Introduction

As a predominant group of invertebrates in terrestrial ecosystems, ants (Hymenoptera: Formicidae) are ubiquitous and abundant organisms, both in biomass and in species richness (Hölldobler & Wilson, 1990). Currently, about 14,225 valid species of ants are registered worldwide, comprising 347 genera and 16 subfamilies (Antwiki, 2022). Studies on Formicidae biogeography have been published at the levels of subfamily, genus, or species (Mikissa et al., 2016) at local (Delabie et al., 2007), regional (Delabie et al., 1997; Human et al., 1998; Resende et al., 2010), or global (McGlynn, 1999) scales. In recent years, great advances have been made in this kind of study, allowing inferring the history of distribution based on phylogenetic analyzes, such as on Myrmicinae (Ward et al., 2014). A range of taxonomic studies by biogeographical regions (Ladino & Feitosa, 2020), new occurrences, and records (Dias & Lattke, 2019; Fernandes & Delabie, 2019; Franco, et al., 2019), invasive species (Chen & Adams, 2018), and diversity (Koch et al., 2020; Silva et al., 2020) can also be found. In addition, modeling distribution studies on Formicidae have gained space, under the following names and techniques: potential distribution modeling (Murphy & Breed, 2007; Koch et al., 2018), niche modeling (Peterson & Nakazawa, 2007), paleodistribution (Cristiano et al., 2016), and projections of future scenarios (Jung et al., 2017). All of these studies assess the potential distribution of species to infer information of several natures on biodiversity.

Abstract

The state of the art of Formicidae biogeographic studies using distribution modeling tools was reviewed. We aimed to evaluate how and for what purpose such tools were used in ant studies, as well as detecting modeling methods, algorithms, and variables selected for these studies. We analyzed papers published from 2001 to 2021 and focused on predicting invasion risks, conservation, and potential distribution of species. We also considered the mechanistic and correlative approaches, types of algorithms, and environmental variables. We observed that modeling is first used to predict invasion risks before conservation. The correlative approach was the most used, although it does not consider biotic or physiological aspects as the mechanistic approach does. The most used algorithm was Maxent, combining data set of occurrences with climatic variables. Nine studies used combinations of algorithms with consensual models. Research using modeling has been conducted more and more. However, it remains still incipient, mainly regarding conservation, as the current distribution of most of the Formicidae species is not well known. Although not frequently used in ant studies, distribution modeling represents an important approach for research in biogeography, ecology, and related areas. Certain perspectives could be useful, for example, for studying climatic changes, since possible variations in ant distributions, if anticipated, could suggest or guide further investigations or decision-making in public policies.
Such biogeographical reconstructions presented a more diverse range of species in tropical climates regions (Jenkins, 2003; Guénard et al., 2010; Moreau & Bell, 2013). However, there are still many gaps regarding the information on the geographic distribution of ants, mainly due to the heterogeneous spatial distribution of experts on this important biological group and data collection bias all over the terrestrial biomes. Thus, many samplings may have occurred in a given region while other zones have not been duly explored. Unexplored regions may host certain species or genera, yet no sampling may have been done, nor data may have been published on them (Guénard et al., 2010).

Thanks to zoological collections, such as the Formicidae Collection from the Cocoa Research Center (CPDC) (Delabie et al., 2020) and internet databases, such as the Global Biodiversity Information Facility (GBIF – https://www.gbif.org/), Antweb (https://www.antweb.org/), and Antwiki (https://www.antwiki.org/wiki/Welcome_to_AntWiki), myrmecologists have access to a variety of information on species occurrences in regions that are incipiently known. Such records are important to understand the range limits of the genera (Guénard et al., 2010). Predictions about potential species occurrences and possible distribution changes caused by different types of impacts, whether anthropic or climatic, have become common in the last 20 years (Peterson et al., 2018). However, as more and more distribution data are being entered into biodiversity databases and made freely available on the internet (Soberón & Peterson, 2004), researchers need to be confident about the correct identification of the species. Many records have not been correctly identified or georeferenced (Peterson et al., 2011). Furthermore, occurrence data collected with inaccurate information may provide incomplete information on their responses to environmental gradients since they may be spatially and environmentally biased (Lobo et al., 2007; Hortal et al., 2008).

The species distribution modeling (SDM) tool uses species occurrence data in addition to abiotic information to estimate the potential distribution of species (correlative modeling) (Peterson et al., 2011; 2015). When the information is added to physiological data, it is called mechanistic modeling, which aims to understand, through detailed biophysical modeling approaches, the environmental requirements that make up the niche of a species. This allows the development of a model of the environmental conditions under which the species may exist (Kearney & Porter, 2009; Kearney et al., 2010; Peterson et al., 2015). When information about organisms is used together with environmental variables, for example, the total environmental range (set of abiotic conditions with different tolerance rates) is estimated. In this space, a species can survive and reproduce even without ideal biotic conditions (Guisan & Zimmermann, 2000; Roura-Pascual & Suarez, 2008; Elith & Leathwick, 2009).

Currently, species distribution modeling (SDM) and ecological niche modeling (ENM) (Warren, 2012; Peterson & Soberón, 2012) are among the most productive and rising research branches in ecology (Zimmermann et al., 2010), with applications in a variety of other disciplines such as biogeography, evolution, and conservation (Guisan & Thuiller, 2005). Applications are found in historical biogeography studies, such as evolutionary processes, the discovery of unknown species, effects of climate change, disease transmissions, species invasions, and conservation (Guisan & Thuiller, 2005; Peterson et al., 2011). Species distribution predictions based on correlative models can help to understand spatial patterns of biological diversity (Jiménez-Valverde et al., 2008). The presence or absence of a given species in a habitat derives from a range of factors (Pulliam, 2000), which means that biogeographic patterns are not just limited by abiotic factors, such as climatic factors, commonly considered in correlative models. These patterns can be shaped by many other elements, such as biotic factors, geographic barriers, anthropogenic effects, stochastic events, and historical factors, among others (Pulliam, 2000; Soberón, 2007). Therefore, the potential distribution can only be discussed as an ideal scenario in which the species distribution is considered in conjunction with the environment, established by favorable abiotic conditions (Jiménez-Valverde et al., 2008). Mechanistic, that is, process-based SDMs can be integrated for additional advanced predictions (Rougier et al., 2015). For example, the effect of temperature on physiological and demographic processes can be used to test a causal effect of temperature on species distribution (Monahan, 2009).

