



## Predicting the distribution of harmful species and their natural enemies in agricultural, livestock and forestry systems: an overview

M. Victoria Lantschner, Gerardo de la Vega & Juan C. Corley

To cite this article: M. Victoria Lantschner, Gerardo de la Vega & Juan C. Corley (2018): Predicting the distribution of harmful species and their natural enemies in agricultural, livestock and forestry systems: an overview, International Journal of Pest Management, DOI: [10.1080/09670874.2018.1533664](https://doi.org/10.1080/09670874.2018.1533664)

To link to this article: <https://doi.org/10.1080/09670874.2018.1533664>



Published online: 29 Nov 2018.



---

Submit your article to this journal [↗](#)



---

View Crossmark data [↗](#)

---



## Predicting the distribution of harmful species and their natural enemies in agricultural, livestock and forestry systems: an overview

M. Victoria Lantschner<sup>a</sup>, Gerardo de la Vega<sup>a</sup> and Juan C. Corley<sup>a,b</sup>

<sup>a</sup>Grupo de Ecología de Poblaciones de Insectos, Instituto de Investigaciones Forestales y Agropecuarias Bariloche, INTA - CONICET, Bariloche, Argentina; <sup>b</sup>Departamento de Ecología, Centro Regional Universitario Bariloche, Universidad Nacional del Comahue, Bariloche, Argentina

### ABSTRACT

Predicting the potential distribution of harmful species to agriculture, livestock and forestry is decisive to prevent their impacts, especially when these are expanding their range due to global change. Recent advances in species distribution modelling (SDM) have made these tools widely used for biosecurity studies. We reviewed the available literature of SDM for pest, weeds, pathogen species and biological-control agents, with the aims of synthesizing and quantifying the available information, and identifying gaps in the knowledge and future perspectives. SDMs for 420 species were collected from 220 publications. Insect pests were the most frequently studied organisms. CLIMEX and MaxEnt were the most commonly used modelling tools, while pure mechanistic approaches were rarely applied. Most studies covered broad scales, and focused on predicting the distribution of invasive species and/or the effects of climate change. The challenge remains for models to include disturbance, resource availability, and biotic factors, as well as to better quantify uncertainty. This future directions will be fundamental to improve the predictive power of SDMs for productive systems in the context of a rapidly changing World.

### ARTICLE HISTORY

Received 30 May 2018  
Accepted 1 October 2018

### KEYWORDS

Species distribution models (SDMs); pests; weeds; pathogens

### Introduction

Agricultural, livestock and forestry systems around the world are threatened by growing numbers of pests and pathogens (Crowl et al. 2008; Aukema et al. 2011; Bebbler et al. 2014). Human activities are largely responsible for changes in the distribution range of many species during the last centuries. Increasing global trade and human movement, as well as increased accessibility to new sources of species pools, has remarkably accelerated the rate of introductions of non-natives around the world (Everett 2000; Levine & D'Antonio 2003; Seebens et al. 2018). The geographic extent of invasions and the proportion of non-native species in different biotas are unprecedented and affect all continents simultaneously (Ricciardi 2007), promoting a process of biotic homogenization through which local specialist species are displaced by global generalists (McKinney and Lockwood 1999). Although not every non-native species has negative impacts in the invaded habitat, the establishment of some can bring about notorious damage to humans, and are so termed pests, weeds or pathogens (Bebber et al. 2014).

Climate change and other components of global change such as land use change can also affect

species distribution (both native and non-native) and resource dynamics. There is ample evidence that recent climatic changes have affected a broad range of organisms with diverse geographical distributions (Walther et al. 2002). In particular, increases in temperature have been found to cause significant range shifts toward higher latitudes and higher elevations (Chen et al. 2011). Land use change, on the other hand, has also promoted important shifts in species distribution through habitat modification. Range expansion of native and non-native generalist herbivores are typically favored by extensive crops which are based mainly on species and varieties selected for increased yield (Bebber et al. 2014). In sum, land use and climate changes interact with human transportation networks to accelerate the range expansion of pest, pathogens and weed species worldwide (Dukes and Mooney 1999; Crowl et al. 2008; Rassati et al. 2018). As a result, the invasion of non-native damaging species as well as range shifts of native ones, represent two of the most important issues concerning international biosecurity planning and pest management (Dukes et al. 2009; Ramsfield et al. 2016).

Predicting the potential distribution of both native and non-native species that threaten food security is crucial to prevent or mitigate their

economic and ecological impacts (Venette et al. 2010; Eyre et al. 2012; Sutherst 2014). While pest, weed and pathogen management and/or eradication is essential once species have established, there is consensus in that prevention is the most cost-effective way to avoid conflicts with human interests and ecosystem services (Eyre et al. 2012; Kriticos 2012; Sutherst 2014). The prediction of the potential distribution of species is very useful in different stages of pest/weed risk assessment: (a) can help allocate and optimize tools and resources in border surveillance, by understanding where a given non-native species is most likely to establish; (b) can help understand where non-natives that have already established could spread to, or for native species, if they can shift their distribution range under global change scenarios and so help focus -monitoring and cross-border surveillance efforts (Kriticos 2012; Poland and Rassati 2018); and (c) can help identify climatically suitable regions (and/or suitable biotypes) in which biological control agent releases are more likely to succeed (Robertson et al. 2008).

Recent advances in statistical and mathematical tools, and the development of large databases, have made species distribution modelling easier, faster, and more widely used by biosecurity researchers and practitioners (Sutherst 2014). Species distribution models (SDMs) aim to predict potential species distributions by matching species' environmental preferences with conditions in physical space (Guisan and Zimmermann 2000). There are two main approaches to SDMs, which can be broadly characterized as: (a) correlative models, which use statistical relationships between the known species distribution and environmental (often climatic) variables to estimate the suitability of other sites based on their similarity to regions where the species already exists (Elith and Leathwick 2009), and (b) mechanistic models, that use eco-physiological information, describing organismal responses to environmental conditions that can either be determined experimentally or inferred from known distributions (Kearney and Porter 2009). In both cases, geographic areas are classified as being more or less suitable for species establishment. Although much debate has been made about the advantages and disadvantages of each approach (Kriticos et al. 2013; Schymanski et al. 2013), both have their strengths and weaknesses, which need to be understood and appreciated (Dormann et al. 2012; Bebbber 2015). In this sense, mechanistic or process-based distribution models are less limited than correlative models in their biological realism and their transferability to novel environments, although they often demand a large number of parameters to be estimated, many

requiring data of limited availability at often high spatiotemporal resolution (Dormann et al. 2012).

There is a substantial body of literature assessing the potential distribution of numerous pest, pathogen, and disease vector species, as well as their biological control agents; in different crops, livestock and forestry systems; at different scales; and under different scenarios of climate and land-use change, with the general purpose of extrapolating risks to the biosecurity sector. Here, we review these studies, with the aim of synthesizing, describing and quantifying, the main studied taxa, the affected systems, the methods used to model species distribution, the location, extent, and temporal scale of the studies, and other general patterns. We discuss the state of the art in the use of SDMs for agricultural, livestock and forestry pest risk assessment, identifying gaps in the knowledge, standing challenges and future perspectives.

## Methods

### Data collection

We conducted a systematic search for papers published since the early days until 2017 that used SDMs to predict the distribution of species regarded as pests (Kingdom Animalia), weeds (Kingdom Plantae), disease vectors (species of any taxa able to act as vector of human or livestock diseases) or pathogens (Kingdom Fungi, Bacteria, Virus or Chromista) of agriculture, forestry or livestock systems, as well as on the biological control agents to manage some of these harmful species. The search included native and non-native species, that are present in a given region but have a limited distribution, or non-native species that are not yet present in an area of concern. The databases used to search for publications were Scopus (<https://www.scopus.com>) and Google Scholar (<https://scholar.google.com/>). We used different combinations of the key words “species distribution model”, “niche modelling”, “pest risk assessment”, “pest”, “pathogen”, “weed”, “disease vector”, “biological control” in the searches. We included in the review only peer-reviewed journal articles published in English. It must be noted here that there is a limited number of papers published in languages other than English, which were not considered in our analysis.

From the extensive set of documents found, we used two specific criteria to determine whether the reference was appropriate for this review. Criteria for inclusion of articles were that the study: (1) explicitly modeled the distribution of one or more species using a niche-based distribution model approach, defined as the process of using computer algorithms to predict the distribution of species in

geographic space, by relating data on species occurrence or eco-physiology to environmental information, and (2) included species have been found to be harmful to agricultural crops, forestry systems, or livestock systems; or are used as biological control agents of any of these species.

