.g. Bacher et al., 2018; Blackburn et al., 2014), these species are perceived to have had serious impacts in at least some of their introduced ranges. Information was required for the analysis on the species' taxonomy, habitat, pathways of introduction, impacts, and global range, and species occurrence records were required to model the distribution of the species. These data were available for most of these species (see below).

1.1.1. Species' taxonomy and habitat

For each of the species, information on taxonomy and habitat was obtained from the GISD. Viruses (n = 1), Protista (n =1) and fungi (n = 5) were excluded from the analysis (see Table S1). Although the GISD recognises six habitat types (brackish water, freshwater, freshwater/terrestrial, marine, marine/terrestrial and terrestrial), for this analysis each species was classified as either terrestrial, marine, or freshwater. The habitat of some species (n = 8) was reclassified using additional habitat information (see Table S2 for details).

1.1.2. Species' pathways of introduction

Information on each species' pathways of introduction was obtained from the GISD. In the GISD the pathways of introduction are classified using the scheme adopted by the Convention on Biological Diversity [CBD (CBD, 2014)] into six categories and 44 sub-categories (see Table S3). At the time of the analysis the pathways of introduction for 11 species had not yet been classified in the GISD using the CBD classification scheme. For some of these species (n = 9), pathway of introduction information was provided in the GISD, and in these instances we used these data to classify the pathways. For the few species (n = 2) for which the GISD did not provide pathway of introduction information, the pathways of introduction were classified using information from other sources (see Table S4 for details).

1.1.3. Species' impacts

Information on each species' recorded impacts was obtained from the GISD. The GISD classifies impacts into mechanisms and outcomes, with the three outcome categories (Environmental ecosystem-habitat, Environmental species-population, Socio-economic) divided into 40 sub-categories (see Table S5).

1.1.4. Species' global range

Information on the countries in which each species occurs as either a native or alien species was obtained from the GISD and the Global Register of Introduced and Invasive Species (Pagad, Genovesi, Carnevali, Schigel, & McGeoch, 2018) using the "originr" package in R (Chamberlain & Bartomeus, 2016).

1.1.5. Species occurrence records

Species occurrence data for each species were obtained from nine online databases (Global Biodiversity Information Facility (GBIF), iNaturalist, eBird, Berkeley's Ecoinformatics Engine, Vertnet, Integrated Digitised Biocollections, AntWeb, Ocean Biogeographic Information System, and the US Geological Survey's Biodiversity Information Serving Our Nation) using the 'spocc' package in R (Chamberlain, 2018). As a large amount of data (more than 100 000 records) were available from GBIF for five species (*Acridotheres tristis*, *Lythrum salicaria*, *Oncorhynchus mykiss*, *Salmo trutta* and *Sturnus vulgaris*), these data were manually downloaded from the GBIF website (http://www.gbif.org/). The species' names as provided in the GISD were used to search for occurrence data, and the R package 'taxize' (Chamberlain & Szocs, 2013) was used to determine if these names were accepted species names. For six species the name provided by the GISD was not an accepted species name. Therefore, for three species, for which an accepted species name was available from the Integrated Taxonomic Information System (ITIS), occurrence data were obtained by searching for both the name listed in the GISD and the accepted species name. The species names as provided by the GISD and those used when searching for species occurrence data are provided in Table S6. We excluded from the analysis seven species for which fewer than 30 species occurrence records were available for species distribution modelling (see below for further information). Therefore, 86 species were included in the analysis.

1.2. Introduction threat

1.2.1. Continuous grid for each pathway

The likelihood of a species being introduced to a new region is often positively related to the prominence of the species' pathways of introduction in that region (Haack, 2001; Levine & D'Antonio, 2003). For example, invertebrates or pathogens that are accidentally introduced along with their host plants when the host is intentionally imported ('contaminant on plants' or 'parasites on plants' pathways of the CBD (2014)), are more likely to be introduced to regions where large quantities of the host are imported than to regions where the host is imported in small quantities (Sikes et al., 2018). As another example, marine alien species that are transported by ships ('hull fouling' or 'ballast water' pathways of the CBD (2014)) are more likely to be introduced to regions with a high shipping intensity than to those where shipping intensity is low (Drake & Lodge, 2004; Kaluza, Kölzsch, Gastner, & Blasius, 2010; Seebens, Gastner, & Blasius, 2013). In order to determine where each species is likely to be introduced global socio-economic data that are related to each of the CBD pathways were collected from various online sources (see sections 1.2.1.1 and 1.2.1.2 below and Table S3 for details). The socio-economic data were used to create a continuous 10-minute global grid for each pathway, where grid cells with high values represent sites where the pathway is prominent and where a species is likely to be introduced through that pathway (see sections 1.2.1.1 and 1.2.1.2 below for details on how these grids were created). We, therefore, assumed that the likelihood of introduction is positively related to the prominence of a species' pathways of introduction (Levine & D'Antonio, 2003), however, while this is sometimes the case, other factors (e.g. the size of the species pool and biosecurity) can also be important (Sikes et al., 2018).

1.2.1.1. Pathways involving the importation of a commodity

Alien species can be introduced through trade when species that are intentionally imported are either released intentionally (e.g. fish for angling), or when they escape from confinement (e.g. pets). Additionally, species can be accidentally introduced through trade as contaminants on imported goods (e.g. insects on their host plants). We assumed that the pathways of introduction for species that are introduced as commodities or as contaminants of commodities would be most prominent in regions where the commodity is imported in large quantities. To create continuous grids for pathways of introduction that are related to the importation of a commodity, we followed a similar methodology to that followed by Early et al. (2016). Country-level import data (value of imports in US dollars) were obtained from the United Nations Comtrade database (UN-Comtrade, 2017). There are various pathways of introduction that involve the import of a commodity, with each of these pathways related to the importation of a specific product. Therefore, data for various types of imports were obtained, and were used to create grids for the pathways to which they are related. For example, live plant import data were used to create grids for pathways involving the intentional importation of plants, and import data for edible vegetables, fruits and nuts were used to create grids for pathways whereby food contaminants are introduced (see Table S3 for details for all pathways). Import data for the period 2012 – 2015 were obtained and for each country an average was calculated. For countries for which import data were not available, we assumed that these countries would import a similar value of goods as countries with comparable incomes, and assigned these countries the average import value of their income group (as classified by the World Bank). We also assumed that once goods were imported into a country they would be transported throughout the country, and that sites with the highest population density would receive the highest quantity of goods and that the pathway of introduction would be prominent at these sites. Therefore, for the trade related pathways that could introduce terrestrial and freshwater species, the country-level import data were distributed spatially within each country using gridded population density data. The average import value for each country was divided by the total population of the country to get a per capita import value. This value was then multiplied by the population density in each 10-minute grid cell to obtain a mean import value per grid cell. For pathways that could introduce marine species, the country-level import data were distributed spatially using gridded coastal population density data (inland countries were not included). Coastal land and sea cells were identified (cells that were within 10-minutes from the coastline) and sea cells were assigned the population density of the closest land cell. As for terrestrial and freshwater species, the import data for marine species were distributed spatially using the population density assigned to the coastal sea cells. For this analysis, gridded global human population density data at a 30 second resolution, obtained from NASA's Socioeconomic Data and Applications Center (SEDAC) (Center for International Earth Science Information Network - CIESIN - Columbia University, 2016), were summed to create a 10-minute resolution grid.

1.2.1.2. Pathways involving the arrival of a transport vector

Alien species can be transported to new regions by ships either through biofouling, within the ballast water carried by ships, or within the ship itself. We assumed that the introduction pathways of species that are introduced through shipping would be most prominent in regions where shipping traffic is high. To create continuous grids for pathways related to shipping we obtained an estimate of the shipping traffic at each global port. Port location data were obtained from the 25th edition of the National Geospatial-Intelligence Agency's World Port Index (National Geospatial-Intelligence Agency, 2016). Gridded cargo ship track data at a one km resolution were obtained (Halpern et al., 2015), and were averaged to create a 10-minute resolution grid. Ship track data for each port were extracted and were assigned to the closest land (for pathways that introduce terrestrial and freshwater species) or sea cell (for pathways that introduce marine species). We, therefore, assumed that the prominence of the pathways that are related to shipping would be greatest at sites closest to maritime ports with high shipping traffic.

Alien species can be transported and introduced to new regions when they stowaway on aeroplanes. We assumed that this pathway of introduction would be most prominent in regions with high air traffic. A continuous grid for this pathway was created by determining air traffic at each airport using airport location and airline data for 2012 from openflights.org (openflights.org, 2016). We estimated air traffic by calculating the number of airline routes to each airport and assigned the results to the grid cell closest to each airport. We, therefore, assumed that this pathway would be most prominent at the sites closest to busy airports. In instances where there was more than one airport associated with a grid cell, we calculated the total number of airline routes associated with the cell. Each year the Airports Council International (ACI) releases airport traffic rankings based on passenger numbers, air cargo volumes and aircraft movements. We obtained these data for the 50 busiest North American airports as ranked by the ACI in 2012 (Airports Council International, 2012), and used these data to determine if the number of routes was a useful proxy for airline traffic. We used the Spearman's rank correlation coefficient to statistically analyse the relationship, for these North American airports, between the number of routes and (1) passenger numbers, (2) air cargo volumes and (3) aircraft movements. We found that for these airports there was a strong and positive correlation between the number of routes and air traffic based on passenger numbers, as well as between the number of routes and number of aircraft movements (see Figure S1). We, therefore, concluded that the number of routes was a useful proxy for airline traffic.

Some marine species (e.g. *Asterias amurensis* and *Caulerpa taxifolia*) have been introduced to new regions as stowaways on fishing equipment, including equipment used for recreational fishing. We assumed that this pathway of introduction would be most prominent in regions where commercial and recreational fishing intensity is high. To create a continuous grid for this pathway, country-level marine fishery data (tons of organisms caught) were obtained from the FishStatJ database of the Food and Agricultural Organisation of the United Nations (FAO) (FAO, 2016). Data for 2011-2014 were obtained and for each country an average was calculated. We assumed that commercial and recreational fishing intensity would be positively related to population density and would be greatest at sites close to the coast, and thus that the pathway of introduction would most prominent in coastal sites with a high population density. Therefore, the average quantity of organisms caught by each country was distributed spatially within each country using the population density values assigned to the coastal sea cells (see section 1.2.1.1 for details).

We assumed for the other pathways of introduction that are related to transport vectors (i.e. container/bulk, machinery/equipment, organic packing material, people and their luggage, vehicles, other means of transport) that most people, goods and vehicles would travel to sites with high population density and, therefore, that these pathways would be most prominent in sites where population density is high. Therefore, we used gridded human population density data (SEDAC data summed to create a 10-minute grid) as continuous grids for these pathways.

1.2.2. Binary grid for each pathway

The continuous grid for each pathway was then converted into a binary grid (1 or 0) using the 75th percentile as a threshold, where cells with values greater than or equal to the 75th percentile were assigned a value of one, and those with values less than the 75th percentile were assigned a value of zero. In these grids cells with a value of one represent sites where the pathway of introduction is prominent and where a species is likely to be introduced through the pathway.