Some studies have discussed important conceptual and methodological parameters of species distribution models, particularly the need for careful delimitation of the analysis coverage (Soberón & Peterson, 2005; Soberón, 2007; Peterson et al., 2011). The BAM diagram (biotic, abiotic, movement) considers biotic factors, abiotic factors, and movement factors to delimitate the geographic distributions of species (Figure 1).

![BAM Diagram](https://www.antwiki.org/wiki/Welcome_to_AntWiki)

**Fig 1.** The ‘BAM diagram’ adapted from Soberón and Peterson (2005): Area G represents the entire region considered, where the species respond to abiotic (A) and biotic (B) and movement/dispersion (M) factors. The G0 intersection represents the actual distribution area of the species. The G1 intersection represents a region that has both biotic and abiotic conditions suitable for the species, which could potentially be invaded if M conditions change (potential distribution).
M corresponds to the area within the dispersal capabilities of the species in question, corresponding to the geographic regions that were accessible to the species within a certain period (Peterson et al., 2011). The intersection of A ∩ B ∩ M is the occupied distributional area and is the subset of the accessible region where abiotic and biotic conditions allow species to maintain populations (Peterson et al., 2011).

The BAM diagram allows the researcher to focus on delimiting the area to be analyzed. Although one of the main functions of modeling in ecology is to estimate yet undescribed diversity, due to insufficient information, global models of species diversity have seldom focused on insects (Guénard et al., 2012), except for species of medical importance (see Ahadji-Dabla et al., 2020; Moo-Llanas et al., 2020). Most of the studies investigate relatively well-known groups, such as vertebrates (see Freeman et al., 2019) or plants (see López Tirado et al., 2018). However, studies involving ants have become more popular in recent decades. Roura-Pascual & Suarez (2008) reviewed climate modeling studies, highlighting applications with correlative and mechanistic methodologies, specifically in forecasting studies on invasive ants, emphasizing future scenarios. Bertelmeier et al. (2016) evaluated the mechanisms by which climate changes could favor future ant invasions at the regional and global levels, as well as in biodiversity hotspots. Both reviews focused only on studies about invasive ants.

In this review, we present qualitative and quantitative approaches to scientific productions related to Formicidae biogeography that used geographic distribution modeling as a tool from 2001 to 2021. We synthesized a diagnosis of the Formicidae biogeography, its history, and degree of development, as well as the types of modeling and the contributions of these studies to the scientific debate. In addition, we analyzed the contribution of these researches to emerging conservation issues and areas in which further research is necessary.

### Material and methods

The search was conducted on the platforms Google Scholar, GBIF, Scielo, and Portal de Periódicos CAPES, in English, by using the keywords “Formicidae” or “ants”, combined with “modeling”, “niche”, “climate change”, “geographic distribution”, “future scenarios”, “paleodistribution”, “potential distribution”, “paleogeography” “palaeoclimatology”, and “bioclimatic envelope”. We considered the studies published from 2001 to 2021, to offer a descriptive analysis and to give an overview of the specific research conducted in this area, using categories such as aims of the study, mechanistic and correlative approaches, algorithms, and variables used. Thus, we evaluated the use of modeling as a tool in ant biogeography studies, describing its applicability and aiming at projecting possible areas of invasion, occurrence, and conservation. We also described the mechanistic and correlative approaches and identified algorithms and variables. The papers were categorized in five-year periods (2001-2005, 2006-2010, 2011-2015, 2016-2021), and classified according to the objectives of the studies as per table 1. Based on descriptive information, we focused on which aspects and dimensions have been highlighted over the years.

### Results

**Evaluation of the applicability of species distribution modeling as a tool in ant biogeography studies**

Forty-four studies of Formicidae published from 2001 to 2021, which used SDM as a tool, were selected. Among these studies, 48% assessed the potential for invasive species to invade new areas (Figure 2). The invasive ant species most studied were: Solenopsis invicta Buren, 1972, with projections to Oklahoma (Leavia & Frost, 2004), global expansion...
Among the studies which opted for consensual models using only a single software/algorithm (Table 1), Identification of algorithms

only a single software/algorithm (Mayr, 1868), with a global prediction (Roura-Pascual et al., 2004), New Zealand (Hartley & Lester, 2003; Harris & Barker, 2007) and Iberian Peninsula (Roura-Pascual et al., 2006; 2009; Abril et al., 2009).

Studies having as the main focus the inference about areas of possible occurrence, predicting habitat or potential distribution unrelated to invasive species, had a frequency of 23%. Studies that used modeling as a conservation tool had a frequency of 11% and were limited to endemic species: Formica exsecta Nylander, 1846, Palearctic Region (Seifert, 2000), Atta robusta Borgmeier, 1939 exclusively for coastal vegetation (restinga) in southeastern Brazil (Fowler 1995; Teixeira et al., 2003; 2004), Lasius balearicus Talavera, Espadaler and Vila 2014, endemic to the Balearic Islands (Spain) (Talavera et al., 2014), and Dinoponera lucida Emery 1901, restricted to a small portion of the Atlantic Forest in parts of the states of Bahia, Minas Gerais, São Paulo, and Espirito Santo, in Brazil (Mariano et al., 2008; Lenhart, et al., 2013; Escarraga et al., 2017). Other purposes, shown in Fig 2, represented only 3%.