### Data processing and analysis

For each paper we recorded information about the publication itself (authors, year of publication, journal, volume and number of pages), the niche-based model/s used to predict species distribution (Climex, Maxent, GLM, BIOCLIM, DOMAIN, etc.), the modelling approach (correlative, semi-mechanistic, mechanistic), the predictive variables used in the SDM (temperature, precipitation, relative humidity, topology, land cover, soil and geology, landscape configuration, human population), the source of the climatic variables (WorldClim, Climond, Climex, local data bases, etc.), the location (region where species distribution was predicted) and the extent of the study area (surface area in km<sup>2</sup>, as well as in terms of political boundaries: global, inter-continental, continental, sub-continental, national, regional or local), the species distribution range used for the construction and for the projection of the model (native or non-native range), the method and dataset used to evaluate the model, and the temporal scale of the prediction (current conditions, future or past conditions). The list of the reviewed articles is available in Supplemental Online Material 1, and the complete database of article characteristics is available in Supplemental Online Material 2.

Additionally, for each studied species we recorded the taxonomic information (kingdom, class, order, family, genus and species), the number of articles that predicted its distribution, the species type (pest, pathogen, disease vector, weed, biological control agent), the affected crop or productive system (cereals, vegetables, fruits, legumes, pastures, forestry, ornamentals, vegetal fibers, oilseeds, livestock, apiculture). The complete data base of studied species information is available in Supplemental Online Material 3.

To assess how variability within the articles was explained by the damaging species category and the modelling approach, we employed a multivariate approach that calculated the distance between categorical variables. We used correspondence analysis (CA) to summarize the dataset in two-dimensional of orthogonal axes. CA employs a double transformation ensuring that all elements present are similarly correlated with the ordination axes (Manly 1994), unlike many other ordination

techniques (e.g., PCA). We first built a contingency table of the counts at each combination of factor levels for damaging species category and the modelling approach. Then, we performed the correspondence analysis, using the “ca” function from ca package (Nenadic and Greenacre 2007) in R (R-Core-Team 2017).

## Results

### General patterns

In total, we collected information on SDMs for 420 species, from 220 publications (Supplemental Online Material 2 and 3). The first articles were published in the late 1980s, and after then we note an exponential growth in the number of publications, where 75% of the articles were published during the last eight years (2010–2017, Figure 1). Among the total publications we recorded, 185 were focused on only one species, 25 analyzed between two and five species, and 11 articles studied more than five species.

### Studied organisms

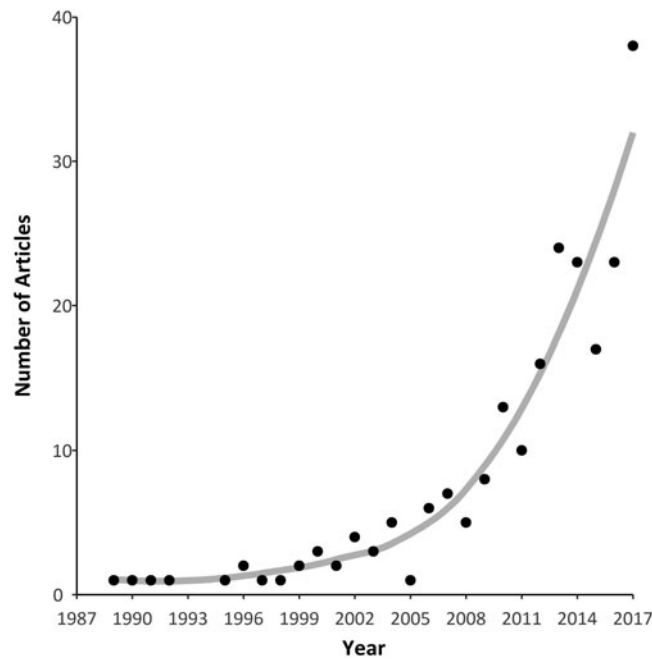
From the 420 studied species, 327 were classified as pests, 33 as pathogens, 25 as weeds, 18 as disease vectors, and 17 as biological control agents (Figure 2a). The studied species comprised a wide range of taxa, dominated by Animalia (86%), from which the majority were Insecta (79%); Fungi (7%), Plantae (6%), and Chromista, Bacteria and Virus (1%). Among the 330 insect species, the most common order was Coleoptera (40%), followed by Hymenoptera (16%), Lepidoptera (14%), Hemiptera (12%), and Diptera (8%). The most commonly studied species were: *Bactrocera dorsalis* (Diptera: Tephritidae, 8 articles), *Ceratitidis capitata* (Diptera: Tephritidae, 5 articles), *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae, 5 articles), *Ostrinia nubilalis* (Lepidoptera: Crambidae, 5 articles), *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae, 5 articles), *Cydia pomonella* (Lepidoptera: Tortricidae, 4 articles), *Diabrotica virgifera* (Coleoptera: Chrysomelidae, 4 articles), *Helicoverpa armigera* (Lepidoptera: Noctuidae, 4 articles), *Lymantria dispar* (Lepidoptera: Lymantriidae, 4 articles), *Phytophthora ramorum* (Peronosporales: Peronosporaceae, 4 articles), *Rhagoletis pomonella* (Diptera: Tephritidae, 4 articles). The most commonly affected systems by the 420 studied species were forestry systems (46%), fruit crops (39%), vegetable crops (38%), cereal crops (30%), and ornamentals (19%; Figure 2b).

### SDMs approaches

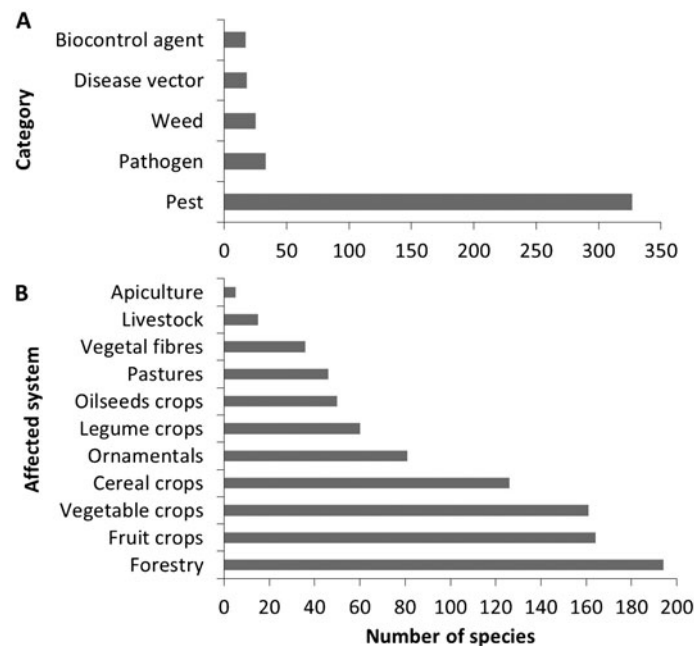
We found that 33 different niche-based model approaches were used among the 220 articles. The most frequently used methodologies were, CLIMEX (57%), MaxEnt (35%), GLM (5%), GARP (5%) and BRT (4%; Table 1). It should be consider, that the frequency of use among methodologies is influenced by the date each model was released. In this sense, MaxEnt is the most recently developed one (in 2006) among the named methodologies, and has really taken over as the predominant correlative method since its creation. When comparing the

modelling tools in terms of correlative or mechanistic approaches we found that semi-mechanistic approaches were the most frequently used ones (56% of the articles), followed by correlative models (43% of the articles), while only a few articles used mechanistic approaches (3% of the articles; Figure 3a).

The first dimension of the correspondence analysis for Category and Modelling approach explain the 85.1% of the variability (*inertia*) and the second the 13.8% (Figure 4). We found that mechanistic modelling approaches are related with disease



**Figure 1.** Number of articles per year modelling the distribution of damaging species of agriculture, livestock and forestry systems. The grey line shows smoothed data.

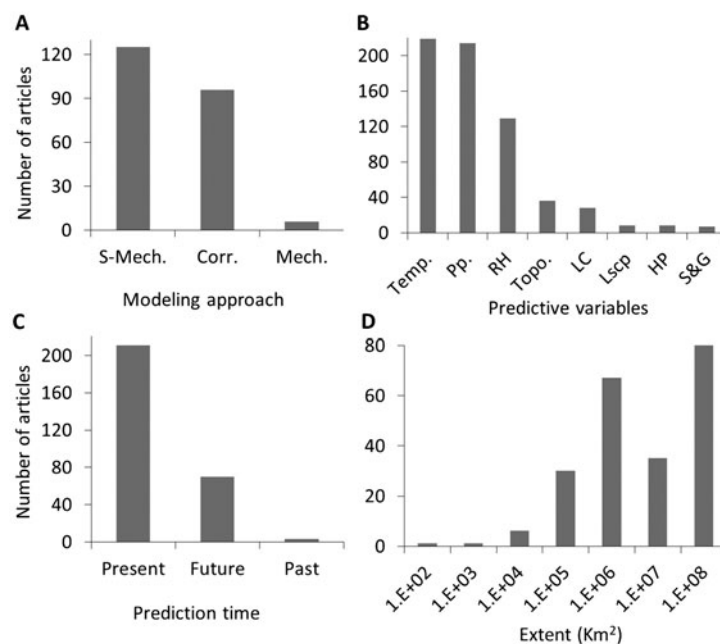


**Figure 2.** Number of species studied in the reviewed articles per (a) category; (b) affected productive system.



**Table 1.** Number of articles reviewed per niche-based modelling method. The modelling approach of each method is described.