1.2.3. Introduction threat grid for each species

Some alien species can be introduced through multiple pathways (Essl et al., 2015; Foxcroft, Spear, van Wilgen, & McGeoch, 2019; Pergl et al., 2017). Using the information collected on the species' pathways of introduction, we identified all of the pathways that had previously facilitated the introduction of each species. For each species, the binary grids for all of the species' pathways of introduction were combined by taking the maximum value for each cell. This resulted in a binary grid (1 or 0) for each species indicating sites where the species is likely to be introduced, or in other words where the species poses an introduction threat.

1.3. Invasion threat

1.3.1. Continuous grid of environmental suitability

Species distribution models (SDMs) were developed for each species and were used to identify parts of the globe that are environmentally suitable for the species to survive and persist. For each species, information on habitat, species occurrence records, and ecologically relevant predictor variables were required for modelling.

1.3.1.1. Preparation of species occurrence data

The quality of the species occurrence records that were collected for each species was assessed. For each species we checked if any records for the wrong species had been included in the dataset. The 'grepl' function in R was used to match the species' names in the occurrence data to those used when searching for the occurrence data. We manually checked instances where the species names did not match but where more than 20 records were downloaded, and in some of these instances the records were for the correct species. In this way true errors in the data were identified (e.g. a record for *Acacia dealbata* was downloaded when searching for data for *Acacia mearnsii*) from instances where synonyms were used (e.g. when searching for data for *Opuntia stricta* records for the synonym *Opuntia dillenii* were downloaded). Hybrid species and records where species identification was uncertain were excluded. The quality of the data were further assessed using the 'Biogeo' package in R (Robertson, Visser, & Hui, 2016). This package was used to identify records that were missing co-ordinate data or for which co-ordinate data were not in decimal degrees, as well as records that were too imprecise for our analysis (at a 10-minute resolution) and those that were in the wrong environment (records in the sea for terrestrial or freshwater species and those on land for marine species). Where possible, records in the wrong environment were moved to the closest cell in the correct environment. We removed from the datasets: imprecise records, records for the wrong species, those without co-ordinate data in decimal degrees, and those that were still in the wrong environment. The data were then checked for duplicate records (i.e. more than one occurrence record in a 10-minute cell), and these were removed. For the marine species *Cercopagis pengoi*, 104 records in the Great Lakes were removed as there were no marine environmental data available for this region. For terrestrial and freshwater species, gridded climate data from the Worldclim dataset (version 1) at a 10-minute resolution were used (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) when assessing the quality of the occurrence records, while for marine species, marine environmental data from the Bio-ORACLE dataset (Tyberghein et al., 2012) were used. As the Bio-ORACLE dataset is available at a 5-minute resolution, these data were aggregated to 10-minutes. For each species, the number of records obtained from the databases, the number of records removed and the reason for their removal are provided in Table S6. Across the species an average of 2283.1 (± 5413.6) clean occurrence records were available for modelling, with a minimum of 4 records and a maximum of 47 594 records. A study by Wisz et al. (2008) demonstrated that SDMs do not predict consistently well with a sample size of fewer than 30 occurrence records and, therefore, we did not model the distributions of seven species for which fewer than 30 occurrence records were available (see Table S6). Therefore, species distribution models were built for 86 species.

1.3.1.2. Environmental predictors

For terrestrial and freshwater species, 9 climatic variables from the Worldclim 10-minute dataset (Hijmans et al., 2005) were considered for modelling: 1) mean annual temperature, 2) maximum temperature of the warmest month, 3) minimum temperature of the coldest month, 4) annual precipitation, 5) precipitation of the wettest month, 6) precipitation of the driest month, 7) temperature seasonality, 8) precipitation seasonality, and 9) mean diurnal temperature range. These variables consider averages and variations in temperature and precipitation, and have previously been used in SDM for a wide range of species (Bellard et al., 2013). Although the distributions of freshwater species have been previously modelled using only climate data (Ruiz-Navarro, Gillingham, & Britton, 2016), for freshwater species we considered an additional topographic variable: topographic wetness index from the Envirem dataset at a resolution of 10-minutes (Title & Bemmels, 2018). This variable has proved to be an important predictor when modelling the distributions of some freshwater species (e.g. Fletcher, Gillingham, Britton, Blanchet, & Gozlan, 2016). Nine marine environmental variables from the Bio-ORACLE dataset (Tyberghein et al., 2012) aggregated to a 10-minute resolution were considered when modelling the distributions of marine species: 1) maximum sea surface temperature, 2) minimum sea surface temperature, 3) mean sea surface temperature, 4) sea surface temperature range, 5) salinity, 6) dissolved oxygen, 7) ph, 8) mean photosynthetically available radiation, and 9) maximum photosynthetically available radiation. These variables were considered as they are important factors that influence marine life. For example, temperature is believed to be the most important physical oceanographic variable determining the abundance, spatial distribution and diversity of marine ectotherms, photosynthetically available radiation provides the energy required by photosynthetic organisms, and ocean acidity (pH) is critical in mediating physiological reactions (Tyberghein et al., 2012). Furthermore, these variables have been used to model the distributions of a wide range of marine organisms including crabs, sea weeds and sea squirts (Crafton, 2015; Januario, Estay, Labra, & Lima, 2015; Verbruggen et al., 2013). For each group (terrestrial, freshwater and marine species), the variables were checked for co-linearity using the Pearson correlation coefficient and variables that were co-linear (correlation $\geq |0.75|$) for more than 50% of the species were not included in the models. These variables were 1) minimum temperature of the coldest month and 2) precipitation of the wettest month for terrestrial species; 1) minimum temperature of the coldest month and 2) annual precipitation for freshwater species; and 1) mean sea surface temperature for marine species. Consequently, seven variables were used to model the distributions of terrestrial species and eight variables were used for freshwater species and for marine species (see Table S7 for details).

1.3.1.3. Species distribution models

The cleaned occurrence records and environmental predictor variables were used to build ensemble species distribution models using the Biomod2 package in R (Thuiller, Lafourcade, Engler, & Araújo, 2009). The SDMs were built using six algorithms: 1) generalised linear models, 2) generalised boosting trees, 3) multivariate adaptive regression splines, 4) random forest, 5) flexible discriminate analysis, and 6) maximum entropy. Five sets of pseudo-absence records were generated for each species by taking 1000 or 10 000 random points from the environments in which the species is found. If the number of occurrence records was less than or equal to 1000, then 1000 pseudo-absence records were generated, if there were greater than 1000 occurrence records then 10 000 pseudo-absence records were generated. It has been argued that using pseudo-absence records that are environmentally distant from presence records artificially increases the rate of well-predicted absences and scores for the area under the receiver operating curve (AUC) (Lobo, Jiménez-Valverde, & Real, 2008). Therefore, for terrestrial and freshwater species, pseudo-absence records were identified from within the bioclimatic regions in which the species is found (as classified by Metzger et al. (2013) and aggregated to a 10-minute resolution). For marine species these records were identified from sites within the coastal shelves (within the 200m isobath) of the latitudinal zones in which the species occurs (as classified by Spalding et al. (2007) and rasterised to a 10-minute resolution). Default modelling options were used, and for model evaluation a four-fold repeated split procedure was used, whereby models were calibrated on 70% of the data and tested on 30% of the data. As recommended by Barbet-Massin and colleagues (Barbet-Massin, Jiguet, Albert, & Thuiller, 2012), equal weighting was given to presences and pseudo-absences (weighted sum of presence equals the weighted sum of pseudo-absence). AUC values are often used to evaluate SDMs (Fielding & Bell, 1997). However, as the use of AUC for model evaluation has been criticised (Lobo et al., 2008), AUC and the true skills statistic (TSS) (Allouche, Tsoar, & Kadmon, 2006) were used to evaluate the models. AUC values range from 0.5 to 1, where a value of 1 indicates a perfect prediction and a value of 0.5 indicates a prediction that is no better than random. TSS values can range between -1 and 1, where a value of +1 indicates a perfect prediction and a value of zero or less indicates a prediction that is no better than random (Allouche et al., 2006). The final ensemble model for each species was generated using all the data and by calculating the weighted mean of the probabilities for each algorithm. Only models with TSS greater than 0.6 were used in the ensemble model and were weighted in proportion to the TSS evaluation. The projected ensemble models performed well for all species, with the TSS values for the models ranging between 0.69 and 0.98, and the AUC values ranging between 0.93 and 0.99 (Figure S2). The output of the species distribution models was a 10-minute grid for each species with values ranging between 0 and 1000.

1.3.2. Binary grid of environmental suitability

For each species the continuous grid of environmental suitability was converted into a binary grid (1 or 0), using the lowest predicted value at which an occurrence record was found as the threshold. Cells with values greater than or equal to the threshold were assigned a value of one, and those with values less than the threshold were assigned a value of zero. Cells with a value of one represent parts of the globe that are environmentally suitable for the species.

1.3.3. Invasion threat grid for each species

As some of the species have already been introduced to parts of the world that were predicted as environmentally suitable, and as the focus of the work was on future rather than current biological invasions, we excluded predicted cells in countries in which the species already occur. The species occurrence data and information on the species' ranges were used to identify any country in which the species occurs, and these grid cells were excluded. For each species the cells that were predicted as environmentally suitable were classified into separate invasions based on whether the cells formed a contiguous group, and whether they fall within the same political boundaries. Therefore, separate contiguous groups of cells found within the same country were classified as the same invasion. Therefore, multiple invasions are possible for each species. For this analysis, the political boundaries used were those of countries, but territories were considered as separate countries from their sovereign states. For example, French Guiana was considered separate from France, and this prevented invasions in geographically distant parts of the world from being classified as one. Invasions that were predicted to occur on land masses where only one country or territory is present were excluded (e.g. invasions on land masses such as Australia and small islands), as were invasions for which there was no introduction threat. It is important to note that as this analysis was performed at a resolution of 10-minutes, invasions predicted on islands that are close to the mainland, or that are close to other island territories were not excluded (e.g. Zanzibar is close to continental Africa). Therefore, we assumed that in these instances the countries or territories in question are close enough to each other to possibly influence each other's biosecurity (i.e. it's possible that alien species could spread from one land mass to the other). This process resulted in a grid for each species which indicates where the environment is suitable for the species to establish in the future, or in other words where the species poses an invasion threat, with all the cells related to a specific invasion uniquely classified.

1.4. Impact

1.4.1. Binary grid for each impact

Alien species can have a wide range of environmental and socio-economic impacts, however, the magnitude of these impacts will partly depend on the properties of the invaded site. Sensitive sites are those where the consequences of the invasion are severe or where the invasion is particularly undesirable (McGeoch et al., 2016). For example, alien species that cause a reduction in native biodiversity are likely to have the greatest impact in regions that have been identified as global conservation priorities. As another example, invasions by species that have impacts on tourism may be particularly undesirable in regions where tourism is economically important. In order to determine where each species is likely to have an especially undesirable negative impact, global environmental and socio-economic data that are related to the 40 types of impact were obtained from various online sources (see sections 1.4.1.1 to 1.4.1.7 below and Table S5 for details). These data were used to create, for each impact, a global binary grid (1 or 0) at a 10-minute resolution, where grid cells with a value of one represent sites that are particularly sensitive to invasions that have this specific impact (see sections 1.4.1.1 to 1.4.1.7 below for details on how each of these grids was created).