Methodological aspects – Detecting the distribution modeling methods used to predict areas of invasion, occurrence, and conservation

Mechanistic versus correlative approaches

We observed that, since 2006, correlative modeling studies with ants have increased while mechanistic studies have decreased considerably (Fig 3). In addition, more recent studies have used correlative modeling to infer potential areas of occurrence (Souza & Delabie, 2013; Cristiano et al., 2016; Simões-Gomes et al., 2017; Koch et al., 2018; Senula et al., 2019), and predict invasion (Bertelsmeier et al., 2015; Jung et al., 2017; Sung et al., 2018; Byeon et al., 2020). Fig 2 presents the frequency of all-purpose correlative modeling studies with Formicidae. In general, studies on Formicidae aimed at assessing the invasive potential of exotic species (Roura-Pascual et al., 2006; Hartley et al., 2006; Roura-Pascual et al., 2009) used mainly correlative modeling. Conservation studies also used correlative modeling, mostly (Dättilo et al., 2012; Talavera et al., 2014; Campiolo et al., 2015). Only one study used the mechanistic approach (Maggini et al., 2002). On the other hand, inference studies of occurrence, habitat, and areas of potential distribution (Solómon et al., 2008; Souza & Delabie, 2013; Cristiano et al., 2016; Sánchez-Restrepo et al., 2019) used only the correlative methodology.

Identification of algorithms

Twenty-five (i.e., 75.8%) out of 33 correlative approach studies used only a single software/algorithm (Table 1). Among the studies which opted for consensual models using more than a single algorithm, five of them used more than four software/algorithms (15.2% of the total). The most used software/algorithms/analysis were Maxent (used in 15 studies), Generalized Linear Model (GLM) (6), Genetic Algorithm for Rule-set Prediction (GARP) (6), Artificial Neural Networks (ANN), Classification Tree Analysis (CTA), Support Vector Machines (SVMs) (used in four studies each), Classification Trees (CT), CLIMEX software, and Generalized Boosting Model (GBM), three studies each (Table 1). The other 14 studies with ants have increased while mechanistic studies have decreased considerably (Fig 3). In addition, more recent studies have used correlative modeling to infer potential areas of occurrence (Souza & Delabie, 2013; Cristiano et al., 2016; Simões-Gomes et al., 2017; Koch et al., 2018; Senula et al., 2019), and predict invasion (Bertelsmeier et al., 2015; Jung et al., 2017; Sung et al., 2018; Byeon et al., 2020). Fig 2 presents the frequency of all-purpose correlative modeling studies with Formicidae. In general, studies on Formicidae aimed at assessing the invasive potential of exotic species (Roura-Pascual et al., 2006; Hartley et al., 2006; Roura-Pascual et al., 2009) used mainly correlative modeling. Conservation studies also used correlative modeling, mostly (Dättilo et al., 2012; Talavera et al., 2014; Campiolo et al., 2015). Only one study used the mechanistic approach (Maggini et al., 2002). On the other hand, inference studies of occurrence, habitat, and areas of potential distribution (Solómon et al., 2008; Souza & Delabie, 2013; Cristiano et al., 2016; Simões-Gomes et al., 2017; Koch et al., 2018; Sánchez-Restrepo et al., 2019; Senula et al., 2019), as well as in two conservation studies (Dättilo et al., 2012; Talavera et al., 2014). The GARP algorithm (Stockwell & Peters, 1999) has been used in some studies mainly to assess the risk of the ant L. humile invasion (Roura-Pascual et al., 2004; 2006; 2009), and was used once in a conservation study (Campiolo et al., 2015), and another time to identify species richness patterns (Chaladze, 2012). GLM was chosen for invasion risk assessment (Hartley et al., 2006; Abril, et al., 2009; Roura-Pascual et al., 2009) and conservation studies (Maggini et al., 2002; Talavera et al., 2014).

We consider that the studies classified here as inferences of area occurrence/ habitat/ potential distribution, without a conservative approach or risk of invasion assessment, aimed to fill in the gaps regarding species distribution. Most of these
Table 1. Main analyzes and software used according to the aims of correlative modeling studies for ants. The codes representing each objective are presented in parentheses.

<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Finality</th>
<th>ANN</th>
<th>CT</th>
<th>CTA</th>
<th>CLIMEX</th>
<th>GARP</th>
<th>GAM</th>
<th>GLM</th>
<th>Maxent</th>
<th>SVM</th>
<th>Other(s)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Leavia &amp; Frost</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Roura-Pascual et al.</td>
<td>Predict invasion (PI)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Roura-Pascual et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2006</td>
<td>Hartley et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2007</td>
<td>Harris &amp; Barker</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Murphy &amp; Breed</td>
<td>Niche requirements (NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Peterson et al.</td>
<td>Evaluate flaws in the modeling method (EF)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ward</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>Solómon et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Steiner et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Roura-Pascual et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (1)</td>
<td>5</td>
</tr>
<tr>
<td>2011</td>
<td>Jenkins et al.</td>
<td>Global density prediction (DP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>2012</td>
<td>Chaladze</td>
<td>Identify spatial pattern of species richness (SP)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Dättlto et al.</td>
<td>Conservation (C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Fitzgerald et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2013</td>
<td>Bertelsmeier et al.</td>
<td>Predict invasion (PI)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (1)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Souza &amp; Delabie</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2014</td>
<td>Bertelsmeier &amp; Courchamp</td>
<td>Predict invasion (PI)</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (4)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Talavera et al.</td>
<td>Conservation (C)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2015</td>
<td>Bertelsmeier et al.</td>
<td>Predict invasion (PI)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (1)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Campiolo et al.</td>
<td>Conservation (C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Kumar et al.</td>
<td>Revision (R)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2016</td>
<td>Cristiano et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2017</td>
<td>Simões-Gomes et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (3)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Jung et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2018</td>
<td>Kock et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sung et al.</td>
<td>Predict invasion (PI)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X (3)</td>
<td>7</td>
</tr>
<tr>
<td>2019</td>
<td>Coulin et al.</td>
<td>Predict invasion (PI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sánchez-Restrepo et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Senula et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2020</td>
<td>Ströer et al.</td>
<td>Investigate phylogeographic patterns (PP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2021</td>
<td>Byeon et al.</td>
<td>Infer areas of occurrence/habitat/potential distribution (O/H/PD)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Jung et al.</td>
<td>Develop predictive model of soil temperature (ST)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>4</td>
<td>7</td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>
studies used the Maxent algorithm. Only one of them (Byeon et al., 2020) used the CLIMEX algorithm, and two generated consensual models from other algorithms, besides Maxent, BIOCLIM, DOMAIN, Generalized Linear Model (GLM), Support Vector Machines (SVM), and Boosted Regression Trees (BRT) (Simões-Gomes et al., 2017), and GARP (Koch et al., 2018). Nine studies used more than a single algorithm, combining the different results, such as in Talavera et al. (2014) who used eight consensual models (Generalized Linear Model – GLM, GBM, Generalized Additive Model – GAM, Flexible Discriminant Analysis – FDA, Multiple Adaptive Regression Splines – MARS, Classification Tree Analysis – CTA, Random Forest – RF, Maxent).