Niche-based modelling methods	Modelling approach	N° articles
Climex	Semi-mechanistic	125
MaxEnt	Correlative	78
Generalized linear model (GLM)	Correlative/Mechanistic	10
Genetic Algorithm for Rule Set Production (GARP)	Correlative	10
Boosted regression trees (BRT)	Correlative	9
Generalized Additive Models (GAM)	Correlative	6
Decision tree analysis	Correlative	4
BIOCLIM	Correlative	5
Ecological Niche Factor Analysis (ENFA)	Correlative	3
Artificial neural network	Correlative	3
Multivariate adaptive regression splines (MARS)	Correlative	3
Random forests	Correlative	3
Climate Space Model	Correlative	2
Envelope Score	Correlative	2
Environmental Distance	Correlative	2
DOMAIN	Correlative	2
Flexible discriminant analysis	Correlative	2
Surface range envelopes	Correlative	2
Fuzzy multi-criteria evaluation	Correlative	2
Mahalanobis Distance and Domain	Correlative	1
Niche mosaic	Correlative	1
Support Vector Machines	Correlative	1
Back propagation neural network	Correlative	1
Logistic regression	Correlative	1
Multi- Model Framework	Correlative	1
Physiologically based demographic model	Mechanistic	1
ECAMON	Mechanistic	1
Insect Life Cycle Model	Mechanistic	1
Thermal niche suitability index	Mechanistic	1
Raster classification	Mechanistic	1
BioSIM	Mechanistic	1

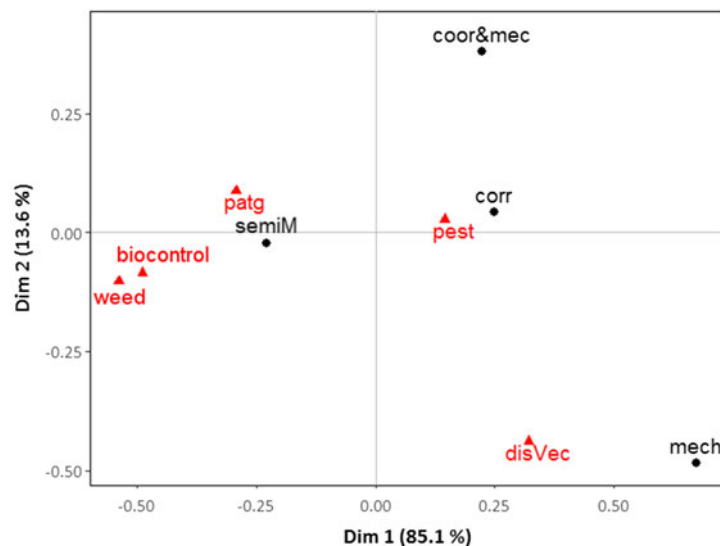


**Figure 3.** Number of articles per (a) modelling approach, S-Mech: semi-mechanistic, Corr.: correlative, Mech.: mechanistic; (b) predictive variables, Temp.: temperature, Pp: precipitation, RH: relative humidity, Topo: topology, LC: land cover, Lscp: landscape configuration, HP: human population, S&G: soil and geology; (c) moment of the prediction; (d) study extent categories. Categories are based on a log scale (i.e., the category “1.E+02” represent all the articles with extensions from 100 to 999 km<sup>2</sup>).

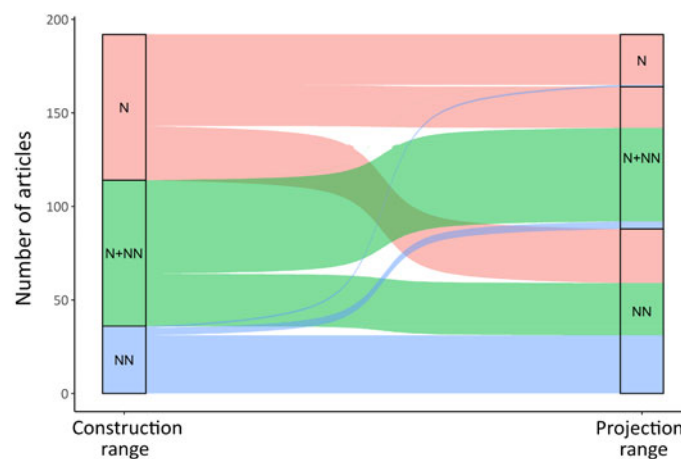
vectors, correlative approaches with pests, and semi-mechanistic approaches with weeds, pathogens and biological control agents. We found a significant relationship between the modelling approach (in the case of studies that used correlative or semi-mechanistic approaches), and the species category ( $X^2 = 11.534$ ,  $df = 4$ ,  $p\text{-value} = 0.02$ ).

### **Range and time of models construction and prediction**

Most of the analyzed articles used species distribution occurrence information to build the models (87%), while 13% were based only on species eco-physiological information. Among the articles that



**Figure 4.** Correspondence Analysis between damaging (triangles) species category and modelling approach (circles). Patg: pathogen, disVec: disease vector, biocontrol: biological control agent, semiM: semi-mechanistic approach, corr: correlative approach, mech: mechanistic approach, coor&mec: correlative and mechanistic approaches.



**Figure 5.** Alluvial diagram showing the frequency of articles that used occurrence data from the native vs. non-native species distribution range for the construction of the model, and the frequency of each of these articles that were projected to the native vs. non-native distribution range of the species. N: native range, NN: non-native range, N+NN: native plus non-native range.

were constructed based on species distribution, relied on species occurrence data that came only from their native distribution range (40.6%), or from both the native and non-native range (40.6%), while a smaller number of articles used only occurrence data from the species non-native range (18.8%). When considering the range of the model projection, most of the 220 reviewed articles (48%) predicted the distribution of species on regions outside of their range, or including both the native and non-native range (35%), while only a low number of articles predicted the distribution of species only on its native range (17%; Figure 5). Regarding the relation between the distribution range from where the model was built, with where it was projected, the most common case was that of models built and projected on both the combination of the native and

non-native ranges (26% of the articles), followed by models constructed and projected on the non-native range (16% of the articles), and models constructed on the native range and projected on the non-native range (15% of the articles, Figure 5).

In terms of the time or scenarios for the predictions, most of the articles predicted the distribution of the species under current environmental conditions (96%), while one third of the studies also predicted the distribution under future conditions (32%), and a few predicted the distribution of species under past conditions (1%; Figure 3c). Among the articles that predicted species distribution under global change scenarios, 63 were based on climate change scenarios, while 10 (all of them CLIMEX models) included land-use change scenarios, particularly the effect of irrigation.

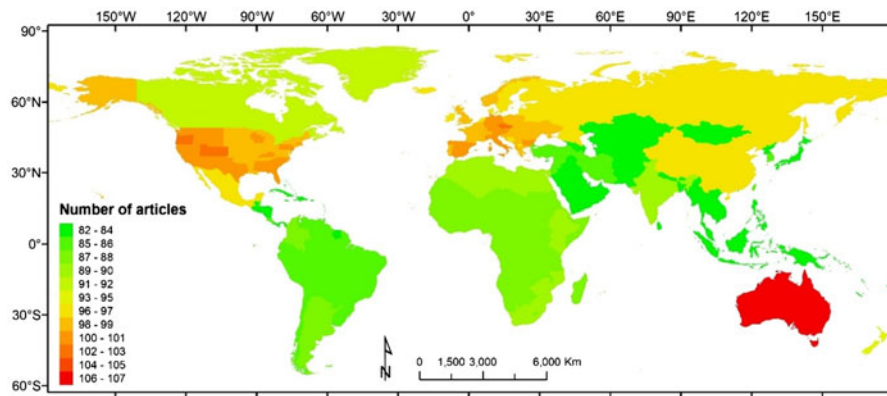


Figure 6. Number of articles per study area location covered by the reviewed SDMs.

### Extent and location of the studies

The spatial extent of the studies varied widely, from 283 km<sup>2</sup> to 148 million km<sup>2</sup>, although there were two extents that were considerably more frequent, those of the order of 100 million km<sup>2</sup> and those in the order of 1 million km<sup>2</sup> (Figure 3d). In terms of political boundaries these extents were distributed between global (36% of the articles) and national (35% of the articles). Regarding their location, all terrestrial regions of the world, except for Antarctica, were represented among the reviewed articles, ranging from 82 to 107 articles per location (Figure 6). The region most frequently modelled was Australia, followed by several countries of Europe, and the USA; the less frequently modelled regions were East and Southern Asia, followed by South America and Africa.

