insects

MDPI

Article

Bioclimatic Modelling Identifies Suitable Habitat for the Establishment of the Invasive European Paper Wasp (Hymenoptera: Vespidae) across the Southern Hemisphere

Matthew W. F. Howse [1,*], John Haywood [2] and Philip J. Lester [1]

[1] School of Biological Sciences, Victoria University of Wellington, Wellington 6140, New Zealand; phil.lester@vuw.ac.nz

[2] School of Mathematics and Statistics, Victoria University of Wellington, Wellington 6140, New Zealand; john.haywood@vuw.ac.nz

* Correspondence: matthew.howse@vuw.ac.nz

Received: 20 October 2020; Accepted: 9 November 2020; Published: 11 November 2020

Simple Summary: The European paper wasp, *Polistes dominula* Christ (Hymenoptera: Vespidae), has become an invasive species across the globe. This wasp can reach high population densities and this, combined with its predatory nature, makes this insect a potential threat to biodiversity. There has been a lot of research conducted on this species throughout the northern hemisphere; however, little is known about their distribution in the southern hemisphere. Our objective was to identify where, in the southern hemisphere, *P. dominula* could become established. Two species distribution modelling approaches were used to make these predictions. Based on these models, there are large areas across southern South America, South Africa, southern Australia, and much of New Zealand that are likely to be at risk of further invasion by this species. These findings can be used to inform biosecurity measures in regions deemed at risk of invasion by this globally important pest.

Abstract: Species distribution models (SDMs) are tools used by ecologists to help predict the spread of invasive species. Information provided by these models can help direct conservation and biosecurity efforts by highlighting areas likely to contain species of interest. In this study, two models were created to investigate the potential range expansion of *Polistes dominula* Christ (Hymenoptera: Vespidae) in the southern hemisphere. This palearctic species has spread to invade North and South America, South Africa, Australia, and more recently New Zealand. Using the BIOCLIM and MAXENT modelling methods, regions that were suitable for *P. dominula* were identified based on climate data across four regions in the southern hemisphere. In South America areas of central Chile, eastern Argentina, parts of Uruguay, and southern Brazil were identified as climatically suitable for the establishment of *P. dominula.* Similarly, southern parts of South Africa and Australia were identified by the model to be suitable as well as much of the North Island and east of the South Island of New Zealand. Based on outputs from both models, significant range expansion by *P. dominula* is possible across its more southern invaded ranges.

Keywords: BIOCLIM; invasive species; MAXENT; *Polistes dominula*; species distribution model

1. Introduction

Species distribution models (SDMs) are becoming increasingly important in ecology, due to their ability to help predict the potential distributions of invasive organisms. These models bring together known species occurrence records and environmental data to provide users with an estimation of the

conditions a species requires to survive. This information can be used to identify locations that could support populations of a particular species [1]. Species distribution models have been used to guide the creation of more effective reserves [2], to project impacts of climate change [3], and to predict the spread of invasive species [4].

Polistes dominula Christ (Hymenoptera: Vespidae) is a well-known and wide-ranging invasive social wasp species [5]. A palearctic species native to Europe, North Africa, and parts of Central Asia, *P. dominula* has spread to both North and South America, South Africa, Australia, and more recently, New Zealand (Figure 1) [5–7]. This species is largely predatory by nature [8]. Similar to its Vespid relatives, *P. dominula* is a frequent predator of Lepidoptera larvae [9–12] though it is thought that this predatory wasp has a more generalist diet than that of other related species [13,14]. *Polistes dominula* has also been shown to have several competitive advantages over closely related species. Studies in North America and South Africa have shown that *P. dominula* nests are more productive than other *Polistes* species, able to produce more offspring over a longer active season [15–17]. This high nest productivity has led *P. dominula* to reach large population densities in their invaded areas.

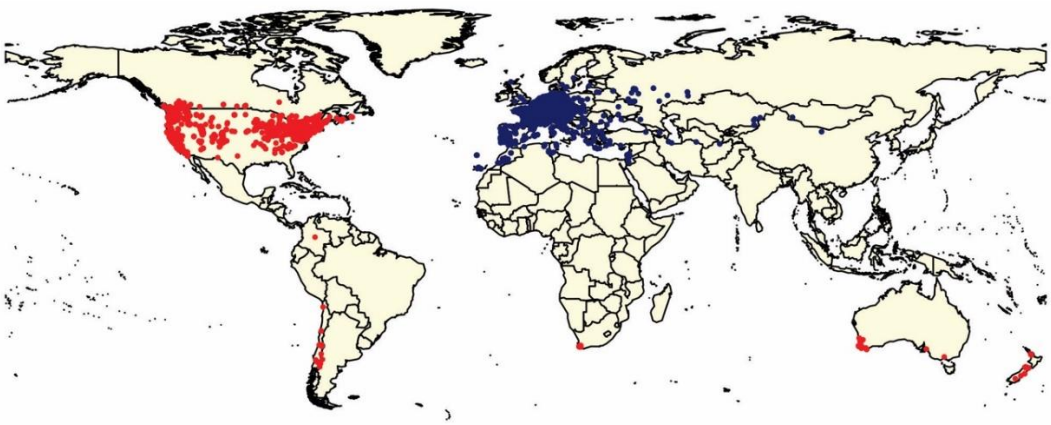

Figure 1. Global distribution of *P. dominula* constructed from data retrieved from Global Biodiversity Information Facility database (GBIF) [18]. Red points indicate occurrences in an invaded range. Blue points indicate occurrences in the assumed native range [5,6]. In total 9246 occurrences were used in this study. Of the total, 3028 occurrences were from invaded ranges and 6218 were from the assumed native range.

The invasion of *P. dominula* across the globe has been fairly well studied but not equally across affected regions. In the northern hemisphere *P. dominula* has famously invaded from the east to west coasts of the United States over the last 50 years [19]. *Polistes dominula* is known to have become established throughout the southern hemisphere but its spread in these regions has been critically understudied. This invasive species appears to have established in the southern hemisphere by the 1980s in Australia [20] and Chile [21,22]. It was found to have established in Argentina by 2003 [23], South Africa by 2008 [24], and in New Zealand by 2016 [7]. Except in South Africa, little research has been conducted on the status of these invasive populations and how they have affected local ecosystems. Many of these invaded areas contain native invertebrate communities that already face threats from other invasive species, habitat alteration, and climate change [25–27]. The establishment of *P. dominula* throughout more of these regions may add to these threats. Thus, predicting and preparing for future invasions or range expansions may help mitigate this effect.

Since their inception, there have been many changes and improvements to how SDMs are formed. One of the earliest and most widely used methods is BIOCLIM [28,29]. BIOCLIM is a profile method of species distribution modelling, whereby the algorithm determines the environmental similarity between a target species' current range and other locations, using a percentile distribution of values [30]. The model will designate a location more suitable if the environmental values are closer to the median values of known occurrence sites. BIOCLIM is a presence-only method and so does not require known

absence data. Another modelling method, known as MAXENT, can also be used with presence-only data, but there are some key differences from BIOCLIM. MAXENT is a machine learning method. The aim of MAXENT is to minimize the relative entropy between the probability density estimated from the occurrence data and the probability density estimated from the rest of the landscape [31]. The MAXENT method was developed more recently than the BIOCLIM approach, is also widely used, and is thought to be one of the better performing modelling methods [32–34].

A species distribution model would identify regions that are conducive to supporting populations of *P. dominula* and provide an insight into where these wasps are likely to spread. In this study two prediction models were produced, respectively using the BIOCLIM and MAXENT methods. Using global occurrence and climate data, these models were used to predict the bioclimatic suitability of four regions across the southern hemisphere for the establishment of *P. dominula*. These findings can inform future biosecurity and control plans for regions anticipated to be impacted by this invasive wasp.

2. Materials and Methods

Polistes dominula occurrence data was downloaded from Global Biodiversity Information Facility database (GBIF) [18] using the "gbif" function in the package dismo [35] in R version 4.0.2 [36]. A total of 20,616 records from 125 published datasets were downloaded. As GBIF data was pooled from a range of sources, from peer reviewed studies to citizen reports, data quality could vary. It has been shown, however, that a combination of data from citizen science and long term expert surveying can still produce robust distribution models [37].

A number of data-cleaning procedures were carried out to ensure the best quality data were used in creating the models. Following the procedures outlined by Hijmans and Elith [30], data were prepared by first removing data points with missing latitude or longitude values. Data were then assessed for the presence of duplicate coordinates which were removed to prevent pseudo replication. This assessment was accomplished using "duplicated", a base function in R [36] that identifies records with identical coordinates to others, which were removed. Remaining data points were then cross checked against a simple world map to identify any coordinates that were located on water. These values were likely to be occurrences recorded with low resolution coordinates and had to be excluded.

A set of climate variables containing temperature and precipitation was used for modelling the distribution of *P. dominula*. Climate data were obtained in the form of 19 environmental layers, each representing a global bioclimatic variable at 2.5-min (5 km^2) resolution (available from: https://worldclim.org/data/worldclim21.html). The WorldClim database is based on global weather station data from 1970 to 2000 and provides high resolution, global layers of monthly climate data [38]. These monthly data were used to create the 19 annual bioclimatic variable layers used in this study (Table 1) and in a variety of other distribution models [39–43].