Variables and projections

Of the 44 studies evaluated, 23 used only climate variables (see for example Chaladze, 2012; Diamond et al., 2012; Koch et al., 2018). The remaining (21) combined climatic variables with a set of further information, such as vegetation, topography, and soil temperature (Peterson & Nakazawa, 2007; Roura-Pascoal et al., 2009; Jenkins et al., 2011; Fitzgerald et al., 2012; Senula et al., 2019; Jung et al., 2021). Climate bias, studies of potential distribution, and scenarios for the future were the majority, while 13 studies estimated the current potential distribution. Seven studies inferred further climate change. A total of 10 studies merged current and future potential distributions, two merged paleodistribution and current potential distribution, one study examined transitions between past, present, and future, and, finally, one study produced only paleodistribution models, comparing with Phylogeography (Table 1).

Discussion

Evaluation of the applicability of species distribution modeling as a tool in ant biogeography studies

Modeling aims to infer the best explanation for a data set and to represent it in a precise and compact way, emphasizing probabilities of where the species may or may not be present (Guisan & Thuiller, 2005). It is a tool that has been widely used in studies that analyze the potential distribution of species, including the evaluation of the impact of global climate change on species distribution, areas susceptible to invasive species, the selection of suitable habitats, and species conservation, besides prioritizing suitable areas for conservation (Guisan & Zimmermann, 2000; Siqueira & Peterson, 2003; Marini et al., 2009).

Approximately half of the studies on Formicidae that used distribution modeling were conducted with invasive species (Ward, 2007; Bertelsmeier et al., 2013; Jung et al., 2017; Sung et al., 2018). Ants are considered a group of organisms that settle easily outside their native distribution area due to their small size, large number of individuals, and being colonial (Bertelsmeier et al., 2013). Thus, anticipating changes in the distribution of invasive species would minimize their impacts. They can be transported by accident, for example, on fruits, ornamental plants, tourism, and trade exchanges, or with agricultural tools (Lofgren et al., 1975; Bertelsmeier & Courchamp, 2014).

A good example is the Argentine ant L. humile, the most studied invasive ant with modeling purposes (Roura-Pascual et al., 2004; 2006; 2009; Hartley et al., 2006; Harris & Barker, 2007; Abril et al., 2009; Fitzgerald et al., 2012; Bertelsmeier et al., 2015). Roura-Pascual et al. (2006) compared native and invasive ecological niches of L. humile, while Hartley et al. (2006) evaluated the uncertainty in the predictions of bioclimatic range, corroborating the potential distribution with the known distribution of L. humile, pointing out the important role of temperature and precipitation in the establishment of this species. Another study evaluated the probable risk of invasion by twelve ant species in New Zealand (Harris & Barker, 2007). These authors deduced that the chances of the temporary establishment of colonies of species such as Solenopsis geminata (Fabricius, 1804), and Anoplolepis gracilipes (Smith, 1857) could be ignored, as temperatures in New Zealand are lower than they can survive (Harris & Barker, 2007).

Other invasive species widely studied were S. invicta 1972 (Levia & Frost, 2004; Bertelsmeier et al., 2015; Sung et al., 2018) and A. gracilipes (Bertelsmeier & Courchamp, 2014; Bertelsmeier et al., 2015; Jung et al., 2017). Native from South America (Vinson & Sorensen, 1986), S. invicta was first introduced and spread in the southern part of the USA and the Caribbean (Morrison et al., 2004), and then was dispersed throughout China, Taiwan, Australia, and Mexico (Valles et al., 2015). In addition, large areas in Mexico, northern South, and Central America, the Caribbean islands, part of the Mediterranean region, as well as some areas close to the Black and Caspian Seas, are at high risk of invasion by this ant (Morrison et al., 2004).

There is no consensus on A. gracilipes original native distribution (Vásquez-Bolaños & Wetterer, 2021). It may have originated in Asia or Africa (Holway et al., 2002; Wetterer, 2005). It is a species that can propagate in humid tropical areas (Wetterer, 2005). As such, it was introduced into regions of Africa, including South Africa and Tanzania; Central, and South America, tropical Asia, and Australia (Wetterer, 2005). It is a quarantine pest in the United States and the Republic of Korea but is not considered as invasive in North America (Csrhess & Hankamer, 2012). Recently, the risk of invasion was analyzed for A. gracilipes and S. invicta (Jung et al., 2017; Sung et al., 2018, respectively). Both studies modeled the potential distribution under current and future scenarios for South Korea and found favorable climatic conditions for these invasive species.