### Distribution and predictive variables databases

Almost all of the reviewed articles used climatic variables to predict the distribution of the studied species. Temperature (99.5% articles) and precipitation (97.3% articles) were the most frequently used variables, followed by relative humidity (58.6% articles), while other non-climatic environmental variables were used to a lesser extent, such as topology (16.4% articles), land cover (12.7% articles), landscape configuration (3.6% articles), human population (3.6% articles), and soil and geology characteristics (3.2% articles; Figure 3b). Among the climatic variables, we found that although several studies used local or national climatic databases, there were two main global climatic databases that were very frequently used: WorldClim (35% articles) and CRU (15% articles), while 17% of the articles used the CliMond database, which is constructed combining data of WorldClim and CRU, and finally, 15% of the remaining articles (the oldest ones) used CLIMEX meteorological station data, which are also based on CRU data. Regarding the source of species distribution data used to construct or validate the

models, most (73%) were based on the combination of published articles, museum collections, and/or field surveys; while nearly 23% also accessed online global databases (mainly GBIF, CABI, and EPPO).

### Model accuracy and validation

Most of the articles (74%) evaluated the accuracy of the models, but a considerable fraction (26%) did not. The most frequently used accuracy analysis methods were AUC (54% of the articles with model evaluation), visual validation (41%), TSS (6%) (Supplemental Online Material 2). The most frequent species distribution data used to evaluate the model was a sub-set from original data base (52% of the articles with model evaluation), while 47.2% of the cases declared to use an independent database, frequently originated from the distribution range where the model was projected. Among the mechanistic models, only one used experimental data to evaluate the model.

## Discussion

### Studied organisms

Our literature review revealed that there is a considerable and increasing amount of articles using SDMs to predict the distribution of damaging species in different productive systems around the world (Figure 1) (Razgour et al. 2016; Bellard et al. 2018; Booth 2018). SDMs have been applied to a great diversity of organisms from several taxa, including pests, pathogens, disease vectors and weeds, as well as species that are used as biological control agents. However, among the different species types, pests, and particularly insect pests (79% of the studied species), have been considerably more studied. This is probably because insects are an abundant and diverse group of organisms and they are easy to transport accidentally become invasive in many cases. Insect herbivores are likely the economically most important pests of agriculture and



forestry (Metcalf and Luckmann 1994; Brockerhoff and Liebhold 2017; Foottit and Adler 2017). This pattern confirms previous findings that show that invasive animal distribution is more frequently studied than that of plants (Pyšek et al. 2006). Additionally, it should be noted that the general biology and distribution of many insect pests is often better known than that of pathogens, for which taxonomy and identification methods are less well resolved (Purse and Golding 2015; Ireland and Kriticos 2018). Furthermore, most of the efforts have been dedicated to model the distribution of harmful species themselves, while much less work has been devoted to study the potential distribution of their biological control agents (4% of the studied species).

The most commonly studied species, with five or more articles modelling their distribution, were all major pests of important crops that have invaded several regions of the world. This is the case of the oriental fruit fly, *Bactrocera dorsalis* and the Mediterranean fruit fly, *Ceratitidis capitata* (Diptera: Tephritidae; 8 and 5 articles respectively), which are very destructive pests of fruit, in areas where they occur. The oriental fruit fly is native to tropical Asia and has invaded much of sub-Saharan Africa, and the United States; while *C. capitata* is native to sub-Saharan Africa, and has invaded Australasia and North and South America (Malacrida et al. 2006). Another well studied case is the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae; 5 articles), which is native to South America, but has become a major potato pest across North America, Europe and Asia (Rafoss and Saethre 2003). The European corn borer, *Ostrinia nubilalis* (Lepidoptera: Crambidae; 5 articles), a pest of grain, particularly corn, is native to Europe and has been introduced to North America (Kocmánková et al. 2011). And finally the North American cotton mealybug, *Phenacoccus solenopsis* (Hemiptera: Pseudococcidae; 5 articles), affects cotton crops everywhere it has invaded (Australasia, Afrotropical, Nearctic, Oriental and Neotropical regions) (Parsa et al. 2012).

### SDMs approaches

The most frequently used species distribution modelling tools among the articles reviewed, were CLIMEX (Sutherst and Maywald 1985) and MaxEnt (Phillips et al. 2004). CLIMEX is a semi-mechanistic approach, which uses a combination of empirically measured parameters, abundance and point distribution records, to examine the relationship between climate, species distributions and patterns of growth (Macfadyen and Kriticos 2012). This tool is very

popular in pest risk assessment probably because it was, in part, created for this goal. Although its use requires a paid license, it has the advantage that it combines distribution and physiological information of the target species, and thus it is useful for situations when the available data abundance, distribution and biology is ‘patchy’ (Hill and Thomson 2015; Jung et al. 2016). In this sense, this approach requires more knowledge about the species being modeled than when using correlative models, which may limit its applicability for species with no such data. MaxEnt, on the other hand, is a machine-learning method that implements the maximum entropy algorithm for modelling species distributions with presence-only data (Phillips et al. 2004). It is a correlative method, and it uses known occurrences of a species together with predictor environment variables to quantify the species’ ecological niche in the examined environmental dimensions. It is not surprising to have found this method is frequently used, as it is freely available and has been ranked among the best-performing SDMs methods yet, especially for small sample sizes (Phillips and Dudík 2008; Wisz et al. 2008; Bradie and Leung 2017) and can be applied to a wide range of species, including those with limited data. Our results are consistent with those found in a review of publications assessing the effects of climate change on invasive alien species distribution (Bellard et al. 2018).

When comparing modelling approaches together, semi-mechanistic models were the most frequently approach used (55% of the articles), since this approach is the one used by CLIMEX. Correlative approaches were also very frequently used (42% of the articles), while pure mechanistic approaches were rarely applied (3% of the articles). Interestingly, we found an association between the modelling approach and species category (Figure 6). Pests were more frequently modelled with correlative approaches, while weeds, pathogens, and biological control agents were associated with semi-mechanistic modelling approaches; and disease vectors with mechanistic approaches. This pattern could be due to the fact that, traditionally, weeds and pathogens distributions have been modeled with CLIMEX, and additionally it is easier to collect eco-physiological data for this group of organisms. Pests, in turn, have been modeled with both semi-mechanistic and a wide variety of correlative models. In the case of mechanistic models, the number of cases is too low to be able to draw conclusions.

There has been much debate about the pros and cons of SDMs approaches (Dormann et al. 2012; Peterson et al. 2015). Correlative models are constrained by the fact that they describe patterns and not the mechanisms behind them. Consequently

predictions based on correlative modelling are usually limited in their biological realism and their application to novel environments (Franklin 2010). Mechanistic models, on the other hand, have the potential for the direct measurement of dimensions of the species' fundamental niche without the confounding effects of accessible areas and interacting species. But, the disadvantage of this approach, and probably the cause of it being less used when compared to correlative ones, is their complexity: the need to identify the key physiological features of an organism and to scale them, often from millimeters to kilometers. Additionally, very few software packages for mechanistic modelling are, as yet, freely available to the scientific community (Peterson et al. 2015). Although semi-correlative models such as CLIMEX use information from correlative and process-based models, it may be fruitful to join these two approaches in a more complex combined workflow (Buckley et al. 2011; Mokany and Ferrier 2011; De la Vega and Corley 2018). Correlative models efficiently sift through descriptive data, thereby generating hypotheses on potentially underlying processes. These hypotheses can then be taken up, along with ecological theory and experimental evidence, by mechanistic models. The resulting mechanistic models can then generate predictions specifically designed for a formal test on independent data (Dormann et al. 2012). In this way, combined models can synergistically progress the field in a way neither correlative nor mechanistic approaches can do so by themselves.

### **Extent and location of the studies**

The spatial extent of the studies varied widely, from regional to global, although most studies focused on broad scales (1-100 million km<sup>2</sup>), being global and national extents (within country limits) the most common ones. Extent together with grain (resolution) are the components of spatial scale, and both are important factors that may affect predictions. Studies with predictions over large extents are often characterized by coarser resolutions or "grain sizes" (e.g., global scales) than data at small extents (e.g., regional) (Guisan et al. 2007; Franklin 2010). Resolution is a crucial ecological issue, and consequently changing the grain size can influence the perception of a phenomenon, such as patterns of presence or abundance, or can even affect the relevance of the output for management applications (Araújo and Luoto 2007). In this sense, the use of broad scales in pest risk SDMs is probably related with the fact that biosecurity practitioners have been usually concerned with extrapolation of risks posed by the potential spread of pest, disease and weed species in very large areas. In fact, 83% of the

reviewed articles predicted the distribution of damaging species on regions outside of their native range. Additionally, the scale of the study is commonly constrained by technical issues such as the grid cell size of available environmental data and the characteristics of the species data -e.g. geographical accuracy, sample size, or autocorrelation structure- (Guisan et al. 2007). It must be also noted that the spatial extent over which species distribution models are usually fitted, often coincides with geopolitical boundaries and it is unclear how this pre-set spatial extent affects the performance of distribution models (Niamir et al. 2016). The selection of the extent of the geographical background (the area encompassing the range of the species plus the space accessible to the species which may not be colonized, for a number of reasons), is also likely to influence SDMs performance; larger geographical backgrounds may be more favorable for increased transferability of the models for some invasive species, but may also overfit environmental conditions (Anderson and Raza 2010; Hill et al. 2017). Future work should focus in developing and integrating more precise models by considering: (i) finer ecological and geographical scales (Kriticos et al. 2015); and (ii) making better background choices around sensible climatic extents at broad scales (Hill et al. 2017). For this, it is necessary to use finer resolution climate datasets, as well as other non-climatic factors affecting species potential distribution at small scales (Kriticos et al. 2015).