1.4.1.1 Environmental impacts

We assumed that the sites that are most sensitive to invasions that have environmental impacts (i.e. any impact on habitats or other species, or those classified in the GISD as having either ‘Environmental ecosystem-habitat’ or ‘Environmental species-population’ outcomes) are those that have been identified as global conservation priorities (see Table S5). Although global conservation priority sites have been identified using a variety of methods (see Brooks et al., 2006), the World Wildlife Fund’s (WWF) Global 200 Ecoregions (Olson & Dinerstein, 1998, 2002) was used in our analysis. Unlike in many other prioritisation schemes [e.g. the biodiversity hotspots (Mittermeier et al., 2004; Myers, Mittermeier, Mittermeier, da Fonseca, & Kent, 2000)], the irreplaceable regions in the Global 200 Ecoregions were not identified solely based on biodiversity, but taxonomic uniqueness, unique phenomena, and the global rarity and intactness of major habitat types were also considered (Olson & Dinerstein, 2002). Furthermore, we selected this prioritisation scheme, as unlike other schemes which tend to focus on one habitat type, the Global 200 Ecoregions includes irreplaceable terrestrial, freshwater and marine habitats, and thus impacts in all three habitats could be considered. Digital data for the Global 200 Ecoregions were obtained from the WWF (World Wildlife Fund, 2012), and these data were rasterised to create three binary 10-minute grids (1 or 0), one for each habitat. In these grids, cells with a value of one indicate the presence a Global 200 Ecoregion and sites that are sensitive to invasions that cause environmental impacts.

1.4.1.2. Impacts on agriculture or forestry

Alien species can have negative impacts on agriculture and forestry, and we assumed that the sites that are sensitive to these impacts are those where crops are grown and where forests are found. Therefore, to identify regions of the globe that are sensitive to these impacts, we used the European Space Agency Climate Change Initiative's Land Cover dataset (version 1.6.1) for the 2010 epoch (European Space Agency: Climate Change Intiative, 2017). These gridded data are at a resolution of 300 m and were resampled to produce a 10-minute grid. Grid cells classified as having rainfed cropland, irrigated cropland or cropland dominated mosaic vegetation were identified and given a value of one, while all other cells were given a value of zero. In the resultant binary grid (1 or 0) cells with a value of one indicate the presence of cropland and sites that are sensitive to invasions that have an impact on agriculture. Similarly, to identify sites that are sensitive to invasions that have an impact on forestry, grid cells classified in the Land Cover dataset as having tree cover were identified and given a value of one, while all other cells were given a value of zero. In the resultant binary grid (1 or 0) cells with a value of one indicate the presence of trees and sites that are sensitive to invasions that have an impact on forestry.

1.4.1.3. Impacts on livestock and livestock products

Some alien species have negative impacts on livestock and related products, and we assumed that the sites that would be most sensitive to these impacts are those where livestock is found at high densities. Therefore, to identify sites that are sensitive to these impacts gridded, global livestock density data at a resolution of 3-minutes were obtained from the Food and Agricultural Organisation of the United Nations (Food and Agricultural Organization of the United Nations, 2010). Gridded density data for cattle, buffaloes, goats, sheep, pigs and poultry were obtained and were summed to obtain a global grid of livestock density. The resultant 3-minute grid was resampled to create a 10-minute grid, and converted to a binary grid (1 or 0) using the 75^{th} percentile as a threshold. Cells with values greater than or equal to the 75^{th} percentile were assigned a value of one, and those with values less than the 75^{th} percentile were assigned a value of zero. In this grid, cells with a value of one indicate sites that are sensitive to invasions that could have an impact on livestock and their products.

1.4.1.4. Impacts on aquaculture, mariculture or fisheries

Alien species can have negative impacts on aquaculture, mariculture and fisheries and we assumed that the sites that are most sensitive to these impacts are those where fishing and aquaculture production are high. Therefore, to identify sites that are sensitive to these impacts we obtained country-level freshwater and marine fishery and aquaculture data (tons of organisms) for 2011 - 2014 from the FishStatJ database (FAO, 2016). For each country these data were averaged to get the mean quantity of marine or freshwater organisms caught or produced through aquaculture. For each country, the mean quantity of freshwater organisms caught and produced through aquaculture were summed, and similarly for marine organisms. As gridded global data on the position of inland aquaculture facilities and fisheries were not available, we assumed that such activities are most likely to occur in areas with high surface water. Therefore, the country-level freshwater aquaculture and fishery data were distributed spatially within each country using gridded runoff data. For each country the value for freshwater organisms was divided by total runoff. This value was then multiplied by the runoff in each 10-minute grid cell. Gridded global runoff data at a 30-minute resolution were obtained (Fekete, Vörösmarty, & Grabs, 2002) and were resampled to create a 10-minute resolution grid. For marine organisms we assumed that fishing and mariculture are most likely to occur in areas with high ocean primary productivity. Therefore, the country-level marine aquaculture and fishery data were distributed spatially within each country using gridded mean chlorophyll data. The value for marine organisms was divided by the total primary productivity along the country's coast (sea grid cells within 10-minutes of the coastline). This value was then multiplied by the primary productivity in each 10-minute grid cell along the coast. Gridded ocean productivity data at a 5-minute resolution (Tyberghein et al., 2012) were obtained and were averaged to create a 10-minute grid. The 10-minute grids for freshwater and marine fisheries and aquaculture were converted, following the methods discussed above into binary grids using the 75th percentile as a threshold. In these grids, cells with a value of one indicate sites that are sensitive to invasions that have an impact on freshwater or marine aquaculture and fisheries.

1.4.1.5. Impacts on recreation and tourism

Alien species can have negative impacts on tourism and recreation, and we assumed that sites that are sensitive to these impacts would be those where tourism is economically important. Sites that are sensitive to these impacts were identified using country-level data, obtained from the World Travel and Tourism Council (World Tourism and Travel Council, 2016), on the contribution of tourism to Gross Domestic Product (GDP). Data for 2012 – 2015 were obtained and for each country an average was calculated. For countries for which data were not available, we assumed that the contribution of tourism to GDP would be similar to that of other countries with a similar income, and thus these countries were assigned the average values of their income group. We also assumed that sites with high population density would have high levels of tourism and recreation, and that at these sites would be sensitive to invasions that have an impact on these activities. Therefore, we followed the method described in section 1.2.1.1, and spatially distributed the country-level tourism data within each country using gridded population density data at a 10-minute resolution. For the marine environment these data were distributed spatially within each country using the population density values assigned to the coastal sea cells (see section 1.2.1.1). Following the methods described above, the resultant 10-minute grid was converted into a binary grid using the 75th percentile as a threshold. In this grid, cells with a value of one indicate sites that are sensitive to invasions that have an impact on recreation and tourism.

1.4.1.6. Impacts on trade and international relations

Alien species can have a negative impact on trade and international relations, and we assumed that the sites that would be sensitive to these impacts would be those where exports are important economically. Therefore, sites that are sensitive to these impacts were identified using country-level data for 2012 – 2015 on the contribution of exports to GDP, obtained from The World Bank (The World Bank, 2016). We followed the methods described in section 1.4.1.5 and assumed that exports would be important in sites with high population density, and that at these sites would be sensitive to invasions that have an impact on these activities. Therefore the country level export data were spatially distributed within each country using gridded population density data at a 10-minute resolution. The resultant 10-minute grid was converted into a binary grid using the 75th percentile as a threshold, which resulted in the creation of a binary grid (1 or 0), where cells with a value of one indicate sites that are sensitive to invasions that have an impact on trade and international relations.

1.4.1.7. Other socio-economic impacts

Alien species can have a range of other socio-economic impacts that are not discussed above. All of these socio-economic impacts are directly related to the presence of humans [e.g. impacts on human health, infrastructure and on access to water or land (see Table S5)] and, therefore, we assumed that sites with a high population density would be sensitive to invasions that cause these types of impacts. Therefore, gridded population density data at a 10-minute resolution were used to identify sites that are sensitive to these impacts. The 10-minute grid was converted into a binary grid (1 or 0) using the 75th percentile as a threshold. In this grid, cells with a value of one indicate sites that are sensitive to invasions that could cause these socio-economic impacts (e.g. impacts on human health, infrastructure and on access to water or land).

1.4.2. Impact grid for each species

An alien species can have multiple impacts (Foxcroft et al., 2019). Using the information collected on the species' impacts, we identified all of the impacts each species had had in their introduced range. For each species, we combined the binary grids for all of the impacts that the species has had by taking the maximum value for each cell. This resulted in a global binary grid (1 or 0) for each species, where cells with a value of one indicate sensitive sites where the species could have an impact if it is introduced.

1.5. Classify the invasions according to scenarios

The introduction threat, invasion threat and impact results for each species were combined and a map of the country boundaries was used to identify the countries for each invasion where the species is likely to first establish, subsequently invade, and have an impact.

We assumed that the country of first establishment for an invasion would be any country with sites where the species is likely to be introduced and subsequently establish. Therefore, for each invasion, sites where the species could first establish were identified by combining the introduction threat and invasion threat grids by calculating the product for each cell. A map of country boundaries was obtained from Natural Earth (version 2.2.0 at a scale of 1:50) and was superimposed onto the resultant grid to identify the country, for each invasion, where the species could first establish.

For each invasion, countries where the species could subsequently invade were identified by superimposing a map of country boundaries onto the invasion threat grid and eliminating the country of first establishment.

We assumed that countries where the species could have an impact would be any country with sensitive sites where the species is likely to establish. Therefore, for each invasion, sites where the species could have an impact were identified by combining the invasion threat and impact grids by calculating the product for each cell. For each invasion, countries where the species could have an impact were identified by superimposing a map of country boundaries onto the resultant grid.

For each invasion, the countries where the species is likely to first establish, subsequently invade, and have an impact were compared and each country was classified as a: 1) country of first establishment where there is no impact, 2) country of first establishment where there is an impact, 3) country of subsequent invasion where there is no impact and 4) country of subsequent invasion where there is an impact. While invasions transcend political boundaries, these country-level results were obtained as invasions are most often managed at a country-level and as information on biosecurity was available at a country-level. The invasions were then classified according the invasion scenario(s) to which they conform. For some invasions there was more than one country where first establishment could occur, and as a consequence, multiple scenarios are possible. In these instances, the invasions were classified according to all of the invasion scenarios that were applicable.