The introduction of A. gracilipes occurred in South Korea through trade routes and S. invicta, although still not found in the country, is a cause of major concern there (Jung et al., 2017; Sung et al., 2018). Countries that are on international trade ways which connect countries from the
The projection of the potential distribution is then justified to Pacific Ocean to the Asian continent, such as South Korea, are
direct collection efforts and define priority areas for
that the generated models can be used to choose areas where
Bahia (Dáttilo et al., 2012). Therefore, the authors suggest
results showing the probability of occurrence (varying from
distribution and climatic variables, the model generated

Although scarce, studies to infer areas of occurrence or
distribution are important tools for research on
Foremiciidae (Solómon et al., 2008; Souza & Delabie, 2013;
Senula et al, 2019; Byeon et al., 2020). Souza and Delabie
(2013) suggested that the occurrence suitability may be an
important and useful parameter in investigating the
biogeography of rare ant taxa. Furthermore, the applicability
of ecological niche models in generating information on the
geographic distribution of pests, providing useful tools for
integrated management, has been described for both genera
Atta Fabricius, 1804 and Acromyrmex Mayr, 1865 (Sánchez-
Restrepo et al., 2019). On the other hand, studies that used
modeling as a tool for the purpose of conservation (Maggini
et al., 2002; Dättilo et al., 2012; Campiolo et al., 2015) are
still incipient. Among these few studies, modeling of the
potential habitat distribution of F. exsecta was conducted in
a conservation area in Switzerland (Maggini et al., 2002), in
order to understand possible reasons for the local distribution
of this species, as it was threatened by extinction. Using
Generalized Linear Models (GLM) (Guisan et al., 2002),
initial local models generated from 160 field samples were
extrapolated to the country scale through a Geographic
Information System (GIS).

The purpose of studies with maps of potential
distribution is, in short, to indicate interesting areas for
further sampling, possible areas for future colonization,
or areas previously occupied, from where the species has disappeared (Maggini et al., 2002). Thus, adjusting a model
with data sampled in a conservation area is an interesting
approach, since it allows for meeting the accurate ecological
requirements of the species (Maggini et al., 2002). Dättilo et
al. (2012) modeled the potential distribution of the leaf-cutting
ant A. robusta, an endemic species with occurrence records
only in restingas in Rio de Janeiro and Espírito Santo, Brazil
(Teixeira et al., 2003). Thus, with a restricted geographical
distribution and climatic variables, the model generated
results showing the probability of occurrence (varying from
low to high) in regions that do not have occurrence records,
such as in the states of Paraná and São Paulo, and in southern
Bahia (Dättilo et al., 2012). Therefore, the authors suggest
that the generated models can be used to choose areas where
to direct collection efforts and define priority areas for
conservation. This was the first study using modeling tools
in Brazil to predict the distribution of an ant species. Also,
the study demonstrated a probable gradient in the probability
of occurrence from the coast towards inland, corroborating
the endemcity characteristics of A. robusta in restinga (Dättilo
et al., 2012).

Another conservation-focused study associated phyl-
genetic inference with modeling the distribution of the endemic
ant L. balearicus in the Spanish Mediterranean, estimating
current and future potential distribution (Talavera et al.,
2014). The latter is based on coupled global climate models
CGCM2 and CGCM3 created by the Canadian Center for
Climate Modeling and Analysis (CCCMA) under three varied
carbon emission scenarios for the years 2050 and 2080. The
study found that the potential of L. balearicus to deal with
climatic change by varying its climate niche is low, added to
the impossibility of dispersion due to its insular situation and
altitude isolation, suggesting that L. balearicus is at risk of
extinction in the short term and the inclusion of the species in
the IUCN Red List of Threatened Species.

Another ant species that was included in the Red Book
of Threatened Brazilian Fauna and that was studied with the
use of modeling was D. lucida. In this study, Campiolo et al.
(2015) used climatic variables to predict past, present, and
future areas of suitability. The generated models demonstrated
that the earlier suitability areas were larger than the current
ones and that those areas would be reduced in the year 2050
due to climate change.

Modeling is a tool that can help in assessing the
distribution of little-known species. However, regarding the
applicability, we observed that modeling studies on Formicidae
have a greater focus on risk assessment in invasive species. The
three most species are L. humile, S. invicta, and A. gracilipes.
This focus on invasive species is due to the substantial
 damages caused by these ants to biodiversity, economy,
and human health that follow invasions (Bertelsmeier et
al., 2015). In fact, ants should be better controlled during
the early stages of the invasion, when the population is still
relatively small and geographically limited (Bertelsmeier
& Courchamp, 2014). Thus, we understand the importance
of these predictions for assessing areas at risk of invasion,
anticipating variations in these distributions, with a view
to the conservation of native species in the studied areas.
Since the financial and environmental costs of a prediction
may prove to be wrong for an invasive species, precaution is
probably the best policy (Harris & Barker, 2007). On the other
hand, studying ant species with little-known distribution, in
addition to species with unknown conservation status is also
important, and could benefit from using the modeling tool.

Methodological aspects – Detecting the distribution modeling
methods used to predict areas of invasion, occurrence, and
conservation.

Mechanistic versus correlative approaches

Depending on the purpose of the model, species
distribution modeling can be conducted according to two
approaches: mechanistic or correlative. In the mechanistic approach, the morphological, physiological, and behavioral requirements are obtained experimentally (Kearney, 2006), and, then, linked to the environmental variables to estimate the geographic distribution of the species (Kearney & Porter 2004; 2009). The formulation of conceptual models based on physiological processes is required. Their predictions are assessed through the theoretical rigor that cause-and-effect relationships are addressed (Guisan & Zimmerman, 2000). Although it is a promising approach, it requires detailed knowledge about the biological aspects (fitness) of species with the environment (Kearney 2006; Buckley et al., 2010).