Table 1. List of bioclimatic variables used in this study. These variables, created by Fick and Hijmans (2017), were downloaded as 2.5 min (5 km^2) resolution environmental layers from worldclim.org. Variables used monthly climate data collected over 30 years between 1970 and 2000. A subset of bioclimatic variables, identified by tick marks, was produced by stepwise regression analysis to be used in the BIOCLIM and MAXENT modelling approaches.

Variable Code	Variable Title	Unit	Inclusion in Final Model
Bio1	Annual Mean Temperature	°C	✓
Bio2	Mean Diurnal Range *(mean of monthly (max temp-min temp))*	°C	✓

Table 1. *Cont.*

Variable Code	Variable Title	Unit	Inclusion in Final Model
Bio3	Isothermality *((Bio2/Bio7) × 100)*	%	✓
Bio4	Temperature Seasonality *(standard deviation × 100)*	°C	✓
Bio5	Max Temperature of Warmest Month	°C	✓
Bio6	Min Temperature of Coldest Month	°C	
Bio7	Temperature Annual Range *(Bio5-Bio6)*	°C	✓
Bio8	Mean Temperature of Wettest Quarter	°C	✓
Bio9	Mean Temperature of Driest Quarter	°C	✓
Bio10	Mean Temperature of Warmest Quarter	°C	✓
Bio11	Mean Temperature of Coldest Quarter	°C	✓
Bio12	Annual Precipitation	mm	
Bio13	Precipitation of Wettest Month	mm	✓
Bio14	Precipitation of Driest Month	mm	✓
Bio15	Precipitation Seasonality *(coefficient of variation)*	%	
Bio16	Precipitation of Wettest Quarter	mm	✓
Bio17	Precipitation of Driest Quarter	mm	
Bio18	Precipitation of Warmest Quarter	mm	✓
Bio19	Precipitation of Coldest Quarter	mm	✓

Pseudoabsence points were created by selecting 10,000 random points from around the globe. This number was chosen to provide an appropriate ratio of presence to pseudoabsence points and maximize model reliability [44]. Climate data was applied to these pseudoabsence points as well as the cleaned occurrence data using the "extract" function in the raster package [45]. Presence data were randomly partitioned into test and training data using the "kfold" function, as recommended by [30] and used by [34]. Model predictions were made using the training data, which contained 7397 of the 9246 occurrence points. The predictions were then tested against the remaining 1849 occurrence points as well as 10,000 random pseudoabsence points.

Variable selection was performed using stepwise logistic regression to produce a subset of variables that would be used in the models (Table 1). The use of an automatic selection method was chosen to remove bias from the selection process. The procedure adds or removes variables from generalized linear models, one by one, checking the significance of all variables in the model each time. If a variable in the new model is deemed nonsignificant it is removed. Variable importance was measured by the Akaike Information Criterion (AIC). Important variables are added to the model while less relevant variables are removed. Using this analysis, the subset of variables that produced the model with the lowest AIC was chosen to perform the subsequent predictions. This process removed unnecessary variables from the model and reduced multicollinearity. The generalized linear model used all presence and background data with associated climatic variables and assumed a binomial distribution for the occurrence of *P. dominula* at each global location.

Two prediction models describing the global distribution of *P. dominula* were made using the worldwide occurrence of *P. dominula* and their associated climatic variables. One model was built using the BIOCLIM modelling method while the other used the MAXENT approach with the default settings [30]. With these models, suitable climatic conditions were predicted and identified across four regions of the southern hemisphere. All these regions have been previously invaded by *P. dominula* and include southern South America, South Africa, Australia, and New Zealand. The models' raw predictive outputs produced maps at 5 km^2 resolution with each cell containing values of habitat suitability. For both models, this predictive value was a number between zero and one. Following the BIOCLIM method, a cell would have a value of one if the environmental variables were equal to the median value for the occurrence data [30]. A zero value is conversely applied to any cell that possesses

climatic values lower than the 10th and higher than the 90th percentile values in the occurrence data. Raw output values produced by the MAXENT modelling method are an approximate probability that the species will be present, given the local environmental conditions, otherwise known as the relative occurrence rate (ROR) [31,46]. Using the "var.importance" function from the package ENMeval [47], values of variable permutation importance to the model produced by MAXENT were identified. Variable permutation importance is a percentage value showing how heavily the model depends on a specific variable. Values of each variable in training, testing and background data are permuted and the resulting model is evaluated. The degree to which the permuted data weakens the MAXENT produced model's performance is normalised to a percentage for each variable and presented as variable permutation importance (Table 2) [34].

Table 2. The 15 variables used in the MAXENT method, listed in order of permutation importance. Variable permutation importance is a value showing how heavily the final model depends on a certain variable. Values of each variable in training, testing and background data are randomized and the resulting model is evaluated. The degree to which the randomly permuted data weakens the model's performance, as originally selected by MAXENT, is normalised to a percentage for each variable and presented as permutation importance [34]. A variable with a high permutation importance is therefore important to the model, since if its values were randomized, the power of the model would decrease significantly.

Variable Code	Variable Title	Permutation Importance
Bio1	Annual Mean Temperature	37.9%
Bio19	Precipitation of Coldest Quarter	15.9%
Bio4	Temperature Seasonality *(standard deviation × 100)*	11.4%
Bio10	Mean Temperature of Warmest Quarter	9.0%
Bio16	Precipitation of Wettest Quarter	4.7%
Bio18	Precipitation of Warmest Quarter	4.2%
Bio11	Mean Temperature of Coldest Quarter	3.8%
Bio3	Isothermality *((Bio2/Bio7) × 100)*	2.8%
Bio9	Mean Temperature of Driest Quarter	2.7%
Bio14	Precipitation of Driest Month	2.3%
Bio2	Mean Diurnal Range *(mean of monthly (max temp–min temp))*	1.8%
Bio5	Max Temperature of Warmest Month	1.4%
Bio13	Precipitation of Wettest Month	1.1%
Bio7	Temperature Annual Range *(Bio5-Bio6)*	0.5%
Bio8	Mean Temperature of Wettest Quarter	0.4%

Presence/absence predictions show cells that contain predicted values of climate similarity above a determined threshold specific to each model. This threshold was determined as the maximum of the sum of the sensitivity (true positive rate) and specificity (true negative rate), also known as the maxSSS method [48]. Any cell with a value of climatic similarity over the threshold was predicted as a presence point, while any cell under this would be considered an absence [30]. This method of threshold selection has been recommended as appropriate when working with presence-only data [48,49].

The models were evaluated looking at the area under the receiver operating characteristic curve (AUROC, abbreviated to AUC) value. Test data with known presence or pseudoabsence status were entered into the models. The degree to which the model could correctly assign these data points to presence or pseudoabsence classes was used to calculate the AUC. This value is a number between zero and one and represents how accurately a model predicts presence/absence. A model with an AUC

value of 1 predicts presence/absence with 100% accuracy while a model with an AUC score of 0.5 is one that predicts a presence or absence correctly 50% of the time [30,50].

To investigate patterns presented in the final models, histograms of the known occurrence climatic values were plotted with the "hist" function [36] and visually compared with the corresponding bioclimatic layer for each of the focus regions. This was conducted for the two bioclimatic variables with the highest permutation importance.

3. Results

After the data cleaning procedure, a total of 9246 occurrence points of *P. dominula* remained for use in the model. Of these points, 3028 occurred in an invaded range while the other 6218 points originate in assumed native range (Figure 1).

Stepwise logistic regression analysis using the 19 bioclimatic explanatory variables identified a subset of variables that produced the best model, as indicated by the lowest AIC value (Table 1). The analysis used generalized linear models predicting presence or absence of *P. dominula*, with variables added or removed based on changes in the resulting AIC values. The final model contained 15 of the 19 WorldClim variables with the analysis removing Minimum Temperature of the Coldest Month (Bio6), Annual Precipitation (Bio12), Precipitation Seasonality (Bio15) and Precipitation of Driest Quarter (Bio17) (Table 1).

Modelling produced by the BIOCLIM method showed a number of regions in southern South America to have a climate conducive to the establishment of *P. dominula* (Figure 2). Once the threshold was applied, the model identified areas of central Chile, central and eastern Argentina as well as parts of Uruguay and southern Brazil as climatically suitable (Figure 3). Records of *P. dominula* are currently restricted to Chile and western Argentina, indicating the potential for a range expansion eastward. The raw output of the model produced by MAXENT highlighted a wider area of potentially suitable habitat than that of BIOCLIM (Figure 2); however, once the threshold was applied a more conservative potential range was predicted (Figure 3). Though not as expansive as the range predicted by BIOCLIM, it follows a similar pattern. Two main clusters of suitable habitats were identified by the MAXENT method with one spreading throughout central Chile and another in eastern Argentina and southern Uruguay (Figure 3).