1.6. Biosecurity

Country-level data on proactive response capacity have been published by Early et al. (2016). These data indicate the likelihood that invasions will be prevented or contained early in the invasion process. Countries that have a high proactive response capacity have comprehensive border control policies and programmes for research, monitoring and public engagement on biological invasions (Early et al., 2016). National reports on the implementation of the Convention on Biological Diversity were used to estimate proactive response capacity, with estimates ranging between zero and three at intervals of 0.5 (Early et al., 2016). These data were used to determine, for each invasion, the proactive response capacity of the country of first establishment and assess if the invasion is likely to be prevented. We classified proactive response capacity into three categories, where a proactive response capacity of 0, 0.5 or 1 was low, 1.5 or 2 was intermediate and 2.5 or 3 was high. For invasions where first establishment could occur in more than one country multiple scenarios are possible, and for each possible scenario the proactive response capacity available to prevent the invasion could vary depending on the country where the species first establishes. In these instances proactive response capacity was assessed for each possible scenario by calculating the minimum proactive response capacity of the countries of first establishment. Furthermore, in instances where first establishment could occur in more than one country, countries of first establishment with a high proactive response capacity could prevent the introduction of the species but will still be at risk if other countries of first establishment in the region have a low proactive response capacity. In order to assess the prevalence of this issue we calculated the minimum and maximum proactive response capacity for every scenario to which an invasion conformed.

References

Airports Council International. (2012). North America Airport Traffic Annual Reports. Retrieved from https://airportscouncil.org/intelligence/north-american-airport-traffic-reports/north-america-airport-traffic-annual-reports/

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*, 1223–1232.

Bacher, S., Blackburn, T. M., Essl, F., Genovesi, P., Heikkilä, J., Jeschke, J. M., … Kumschick, S. (2018). Socio-economic impact classification of alien taxa (SEICAT). *Methods in Ecology and Evolution*, *9*(1), 159–168.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*, 327–338.

Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M., & Courchamp, F. (2013). Will climate change promote future invasions? *Global Change Biology*, *19*(12), 3740–3748.

Blackburn, T. M., Essl, F., Evans, T., Hulme, P. E., Jeschke, J. M., Kühn, I., … Bacher, S. (2014). A unified classification of alien species based on the magnitude of their environmental impacts. *PLoS Biology*, *12*(5), e1001850.

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., … Rodrigues, A. S. L. (2006). Global biodiversity conservation priorities. *Science*, *313*, 58–61.

CAB International. (2018). *Eleutherodactylus coqui* (Caribbean tree frog). Retrieved from https://www.cabi.org/ISC/datasheet/84734

CAB International. (2019). *Cinara cupressi sensu lato* (Cypress aphid). Retrieved from https://www.cabi.org/ISC/datasheet/13579

CBD. (2014). Pathways of introduction of invasive species, their prioritization and management. Retrieved from https://www.cbd.int/doc/meetings/sbstta/sbstta-18/official/sbstta-18-09-add1-en.pdf

Center for International Earth Science Information Network - CIESIN - Columbia University. (2016). Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision of UN WPP Country Totals. Retrieved from http://sedac.ciesin.columbia.edu

Chamberlain, S. (2018). spocc: Interface to species occurrence data sources. R package. Retrieved from https://cran.r-project.org/package=spocc

Chamberlain, S., & Bartomeus, I. (2016). originr: Fetch species origin data from the web. R package. Retrieved from https://cran.r-project.org/package=originr

Chamberlain, S., & Szocs, E. (2013). taxize - taxonomic search and retrieval in R. *F1000Research*, *2*(191).

Cowan, P. E. (1992). The eradication of introduced Australian brushtail possums, *Trichosurus vulpecula*, from Kapiti Island, a New Zealand nature reserve. *Biological Conservation*, *61*, 217–226.

Crafton, R. E. (2015). Modeling invasion risk for coastal marine species utilizing environmental and transport vector data. *Hydrobiologia*, *746*, 349–362.

Drake, J. M., & Lodge, D. M. (2004). Global hot spots of biological invasions: Evaluating options for ballast − water management. *Proceedings of the Royal Society of London B*, *271*, 575–580.

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., … Tatem, A. J. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications*, *7*, 12485.

Essl, F., Bacher, S., Blackburn, T. M., Booy, O., Brundu, G., Brunel, S., … Jeschke, J. M. (2015). Crossing frontiers in tackling pathways of biological invasions. *BioScience*, *65*, 769–782.

European Space Agency: Climate Change Intiative. (2017). Land cover. Retrieved from https://www.esa-landcover-cci.org/

FAO. (2016). FishStatJ - software for fishery statistical time series. Retrieved from http://www.fao.org/fishery/statistics/software/fishstatj/en

Fekete, B. M., Vörösmarty, C. J., & Grabs, W. (2002). High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biochemical Cycles*, *16*(3), 1042.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49.

Fletcher, D. H., Gillingham, P. K., Britton, J. R., Blanchet, S., & Gozlan, R. E. (2016). Predicting global invasion risks: A management tool to prevent future introductions. *Scientific Reports*, *6*, 26316.

Food and Agricultural Organization of the United Nations. (2010). Livestock systems. Retrieved from http://www.fao.org/livestock-systems/global-distributions/en/

Foxcroft, L. C., Spear, D., van Wilgen, N. J., & McGeoch, M. A. (2019). Assessing the association between pathways of alien plant invaders and their impacts in protected areas. *NeoBiota*, *43*, 1–25.

Global Invasive Species Database. (2019a). Species profile: *Carcinus maenas*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Carcinus+maenas

Global Invasive Species Database. (2019b). Species profile: *Cercopagis pengoi*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Cercopagis+pengoi

Global Invasive Species Database. (2019c). Species profile: *Eichhornia crassipes*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Eichhornia+crassipes

Global Invasive Species Database. (2019d). Species profile: *Eleutherodactylus coqui*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Eleutherodactylus+coqui

Global Invasive Species Database. (2019e). Species profile: *Lithobates catesbeianus*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Lithobates+catesbeianus

Global Invasive Species Database. (2019f). Species profile: *Myocastor coypus*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Myocastor+coypus

Global Invasive Species Database. (2019g). Species profile: *Rhinella marina*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Rhinella+marina

Global Invasive Species Database. (2019h). Species profile: *Trachemys scripta elegans*. Retrieved from http://www.iucngisd.org/gisd/speciesname/Trachemys+scripta+elegans

Haack, R. A. (2001). Intercepted Scolytidae (Coleoptera) at U.S. ports of entry: 1985–2000. *Integrated Pest Management Reviews*, *6*, 253–282.

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., … Walbridge, S. (2015). Cumulative human impacts: Raw stressor data (2008 and 2013). Retrieved from https://knb.ecoinformatics.org/view/doi:10.5063/F1S180FS

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hulme, P. E., Bacher, S., Kenis, M., Klotz, S., Kühn, I., Minchin, D., … Vilà, M. (2008). Grasping at the routes of biological invasions: A framework for integrating pathways into policy. *Journal of Applied Ecology*, *45*, 403–414.

Januario, S. M., Estay, S. A., Labra, F. A., & Lima, M. (2015). Combining environmental suitability and population abundances to evaluate the invasive potential of the tunicate

Ciona intestinalis along the temperate South American coast. *PeerJ*, *3*, e1357.

Kaluza, P., Kölzsch, A., Gastner, M. T., & Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society, Interface*, *7*, 1093–1103.

Levine, J. M., & D'Antonio, C. M. (2003). Forecasting biological invasions with increasing international trade. *Conservation Biology*, *17*(1), 322–326.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*, 145–151.

Lowe, S., Browne, M., Boudjelas, S., & De Poorter, M. (2000). *100 of the world's worst invasive alien species: A selection from the Global Invasive Species Database*. Auckland, New Zealand: The Invasive Species Specialist Group.

Luque, G. M., Bellard, C., Bertelsmeier, C., Bonnaud, E., Genovesi, P., Simberloff, D., & Courchamp, F. (2014). The 100th of the world's worst invasive alien species. *Biological Invasions*, *16*(5), 981–985.

McGeoch, M. A., Genovesi, P., Bellingham, P. J., Costello, M. J., McGrannachan, C., & Sheppard, A. (2016). Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biological Invasions*, *18*(2), 299–314.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A., & Zomer, R. (2013). A high-resolution bioclimate map of the world: A unifying framework for global biodiversity research and monitoring. *Global Ecology and Biogeography*, *22*, 630–638.

Mittermeier, R. A., Gil, P. R., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, C. G., … da Fonseca, G. A. B. (2004). *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions*. CEMEX.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–858.

National Geospatial-Intelligence Agency. (2016). *World port index: Twenty-fifth edition*. Springfield, Virginia: National Geospatial-Intelligence Agency.

Olson, D. M., & Dinerstein, E. (1998). The global 200: A representation approach to conserving the earth's most biologically valuable ecoregions. *Conservation Biology*, *12*(3), 502–515.

Olson, D. M., & Dinerstein, E. (2002). The global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, *89*(2), 199–224.

openflights.org. (2016). Airport, airline and route data. Retrieved from https://openflights.org/data.html

Pagad, S., Genovesi, P., Carnevali, L., Schigel, D., & McGeoch, M. A. (2018). Data descriptor: Introducing the Global Register of Introduced and Invasive Species. *Scientific Data*, *5*, 170102.

Pergl, J., Pyšek, P., Bacher, S., Essl, F., Genovesi, P., Harrower, C. A., … Nentwig, W. (2017). Troubling travellers: Are ecologically harmful alien species associated with particular introduction pathways? *NeoBiota*, *32*, 1–20.

Robertson, M. P., Visser, V., & Hui, C. (2016). Biogeo: An R package for assessing and improving data quality of occurrence record datasets. *Ecography*, *39*(4), 394–401.

Ruiz-Navarro, A., Gillingham, P. K., & Britton, J. R. (2016). Predicting shifts in the climate space of freshwater fishes in Great Britain due to climate change. *Biological Conservation*, *203*, 33–42.

Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, *16*(6), 782–790.

Sikes, B. A., Bufford, J. L., Hulme, P. E., Cooper, J. A., Johnston, P. R., & Duncan, R. P. (2018). Import volumes and biosecurity interventions shape the arrival rate of fungal pathogens. *PLoS Biology*, *16*(5), e2006025.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., … Robertson, J. (2007). Marine Ecoregions of the World: A bioregionalization of coastal and shelf areas. *BioScience*, *57*(7), 573–583.

The World Bank. (2016). World Development Indicators. Retrieved from https://databank.worldbank.org/data/home.aspx

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD - A platform for ensemble forecasting of species distributions. *Ecography*, *32*, 369–373.

Title, P. O., & Bemmels, J. B. (2018). ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, *41*, 291–307.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, *21*(2), 272–281.

UN-Comtrade. (2017). Comtrade Database—beta trade data extraction interface. Retrieved from https://comtrade.un.org/data/

Verbruggen, H., Tyberghein, L., Belton, G. S., Mineur, F., Jueterbock, A., Hoarau, G., … De Clerck, O. (2013). Improving transferability of introduced species' distribution models: New tools to forecast the spread of a highly invasive seaweed. *PLoS ONE*, *8*(6), e68337.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*, 763–773.

World Tourism and Travel Council. (2016). Data gateway. Retrieved from https://www.wttc.org/datagateway/

World Wildlife Fund. (2012). Global 200. Retrieved from https://www.worldwildlife.org/publications/global-200

Supplementary tables and figures

Table S1. Details on the fungi, Protista and viruses that are listed as 100 of the world's worst invasive species, but which were not included in the analysis.