Following a mechanistic approach, Korzukhin et al. (2001) and Sutherst and Maywald (2005) on a regional scale, and Morrison et al. (2004) on a global scale, estimated the expansion of the potential range of the invasive *S. invicta*, using dynamic models, and the ecophysiological aspect of colony growth, evaluating soil temperature. This species is currently spread over much of the southern United States, besides its natural distribution in South America. Studies predicted its expansion to the North (Korzukhin et al., 2001; Sutherst & Maywald, 2005), and large portions of Europe, Asia, Africa, and Australia (Morrison et al., 2004), as the adaptation of their populations to cooler or drier climates could increase the area of their potential range (Morrison et al., 2004). A mechanistic study, which measured air and soil temperature, was conducted with the Argentine ant *L. humile* (Hartley & Leter, 2003), suggesting locations in New Zealand that meet the appropriate conditions for the species. In another mechanistic study, Diamond et al. (2012) observed that the vast majority of ant genera are within the region where their heating tolerance is lower. Such mechanistic models incorporate relations between environmental conditions and the organisms’ performance (Buckley et al., 2010). Although there are still problems related to the adequacy of the limits established experimentally in view of the reality of species in nature, the mechanistic approach brings a better understanding of the factors that determine the patterns of species distribution at large spatial scales (Deutsch et al., 2008; Hofmann & Todgham, 2009). Models based on a mechanistic approach allow a more direct view of the fundamental ecological niche since they can be developed regardless of access restrictions or biotic environments (Peterson et al., 2018). The low number of studies with mechanistic models is possibly because they are more laborious to generate than correlative models. In addition, they require the collection of a lot of physiological data, which may not be available.

In contrast, according to the correlative approach, the environmental conditions of the species are estimated by the spatial superposition between occurrences and environmental variables (Elith & Leathwick, 2009). Koch et al. (2018), for example, modeled the potential distribution of *Gracillidris pombero* Wild and Cuezzo (2006), in order to identify probable areas of occurrences. In South Korea, a country that presents a high risk of biotic invasion, since it connects countries in the Pacific Region with Asia for international trade, two potential invasive species, namely *A. gracilipes* and *S. invicta*, were studied by applying species distribution models, (Jung et al., 2017; Sung et al., 2018). Most studies on Formicidae aiming to predict suitable areas for invasive species (Roura-Pascual et al., 2006; Hartley et al., 2006; Roura-Pascual et al., 2009) used correlative modeling, considering only abiotic variables and omitting the effects that other species may have on their distribution. The same happens with conservation studies with correlative modeling (Dátillo et al., 2012; Talavera et al., 2014; Campiolo et al., 2015), in which only one used the mechanistic approach (Maggini et al., 2002). On the other hand, all studies of inference of occurrence, habitat, and potential distribution areas (Solómon et al., 2008; Souza & Delabie, 2013; Cristiano et al., 2016; Sánchez-Restrepo et al., 2019) used the correlative methodology.

Studying *L. humile*, Roura-Pascual et al. (2009) calibrated the models differently from previous studies (Roura-Pascual et al., 2006; Hartley et al., 2006), distinguishing native and invaded areas. Thus, they focused on invaded area records, using consensual models, which are important in this case, since such records may overestimate or underestimate the potential range of occupancy by the ant. They also suggested that future studies should pay special attention to areas of maximum uncertainty between the different models, aiming to elucidate the determinants of species distribution. Bertelsmeier and Courchamp (2014) and Bertelsmeier et al. (2015) also used correlative modeling to infer the potential distribution of invasive species, combining several kinds of predictions in a single consensus model.

Correlative modeling is the most used approach among the studies since it does not make use of prior knowledge of the fundamental niche of species (Kearney, 2006). In addition, there is a range of occurrence information in databases, such as SpeciesLink, GBIF, and Antweb, which favors the application of this modeling method for ants (Pearson, 2010). Even so, when it comes to insects, for most taxonomic groups, the geographic distributions are poorly known and have many gaps, due to the so-called Wallacean deficit (Bini et al., 2006). Thus, in many cases, the choice of method can be restricted due to the lack of verified occurrences.

In the correlative approach, the model is empirical. In other words, it does not have the attribution of a cause-effect relationship (Guisan & Zimmerman, 2000). There are limitations and criticisms regarding this model since physiological variables are not applied and possible biotic interactions are not considered (Dormann, 2007; Kearney et al., 2010; Buckley et al., 2010). Therefore, exploring the physiological mechanisms that establish geographic occurrence is not possible as it is in the mechanistic approach. Although it presents only a statistical approximation of reality (Guisan & Zimmermann, 2000), the mechanistic approach contributes to the formulation of new hypotheses about the mechanisms.
that determine the distribution of species, inferring areas of distribution with a greater (or smaller) degree of environmental suitability (Guisan & Zimmermann, 2000).

Identification of algorithms

Currently, a variety of species distribution modeling techniques are available. Good knowledge of the performance of these techniques becomes extremely important to help researchers to select the most appropriate approach for their particular purposes (Jiménez-Valverde et al., 2008). Modeling studies should first test a set of algorithms, regarding their predictive capacity. Studies that do not take this first step may use inappropriate algorithms (Qiao et al., 2015).

The algorithms used by the correlative models aimed to establish non-random relationships between species occurrence data and data on relevant environmental variables. In essence, the methods extrapolate associations between occurrence points and the set of environmental data to identify predicted areas of occurrence using maps (Pereira & Siqueira, 2007). The choice of an algorithm must be based on the availability of occurrence data (number of records and presence/absence data), as well as based on the study question. When only presence data exists, algorithms such as DOMAIN and BIOCLIM can be used (see Ward, 2007). When the data available are presence/absence, distribution modeling can be performed using statistical methods (Guisan & Zimmermann, 2000), such as the GLM (Jenkis et al., 2011; Abril et al., 2019) and the GAM (Talavera et al., 2014; Sung et al., 2018). However, absence data are difficult to verify and do not always reflect the true absence of the species at that location. The absence of the species may be due to poor sampling, unavailable records, low detectability of the method, or the impossibility of dispersing the species to the site, among other factors (Peterson et al., 2011). Therefore, the absence of data should be used with caution, as they may underestimate the occurrence of the species.