In southern Africa, the BIOCLIM method identified an extensive range of climatically suitable habitat. The raw output of the BIOCLIM method highlighted areas of South Africa from the southwest of the country, eastward into the interior, encompassing much of Lesotho and into parts of Eswatini (Figure 2). With the threshold applied, much of this described area was predicted to be suitable for *P. dominula* (Figure 3). The raw output of the MAXENT method highlighted a wider area of the region but followed a pattern much the same as that of the BIOCLIM prediction (Figure 2). With the threshold applied, however, the MAXENT model identified a much smaller potential range than that of BIOCLIM. Only areas in the southwest of South Africa were predicted to be climatically suitable for the establishment of *P. dominula* (Figure 3). These areas fell within the South Cape province, which contains all known *P. dominula* occurrences within the southern Africa region.

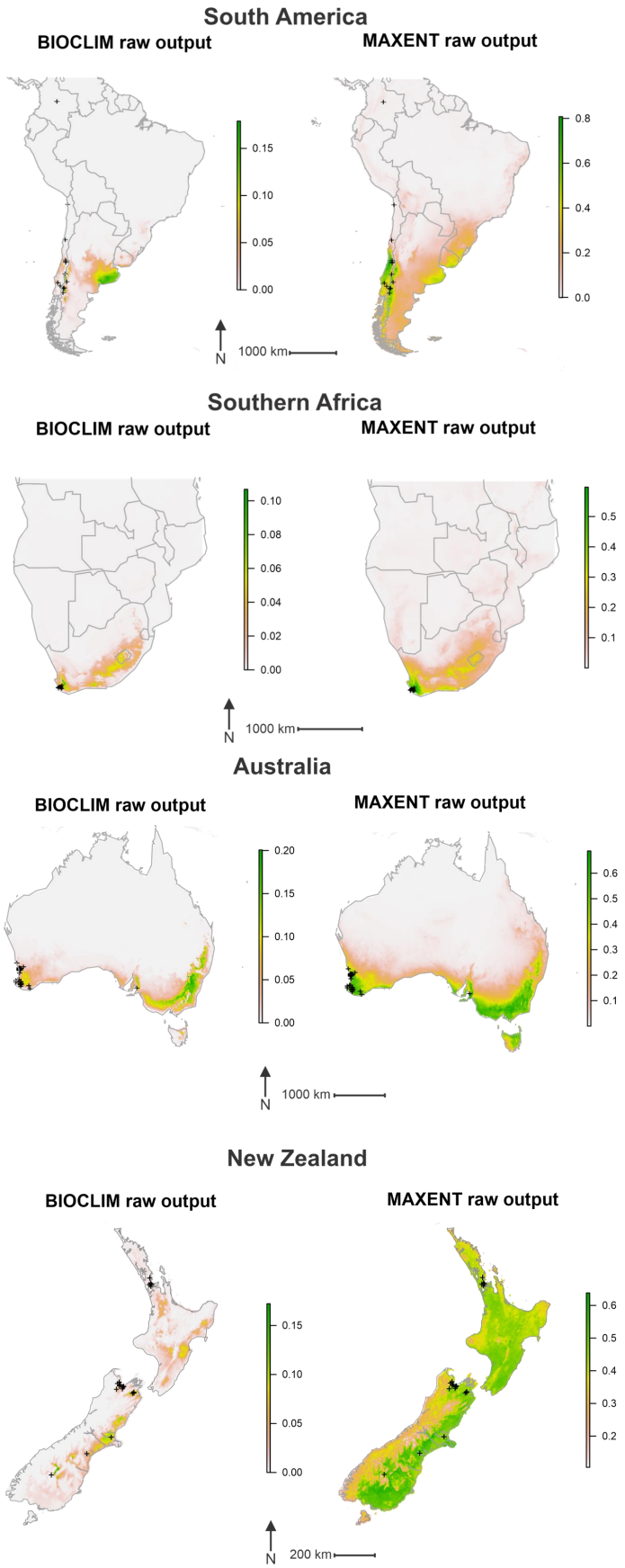

Figure 2. Raw outputs of both prediction models. The left images show the raw output given by BIOCLIM. The raw output of the BIOCLIM prediction is a value between 0 and 1 where the higher the number the more suitable the environment is. The BIOCLIM algorithm compares the environmental values of a cell to the median values of the environmental values of cells containing known occurrences

of the target species. Percentile scores closest to 0.5 are most suitable so values over this are subtracted from 1. The resulting score is multiplied by 2 to get a final value between 0 and 1 [30]. The right images show the raw output given by MAXENT where the value is an approximate probability that the species will be present given the local environmental conditions [31,46]. The black crosses indicate known occurrences of *P. dominula*.