Kingdom	Species	Habitat	Description
Fungi	*Aphanomyces astaci*	Freshwater	Water mould that infects only crayfish species
Fungi	*Batrachochytrium dendrobatidis*	Freshwater/terrestrial	Fungus that has been associated with population declines in endemic amphibian species
Fungi	*Cryphonectria parasitica*	Terrestrial	*Cryphonectria parasitica* is a fungus that attacks primarily *Castanea* spp. but also has been known to cause damage to various *Quercus* spp. along with other species of hardwood trees.
Fungi	*Ophiostoma ulmi sensu lato*	Terrestrial	Fungus disseminated by specialised bark beetles that causes Dutch Elm Disease
Fungi	*Phytophthora cinnamomi*	Terrestrial	A widespread soil-borne pathogen that infects woody plants causing root rot and cankering
Protista	*Plasmodium relictum*	Terrestrial	The protozoa, *Plasmodium relictum*, is one of the causative parasites of avian malaria
Virus	Banana bunchy top virus (BBTV)	Terrestrial	A deadly pathogen which affects many areas of the world-wide banana industry

Table S2. The species for which habitat was reclassified, the habitat of these species as classified in the Global Invasive Species Database (GISD), the information used to reclassify habitat, and the habitat classification used in the analysis.

Species	Class	Habitat in GISD	Habitat information	Habitat used in analysis
Cercopagis pengoi	Branchiopoda	Brackish	*Cercopagis pengoi* is euryhaline and eurythermic, and can tolerate a wide of salinities and temperatures. In some regions, *C. pengoi* abundance increases with distance from shore, suggesting that this is a typical pelagic species, which lives in the open sea, away from the littoral zone (Global Invasive Species Database, 2019b).	Marine
Carcinus maenas	Malacostraca	Marine/ terrestrial	Adult *Carcinus maenas* (European shore crab) can tolerate temperatures ranging from 0 to 33°C and salinities from 4 to 54. As an adult the species can occupy a variety of substrates such as mud, sand, rock, and eelgrass, and it can also occupy depths ranging from high tide to 6 meters, with some records of up to 60 meters (Global Invasive Species Database, 2019a). The species has been introduced through shipping (Global Invasive Species Database, 2019a).	Marine
Myocastor coypus	Mammalia	Freshwater/ terrestrial	*Myocastor coypus* (coypu) are generally found near permanent water, particularly reed beds and swamp/marsh. Also found in rivers, streams, lakes, ponds and brackish marsh in coastal areas (Global Invasive Species Database, 2019f).	Freshwater
Trachemys scripta elegans	Reptilia	Freshwater/ terrestrial	Within its native range *Trachemys scripta* lives in a wide variety of freshwater habitats including rivers, ditches, swamps, lakes and ponds (Global Invasive Species Database, 2019h).	Freshwater
Eleutherodactylus coqui	Amphibia	Freshwater/ terrestrial	*Eleutherodactylus coqui* is a habitat generalist (Global Invasive Species Database, 2019d). Natural and semi-natural environments inhabited are natural forests, riverbanks and wetlands (CAB International, 2018).	Freshwater
Lithobates catesbeianus	Amphibia	Freshwater/ terrestrial	Bullfrogs occupy a wide range of aquatic habitats including lakes, ponds, swamps, bogs and backwaters well as reservoirs, marshes, brackish ponds (in Hawaii), streams irrigation ponds and ditches (Global Invasive Species Database, 2019e).	Freshwater
Rhinella marina	Amphibia	Freshwater/ terrestrial	Cane toads are found in many places, such as man-made ponds, gardens, drain pipes, debris, under cement piles and beneath houses. Cane toads will usually stay on dry land and reproduce in any shallow water near its surroundings (Global Invasive Species Database, 2019g).	Freshwater
Eichhornia crassipes	Liliopsida	Terrestrial	*E. crassipes* grows in shallow temporary ponds, wetlands and marshes, sluggish flowing waters, lakes, reservoirs and rivers (Global Invasive Species Database, 2019c).	Freshwater

Table S3. The mechanisms through which alien species can enter a new region (Hulme et al., 2008), the related pathways of introduction (categories and sub-categories) as classified by the Convention on Biological Diversity (CBD, 2014), and the data used to create continuous grids for each pathway.

Mechanism	Pathway category	Pathway sub-category	Data used to create a continuous grid
Commodity	Release in nature	Biological control	Country-level import data (live plant, animal or fish imports depending on organism type) distributed spatially within each country based on population density
Commodity	Release in nature	Erosion control/ dune stabilization	Same as above
Commodity	Release in nature	Fishery in the wild	Same as above
Commodity	Release in nature	Hunting in the wild	Same as above
Commodity	Release in nature	Landscape/flora/fauna improvement	Same as above
Commodity	Release in nature	Release in nature for use	Same as above
Commodity	Release in nature	Conservation introduction	No species introduced through this pathway
Commodity	Release in nature	Other Intentional release	No species introduced through this pathway
Commodity	Escape from confinement	Agriculture	Country-level import data (live plant, animal or fish imports depending on organism type) distributed spatially within each country based on population density
Commodity	Escape from confinement	Aquaculture/mariculture	Same as above
Commodity	Escape from confinement	Botanical garden/zoo/aquaria	Same as above
Commodity	Escape from confinement	Farmed animals	Same as above
Commodity	Escape from confinement	Forestry	Same as above
Commodity	Escape from confinement	Fur farms	Same as above
Commodity	Escape from confinement	Horticulture	Same as above
Commodity	Escape from confinement	Live food and live baits	Same as above
Commodity	Escape from confinement	Ornamental purpose	Same as above
Commodity	Escape from confinement	Pet/aquarium/terrarium species	Same as above
Commodity	Escape from confinement	Other escape from confinement	Same as above
Commodity	Escape from confinement	Research (in facilities)	No species introduced through this pathway
Commodity	Transport - Contaminant	Food contaminant	Country-level import data for edible vegetables and edible fruits and nuts distributed spatially within each country based on population density
Commodity	Transport - Contaminant	Seed contaminant	Country-level import data for seeds for sowing distributed spatially within each country based on population density

Mechanism	Pathway category	Pathway sub-category	Data used to create a continuous grid
Commodity	Transport - Contaminant	Contaminant on animals	Country-level live animal import data distributed spatially within each country based on population density
Commodity	Transport - Contaminant	Parasites on animals	Same as above
Commodity	Transport - Contaminant	Contaminant on plants	Country-level live plant import data distributed spatially within each country based on population density
Commodity	Transport - Contaminant	Contaminant nursery material	Same as above
Commodity	Transport - Contaminant	Parasites on plants	Same as above
Commodity	Transport - Contaminant	Transportation of habitat material	Same as above
Commodity	Transport - Contaminant	Timber trade	Country-level wood and wood articles import data distributed spatially within each country based on population density
Commodity	Transport - Contaminant	Contaminated bait	No species introduced through this pathway
Vector	Transport - Stowaway	Hitchhikers in or on plane	Air traffic at each airport
Vector	Transport - Stowaway	Angling/fishing aquaculture equipment	Country-level marine fishery data distributed spatially within each country based on coastal population density
Vector	Transport - Stowaway	Container/bulk	Human population density
Vector	Transport - Stowaway	Machinery/equipment	Human population density
Vector	Transport - Stowaway	Organic packing material	Human population density
Vector	Transport - Stowaway	Other means of transport	Human population density
Vector	Transport - Stowaway	People and their luggage	Human population density
Vector	Transport - Stowaway	Vehicles	Human population density
Vector	Transport - Stowaway	Hitchhikers on ship/boat	Shipping traffic at each maritime port
Vector	Transport - Stowaway	Ship/boat ballast water	Shipping traffic at each maritime port
Vector	Transport - Stowaway	Ship/boat hull fouling	Shipping traffic at each maritime port
Dispersal	Corridors	Interconnected waterways/basins/seas	No species introduced through this pathway
Dispersal	Corridors	Tunnels and land bridges	No species introduced through this pathway
Dispersal	Unaided	Natural dispersal across borders	No species introduced through this pathway

Table S4. The species for which information on pathways of introduction was not provided in the Global Invasive Species Database, the assigned pathways of introduction (categories and sub-categories) as classified by the Convention on Biological Diversity (CBD, 2014) and justification for these assignments.

Species	Assigned CBD pathway category	Assigned CBD pathway sub-category	Justification
Cinara cupressi	Transport - Contaminant	Contaminant on plants (Transport - Contaminant), Transportation of habitat material (Transport - Contaminant)	The cypress aphid can be transported on imported plant material (Remaudière & Binazzi, 2003 in CAB International, 2019).
Trichosurus vulpecula	Escape	Fur farms (Escape)	The Australian brushtail possum *Trichosurus vulpecula* was introduced in about 1840 to establish a fur industry which has continued to the present day (Cowan, 1992)

Table S5. The impact outcomes of alien species as classified by the Global Invasive Species Database (categories and sub-categories), and the data used to create global, binary grids showing sites that are sensitive to each impact.

Impact outcome category	Impact outcome sub-category	Data used to create a global grid of sensitive sites
Environmental Ecosystem - Habitat	Modification of hydrology/water regulation, purification and quality /soil moisture	Global 200 ecoregions (terrestrial, freshwater or marine ecoregions depending on the organism's habitat). Impact in sites where ecoregions are found
Environmental Ecosystem - Habitat	Primary production alteration	Same as above
Environmental Ecosystem - Habitat	Modification of nutrient pool and fluxes	Same as above
Environmental Ecosystem - Habitat	Modification of natural benthic communities	Same as above
Environmental Ecosystem - Habitat	Modification of food web	Same as above
Environmental Ecosystem - Habitat	Reduction in native biodiversity	Same as above
Environmental Ecosystem - Habitat	Unspecified ecosystem modification	Same as above
Environmental Ecosystem - Habitat	Habitat degradation	Same as above
Environmental Ecosystem - Habitat	Habitat or refugia replacement/loss	Same as above
Environmental Ecosystem - Habitat	Physical disturbance	Same as above
Environmental Ecosystem - Habitat	Modification of fire regime	Same as above
Environmental Ecosystem - Habitat	Modification of successional patterns	Same as above
Environmental Ecosystem - Habitat	Soil or sediment modification: erosion	Same as above
Environmental Ecosystem - Habitat	Soil or sediment modification: accretion/bioaccumulation	Same as above
Environmental Ecosystem - Habitat	Soil or sediment modification: modification of structure	Same as above
Environmental Ecosystem - Habitat	Soil or sediment modification: modification of pH, salinity or organic substances	Same as above
Environmental Ecosystem - Habitat	Other	Same as above
Environmental Species - Population	Population size decline	Same as above
Environmental Species - Population	Species range change (i.e. contraction, shift)	Same as above
Environmental Species - Population	Reduces/inhibits the growth of other species	Same as above
Environmental Species - Population	Alteration of genetic resources	Same as above
Environmental Species - Population	Indirect mortality	Same as above
Environmental Species - Population	Plant/animal health	Same as above