The algorithms that stand out in the modeling studies in Formicidae are firstly Maxent and secondly GARP. These algorithms show good results with a low number of occurrence points (Wisz et al., 2008). Both fall into an intermediate category with respect to occurrence points, as they use presence and pseudo-absence (background) data to generate the SDMs (see Stockwell & Peters, 1999; Phillips et al. 2006). Maxent (Phillips et al., 2006) is based on the principle of maximum entropy. This method generates predictions from incomplete information regarding the target distribution. Overall, Maxent has outperformed other modeling methods, hence it is more popular than others (Elith et al., 2006; Wisz et al., 2008).

GARP, the second most used method, makes use of a genetic algorithm to search for non-random associations between environmental variables and known occurrences, in contrast to the environmental characteristics of the general study area (Roura-Pascual et al., 2004; Peterson & Nakazawa, 2007; Campiolo et al., 2015). Peterson and Nakazawa (2007) used GARP to model the potential distribution of S. invicta and Solenopsis richteri Forel 1909. The authors depicted the effects of different environmental data sets on the model quality. However, they emphasized that using GARP was only a methodological option and similar results were found in preliminary tests using Maxent (Phillips et al., 2006).

Roura-Pascual et al. (2009) also used correlative modeling with the selection of consensual areas, but with five different modeling techniques: GLM, GAM, GBM, GARP, and Maxent. Given the results, a lack of geographic congruence between predictions from different approaches is evident. They are also divergent about the usefulness of group predictions in identifying areas of uncertainty on the potential invasiveness of some species. Ward (2007) used three algorithms: DOMAIN, BIOCLIM, and Maxent, to model the virtual distribution of invasive ants in New Zealand. The study found that among the six modeled species, BIOCLIM performance was worse than the other two modeling methods. The consensual areas indicate the environmental areas in which all models offer the conditions allowing the species occurrence. Different consensus methods are currently available, such as PCA, which focuses on different algorithms or environmental layers (Araújo et al., 2006), weighted averages of the results obtained through the accuracy values (AUC, TSS or Kappa) (Thuiller et al., 2009), the combination of the resulting maps (Diniz-Filho et al., 2009), etc. The most used technique is ensemble forecasting, which consists of generating a consensual model from the results of different algorithms or different scenarios (Araújo & New, 2007). The understanding of the use of consensual models is based on the criteria that single predictions are not reliable, as well as that the whole models are incomplete at some point, although they carry useful information (Araújo et al., 2005; Araújo & New, 2007). The practice of combining results should not be an alternative to the traditional approach to building ever more accurate models (Araújo & New, 2007). However, combining results from different models still depends on individual predictions, although it may improve the quality of multiple predictions. Therefore, if better individual predictions can be reached, a more confident consensus may occur (Araújo et al., 2005).

Currently, there is a wide discussion in the literature regarding the factors that can affect the ability to develop robust predictive models (Boria et al., 2016; Peterson et al., 2018). Two of these factors about ecological niche models used to anticipate possible distribution patterns are climate uncertainty and algorithmic uncertainty. The former concerns the current existing GCMs, as they do not capture all future details. The latter covers many comparative studies, which combine results from algorithms with consensual models, to identify the more appropriate projections. However, there are criticisms regarding consensual models since the different
algorithms evaluate different parameters. Thus, the consensus should be designed with a single algorithm, when the aim is to assess the variability after projecting the provisions for different periods with the same parameters (Boria et al., 2016). Possibly, the development of integrative models, which aggregate climatic, ecological, and evolutionary variables, would allow more accurate inferences about responses and suitability of species to climatic fluctuations.

**Variables and projections**

Although modeling studies can be developed with variables related to climate, such as soil, vegetation, and topography, among others, in recent years significant advances have occurred in methods that seek to estimate changes in the distribution of species considering climate changes (Franklin, 2010). These advances also happened in Formicidae studies.

One of the central demands of modern ecology is to understand how current climate changes will affect species, and efforts have been made to predict and mitigate such effects (Araújo et al., 2004; Araújo & Rahbek, 2006). The growing availability of environmental variables that report to the past (e.g., 6,000, 21,000 years ago), as well as to the future (e.g., in the years 2050, 2070), allows us to perform temporal transfers, and predict spatial responses of organisms to climate change (Faleiro et al., 2013). An important research field in ant studies aims at predicting possible distributions for species under altered climatic conditions (Peterson et al., 2018). Such ‘bioclimatic’ variables are provided from databases such as WorldClim (Hijmans et al., 2005) and Chelsa (Karger et al., 2017). These databases provide free access to high spatial resolution global climate data, including climate information for various time periods, ranging from paleoclimatic periods to current or future scenarios. The climatic data result from global circulation models (GCMs), which represent scenarios of emission levels, and greenhouse gas trajectories for the climate in the future, the Representative Concentration Pathways (RCPs) (Van Vuuren et al., 2011; Aguilíar et al., 2015).

Computational tools that include bioclimatic models have been developed to define the relations (also called correlative models) between data on the confirmed occurrence of species and their spatial variation in different environmental conditions (Guisan & Thuiller, 2005). These correlative models establish a relationship between species occurrence and climatic variables in space and time in order to redesign the species’ geographic distribution after climatic changes, based on the assumption that the species will stay in equilibrium with the environment (Pearson & Dawson, 2003; Hartley et al., 2010). Leiva and Frost (2004) assessed the climatic suitability for the expansion of *S. invicta* Buren in Oklahoma under the current climate and with the duplication of atmospheric CO₂ using three general circulation models (GCMs) (GFDL R30, OSU, UKMO). Roura-Pascual et al. (2004) drew ecological niche models from four general climate model scenarios for the future (horizon 2050), which strongly suggested the potential expansion of *L. humile* distribution in hot climates.