Regarding the location of the studies, all terrestrial regions of the World, except for the Antarctica, were represented among the reviewed articles. However, we found a geographic bias towards Australia, several counties of Europe, and the USA, which were the most frequently modelled regions. Countries from South Asia, South America and Africa were the less studied (Figure 5). These results, in accordance with other reviews of invasive species (Lowry et al. 2013; Bellard et al. 2018), reflect that the distribution of damaging species have been more frequently modeled in productive systems of developed economies, probably because more research efforts and resources are invested there. Interestingly, Australia was by far the country where species distribution were most frequently modeled. This probably reflects the magnitude of the biological invasions issue in this country (Pyšek and Richardson 2010). Also several SDMs methods, such as CLIMEX (Sutherst and Maywald 1985) and Bioclim (Nix 1986), have been developed in Australia.

### **Data availability and selection of predictive variables**

Nearly all works analyzed used bioclimatic variables to predict the distribution of the studied species, as it

was found in SDMs in general (Bradie and Leung 2017). Bioclimatic variables are commonly derived from monthly temperature and precipitation values in order to generate biologically meaningful variables, including annual values, seasonality, and extreme or limiting environmental factors (Hijmans et al. 2005; Kriticos et al. 2011; Booth et al. 2014). The set of 19 bioclimatic variables was first adopted for BIOCLIM in about 1996 (Booth 2018) and then it was adopted for use in the WorldClim database (Hijmans et al. 2005) and has been widely accepted in SDMs in general. Although it is generally agreed that climatic regimes are critical in determining species distribution, the inclusion of other relevant variables in the models is rarely considered (less than 25% of the papers included non-climatic predictive variables). In this sense, which environmental variables are chosen depends largely on our understanding of the causal mechanisms behind the distribution of the studied species, as well as on data availability (Franklin 2010). Consequently, most studies use only climate, probably because other important causal mechanisms are not as readily quantified or available, and because in most cases insects (whose activity, development and mortality are closely dependent on climate) have been the focus of most works (Kriticos and Leriche 2010).

Our overview reflects how the availability of predictor variables, and particularly large climatic databases, has influenced the development of pest, weed and pathogen SDMs. The first studies that mapped species potential ranges relied directly upon meteorological station data, as typified by the dot maps of early CLIMEX models (Sutherst and Maywald 1985; Sutherst 2014). With the development of high-resolution geospatial databases of average monthly climate data since the 2000s, such as those from the Climatic Research Unit “CRU” (New et al. 2002), WorldClim (Hijmans et al. 2005), and CliMond (Kriticos et al. 2011), the number and resolution of SDMs has grown exponentially. However, it must be noted that species distribution models that rely only on climatic predictors can produce only approximation of distributions at broad scales (Franklin 2010). In the cases where more precise SDMs are needed, it is necessary to include both broad-scale climate variables, and finer-scale predictors that capture terrain-mediated variation in water, energy, and nutrient availability. Further development of high-resolution databases of environmental factors that determine species distribution, such as those limiting factors that control eco-physiology (e.g., eco-climatic variables), as well as disturbance factors (e.g., natural and anthropogenic disturbances), resource factors (e.g., host distribution, landscape configuration) and biotic components such as the presence and abundance of natural enemies, are needed to improve the predictive

power of SDMs (Bradie and Leung 2017). In this sense, considering single species studies is one of the main limitations of SDMs (Hill and Thomson 2015), as species interactions such as competition, predation, parasitism and facilitation are known to influence species distribution. Incorporating biotic interactions in species distribution models has proved challenging, although there is great potential for the development of novel approaches that incorporate multispecies interactions to the projection of species distributions and community structure at large spatial extents (Kissling et al. 2012). This is particularly relevant in terms of biological control applications, since natural enemies (predators and parasitoids) may strongly influence pest and weed distribution; and on the other hand, their distribution may respond to global change in completely different ways than those of the species they attack (Thomson et al. 2010). Consequently, joint distribution models should be performed, to consider the influence of biotic interactions and their success under a changing scenarios.

The availability of species distribution locations is also of great importance for developing accurate SDMs. Compilation of referenced specimen records from natural history collections, regional and national surveys and monitoring, and other sources, into big digital databases has been growing during the last decades. Most of the articles reviewed in this study are based on such sources of information, which have a great potential to contribute to our understanding and prediction of the distributions of harmful species. These databases, however, are known to be affected by strong geographical biases (Beck et al. 2014) and are prone to data errors, related to incomplete or erroneous information at the data gathering level, during the publishing processes, and during the central harvesting and indexing procedures (Gueta and Carmel 2016). For damaging species particularly, there are also biases in reporting due to the status of a species, since some species may not be regarded as pests throughout their distribution range. Consequently, further development of more precise and complete databases, as well as methodologies for handling and correcting biases are required to improve the predictions of pest risk assessment models.

### ***SDMs and the effect of global change***

Our review reflects the fact that a key aim of pest SDMs is to assess the effect of global changes on agricultural, forestry and livestock production systems. Most articles focus on determining the establishment probability of invasive species. The introduction of non-native species is likely the most important global change affecting the health of

crops and other land management activities (Pimentel et al. 2000; Aukema et al. 2011). In consequence, it is reasonable to find that considerable efforts in pest risk assessment have focused on predicting the risk of establishment of invasive species. Currently, more than one-tenth of all declared pests worldwide, have reached more than half of the countries that grow their host crops, a tendency that is expected to increase in the near future (Bebber et al. 2014).

An important part of the studies (about 30%) focused on predicting the effects of climate change on species distribution of pests, weeds, pathogens, and disease vectors. Climate change is expected to cause important shifts in the geographic distribution of species, as a result of a rearrangement of climate zones (Walther et al. 2002). Because invasive pest species are often generalists with broad climatic tolerances, they are generally considered to cope better with climate change than native species, and consequently expand into new areas (Walther et al. 2009; Bebber 2015). Furthermore, it has been found that SDMs of most of the studied pests and pathogens species, predict that their distribution is more likely to increase under climate change scenarios (Bellard et al. 2018).

Only a few articles (less than 5%) considered the effect off other global change scenarios such as land-use change, and these studies were mostly focused on irrigation scenarios. Land use change is strongly affecting the distribution of species, particularly in agricultural and forestry systems. These systems, are some of the most important land use modes, in several regions of the world, and both have changed considerably in the last decades (Rounsevell et al. 2006). Consequently, further work is needed to understand the effects of different land use scenarios on the distribution of damaging species. A cause for the under-representation of SDMs considering the effect of land use change may be that these impacts usually become more apparent than climate at smaller spatial scales (Dormann 2007), although it has been found that human impacted environments are also relevant at broad scales (Hill et al. 2017). It has been suggested that the integration of dynamic climate and land use variables in the same modelling frameworks may avoid producing unrealistic projections of future species distributions (Brook et al. 2008); yet few studies have addressed this issue so far (Schweiger et al. 2011; Martin et al. 2013).

### ***SDMs accuracy and uncertainty***

An important issue in the development of SDMs is the assessment of predictive accuracy and

uncertainty, which helps to determine the suitability of the model for specific applications, and to identify those aspects of the model that need improvement (Vaughan and Ormerod 2005; Barry and Elith 2006). Among the revised articles, 26% did not report model performance analysis, while 41% informed a visual validation of the models, but no quantitative assessment of performance. This reflects the fact that model accuracy evaluation has been ignored or poorly quantified by most SDMs carried out for pests, weeds, pathogens and vectors.

Validation refers to the assessment of the correctness of model predictions using data not used for the building or calibration of the model. Among the studies that evaluated the SDMs, most (52%) used a sub-set from original distribution database, but 47.2% used a completely independent database. The use of a subset of the data for validation of correlative models is intrinsically optimistic compared with external validation, because it only validates the model for data from the same region and time, and consequently, the generality of the model remains unassessed (Dormann et al. 2012). In this sense, whether the evaluation data are derived from partitioning of the original data or else an independent source, it is important that the sample truly represents the environments that are being predicted by the model. This can be accomplished by using a sample that is geographically and environmentally representative of the study area (Franklin 2010).