Impact outcome category	Impact outcome sub-category	Data used to create a global grid of sensitive sites
Environmental Species - Population	Interference with reproduction	Same as above
Environmental Species - Population	Other	Same as above
Socio-Economic	Damage to agriculture	Global landcover data indicating the presence of cropland. Impacts in sites where cropland is found
Socio-Economic	Damage to forestry	Global landcover data indicating the presence of tree cover. Impacts in sites where trees are found
Socio-Economic	Damage on aquaculture/mariculture/fishery	Country-level data on the quantity of organisms caught through fishing or produced through aquaculture distributed spatially within each country based on inland water presence for freshwater organisms and ocean primary productivity for marine organisms. Impact in sites with high values (≥ 75th percentile)
Socio-Economic	Reduce/damage livestock and products	Livestock density data. Impact in sites with high values (≥ 75th percentile)
Socio-Economic	Alteration of recreational use and tourism	Country-level data on the contribution of tourism to GDP distributed spatially within each country based on population density. Impact in sites with high values (≥ 75th percentile)
Socio-Economic	Impact on trade/international relations	Country-level data on the contribution of exports to GDP distributed spatially within each country based on population density. Impact in sites with high values (≥ 75th percentile)
Socio-Economic	Human health	Population density data. Impact in sites with high values (≥ 75th percentile)
Socio-Economic	Human nuisance	Same as above
Socio-Economic	Modification of landscape	Same as above
Socio-Economic	Damage to infrastructures	Same as above
Socio-Economic	Damage to ornamentals	Same as above
Socio-Economic	Modification of cultural, educational, aesthetic, religious values	Same as above
Socio-Economic	Limited access to water, land and other	Same as above
Socio-Economic	Other economic impact	Same as above
Socio-Economic	Other livelihoods	Same as above

Table S6. The species names as provided by the Global Invasive Species Database (GISD), and those used when searching for species occurrence data, the habitat in which the species occurs, the number of species occurrence records downloaded, the number with errors and the number of clean records used in the species distribution models. The distributions of seven species for which less than 30 clean occurrence records were available were not modelled (in grey).

GISD species name	Searched species name	Habitat	Records	Not in decimal degrees	Wrong species	Imprecise	Wrong environment	Wrong environment moved	Wrong environment removed	Duplicate	Errors	Clean records
Acridotheres tristis	*Acridotheres tristis*	Terrestrial	252804	0	0	17114	227	128	99	228448	245596	7208
Euglandina rosea	*Euglandina rosea*	Terrestrial	1413	2	0	627	269	225	44	493	1125	286
Oryctolagus cuniculus	*Oryctolagus cuniculus*	Terrestrial	77153	188	2	78	232	155	77	67959	68114	8851
Cinchona pubescens	*Cinchona pubescens*	Terrestrial	1474	167	0	4	0	0	0	1042	1046	261
Vespula vulgaris	*Vespula vulgaris*	Terrestrial	10601	45	12	7	2	2	0	8759	8778	1778
Acacia mearnsii	*Acacia mearnsii*	Terrestrial	5795	95	21	116	12	1	11	4888	5036	664
Achatina fulica	*Achatina fulica*	Terrestrial	222	0	0	13	32	22	10	99	120	102
Aedes albopictus	*Aedes albopictus*	Terrestrial	25885	0	0	64	12	11	1	20321	20386	5499
Anopheles quadrimaculatus	*Anopheles quadrimaculatus*	Terrestrial	402	5	0	0	0	0	0	332	332	65
Anoplolepis gracilipes	*Anoplolepis gracilipes*	Terrestrial	1655	62	0	3	91	54	37	1359	1399	194
Anoplophora glabripennis	*Anoplophora glabripennis*	Terrestrial	6	0	1	0	0	0	0	1	2	4
Ardisia elliptica	*Ardisia elliptica*	Terrestrial	459	7	0	17	21	17	4	267	288	164
Arundo donax	*Arundo donax*	Terrestrial	22752	94	8	57	101	68	33	20266	20361	2297
Asterias amurensis	*Asterias amurensis*	Marine	478	0	0	12	11	7	4	270	282	196
Bemisia tabaci	*Bemisia tabaci*	Terrestrial	1066	0	0	8	5	5	0	832	840	226
Boiga irregularis	*Boiga irregularis*	Terrestrial	5557	8	0	42	17	1	16	4741	4792	757
Rhinella marina	*Rhinella marina*	Freshwater	22141	68	0	80	8	2	6	19452	19538	2535
Capra hircus	*Capra hircus*	Terrestrial	13808	19	6	9	33	33	0	12441	12456	1333
Carcinus maenas	*Carcinus maenas*	Marine	18786	0	0	8	237	224	13	17676	17695	1091

GISD species name	Searched species name	Habitat	Records	Not in decimal degrees	Wrong species	Imprecise	Wrong environment	Wrong environment moved	Wrong environment removed	Duplicate	Errors	Clean records
Caulerpa taxifolia	*Caulerpa taxifolia*	Marine	1253	8	0	5	41	38	3	882	890	355
Cercopagis pengoi	*Cercopagis pengoi*	Marine	112	0	0	2	105	0	105	0	107	5
Cervus elaphus	*Cervus elaphus*	Terrestrial	27582	95	3	53	36	15	21	22954	23031	4456
Chromolaena odorata	*Chromolaena odorata*	Terrestrial	5640	391	18	79	18	3	15	3539	3650	1599
Cinara cupressi	*Cinara cupressi*	Terrestrial	34	0	0	0	0	0	0	23	23	11
Clarias batrachus	*Clarias batrachus*	Freshwater	1087	9	0	4	6	5	1	834	839	239
Clidemia hirta	*Clidemia hirta*	Terrestrial	4356	693	0	53	40	8	32	2396	2479	1184
Coptotermes formosanus	*Coptotermes formosanus*	Terrestrial	10	0	0	0	0	0	0	3	3	7
Cyprinus carpio	*Cyprinus carpio*	Freshwater	68596	9	0	93	292	48	244	59825	60137	8450
Dreissena polymorpha	*Dreissena polymorpha*	Marine	9802	0	1	8	8452	1692	6760	2895	9658	144
Eichhornia crassipes	*Eichhornia crassipes*	Freshwater	7208	257	0	121	15	8	7	5444	5570	1381
Eleutherodactylus coqui	*Eleutherodactylus coqui*	Freshwater	5929	0	27	1	3	3	0	5825	5853	76
Eriocheir sinensis	*Eriocheir sinensis*	Freshwater	10155	0	0	3	16	6	10	9861	9874	281
Euphorbia esula	*Euphorbia esula*	Terrestrial	23784	53	1	13	7	2	5	20523	20540	3191
Felis catus	*Felis catus*	Terrestrial	998	8	6	16	60	5	55	582	655	335
Gambusia affinis	*Gambusia affinis*	Freshwater	32317	45	0	25	7	2	5	27387	27417	4855
Hedychium gardnerianum	*Hedychium gardnerianum*	Terrestrial	1560	2	0	1	0	0	0	1352	1353	205
Hiptage benghalensis	*Hiptage benghalensis*	Terrestrial	398	1	1	12	21	3	18	260	286	111
Lantana camara	*Lantana camara*	Terrestrial	27946	1213	2	866	143	65	78	21943	22887	3846
Lates niloticus	*Lates niloticus*	Freshwater	430	8	0	100	0	0	0	213	313	109
Leucaena leucocephala	*Leucaena leucocephala*	Terrestrial	8387	1209	119	111	66	16	50	5499	5774	1404
Ligustrum robustum	*Ligustrum robustum*	Terrestrial	128	3	0	0	1	1	0	79	79	46
Linepithema humile	*Linepithema humile*	Terrestrial	28139	196	0	6	24	22	2	27525	27533	410
Lymantria dispar	*Lymantria dispar*	Terrestrial	2558	2	0	66	15	15	0	1734	1800	756
Lythrum salicaria	*Lythrum salicaria*	Terrestrial	196795	8	0	55	588	422	166	187555	187769	9018
Macaca fascicularis	*Macaca fascicularis*	Terrestrial	930	39	0	68	16	15	1	654	722	169
Melaleuca quinquenervia	*Melaleuca quinquenervia*	Terrestrial	4831	8	18	31	9	5	4	4341	4394	429

GISD species name	Searched species name	Habitat	Records	Not in decimal degrees	Wrong species	Imprecise	Wrong environment	Wrong environment moved	Wrong environment removed	Duplicate	Errors	Clean records
Miconia calvescens	*Miconia calvescens*	Terrestrial	4349	173	0	33	3	2	1	3757	3791	385
Micropterus salmoides	*Micropterus salmoides*	Freshwater	15002	6	0	6	0	0	0	10340	10346	4650
Mikania micrantha	*Mikania micrantha*	Terrestrial	904	39	204	9	4	0	4	238	454	411
Mimosa pigra	*Mimosa pigra*	Terrestrial	4567	300	1	128	22	9	13	2694	2830	1437
Mnemiopsis leidyi	*Mnemiopsis leidyi*	Marine	30528	0	3	4	24	20	4	30459	30470	58
Mus musculus	*Mus musculus*	Terrestrial	94308	615	0	228	260	119	141	82508	82874	10819
Mustela erminea	*Mustela erminea*	Terrestrial	26956	118	0	583	28	14	14	21168	21765	5073
Myocastor coypus	*Myocastor coypus*	Freshwater	7891	48	0	35	3	0	3	6027	6065	1778
Morella faya	*Morella faya*	Terrestrial	953	1	0	1	7	4	3	878	882	70
Mytilus galloprovincialis	*Mytilus galloprovincialis*	Marine	1661	0	5	158	80	74	6	1140	1309	352
Oncorhynchus mykiss	*Oncorhynchus mykiss*	Freshwater	114585	6	0	88	628	593	35	108467	108589	5990
Opuntia stricta	*Opuntia stricta*	Terrestrial	6333	11	1	19	116	34	82	5164	5266	1056
Oreochromis mossambicus	*Oreochromis mossambicus*	Freshwater	4116	3	2	22	22	10	12	2948	2980	1133
Pheidole megacephala	*Pheidole megacephala*	Terrestrial	6151	545	3	7	633	93	540	4574	5121	485
Pinus pinaster	*Pinus pinaster*	Terrestrial	29820	13	7	21	31	29	2	27531	27561	2246
Pomacea canaliculata	*Pomacea canaliculata*	Freshwater	372	0	2	18	15	8	7	203	227	145
Potamocorbula amurensis	*Potamocorbula amurensis*	Marine	149	0	1	0	39	36	3	126	130	19
Prosopis glandulosa	*Prosopis glandulosa*	Terrestrial	2328	94	0	105	0	0	0	1356	1461	773
Psidium cattleianum	*Psidium cattleianum*	Terrestrial	2643	633	3	26	27	5	22	1593	1644	366
Pueraria montana var. lobata	*Pueraria montana*	Terrestrial	1839	57	0	51	23	6	17	1144	1211	571
Pycnonotus cafer	*Pycnonotus cafer*	Terrestrial	54814	0	45	38	21	13	8	51936	52027	2787
Lithobates catesbeianus	*Lithobates catesbeianus*	Freshwater	26126	141	0	11	9	1	8	22317	22336	3649
Rattus rattus	*Rattus rattus*	Terrestrial	39822	285	0	419	298	130	168	34452	35025	4512
Rubus ellipticus	*Rubus ellipticus*	Terrestrial	361	7	0	1	1	1	0	211	212	142
Salmo trutta	*Salmo trutta*	Freshwater	216533	46	19	114	481	318	163	203663	203947	12540
Salvinia molesta	*Salvinia molesta*	Terrestrial	980	5	0	59	6	2	4	641	704	271