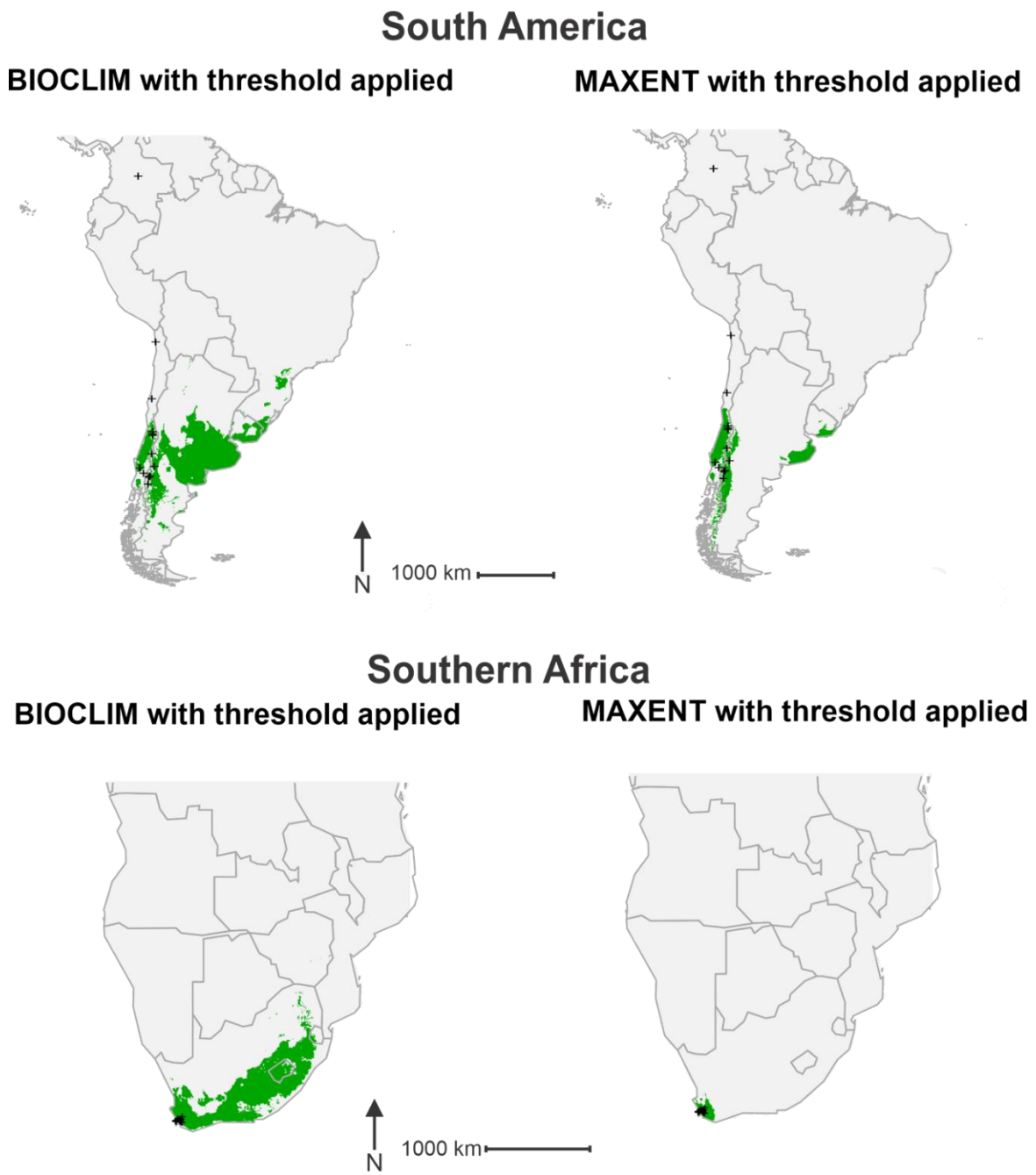

Figure 3. *Cont.*

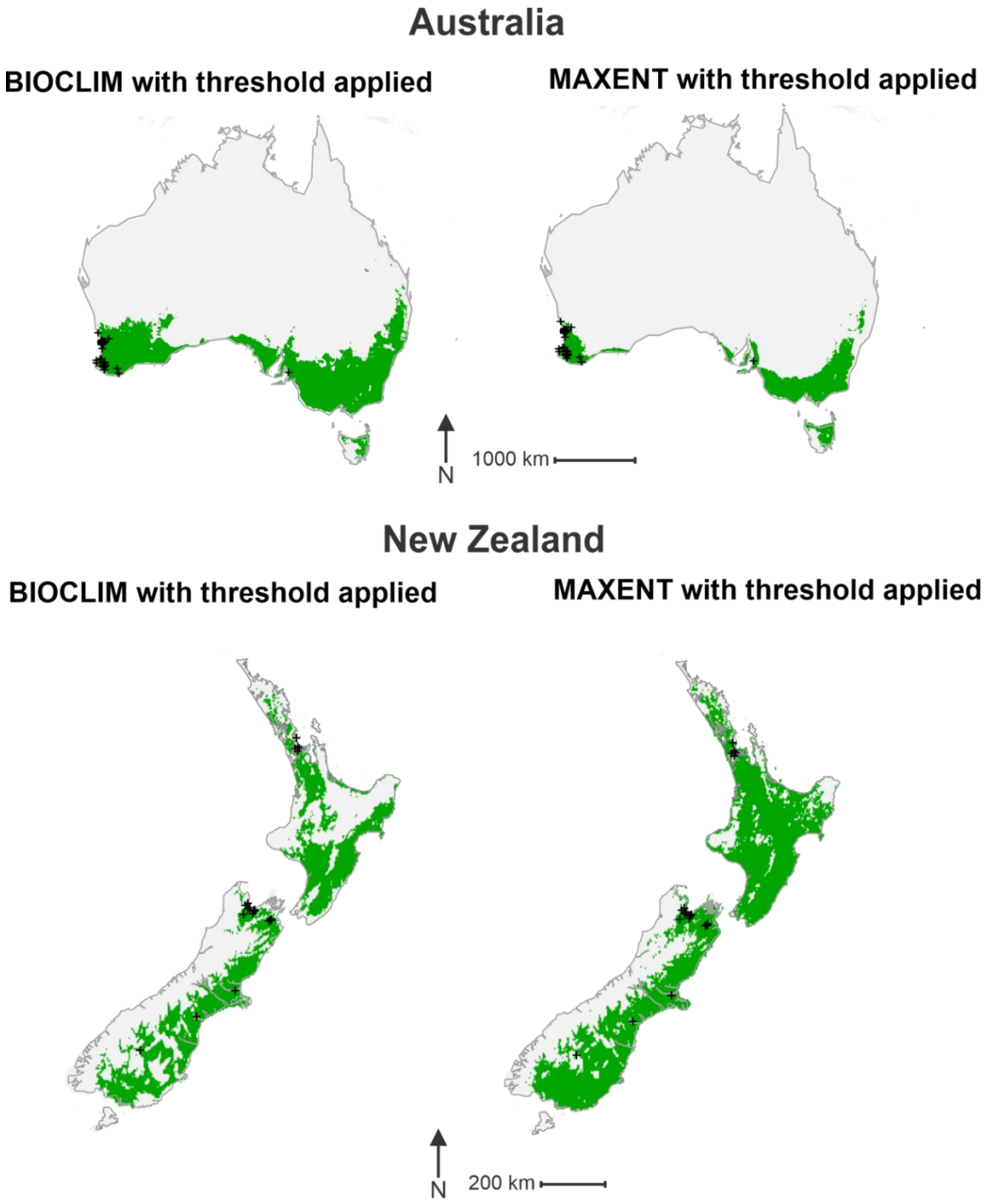

Figure 3. Predicted presence/absence plots for each region based on the two model outputs. For both models, thresholds (0.005 and 0.38, respectively) are calculated as the raw output value at which the sum of the true positive and true negative rates is maximized. Locations where the raw output values are over these thresholds are denoted a present status (1) and highlighted in green. Locations where raw output values are lower than the threshold are denoted an absent status (0) and remain grey. Black crosses indicate known occurrences of *P. dominula*.

Both prediction models identified much of the southern portion of Australia as climatically suitable for the establishment of *P. dominula.* The raw output of the BIOCLIM method highlighted most of the country south of approximately −30 degrees latitude (Figure 2). Once the threshold was applied the model identified most of this area as climatically suitable, with most of southwest and southeast of

the country denoted a potential present value (Figure 3). The raw output of the MAXENT method again highlighted a larger area of Australia than the BIOCLIM method, focused around the south and east of the country (Figure 2). When the threshold was applied the model identified two main clusters of suitable habitat in the south of Australia. One cluster was predicted around the southern parts of the state of Western Australia expanding eastward from where *P. dominula* is currently known to occur. The other cluster is focused around the south-east of the country in a pattern similar to, but more conservative than that of the BIOCLIM method (Figure 3).

In New Zealand, the raw output produced by the BIOCLIM method indicated suitable climatic conditions throughout the upper and lower parts of the North Island, with central and western areas deemed less suitable. Much of the eastern side of the South Island was highlighted by the prediction model as climatically suitable, while the west coast was not (Figure 2). Once the threshold was applied, much of the country highlighted in the raw output was predicted as climatically suitable for *P. dominula* (Figure 3). Similarly, the raw output produced by the MAXENT method highlighted an extensive area across New Zealand with much of the North Island and east of the South Island receiving the highest values of suitability (Figure 2). With the threshold applied, the MAXENT method predicted that much of the North Island and the east of the South Island is climatically suitable for *P. dominula* (Figure 3). The potential invadable range identified by both the BIOCLIM and MAXENT methods indicate that known populations of *P. dominula* in New Zealand are likely to expand their range.

The BIOCLIM method's performance as evaluated by AUC was 0.970 while the MAXENT method's performance was slightly higher with a value of 0.982 (Figure 4). These can be considered high AUC values as both values are close to 1, indicating the BIOCLIM and MAXENT methods were able to discriminate between the test-presence and background points 97% and 98.2% of the time respectively [50]. For the MAXENT method, variable permutation importance showed that annual mean temperature (Bio1) was the most important variable with a value of 37.9%, followed by precipitation of the coldest quarter (Bio19) at 15.9%, temperature seasonality (Bio4) at 11.4%, and mean temperature of the warmest quarter (Bio10) at 9.0% (Table 2).

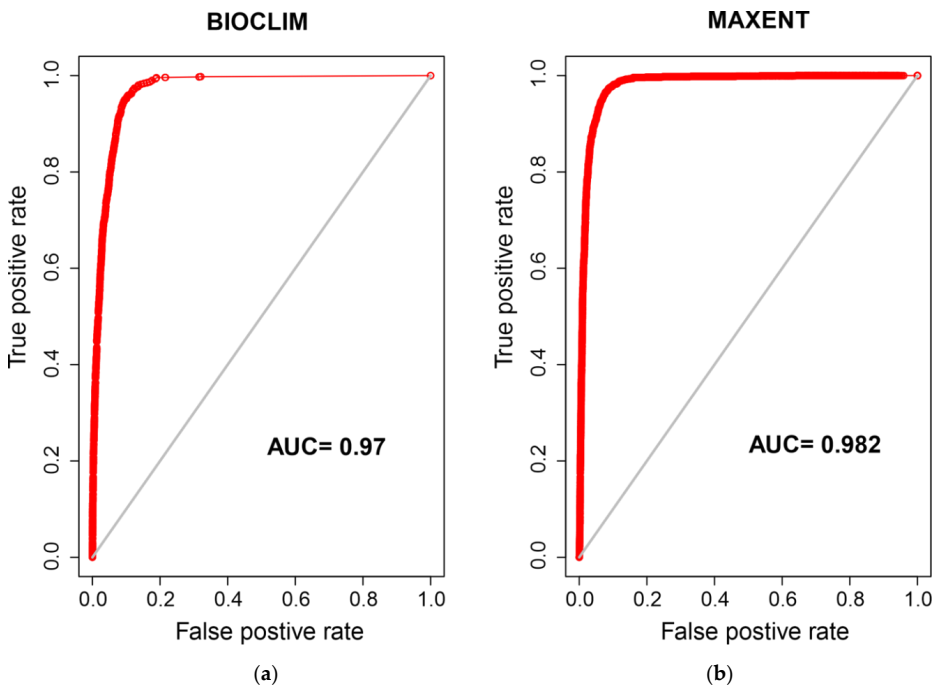

Figure 4. Area under the receiver operating characteristic curve (AUC) plots for BIOCLIM prediction (**a**) and MAXENT prediction (**b**). This value is a number between zero and one and represents how accurately a model predicts presence/absence. A model with an AUC value of 1 predicts presence/absence with 100% accuracy while a model with an AUC score of 0.5 is one that predicts presence or absence correctly 50% of the time [30].

Histograms were produced (Figure 5) for the highest-ranking climatic variables, as measured by permutation importance (Table 2). Known occurrences of *P. dominula* were found to experience average annual temperatures (Bio1) between −1 °C and 26.6 °C, with a sharp peak between 8 °C and 10 °C (Figures 5 and 6). Known occurrences of *P. dominula* were found to experience values of precipitation during the coldest quarter (Bio19) of 0–1057 mm, with most falling between 100 mm and 300 mm (Figures 5 and 7).