The correlative studies conduct the calibration and evaluation of ecological niche models in the whole current species distribution and, subsequently, the transfer of the model to climatic conditions for the years 2050 - 2070 (Peterson et al., 2018). Climate transferability can also occur under scenarios with past climatic conditions, resulting in important information about the history of species distribution. Although only a few studies have been conducted on this issue, paleodistribution was estimated for *Atta* spp. (Solómon et al., 2008), *Acromyrmex striatus* (Roger, 1863) (Cristiano et al., 2016), *Odontomachus meieri* Forel, 1905, *Octostruma* spp., and *Strumigenys* spp. (Ströer et al., 2019). These correlative studies used Maxent and, all of them, integrated modeling and phylogeography. Solómon et al. (2008) used the last glacial maximum (LGM) to test their hypotheses about the biogeography of speciation in the Amazon basin, suggesting that marine incursions into the Miocene or climatic changes during the Pleistocene shaped the population structure observed today in the three species evaluated. A pioneer paleoclimatic study of *Acromyrmex* spp. was conducted by Bigarella et al. (1975), who speculated about the paleoenvironmental conditions that characterized the Brazilian Pleistocene. Recently, in their assessment of the paleodistribution of *A. striatus* to the last interglacial (LIG) and the last glacial maximum (LGM), Cristiano et al. (2016) found that, in general, the past potential distribution included the current potential distribution of the species, showing constancy over time. Ströer et al. (2019) transferred the calibrated models to LGM and LIG. Their results support the traditional north/south division of the Brazilian Atlantic Forest, in addition to substantial differences between species in the location of genetic divisions and patterns of genetic variation within areas.

Studies have shown that temperature and precipitation are the two factors that most influence ant diversity patterns (Kaspari et al., 2003; Dunn et al., 2009; Sánchez-Restrepo et al., 2019), and are relevant to explain the models. Climatic variables were also used for mechanistic studies. Annual maximum temperature, seasonality, and aridity were the strongest predictors in the analysis of the warming tolerance of ant assemblages (CTMax) (Diamond et al., 2012). Still at the assemblage level, predicting potential changes in their relative abundance, a trait-based community selection model (CATS), was used to assess the relationship between temperature and UV-B (Bishop et al., 2019). This last study suggests that many more species will be present in higher elevation sites in the future and highlights the importance of environmental analyzes mediated by characteristics such as body color and size, as these can have consequences on thermoregulation and protection (see Bishop et al., 2016; 2019). Tropical organisms are more vulnerable to climate warming than temperate ones, especially when other factors, such as phylogenetic history and ecological characteristics, are accounted for (Diamond et al., 2012).
Some studies have suggested that the climate change effect can overcome habitat loss as the greatest threat to biodiversity (Pearson & Dawson, 2003; Lorenzen et al., 2011). Although the risk of invasion stems from biotic and abiotic factors, the climate seems to be primarily responsible for determining the distribution of ants on a global scale (Sanders et al., 2007; Jenkins et al., 2011; Roura-Pascual et al., 2011). Ants are ectothermic organisms, sensitive to temperature, and humidity, requiring adequate climatic conditions for their establishment (Diamond et al., 2012; Bertelsmeier and Courchamp, 2014). The Argentine ant distribution, for example, seems to be influenced mainly by altitude, average temperature, and precipitation (Roura-Pascoa et al., 2009). For Harris and Becker (2007), the average annual temperature and precipitation would be sufficient factors to highlight invasion risks, suggesting that the chances of a successful establishment would be reduced in cases in which these parameters were close to their limits. However, the lack of data related to human beings in fine resolution prevents the approach of anthropogenic influences, which is perhaps a better indicator of the establishment and spreading of the Argentine ant in some areas than the climatic features (Roura-Pascual et al. 2006).

The non-climatic variables used were mainly topography, soil temperature, vegetation, and soil inclination angle. As an example, solar radiation is considered one of the most important F. exsecta requirements (Maggini et al., 2002). In this case, terrain slope can be used as a surrogate to give a good idea of solar radiation, as this variable is challenging to be measured accurately.

Models based on the vegetation index (NDVI) can predict wider potential distributions than models that include only topographic information (Roura-Pascual et al., 2006). The addition of non-climatic data sets, such as soil characteristics (soil temperature, see Jung et al., 2021), landscape configuration, and land use/cover, would likely refine predictions considerably (Peterson & Nakazawa, 2007). The role that climate change has played in species diversification must be then assessed, but other mechanisms possibly synergetic must also be considered (Solómon et al., 2008).

Conclusion

A summary of the current panorama of species distribution modeling in Formicidae is represented in Figure 4. Most studies on Formicidae focused on invasive species, and how climate change can act on their distribution and occurrence. The correlative models were the most used to estimate changes in their potential ranges, also using different global warming scenarios. This useful tool makes these studies necessary for investigations aiming to mitigate the effect of invasive species on biodiversity. However, we suggest further studies be conducted, especially for conservation purposes, since the distribution of many ant species is incipiently known, as well as their conservation status.

Modeling is a tool that can be used in biological management and conservation strategies. However, it does not substitute the need for original field records, since they guarantee the construction of more robust, predictive models and their respective validation. In addition, in the future,
the risk of extinction due to loss of habitat, likewise, may be inferred based on the area of the potential occurrence of species, as long as maps of native vegetation remnants are available. In this case, an interesting approach to be studied would be the approximation of real occurrence areas of species from those of potential occurrence.

**Acknowledgment**

We are very grateful to the following institutions and funding agencies: Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), for PSS granting a scholarship, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), for JHCD research scholarship, and Universidade Federal do Sul da Bahia (UFSB), for Financial Support.

**Authors’ Contributions**

PSS: conceptualization, methodology, conceptualization, methodology, writing and editing the manuscript

JHCD: conceptualization, methodology, writing and editing the manuscript

AA: conceptualization, methodology, writing and editing the manuscript

EBAK: Formal analysis, writing and editing the manuscript.

**REFERENCES**


Guénard, B., Weiser, M.D. & Dunn, R.R. (2012). Global models of ant diversity suggest regions where new discoveries are most likely are under disproportionate deforestation threat. PNAS, 109: 7368-7373. doi: 10.1073/pnas.1113867109