GISD species name	Searched species name	Habitat	Records	Not in decimal degrees	Wrong species	Imprecise	Wrong environment	Wrong environment moved	Wrong environment removed	Duplicate	Errors	Clean records
Schinus terebinthifolius	*Schinus terebinthifolius*	Terrestrial	890	0	2	4	24	6	18	450	474	416
Sciurus carolinensis	*Sciurus carolinensis*	Terrestrial	29265	43	0	15	7	3	4	26612	26631	2591
Solenopsis invicta	*Solenopsis invicta*	Terrestrial	3617	4	0	1	1	0	1	3072	3074	539
Spartina anglica	*Spartina anglica*	Terrestrial	9069	2	0	2	67	63	4	8743	8749	318
Spathodea campanulata	*Spathodea campanulata*	Terrestrial	1152	55	2	8	6	4	2	616	628	469
Sphagneticola trilobata	*Sphagneticola trilobata*	Terrestrial	3186	474	0	16	63	18	45	1731	1791	921
Sus scrofa	*Sus scrofa*	Terrestrial	72823	160	1	64	75	15	60	65581	65706	6957
Tamarix ramosissima	*Tamarix ramosissima*	Terrestrial	2745	25	0	38	4	3	1	1801	1840	880
Trachemys scripta elegans	*Trachemys scripta elegans*	Freshwater	6716	0	36	4	25	3	22	5282	5344	1372
Trichosurus vulpecula	*Trichosurus vulpecula*	Terrestrial	2520	37	0	7	10	10	0	2145	2152	331
Trogoderma granarium	*Trogoderma granarium*	Terrestrial	17	0	0	1	0	0	0	9	10	7
Ulex europaeus	*Ulex europaeus*	Terrestrial	49985	173	0	24	131	101	30	46498	46552	3260
Vulpes vulpes	*Vulpes vulpes*	Terrestrial	3400	228	0	58	25	12	13	2123	2194	978
Wasmannia auropunctata	*Wasmannia auropunctata*	Terrestrial	13631	0	0	4	27	8	19	13213	13236	395
Imperata cylindrica	*Imperata cylindrica*	Terrestrial	19743	47	0	376	56	40	16	16924	17316	2380
Platydemus manokwari	*Platydemus manokwari*	Terrestrial	15	0	0	0	0	0	0	7	7	8
Undaria pinnatifida	*Undaria pinnatifida*	Marine	380	0	0	0	5	4	1	275	276	104
Sturnus vulgaris	*Sturnus vulgaris*	Terrestrial	5357922	26	0	24892	23810	21619	2191	5283232	5310302	47594
Cecropia peltata	*Cecropia peltata & Cecropia schreberiana antillarum*	Terrestrial	2011	34	9	41	7	6	1	1355	1410	567
Herpestes auropunctatus	*Herpestes auropunctatus & Herpestes javanicus auropunctatus*	Terrestrial	7512	0	12	0	67	67	0	7369	7448	64
Polygonum cuspidatum Sieb. & Zucc. (=*Fallopia japonica* (Houtt. Dcne.)	*Fallopia japonica & Polygonum cuspidatum*	Terrestrial	29133	10	0	15	13	8	5	25087	25251	3872

Table S7. The datasets and variables used to model the distributions of the terrestrial, freshwater and marine species.

Habitat	Datasets	Variables
Terrestrial	Worldclim	Mean annual temperature, maximum temperature of the warmest month, annual precipitation, precipitation of the driest month, temperature seasonality, precipitation seasonality and mean diurnal temperature range
Freshwater	Worldclim and Envirem	Mean annual temperature, maximum temperature of the warmest month, precipitation of the wettest month, precipitation of the driest month, temperature seasonality, precipitation seasonality, mean diurnal temperature range and topographic wetness index
Marine	Bio-ORACLE	Maximum sea surface temperature, minimum sea surface temperature, sea surface temperature range, salinity, dissolved oxygen, pH, mean photosynthetically available radiation, and maximum photosynthetically available radiation

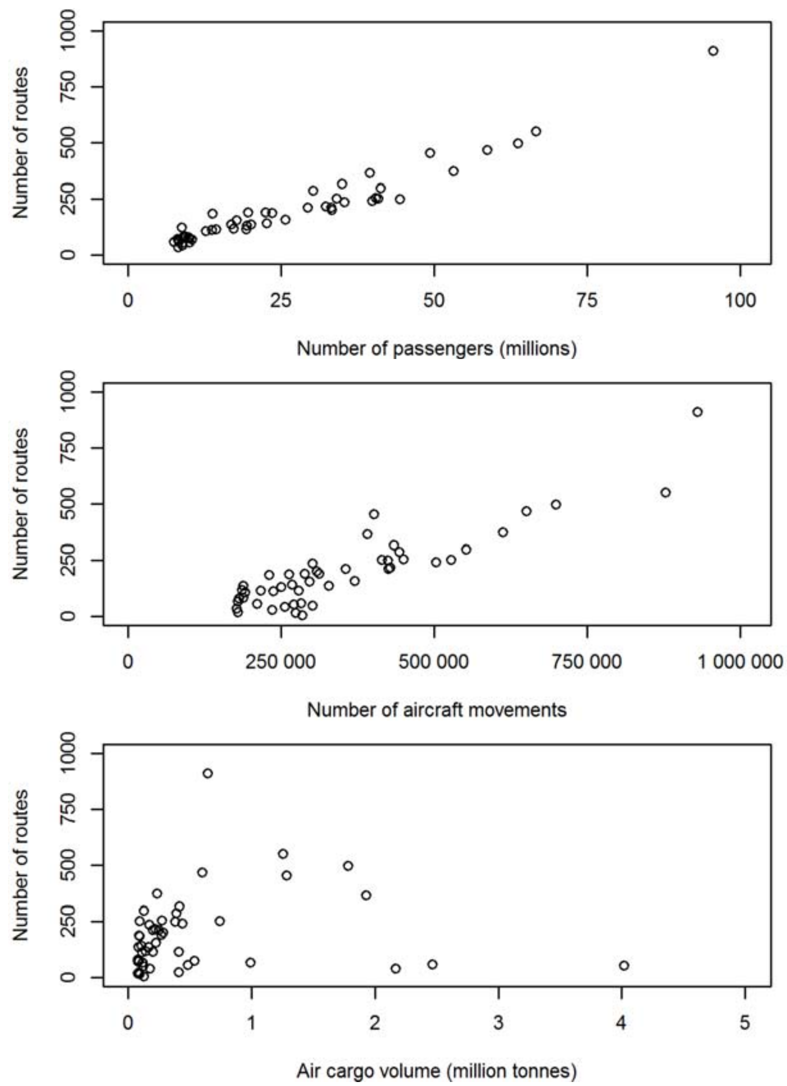

Figure S1. The relationship, for the 50 busiest North American airports as ranked by Airports Council International (ACI), between the number of routes and passenger numbers (top panel: Spearman's rho = 0.95, $p < 0.001$), number of aircraft movements (middle panel: Spearman's rho = 0.83, $p < 0.001$), and air cargo volumes (bottom panel: Spearman's rho = 0.38, $p = 0.007$). Data for the number of passengers, number of aircraft movements and air cargo volumes were obtained from ACI, and data for the number of routes were obtained from openflights.org.

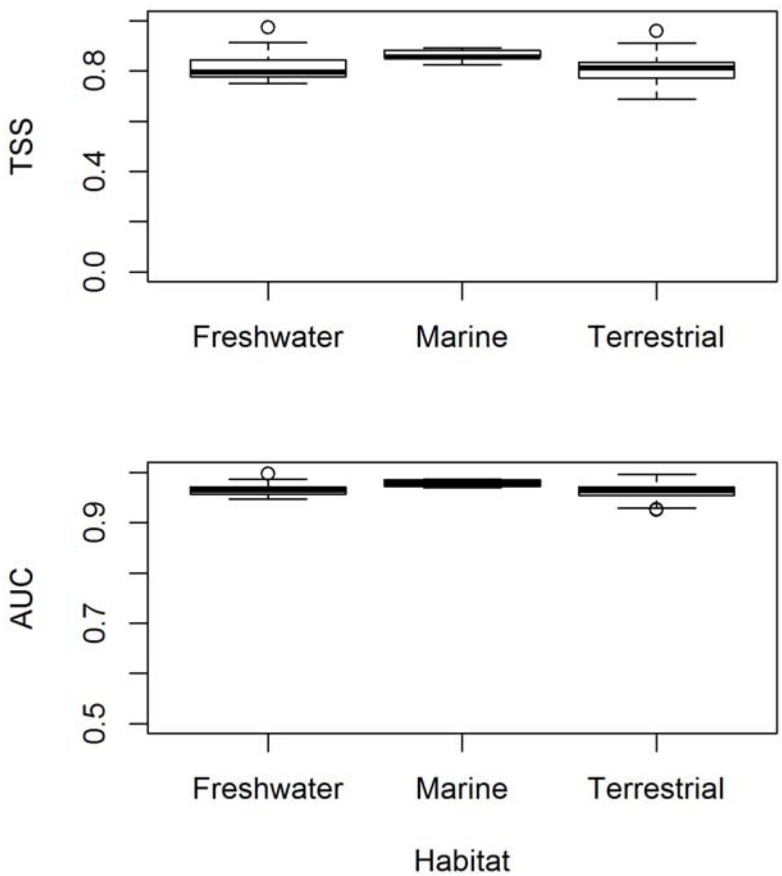

Figure S2. Boxplots of the true skills statistic (top panel) and area under the receiver operating curve (bottom panel) values for the projected ensemble models for freshwater, marine and terrestrial species.