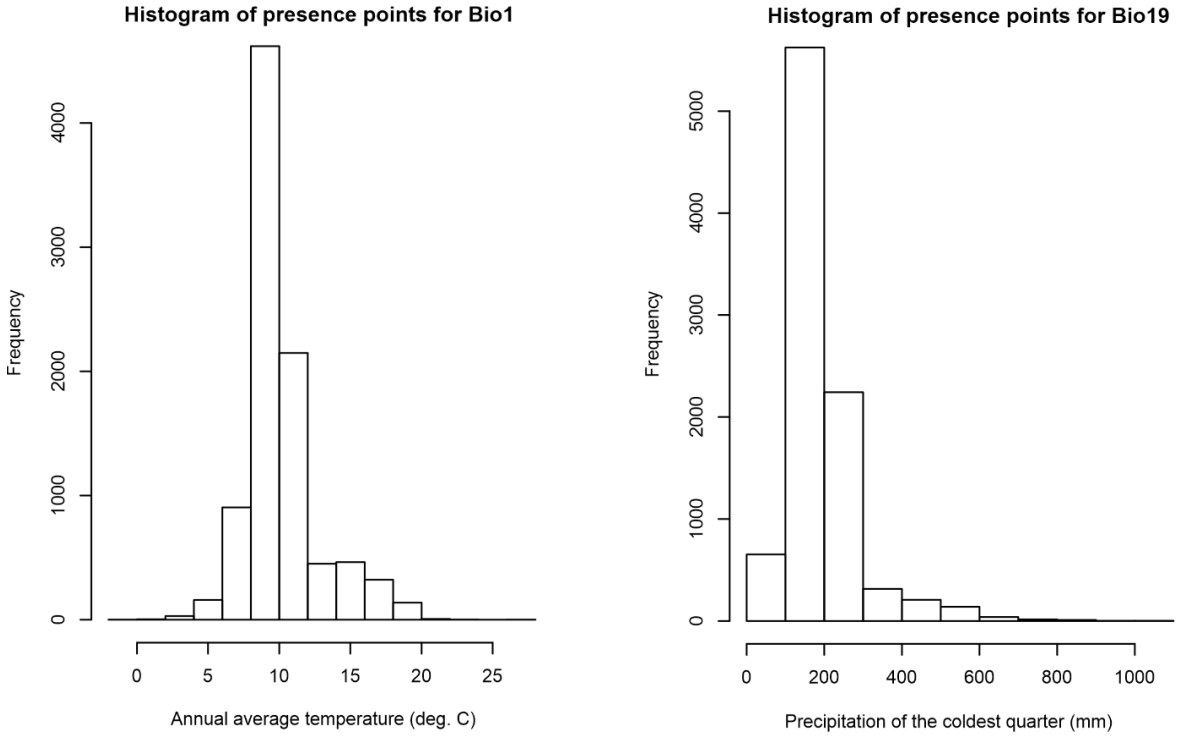

Figure 5. Histograms showing the distribution of bioclimatic values at known occurrences of *P. dominula.* Left is the distribution of annual average temperature (Bio1) and on the right the distribution of precipitation of the coldest quarter (Bio19). These two variables were chosen from the full list of 15, based on their high permutation importance in the MAXENT modelling approach.

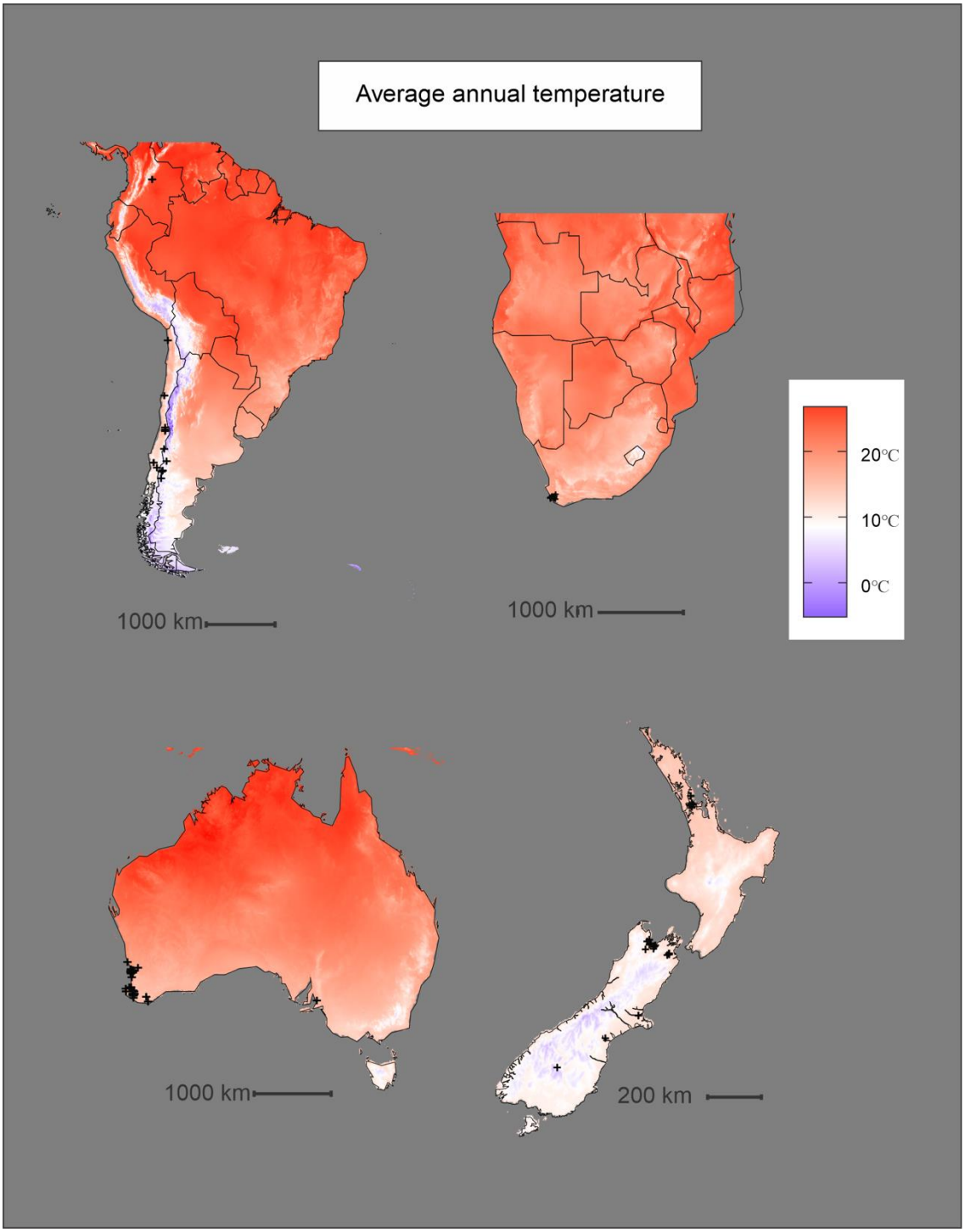

Figure 6. Plot of the average annual temperature (Bio1) bioclimatic layer at each region highlighted in this study. This variable had the highest permutation importance of 37.9%. Black crosses indicate known occurrences of *P. dominula.*