Among the articles that quantitatively assess model performance (33%), the area under the receiver operating characteristic curve (AUC) is the most frequently used measure (81% of these articles). The AUC of a model is equivalent to the probability that the model will rank a randomly chosen presence site higher than a randomly chosen absence site. Although the AUC has been considered a highly effective measure for the performance of ordinal score models (McPherson et al. 2004), there are several recognized features that prevent its use as a measure of SDMs accuracy (Lobo et al. 2008). Some of the most important problems of this method, in the context of pest risk assessment, are that it does not provide information about the spatial distribution of model errors, and that species distribution data used for the construction of the model are referred to a specific geographical extent, and also that increasing the geographical extent outside the presence environmental domain, entails obtaining higher AUC scores (Lobo et al. 2008). Other less frequently used methods are sensitivity scores (2,8% of the articles), a basic index that assesses the probability that the model correctly predicts an observation for a species at a given site; the Cohen's kappa (0,9% of the articles), which is a



popular measure for the accuracy of presence–absence predictions; and the true skill statistic (TSS, 9.3% of the articles), which is a measure for the performance of presence–absence distribution models that corrects for this dependency while still keeping all of the advantages of kappa (Liu et al. 2011). The TSS is gaining acceptance over the AUC, however, evaluations of a model accuracy solely based on this statistic may also be misleading. The main problem of evaluating the performance of SDMs for species that are changing their distribution range is that the potential distribution in the invaded range or under climate change is unknown, and test data are not available. In this sense, it is unknown whether the model fitted under current environmental conditions is relevant to the future range, and distributional data in the invaded range are unlikely to provide a reliable test of model performance because the species is likely to be invading (Elith 2017). Consequently, models (both correlative and mechanistic ones) need to be assessed for their ecological relevance, by using expert knowledge, by sourcing additional data including experimentally derived physiological information, by sampling within the predicted ranges, or by comparison with completely independent models that do not use distributional records (Hill et al. 2013; Elith 2017). Although evaluation of the SDMs have almost only focused on predictive performance as the most important measure of model validity, predictive performance is only one aspect of model evaluation, while ecological realism and credibility are also important evaluation criteria that should be considered (Franklin 2010).

It must be noted that other methodological limitations have arisen from using correlative methods to model the distribution of species outside of the current known geographic extent or under climate change scenarios. As we have shown, most of the SDMs were developed from the species current distribution, and were extrapolated in space and/or time to forecast potential changes in distribution of species. However, none of the commonly used correlative SDMs were originally designed to model global change effects. Two assumptions of correlative SDMs do not hold, when modelling invasive species or the effect of changes in environmental conditions: (i) species are not in equilibrium with their environment, and (ii) niche quantification is transferable in space and time (Elith and Leathwick 2009; Gallien et al. 2012). Additionally, most of the articles we analyzed here (59%), have used species occurrences from the invaded range to construct the model (Figure 5). However, distribution data outside of the native range are only reliable in regions that have been occupied long enough for the species to

have had the opportunity to persist or else to die out at given locations, as well as to spread and occupy the suitable areas in the considered region (Hill and Thomson 2015; Elith 2017). If this does not happen, the assumptions of the species being in equilibrium with the environment is also being violated in this cases.

Invasive species are able to expand into parts of their fundamental niche that are not available in their native range. The same occurs with native and non-native species under the effect of climate or land use change. Consequently, methods best suited for modelling the potential distribution of these species are those that directly estimate the fundamental niche (mechanistic methods), although they usually overestimate the final distribution (but for models that consider biotic restrictions; (Elith 2017). Semi-mechanistic methods (i.e., CLIMEX), are also considered more adequate for this goal, because they have been specifically developed for invasive species and have some features that make them safer to use in this context (e.g., the way their indices can be controlled to extrapolate beyond the realized niche) (Elith 2017). Novel frameworks have been also proposed to produce more reliable predictions of the distribution of an invasive species outside of their native range, based on correlative models (Gallien et al. 2012), however very few studies have addressed this issue directly.

In sum, there is need to further develop methods to better describe and quantify the uncertainty, both statistically and spatially, of correlative SDMs that attempt to predict the probabilities of establishment of species in new regions or in areas under changing environmental conditions. These approaches should therefore consider the fact this new or changing regions will combine environmental conditions that do not exist in their current distribution range.

### ***Comparing SDMs characteristics according to their goals***

Spatial Distribution Models have been developed with very different goals. Two main groups of objectives may be identified: pest management and biosecurity on the one hand, and biogeography and biodiversity conservation on the other (Franklin 2010; Sutherst 2014). The first group, the one reviewed in the present work, is concerned with the extrapolation of risks posed by the potential spread of pest, disease and weed species and their biocontrol agents. SDMs for biogeography and biodiversity conservation, in contrast, are mainly used to describe patterns of species distribution and identify the environmental requirements of given species (Franklin 2010; Sutherst 2014).



**Table 2.** Main findings for each of the analyzed issues among the revised SDMs, and suggestions of future directions.

Issue	Main findings	Future directions
Studied organisms	✓ Insect pests were by far the most frequently studied organisms.	✓ Develop more precise databases of distribution and/or eco-physiological variables for poorly studied taxa, like pathogens and diseases, to be able to better assess their potential distribution.
SDMs approaches	✓ CLIMEX and MaxEnt (semi-mechanistic and correlative approaches) were the most commonly used modelling tools, while pure mechanistic approaches were rarely applied	✓ Further development of models that join correlative and mechanistic approaches in a combined workflow, to generate hybrid models that incorporate the advantages of each approach.
Extent and location of the studies	✓ Most studies covered broad scales (national and global). Studies in developed countries were more frequent than in developing ones.	✓ Develop more precise models by considering finer ecological and geographical scales; as well as making better background choices around sensible climatic extents at broad scales. ✓ Further study the potential distribution of pests of developing countries.
Data availability and predictive variables	✓ Most works used only bioclimatic variables to predict the distribution of the studied species. ✓ Species interactions, which are particular relevant for biological control applications, were very poorly considered. ✓ Species distribution databases are usually incomplete and are affected by geographical biases.	✓ Further development of high-resolution databases and methods to incorporate more complex environmental factors that determine species distribution, such as disturbance factors, resource factors, and biotic interactions. ✓ Further development of more precise and complete species distribution databases, as well as methodologies for handling and correcting biases.
Effect of global change	✓ Most studies focused on predicting the distribution of invasive species and/or the effects of climate change.	✓ Further study the effects of less represented global change factors, such as different land use scenarios, on the distribution of damaging species.
Accuracy and uncertainty	✓ Most of the studies did not quantitatively assess the performance of SDMs. ✓ AUC was the most frequently used measure, however there are several features that prevent its use as a measure of SDMs accuracy for species expanding their distribution.	✓ Develop methods to better describe and quantify the uncertainty, as well as the ecological relevance of SDMs that attempt to predict the probabilities of establishment of species under the combination of environmental conditions that do not exist in their current distribution range.

One of the main differences we found among these two group of modelling goals, is the most commonly studied taxonomic groups. While, pest management and biosecurity SDMs are mainly focused on insects, biogeographic or conservation studies are usually focused on other taxa such as birds, reptiles, amphibians and plants (Porfirio et al. 2014; Razgour et al. 2016; Weber et al. 2017). Conservation and biogeography studies also differ in the most commonly used modelling methods, which are MaxEnt, GARP and GLM (Razgour et al. 2016; Weber et al. 2017), while CLIMEX, the most commonly modelling method used for pest risk assessment is rarely employed for these other aims. Additionally, conservation and biogeography studies usually include larger species pools (Weber et al. 2017), and they usually focus on smaller extents, particularly conservation studies (i.e., regional scale) (Porfirio et al. 2014; Sutherst 2014; Razgour et al. 2016).

### **Integrating SDMs with other bio-surveillance tools**

The use of SDMs, especially in the context of biosecurity, is particularly efficient if integrated with other tools and approaches currently used in bio-

surveillance. An example of this is the use of sentinel trees (Roques et al. 2015), an approach aimed at identifying which species can become forest pests if introduced into areas in which these same trees are cultivated or are native to. In this sense, SDMs applied on such species represent a key step to improve pre-border surveillance activities of a country (Poland and Rassati 2018). SDMs can also help to improve guidelines for visual inspections –regular checks carried out by inspectors at points of entry or goods to a country or region–, by defining which non-native species have higher risk of establishing at a determined location (Eschen et al. 2015). Similarly, the information obtained from SDMs can be integrated with border and post-border surveillance tools, such as baited traps (Poland and Rassati 2018). Traps lured with attractants are commonly used for early detection of non-native species, and can be used in the context of surveillance at points-of-entry for imported commodities and of surveillance in natural areas (Brockerhoff et al. 2006; Wylie et al. 2008). In this context, SDMs can help to optimize trappings by defining specific surveillance needs to capture a certain target species (Poland and Rassati 2018). In summary, SDMs can help us better understand where a given non-native species is most likely to arrive and

establish, allowing to concentrate surveillance efforts in such area.