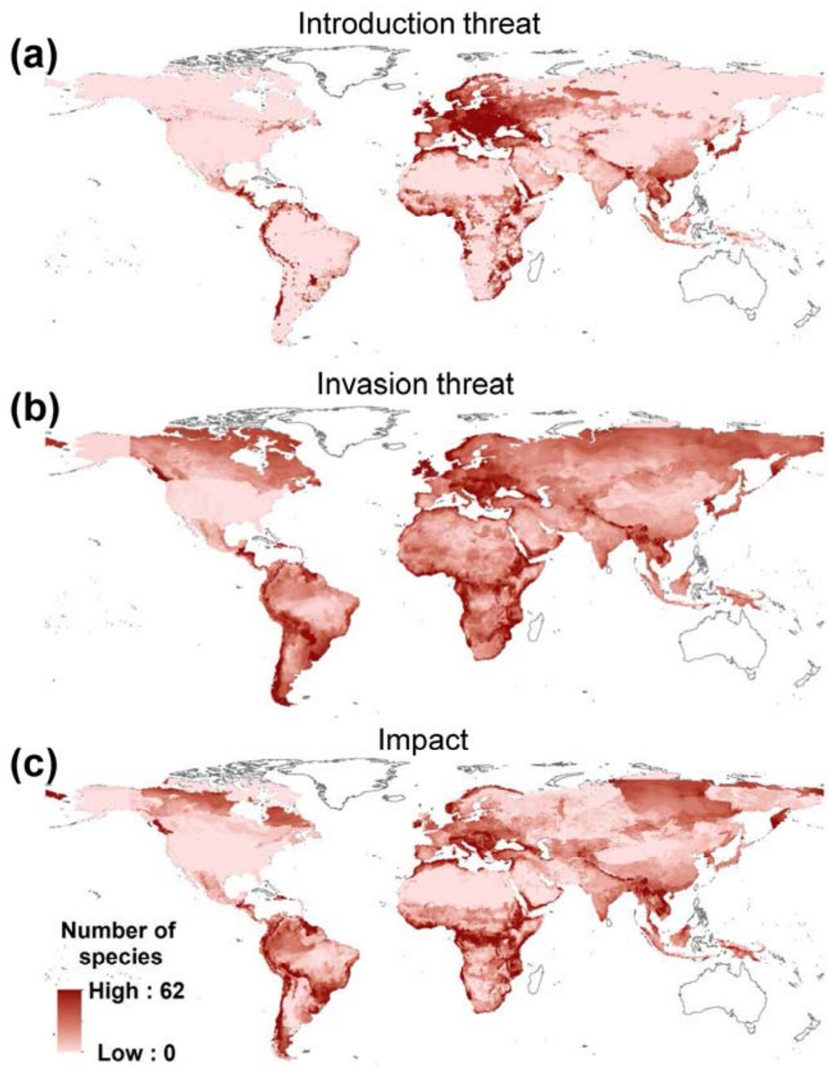

Figure S3. The number of species that could (a) be introduced, (b) invade and (c) have an impact in regions with contiguous countries. These maps are based on data for 86 species, and all maps have the same scale. Isolated countries (e.g. Australia) were not included in the analysis and are shown in white on the map.

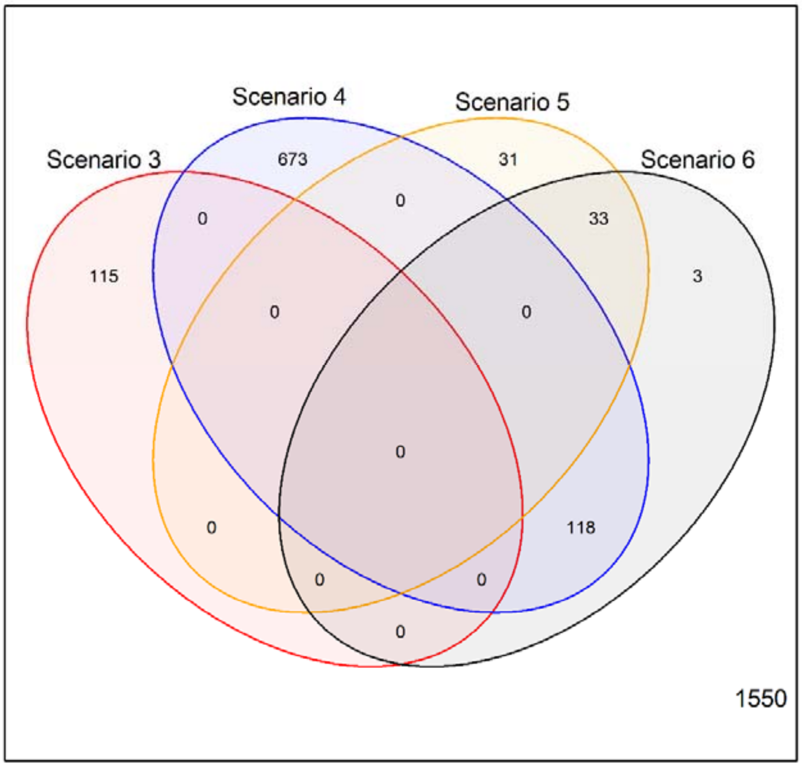

Figure S4. Venn diagram showing the number of invasions which conform to the four scenarios that describe invasions where a species spreads from the country of first establishment into neighbouring countries. For invasions with more than one potential country of first establishment, multiple scenarios are possible and these are shown in the diagram where the circles for the scenarios overlap. For example, 791 invasions could conform to Scenario 4, but of these invasions 118 could also conform to Scenario 6. 1550 invasions do not conform to any of the four scenarios and in these instances the species will not spread from the country of first establishment into neighbouring countries (i.e. they conform to either scenario 1 or 2).

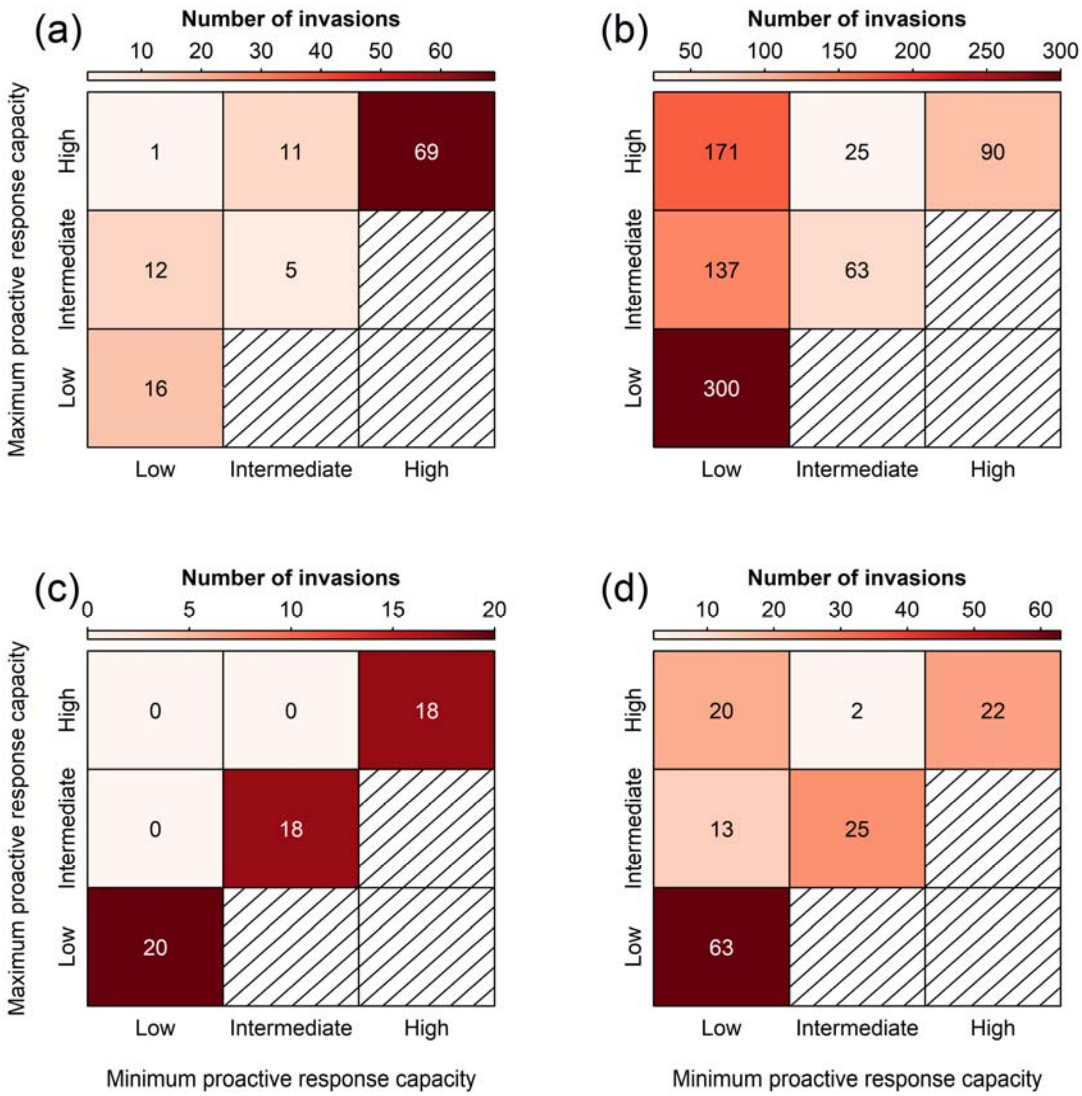

Figure S5. The minimum and maximum proactive response capacity of all of the potential countries of first establishment (FE) for invasions that conform to (a) scenario 3: an invasive species has no impact in the country of first establishment, and spreads into countries of subsequent invasion, where it also has no impact, (b) scenario 4: an invasive species has an impact in the country of first establishment, and spreads into countries of subsequent invasion, where it also has an impact, (c) scenario 5: an invasive species has an impact in the country of first establishment, and spreads into countries of subsequent invasion, where it has no impact, and (d) scenario 6: an invasive species has no impact in the country of first establishment, and spreads into countries of subsequent invasion, where it has an impact.

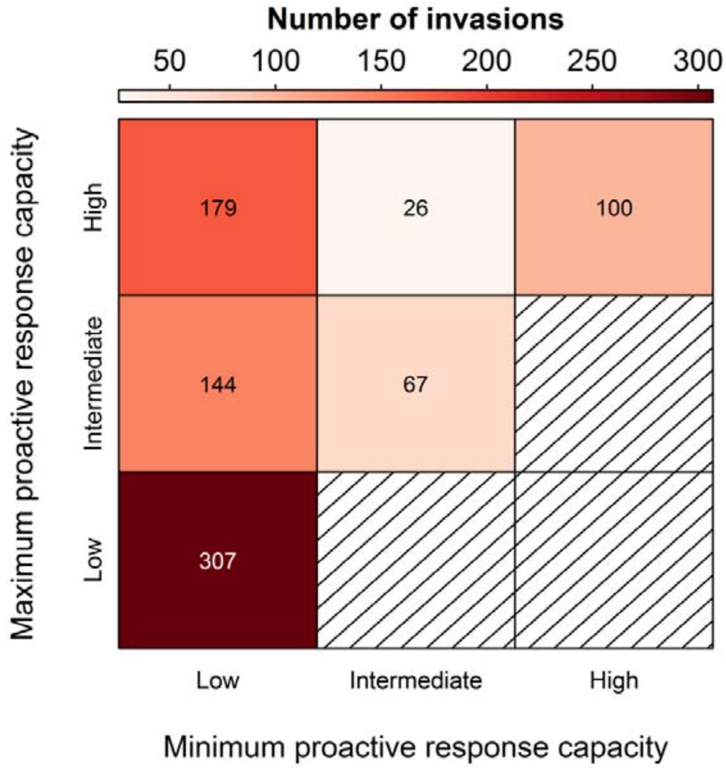

Figure S6. The minimum and maximum proactive response capacity of all of the potential countries of first establishment (FE) for invasions that conform to either scenario 4 or scenario 6. Scenario 4 occurs when an invasive species has an impact in the country of first establishment, and spreads into countries of subsequent invasion, where it also has an impact. Scenario 6 occurs when an invasive species has no impact in the country of first establishment, and spreads into countries of subsequent invasion, where it has an impact. Information on proactive response capacity was not available for 4 invasions.