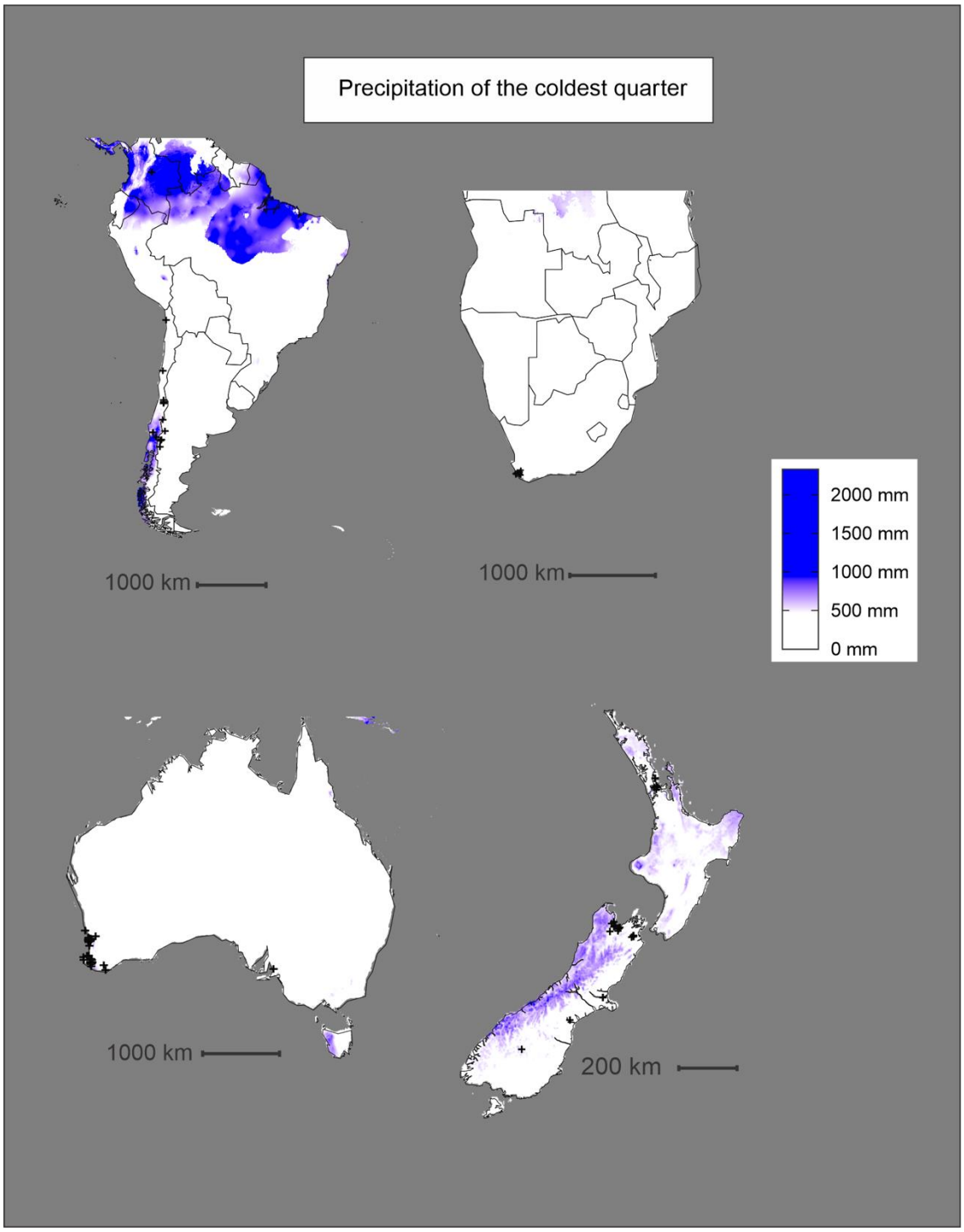

Figure 7. Plot of the precipitation of the coldest quarter (Bio19) bioclimatic layer at each region highlighted in this study. This variable had the second highest permutation importance of 15.9%. Black crosses indicate known occurrences of *P. dominula.*

4. Discussion

Polistes dominula is a widespread invader with introduced populations extending throughout North America, parts of South America, South Africa, Australia, and, more recently, New Zealand [5–7]. This species has the potential to reach higher densities than other paper wasps [17], and their preference to live in close proximity to human habitation [51] makes that an issue for human health. Invasive wasps have been linked to declines of native invertebrate species across the regions studied here [52–54].

Populations of *P. dominula* have already established across the four regions in this study but the estimated models predict a range expansion that could result in this species becoming present throughout more of the southern hemisphere.

The prediction models based on the BIOCLIM and MAXENT methods share some similarities in their predicted range of *P. dominula* across the four regions. In Australia, predicted distributions of *P. dominula* produced by both models closely follow patterns of average annual temperature. The majority of *P. dominula* occurrences fell in regions where the annual average temperature (Bio1) is between 8 °C and 12 °C with most of the country north of the predicted range averaging over 15 °C (Figure 6). This same pattern can be seen in South America where the potential range of *P. dominula* appears to be constrained by cooler temperatures further south (Figures 3 and 6). The models incorporate the contribution of other variables too; however, it is apparent that some variables influence *P. dominula* distribution more than others. For example, the east coast of the North and South Islands of New Zealand were identified as climatically suitable for the establishment of *P. dominula* while areas such as the West Coast region of the South Island appear to be much less suitable (Figure 3). The limitation on their distribution there is likely due to the high precipitation. Mean precipitation of the coldest quarter (Bio19) was shown to be the second most important variable contributing to the MAXENT model's accuracy (Table 2). The mean value of Bio19 from all known occurrence data was 195 mm of precipitation. The West Coast receives over 500 mm of rain in the coldest quarter [55] and so is considerably wetter than most places known to contain *P. dominula*. Similar patterns are observed across the other regions with the west coast of Tasmania, Australia also receiving high precipitation, well above the levels preferred by *P. dominula* (Figures 3 and 7).

The two distribution models did produce slightly different predictions. Both models had high AUC values (Figure 4) indicating robust models. The BIOCLIM method may be prone to overfitting when using many variables, leading to a narrower potential distribution [56] and that method can be prone to underpredicting potential distributions, possibly explaining the difference in raw model outputs (Figure 2) [40]. MAXENT, by contrast, has been repeatedly identified as a more reliable method, classed as a high performing, stable modelling approach compared to the BIOCLIM method [32,33]. Unlike the BIOCLIM modelling method, MAXENT weights variables differently depending on how their inclusion affects the models' AUC. This fundamental difference in modelling strategy leads to differences in the model outputs and hence explains some of the differences between predictions.

Threshold selection is another area that may explain differences between the BIOCLIM and MAXENT predictions. Thresholds were chosen using the maxSSS method, necessarily producing a different threshold for each method. The BIOCLIM method fitted models with a comparatively low threshold of 0.005 and so ranked nearly all the cells highlighted in the raw output plot as a presence value. The threshold for the MAXENT method was set at a comparatively higher 0.38, meaning that only the most climatically suitable of the areas highlighted in the raw MAXENT output were denoted a presence value. This difference in threshold explains the pattern we see where MAXENT presence/absence predictions appear to be more conservative for Australia, South Africa, and South America (Figure 3). By contrast, in New Zealand the MAXENT method's raw values of habitat suitability were so high that even with a more restrictive threshold, the areas denoted a presence status were larger than those resulting from the BIOCLIM method. Despite these differences in the final outputs both models predict areas that experience more mild average annual temperatures and are drier in the cooler months to be more suitable for the establishment of *P. dominula.* Both models predict a range expansion of *P. dominula* throughout the southern hemisphere.

Polistes dominula has been shown to be an important predator of invertebrates. They have been linked to the decline of invertebrates in invaded ranges around the world [9,57,58]. This is a pattern seen in invasive social wasp species due to their predatory behaviour and ability to reach high population densities [59,60]. With this species' continued range expansion, it is likely that *P. dominula* will compound existing pressures to native invertebrate fauna across the southern hemisphere. Molecular diet analysis performed on related *Polistes* species, also invasive to New Zealand, showed that

both native and introduced Lepidopteran species made up the largest portion of their diets [11,61]. Over half of all of New Zealand's threatened native Lepidopterans are found on the east coast of the South Island [62], a region that was identified as climatically suitable for *P. dominula* by both models. The seasonal nature of the *P. dominula* lifecycle [17,63] means only those prey species that are abundant or breed during the summer months are likely to be targeted [52]. This seasonality will still likely impact assemblages of invertebrates in the predominantly temperate regions that are predicted to be most suitable for establishment of this species.

Both models predict rather significant areas of bioclimatically suitable habitat across the southern hemisphere; however, these models have only considered the effects of temperature and precipitation on *P. dominula* establishment. Observations from the field in both native and invaded ranges show that *P. dominula* appear to preferentially nest near human habitation [15,63]. This pattern of synanthropy is seen in related species [59] and has been shown to allow other taxa to invade regions where environmental conditions are unsuitable [64]. Additional modelling including human habitation and land-use as variables may be able to produce a more realistic prediction of *P. dominula* potential range. Physical barriers, potential corridors, long-distance dispersal events (natural or human-related), and climatic change may influence colonization patterns. *Polistes dominula* was the most commonly reported nuisance wasp species in Colorado only four years after it was first identified in the state [65]. This species' ability to attain high densities close to human habitation could lead to the potential for an increased rate of human–wasp conflict in these predicted ranges.

Climate-based models such as the ones presented in this paper have been criticized for not considering biotic interactions when predicting habitat suitability [66–68]. Other studies consider the presence or absence of competitors [69–71]. This could be an important factor to consider in the spread of *P. dominula*, as other *Polistes* species are found in many of these regions. In southern South America there are 21 other species of *Polistes*, South Africa is home to 6 native species, Australia contains 15 other species while New Zealand contains two other invasive species of paper wasp [6]. These other, closely related species represent potential competitors that may impact the likelihood of establishment by *P. dominula* regardless of habitat suitability. While *P. dominula* has been shown to initially outcompete congener *Polistes* species, total displacement and replacement has not been shown to occur. It had been suggested that *P. dominula* was in the process of replacing the native *Polistes fuscatus* in the north-eastern United Sates [17,19,72]. A ten-year study of the two species in Michigan, USA, showed that despite initial displacement of *P. fuscatus* by *P. dominula*, populations eventually stabilized, likely due to the presence of a common parasitoid [73]. In South Africa, however, *P. dominula* continued to outnumber its native counterpart, *Polistes marginalis*, and maintain more productive nests despite the higher infection rates by a parasitoid [15]. Of the regions discussed in this paper, South America and Australia already contain a diverse range of *Polistes* species and so too presumably *Polistes* parasites and pathogens, which could have implications on invasion success.