### Concluding remarks

Increasingly, managers of agriculture and forestry systems, are concerned with predicting the range expansions of important pests, weeds, pathogens and disease vectors. SDMs have shown to be a valuable approach for modelling the distribution of harmful species across different scales, locations and also considering global change scenarios. We found that an exponential growth in the availability of pest SDMs has occurred during the last decade, together with the development of statistical and mathematical tools, and the development of large, reliable and accessible databases. We expect that this tendency will continue, and further development of more precise process-based SDMs, as well as better methods to evaluate and inform their uncertainty (see Table 2 for more detail on main findings and future directions), will play a key role in the future of studies predicting the potential distribution of species that are harmful to food security and land management, and will be crucial for preventing or mitigating their social, economic and ecological impacts, in the context of a rapidly changing World.

### Acknowledgements

We thank two anonymous reviewers for their thoughtful and helpful comments on a previous version of this paper.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

This study was financed by grant from Agencia Nacional de Promoción Científica y Tecnológica, PICT 2016-0705.

### References

- Anderson RP, Raza A. 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *J Biogeogr.* 37:1378–1393.
- Araújo MB, Luoto M. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecol Biogeogr.* 16:743–753.
- Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM, et al. 2011. Economic impacts of non-native forest insects in the continental United States. *PLoS One.* 6: e24587.
- Barry S, Elith J. 2006. Error and uncertainty in habitat models. *J Appl Ecol.* 43:413–423.
- Bebber DP. 2015. Range-expanding pests and pathogens in a warming world. *Ann Rev Phytopathol.* 53:335–356.
- Bebber DP, Holmes T, Gurr SJ. 2014. The global spread of crop pests and pathogens. *Glob Ecol Biogeogr.* 23: 1398–1407.
- Beck J, Böller M, Erhardt A, Schwanghart W. 2014. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol Informat.* 19:10–15.
- Bellard C, Jeschke JM, Leroy B, Mace GM. 2018. Insights from modeling studies on how climate change affects invasive alien species geography. *Ecol Evol.* 8: 5688–5700.
- Booth TH. 2018. Why understanding the pioneering and continuing contributions of BIOCLIM to species distribution modelling is important. *Austral Ecol.* doi: 10.1111/aec12628.
- Booth TH, Nix HA, Busby JR, Hutchinson MF. 2014. Bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Divers Distrib.* 20:1–9.
- Bradie J, Leung B. 2017. A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. *J Biogeogr.* 44:1344–1361.
- Brockerhoff EG, Bain J, Kimberley M, Knížek M. 2006. Interception frequency of exotic bark and ambrosia beetles (Coleoptera: Scolytinae) and relationship with establishment in New Zealand and worldwide. *Can J For Res.* 36:289–298.
- Brockerhoff EG, Liebhold AM. 2017. Ecology of forest insect invasions. *Biol Invasions.* 19:3141–3159.
- Brook BW, Sodhi NS, Bradshaw CJA. 2008. Synergies among extinction drivers under global change. *Trends Ecol Evol.* 23:453–460.
- Buckley LB, Waaser SA, MacLean HJ, Fox R. 2011. Does including physiology improve species distribution model predictions of responses to recent climate change? *Ecology.* 92:2214–2221.
- Chen I-C, Hill JK, Ohlemüller R, Roy DB, Thomas CD. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science.* 333:1024–1026.
- Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. *Front Ecol Environ.* 6:238–246.
- De la Vega G, Corley JC. 2018. *Drosophila suzukii* distribution modelling improves our understanding of pest limit-range. *Int J Pest Manage.* This issue.
- Dormann CF. 2007. Promising the future? Global change projections of species distributions. *Basic Appl Ecol.* 8: 387–397.
- Dormann CF, Schymanski SJ, Cabral J, Chuine I, Graham C, Hartig F, Kearney M, Morin X, Römermann C, Schröder B, Singer A. 2012. Correlation and process in species distribution models: bridging a dichotomy. *J Biogeogr.* 39:2119–2131.
- Dukes JS, Mooney HA. 1999. Does global change increase the success of biological invaders? *Trends Ecol Evol.* 14:135–139.
- Dukes JS, Pontius J, Orwig D, Garnas JR, Rodgers VL, Brazee N, Cooke BJ, Theoharides KA, Strange EE, Harrington R, et al. 2009. Responses of insect pests, pathogens, and invasive plant species to climate change in the forests of northeastern North America: what can we predict? *Can J For Res.* 39:231–248.

- Elith J. 2017. Predicting distributions of invasive species. In: Robinson AP, Walshe T, Burgman MA, Nunn M, editors. *Invasive species: risk assessment and management*. Cambridge (UK): Cambridge University Press; p. 93–129.
- Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. *Ann Rev Ecol Evol Syst.* 40:677–697.
- Eschen R, Rigaux L, Sukovata L, Vettraino AM, Marzano M, Grégoire J-C. 2015. Phytosanitary inspection of woody plants for planting at European Union entry points: a practical enquiry. *Biol Invasions.* 17: 2403–2413.
- Everett RA. 2000. Patterns and pathways of biological invasions. *Trends Ecol Evol.* 15:177–178.
- Eyre D, Baker RHA, Brunel S, Dupin M, Jarošik V, Kriticos DJ, Makowski D, Pergl J, Reynaud P, Robinet C, Worner S. 2012. Rating and mapping the suitability of the climate for pest risk analysis\*. *EPPO Bull.* 42: 48–55.
- Footitt RG, Adler PB. 2017. *Insect biodiversity: science and society*. Oxford (UK): John Wiley & Sons, Ltd.
- Franklin J. 2010. *Mapping species distributions: spatial inference and prediction*. Cambridge (UK): Cambridge University Press.
- Gallien L, Douzet R, Pratte S, Zimmermann NE, Thuiller W. 2012. Invasive species distribution models – how violating the equilibrium assumption can create new insights. *Glob Ecol Biogeogr.* 21:1126–1136.
- Gueta T, Carmel Y. 2016. Quantifying the value of user-level data cleaning for big data: a case study using mammal distribution models. *Ecol Informat.* 34: 139–145.
- Guisan A, Graham CH, Elith J, Huettmann F, null null. 2007. Sensitivity of predictive species distribution models to change in grain size. *Divers Distrib.* 13:332–340.
- Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. *Ecol Model.* 135:147–186.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 25:1965–1978.
- Hill MP, Chown SL, Hoffmann AA. 2013. A predicted niche shift corresponds with increased thermal resistance in an invasive mite, *Halotydeus destructor*. *Glob Ecol Biogeogr.* 22:942–951.
- Hill MP, Gallardo B, Terblanche JS. 2017. A global assessment of climatic niche shifts and human influence in insect invasions. *Glob Ecol Biogeogr.* 26:679–689.
- Hill MP, Thomson LJ. 2015. Species distribution modelling in predicting response to climate change. In: Bjorkman C, Niemela P, editors. *Climate change and insect pests*. Wallingford (UK): CAB International; p. 16–37.
- Ireland K, Kriticos DJ. 2018. Why are plant pathogens under-represented in eco-climatic niche modelling? *Int J Pest Manage.* This issue.
- Jung J, Lee W, Jung S. 2016. Insect distribution in response to climate change based on a model: review of function and use of CLIMEX. *Entomol Res.* 46: 223–235.
- Kearney M, Porter W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecol Lett.* 12:334–350.
- Kissling WD, Dormann CF, Groeneveld J, Hickler T, Kühn I, McNerny GJ, Montoya JM, Römermann C, Schiffers K, Schurr FM, et al. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *J Biogeogr.* 39:2163–2178.
- Kocmánková E, Trnka M, Eitzinger J, Dubrovský M, Štěpánek P, Semerádová D, Balek J, Skalák P, Farda A, Juroch J, et al. 2011. Estimating the impact of climate change on the occurrence of selected pests at a high spatial resolution: a novel approach. *J Agric Sci.* 149: 185–195.
- Kriticos DJ. 2012. Regional climate-matching to estimate current and future sources of biosecurity threats. *Biol Invasions.* 14:1533–1544.
- Kriticos DJ, Brunel S, Ota N, Fried G, Oude Lansink AGJM, Panetta FD, Prasad TVR, Shabbir A, Yaacoby T. 2015. Downscaling pest risk analyses: identifying current and future potentially suitable habitats for *Parthenium hysterophorus* with particular reference to Europe and North Africa. *PLoS One.* 10:e0132807.
- Kriticos DJ, Leriche A. 2010. The effects of climate data precision on fitting and projecting species niche models. *Ecography.* 33:115–127.
- Kriticos DJ, Le Maitre DC, Webber BL. 2013. Essential elements of discourse for advancing the modelling of species' current and potential distributions. *J Biogeogr.* 40:608–611.
- Kriticos DJ, Webber BL, Leriche A, Ota N, Macadam I, Bathols J, Scott JK. 2011. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods Ecol Evol.* 3:53–64.
- Levine JM, D'Antonio CM. 2003. Forecasting biological invasions with increasing international trade. *Conserv Biol.* 17:322–326.
- Liu C, White M, Newell G. 2011. Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography.* 34:232–243.
- Lobo JM, Jiménez-Valverde A, Real R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr.* 17:145–151.
- Lowry E, Emily R, Adam L, Tracy S, Matthew A, Sarah G, Mickley J, Gurevitch J. 2013. Biological invasions: a field synopsis, systematic review, and database of the literature. *Ecol Evol.* 3:182–196.
- Macfadyen S, Kriticos DJ. 2012. Modelling the geographical range of a species with variable life-history. *PLoS One.* 7:1–12.
- Malacrida AR, Gomulski LM, Bonizzoni M, Bertin S, Gasperi G, Guglielmino CR. 2006. Globalization and fruitfly invasion and expansion: the medfly paradigm. *Genetica.* 131:1–9.
- Martin Y, Van Dyck H, Dendoncker N, Titeux N. 2013. Testing instead of assuming the importance of land use change scenarios to model species distributions under climate change. *Glob Ecol Biogeogr.* 22:1204–1216.
- McKinney ML, Lockwood JL. 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends Ecol Evol.* 14:450–453.
- McPherson JM, Jetz W, Rogers DJ. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J Appl Ecol.* 41:811–823.
- Metcalfe RL, Luckmann WH. 1994. *Introduction to insect pest management*. New York (NY): John Wiley & Sons, Ltd.
- Mokany K, Ferrier S. 2011. Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Divers Distrib.* 17:374–380.