5. Conclusions

Despite their relative simplicity, the use of climate based SDMs to predict potential ranges of species has been shown to provide strong predictive power [74]. Simple models such as these are still useful for making predictions over the broader scale and where more of the finer details of a species' ecology are not well understood. This paper presents the first attempt to use species distribution models to identify potentially suitable habitat for the establishment of the invasive *P. dominula*. Information gathered from these SDMs can help inform governments and conservation groups about the likelihood of *P. dominula* establishing in their respective regions. Regions with no known populations of *P. dominula* that were highlighted as suitable may invest in early detection and control to prevent the species from becoming established. Methods used here could be used and built on in future work to predict the range of other invasive species.

Author Contributions: Conceptualization, P.J.L. and M.W.F.H.; methodology, data curation, and formal analysis, J.H. and M.W.F.H.; visualization and writing—original draft preparation, M.W.F.H.; writing—review and editing, J.H. and P.J.L.; supervision, P.J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Thank you to James Baty, Mariana Bulgarella, Jana Dobelmann, Antoine Felden, and Rose McGruddy for their comments and feedback in the preparation of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pearson, R.G. Species' distribution modeling for conservation educators and practitioners. *Lessons Conserv.* **2007**, *3*, 54–89.
2. Leathwick, J.; Rowe, D.; Richardson, J.; Elith, J.; Hastie, T. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshw. Biol.* **2005**, *50*, 2034–2052. [CrossRef]
3. Moor, H.; Hylander, K.; Norberg, J. Predicting climate change effects on wetland ecosystem services using species distribution modeling and plant functional traits. *Ambio* **2015**, *44*, 113–126. [CrossRef]
4. Curry, P.A.; Yeung, N.W.; Hayes, K.A.; Cowie, R.H. The potential tropical island distribution of a temperate invasive snail, *Oxychilus alliarius*, modeled on its distribution in Hawaii. *Biol. Invasions* **2020**, *22*, 307–327. [CrossRef]
5. Buck, M.; Marshall, S.A.; Cheung, D.K. Identification atlas of the Vespidae (Hymenoptera, Aculeata) of the northeastern Nearctic region. *Can. J. Arthropod Identif.* **2008**, *5*, 1–492. [CrossRef]
6. Carpenter, J.M. Distributional checklist of species of the genus *Polistes* (Hymenoptera: Vespidae; Polistinae, Polistini). *Am. Mus. Novit.* **1996**, *3188*, 1–39.
7. Ministry for Primary Industries. Pest Watch. In *Surveillance*; Ministry for Primary Industries: Wellington, New Zealand, 2016; Volume 43, p. 26. ISSN 1176-5305.
8. Brown, R.; Payne, A.; Graham, K.K.; Starks, P.T. Prey capture and caste-specific payload capacities in the European paper wasp Polistes dominulus. *Insectes Sociaux* **2012**, *59*, 519–525. [CrossRef]
9. Baker, A.M.; Potter, D.A. Invasive paper wasp turns urban pollinator gardens into ecological traps for monarch butterfly larvae. *Sci. Rep.* **2020**, *10*, 1–7. [CrossRef] [PubMed]
10. Jeon, M.G.; Kim, T.G.; Jung, J.C.; Choi, M.B. Prey diversity of *Polistes rothneyi koreanus* in different landscapes using DNA barcoding. *J. Appl. Entomol.* **2019**, *143*, 1052–1063. [CrossRef]
11. Ward, D.F.; Ramón-Laca, A. Molecular identification of the prey range of the invasive Asian paper wasp. *Ecol. Evol.* **2013**, *3*, 4408–4414. [CrossRef] [PubMed]
12. Kasper, M.L.; Reeson, A.F.; Cooper, S.J.; Perry, K.D.; Austin, A.D. Assessment of prey overlap between a native (*Polistes humilis*) and an introduced (*Vespula germanica*) social wasp using morphology and phylogenetic analyses of 16S rDNA. *Mol. Ecol.* **2004**, *13*, 2037–2048. [CrossRef] [PubMed]
13. Cervo, R.; Zacchi, F.; Turillazzi, S. *Polistes dominulus* (Hymenoptera, Vespidae) invading North America: Some hypotheses for its rapid spread. *Insectes Sociaux* **2000**, *47*, 155–157. [CrossRef]
14. Schenk, D.; Bacher, S. Functional response of a generalist insect predator to one of its prey species in the field. *J. Anim. Ecol.* **2002**, *71*, 524–531. [CrossRef]
15. Roets, F.; Benadé, P.; Samways, M.; Veldtman, R. Better colony performance, not natural enemy release, explains numerical dominance of the exotic *Polistes dominula* wasp over a native congener in South Africa. *Biol. Invasions* **2019**, *21*, 925–933. [CrossRef]
16. Pickett, K.; Wenzel, J. High productivity in haplometrotic colonies of the introduced paper wasp *Polistes dominulus* (Hymenoptera: Vespidae; Polistinae). *J. N. Y. Entomol. Soc.* **2000**, *108*, 314–325. [CrossRef]
17. Gamboa, G.; Greig, E.; Thom, M. The comparative biology of two sympatric paper wasps, the native *Polistes fuscatus* and the invasive *Polistes dominulus* (Hymenoptera, Vespidae). *Insectes Sociaux* **2002**, *49*, 45–49. [CrossRef]
18. GBIF.org. GBIF Occurence Download. 2020. Available online: https://doi.org/10.15468/dl.jzezgu (accessed on 28 February 2020).

19. Liebert, A.E.; Gamboa, G.J.; Stamp, N.E.; Curtis, T.R.; Monnet, K.M.; Turillazzi, S.; Starks, P.T. Genetics, behavior and ecology of a paper wasp invasion: *Polistes dominulus* in North America. *Ann. Zool. Fenn.* **2006**, *43*, 595–624.
20. Taylor, R.W.; Brown, D.R.; Cardale, J.C. *Hymenoptera: Formicoidea, Vespoidea and Sphecoidea*; Australian Government Publishing Service: Canberra, Australia, 1985; Volume 2, ISSN 0644039221.
21. González, R.H. *Insectos y Ácaros de Importancia Agrícola y Cuarentenaria en Chile*; Universidad de Chile Santiago: Santiago, Chile, 1989; ISBN 9567049016.
22. Elgueta, M. Antecedentes sobre especies de reciente introducción a Chile (Diptera: Stratiomyidae; Hymenoptera: Vespidae). *Rev. Chil. Entomol.* **1989**, *17*, 97–98.
23. Sackmann, P.; Villacide, J.M.; Corley, J. Presencia de una nueva avispa social exótica, *Polistes dominulus* (Hymenoptera: Vespidae) en la Patagonia Argentina. *Rev. Soc. Entomológica Argent.* **2003**, *62*, 68–70.
24. Eardley, C.; Koch, F.; Wood, A. *Polistes dominulus* (Christ, 1791)(Hymenoptera: Polistinae: Vespidae) newly recorded from South Africa. *Afr. Entomol.* **2009**, *17*, 226–227. [CrossRef]
25. Stringer, I.; Hitchmough, R. Assessing the conservation status of New Zealand's native terrestrial invertebrates. *N. Z. Entomol.* **2012**, *35*, 77–84. [CrossRef]
26. Lester, P.; Brown, S.; Edwards, E.; Holwell, G.; Pawson, S.; Ward, D.; Watts, C. Critical issues facing New Zealand entomology. *N. Z. Entomol.* **2014**, *37*, 1–13. [CrossRef]
27. Sands, D.P.A. Important issues facing insect conservation in Australia: Now and into the future. *Austral Entomol.* **2018**, *57*, 150–172. [CrossRef]
28. Booth, T.H.; Nix, H.A.; Busby, J.R.; Hutchinson, M.F. bioclim: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Divers. Distrib.* **2014**, *20*, 1–9. [CrossRef]
29. Nix, H.A. A biogeographic analysis of Australian elapid snakes. In *Atlas of Elapid Snakes of Australia*; Longmore, R., Ed.; Australian Government Publishing Service: Canberra, Australian, 1986; Volume 7, pp. 16–77.
30. Hijmans, R.J.; Elith, J. Species Distribution Modeling with R. Available online: http://cran.r-project.org/web/packages/dismo/vignettes/dm.pdf (accessed on 8 February 2020).
31. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [CrossRef]
32. Duan, R.-Y.; Kong, X.-Q.; Huang, M.-Y.; Fan, W.-Y.; Wang, Z.-G. The predictive performance and stability of six species distribution models. *PLoS ONE* **2014**, *9*, e112764. [CrossRef]
33. Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [CrossRef]
34. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
35. Hijmans, R.J.; Phillips, S.; Leathwick, J.; Elith, J. Dismo: Species distribution modeling. In *R Package Version 1.1-4*; 2017; Available online: https://CRAN.R-project.org/package=dismo (accessed on 8 February 2020).
36. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
37. Sumner, S.; Bevan, P.; Hart, A.G.; Isaac, N.J.B. Mapping species distributions in 2 weeks using citizen science. *Insect Conserv. Divers.* **2019**, *12*, 382–388. [CrossRef]
38. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]
39. Beckham, J.L.; Atkinson, S. An updated understanding of Texas bumble bee (Hymenoptera: Apidae) species presence and potential distributions in Texas, USA. *PeerJ* **2017**, *5*, e3612. [CrossRef] [PubMed]
40. Ward, D.F. Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biol. Invasions* **2007**, *9*, 723–735. [CrossRef]
41. Rodder, D.; Schmidtlein, S.; Veith, M.; Lotters, S. Alien invasive slider turtle in unpredicted habitat: A matter of niche shift or of predictors studied? *PLoS ONE* **2009**, *4*, e7843. [CrossRef] [PubMed]
42. Raghavan, R.K.; Peterson, A.T.; Cobos, M.E.; Ganta, R.; Foley, D. Current and future distribution of the lone star tick, *Amblyomma americanum* (L.)(Acari: Ixodidae) in North America. *PLoS ONE* **2019**, *14*, e0209082. [CrossRef]

43. Dudaniec, R.Y.; Yong, C.J.; Lancaster, L.T.; Svensson, E.I.; Hansson, B. Signatures of local adaptation along environmental gradients in a range-expanding damselfly (*Ischnura elegans*). *Mol. Ecol.* **2018**, *27*, 2576–2593. [CrossRef]
44. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
45. Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. In *R Package Version 3.3-13*; 2020; Available online: https://CRAN.R-project.org/package=raster (accessed on 8 February 2020).
46. Merow, C.; Smith, M.J.; Silander, J.A., Jr. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **2013**, *36*, 1058–1069. [CrossRef]
47. Muscarella, R.; Galante, P.J.; Soley-Guardia, M.; Boria, R.A.; Kass, J.; Uriarte, M.; Anderson, R.P. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods Ecol. Evol.* **2014**, *5*, 1198–1205. [CrossRef]
48. Liu, C.; Newell, G.; White, M. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecol. Evol.* **2016**, *6*, 337–348. [CrossRef]
49. Liu, C.; White, M.; Newell, G. Selecting thresholds for the prediction of species occurrence with presence-only data. *J. Biogeogr.* **2013**, *40*, 778–789. [CrossRef]
50. Hosmer, D.W.; Lemeshow, S.; Sturdivant, R.X. Assessing the fit of the model. In *Applied Logistic Regression*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013; pp. 153–226.
51. Benadé, P.C.; Veldtman, R.; Samways, M.J.; Roets, F. Rapid range expansion of the invasive wasp *Polistes dominula* (Hymenoptera: Vespidae: Polistinae) and first record of parasitoids on this species and the native *Polistes marginalis* in the Western Cape province of South Africa. *Afr. Entomol.* **2014**, *22*, 220–225. [CrossRef]
52. Beggs, J.R.; Rees, J.S. Restructuring of Lepidoptera communities by introduced *Vespula* wasps in a New Zealand beech forest. *Oecologia* **1999**, *119*, 565–571. [CrossRef] [PubMed]
53. Toft, R.J.; Rees, J.S. Reducing predation of orb-web spiders by controlling common wasps (*Vespula vulgaris*) in a New Zealand beech forest. *Ecol. Entomol.* **1998**, *23*, 90–95. [CrossRef]
54. Potter-Craven, J.; Kirkpatrick, J.B.; McQuillan, P.B.; Bell, P. The effects of introduced vespid wasps (*Vespula germanica* and *V. vulgaris*) on threatened native butterfly (*Oreixenica ptunarra*) populations in Tasmania. *J. Insect Conserv.* **2018**, *22*, 521–532. [CrossRef]
55. Macarra, G.R. The climate and weather of West Coast. In *NIWA Science and Technology Series*; NIWA: Christchurch, New Zealand, 2016; ISSN 1173-0382.
56. Beaumont, L.J.; Hughes, L.; Poulsen, M. Predicting species distributions: Use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecol. Model.* **2005**, *186*, 251–270. [CrossRef]
57. Stamp, N.; Bowers, M. Direct and indirect effects of predatory wasps (*Polistes* sp.: Vespidae) on gregarious caterpillars (*Hemileuca lucina*: Saturniidae). *Oecologia* **1988**, *75*, 619–624. [CrossRef]
58. Pilowsky, J.; Starks, P. Displacement and replacement in real time: *Polistes dominula*'s impact on *P. fuscatus* in the northeastern U.S. *Biol. Invasions* **2018**, *20*, 1161–1169. [CrossRef]
59. Parent, C.E.; Peck, S.B.; Causton, C.E.; Roque-Albelo, L.; Lester, P.J.; Bulgarella, M. *Polistes versicolor* (Hymenoptera: Vespidae), an introduced wasp in the Galapagos Islands: Its life cycle and ecological impact. *Environ. Entomol.* **2020**. [CrossRef]
60. Beggs, J.R.; Brockerhoff, E.G.; Corley, J.C.; Kenis, M.; Masciocchi, M.; Muller, F.; Rome, Q.; Villemant, C. Ecological effects and management of invasive alien Vespidae. *BioControl* **2011**, *56*, 505–526. [CrossRef]
61. Lefort, M.-C.; Beggs, J.R.; Glare, T.R.; Saunders, T.E.; Doyle, E.J.; Stephane, B. A Molecular Approach to Studying Hymenoptera Diets Using Polistine Wasps. Available online: https://www.biorxiv.org/content/10.1101/2020.04.06.024422v1 (accessed on 17 August 2020).
62. Stringer, I.; Hitchmough, R.; Dugdale, J.; Edwards, E.; Hoare, R.; Patrick, B. The conservation status of New Zealand Lepidoptera. *N. Z. Entomol.* **2012**, *35*, 120–127. [CrossRef]
63. Höcherl, N.; Tautz, J. Nesting behavior of the paper wasp *Polistes dominula* in Central Europe-a flexible system for expanding into new areas. *Ecosphere* **2015**, *6*, 1–11. [CrossRef]
64. Schapheer, C.; Sandoval, G.; Villagra, C.A. Pest Cockroaches May Overcome Environmental Restriction Due to Anthropization. *J. Med Entomol.* **2018**, *55*, 1357–1364. [CrossRef] [PubMed]

65. Cranshaw, W.S.; Larsen, H.J.; Zimmerman, R.J. Notes on fruit damage by the European paper wasp, *Polistes dominula* (Christ) (Hymenoptera: Vespidae). *Southwest. Entomol.* **2011**, *36*, 103–105. [CrossRef]
66. Wisz, M.S.; Pottier, J.; Kissling, W.D.; Pellissier, L.; Lenoir, J.; Damgaard, C.F.; Dormann, C.F.; Forchhammer, M.C.; Grytnes, J.A.; Guisan, A. The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biol. Rev.* **2013**, *88*, 15–30. [CrossRef]
67. Giannini, T.C.; Chapman, D.S.; Saraiva, A.M.; Alves-dos-Santos, I.; Biesmeijer, J.C. Improving species distribution models using biotic interactions: A case study of parasites, pollinators and plants. *Ecography* **2013**, *36*, 649–656. [CrossRef]
68. Godsoe, W.; Harmon, L.J. How do species interactions affect species distribution models? *Ecography* **2012**, *35*, 811–820. [CrossRef]
69. Bulgarella, M.; Trewick, S.A.; Minards, N.A.; Jacobson, M.J.; Morgan-Richards, M. Shifting ranges of two tree weta species (*Hemideina spp*.): Competitive exclusion and changing climate. *J. Biogeogr.* **2014**, *41*, 524–535. [CrossRef]
70. Mpakairi, K.S.; Ndaimani, H.; Tagwireyi, P.; Gara, T.W.; Zvidzai, M.; Madhlamoto, D. Missing in action: Species competition is a neglected predictor variable in species distribution modelling. *PLoS ONE* **2017**, *12*, e0181088. [CrossRef]
71. Pellissier, L.; Anne Bråthen, K.; Pottier, J.; Randin, C.F.; Vittoz, P.; Dubuis, A.; Yoccoz, N.G.; Alm, T.; Zimmermann, N.E.; Guisan, A. Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. *Ecography* **2010**, *33*, 1004–1014. [CrossRef]
72. Gamboa, G.; Noble, M.; Thom, M.; Togal, J.; Srinivasan, R.; Murphy, B. The comparative biology of two sympatric paper wasps in Michigan, the native *Polistes fuscatus* and the invasive *Polistes dominulus* (Hymenoptera, Vespidae). *Insectes Sociaux* **2004**, *51*, 153–157. [CrossRef]
73. Miller, G.L.; Donnelly, C.R.; Gamboa, G.J. A ten-year comparative study of the population dynamics and parasitoidism in the native paper wasp *Polistes fuscatus* and the invasive *P. dominulus*. *Insectes Sociaux* **2013**, *60*, 49–56. [CrossRef]
74. Kearney, M.R.; Wintle, B.A.; Porter, W.P. Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. *Conserv. Lett.* **2010**, *3*, 203–213. [CrossRef]