- Nenadic O, Greenacre M. 2007. Correspondence analysis in R, with two- and three-dimensional graphics: the ca package. *J Stat Softw.* 20:1–13.
- New M, Lister D, Hulme M. 2002. A high-resolution data set of surface climate over global land areas. *Clim Res.* 21:1–25.
- Niamir A, Skidmore AK, Toxopeus AG, Real R. 2016. Use of taxonomy to delineate spatial extent of atlas data for species distribution models. *Glob Ecol Biogeogr.* 25:227–237.
- Nix HA. 1986. A biogeographic analysis of Australian elapid snake. In: Longmore R, editor. *Atlas of elapid snakes of Australia: Australian flora and fauna series 7.* Canberra (Australia): Bureau of Flora and Fauna; p. 4–15.
- Parsa S, Kondo T, Winotai A. 2012. The cassava mealybug (*Phenacoccus manihoti*) in Asia: first records, potential distribution, and an identification key. *PLoS ONE.* 7:1–11.
- Peterson AT, Papeş M, Soberón J. 2015. Mechanistic and correlative models of ecological niches. *Eur J Ecol.* 1: 28.
- Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography.* 31:161–175.
- Phillips SJ, Dudík M, Schapire RE. 2004. A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning.* Banff, Canada: ACM; p. 8.
- Pimentel D, Lach L, Zuniga R, Morrison D. 2000. Environmental and economic costs of nonindigenous species in the United States. *BioScience.* 50:53–64.
- Poland TM, Rassati D. 2018. Improved biosecurity surveillance of non-native forest insects: a review of current methods. *J Pest Sci.* doi 10.1007/s10340-018-1004-y.
- Porfirio LL, Harris RMB, Lefroy EC, Hugh S, Gould SF, Lee G, Bindoff NL, Mackey B. 2014. Improving the use of species distribution models in conservation planning and management under climate change. *PLoS ONE.* 9: 1–21.
- Purse B V, Golding N. 2015. Tracking the distribution and impacts of diseases with biological records and distribution modelling. *Biol J Linn Soc.* 115:664–677.
- Pyšek P, Richardson DM. 2010. Invasive species, environmental change and management, and health. *Ann Rev Environ Res.* 35:25–55.
- Pyšek P, Richardson DM, Jarošík V. 2006. Who cites who in the invasion zoo: insights from an analysis of the most highly cited papers in invasion ecology. *Preslia.* 78:437–468.
- R-Core-Team. 2017. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. URL <https://www.R-project.org>.
- Rafoss T, Saethre M. 2003. Spatial and temporal distribution of bioclimatic potential for the codling moth and the Colorado potato beetle in Norway: model predictions versus climate and field data from the 1990s. *Agric For Entomol.* 5:75–86.
- Ramsfield TD, Bentz BJ, Faccoli M, Jactel H, Brockerhoff EG. 2016. Forest health in a changing world: effects of globalization and climate change on forest insect and pathogen impacts. *For Int J For Res.* 89:245–252.
- Rassati D, Haack RA, Knížek M, Faccoli M. 2018. National trade can drive range expansion of bark- and wood-boring beetles. *J Econ Entomol.* 111:260–268.
- Razgour O, Rebelo H, Di Febbraro M, Russo D. 2016. Painting maps with bats: species distribution modelling in bat research and conservation. *Hystrix.* 27:1–8.
- Ricciardi A. 2007. Are modern biological invasions an unprecedented form of global change? *Conserv Biol.* 21:329–336.
- Robertson MP, Kriticos DJ, Zachariades C. 2008. Climate matching techniques to narrow the search for biological control agents. *Biol Control.* 46:442–452.
- Roques A, Fan J, Courtial B, Zhang Y, Yart A, Auger-Rozenberg M-A, Denux O, Kenis M, Baker R, Sun J. 2015. Planting sentinel European trees in eastern Asia as a novel method to identify potential insect pest invaders. *PLoS ONE.* 10:1–19.
- Rounsevell MDA, Reginster I, Araújo MB, Carter TR, Dendoncker N, Ewert F, House JI, Kankaanpää S, Leemans R, Metzger MJ, et al. 2006. A coherent set of future land use change scenarios for Europe. *Agric Ecosyst Environ.* 114:57–68.
- Schweiger O, Heikkinen RK, Harpke A, Thomas H, Klotz S, Kudrna O, Kühn I, Pöyry J, Settele J. 2011. Increasing range mismatching of interacting species under global change is related to their ecological characteristics. *Glob Ecol Biogeogr.* 21:88–99.
- Schymanski SJ, Dormann CF, Cabral J, Chuine I, Graham CH, Hartig F, Kearney M, Morin X, Römermann C, Schröder B, Singer A. 2013. Process, correlation and parameter fitting in species distribution models: a response to Kriticos et al. *J Biogeogr.* 40:612–613.
- Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, Pagad S, Pyšek P, van Kleunen M, Winter M, et al. 2018. Global rise in emerging alien species results from increased accessibility of new source pools. *Proc Natl Acad Sci USA.* 115: E2264–E2273.
- Sutherst RW. 2014. Pest species distribution modelling: origins and lessons from history. *Biol Invasions.* 16: 239–256.
- Sutherst RW, Maywald GF. 1985. A computerised system for matching climates in ecology. *Agric Ecosyst Environ.* 13:281–299.
- Thomson LJ, Macfadyen S, Hoffmann AA. 2010. Predicting the effects of climate change on natural enemies of agricultural pests. *Biol Control.* 52:296–306.
- Vaughan IP, Ormerod SJ. 2005. The continuing challenges of testing species distribution models. *J Appl Ecol.* 42:720–730.
- Venette RC, Kriticos DJ, Magarey RD, Koch FH, Baker RHA, Worner SP, Raboteaux NNG, McKenney DW, Dobsberger EJ, Yemshanov D, et al. 2010. Pest risk maps for invasive alien species: a roadmap for improvement. *BioScience.* 60:349–362.
- Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebe TJ, Fromentin J-M, Hoegh-Guldberg O, Bairlein F. 2002. Ecological responses to recent climate change. *Nature.* 416:389.
- Walther G-R, Roques A, Hulme PE, Sykes MT, Pyšek P, Kühn I, Zobel M, Bacher S, Botta-Dukát Z, Bugmann H, et al. 2009. Alien species in a warmer world: risks and opportunities. *Trends Ecol Evol.* 24:686–693.
- Weber MM, Stevens RD, Diniz-Filho JAF, Grelle CEV. 2017. Is there a correlation between abundance and



- environmental suitability derived from ecological niche modelling? A meta-analysis. *Ecography*. 40: 817–828.
- Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. *Divers Distrib*. 14:763–773.
- Wylie FR, Griffiths M, King J. 2008. Development of hazard site surveillance programs for forest invasive species: a case study from Brisbane, Australia. *Aust For*. 71:229–235.