


## Projecting the suitability of global and local habitats for myrtle rust (*Austropuccinia psidii*) using model consensus

H. A. Narouei-Khandan<sup>a\*</sup> , S. P. Worner<sup>b</sup>, S. L. H. Viljanen<sup>c</sup>, A. H. C. van Bruggen<sup>d</sup> and E. E. Jones<sup>b</sup>

<sup>a</sup>Bio-Protection Research Centre, Lincoln University, PO Box 85084; <sup>b</sup>Department of Pest-management and Conservation, Faculty of Agriculture and Life Sciences, Lincoln University, PO Box 85084; <sup>c</sup>The New Zealand Institute for Plant & Food Research Limited, Private Bag 4704, Christchurch 8140, New Zealand; and <sup>d</sup>Department of Plant Pathology and Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611-0680, USA

Myrtle rust (caused by *Austropuccinia psidii*) affects more than 500 known host species in the Myrtaceae family. Three different modelling approaches (CLIMEX, MaxEnt and Multi-Model Framework) were used to project the habitat suitability for myrtle rust at both global and local scales. Current data on the global occurrence of myrtle rust were collected from online literature and expert solicitation. Long-term averages of climate data (1960–1990) were sourced from WorldClim and CliMond websites. Recent reports of myrtle rust in New Zealand were used for validation of model outputs but not in model training and testing. The model outputs were combined into a consensus model to identify localities projected to be suitable for myrtle rust according to two or three models (hotspots). In addition to the locations where the pathogen is currently present, all models successfully projected independent occurrence data in New Zealand suitable for establishment of the pathogen. Climate suitability for the pathogen was primarily related to temperature followed by rainfall in MaxEnt and the CLIMEX model. The results confirmed the optimum temperature range of this pathogen in the literature (15–25 °C). Additional analysis of the precipitation variables indicated that excessive rain (more than 2000 mm in warmest quarter of the year) combined with high temperatures (>30 °C) constrain pathogen establishment. The results of the current study can be useful for countries such as New Zealand, China, South Africa and Singapore where the pathogen has not fully spread or established.

**Keywords:** *Austropuccinia psidii*, guava rust, model consensus, myrtle rust, species distribution models

### Introduction

Myrtle rust, also known as eucalyptus rust, 'ōhi'a rust and guava rust, is an important plant disease affecting more than 500 known host species in the Myrtaceae family (Berthon *et al.*, 2018; Yong *et al.*, 2019). The taxonomic position of the causal agent has been unclear until a recent study based on phylogenetic analyses revealed that the fungus belongs to a new genus, *Austropuccinia psidii*, within the family Sphaerophragmiaceae (Pucciniales; Beenken, 2017). *Austropuccinia psidii* can infect stems, leaves and fruit of its hosts, especially on young, actively growing eucalypts (Coutinho *et al.*, 1998; Pegg *et al.*, 2017). The infection may cause twisting or bending of the leaves and leaf drop (Coutinho *et al.*, 1998; Pegg *et al.*, 2017). The impact of myrtle rust on native Myrtaceae in Australia can be very serious, resulting in infected shoots and crowns and ultimately leading to tree death (Carnegie

*et al.*, 2016; Pegg *et al.*, 2017). For example, after repeated and severe infections by *A. psidii*, brown malletwood trees (*Rhodamnia rubescens*) and native guava trees (*Rhodomyrtus psidioides*) were killed in less than four years and the population of the native guava declined more than 50% in less than five years (Carnegie *et al.*, 2016; Pegg *et al.*, 2017). Similar impacts have been reported on nioi (*Eugenia koolauensis*), an endangered species in Hawaii (Loope, 2010). Impacts on different Myrtaceae species vary greatly from resistant to highly susceptible reactions, even within species, and similarly, *A. psidii* biotypes from different hosts and locations vary greatly in their ability to infect different host plants (Marlatt & Kimbrough, 1979; Berthon *et al.*, 2018).

The first serious outbreak of myrtle rust was reported from nurseries in Brazil in 1973 where it affected non-native (more susceptible) *Eucalyptus* species (Ferreira, 1981). Later the pathogen was reported from other parts of Latin America. The disease appeared in Florida in 1977 (Marlatt & Kimbrough, 1979) and in Hawaii in 2005 (Uchida & Loope, 2009). Thereafter, the pathogen quickly spread into other countries (Fig. 1) such as Australia in 2010 (Carnegie *et al.*, 2010) and New Caledonia and South Africa in 2013 (Roux *et al.*, 2013; Soewarto

\*E-mail: hakhandan@gmail.com

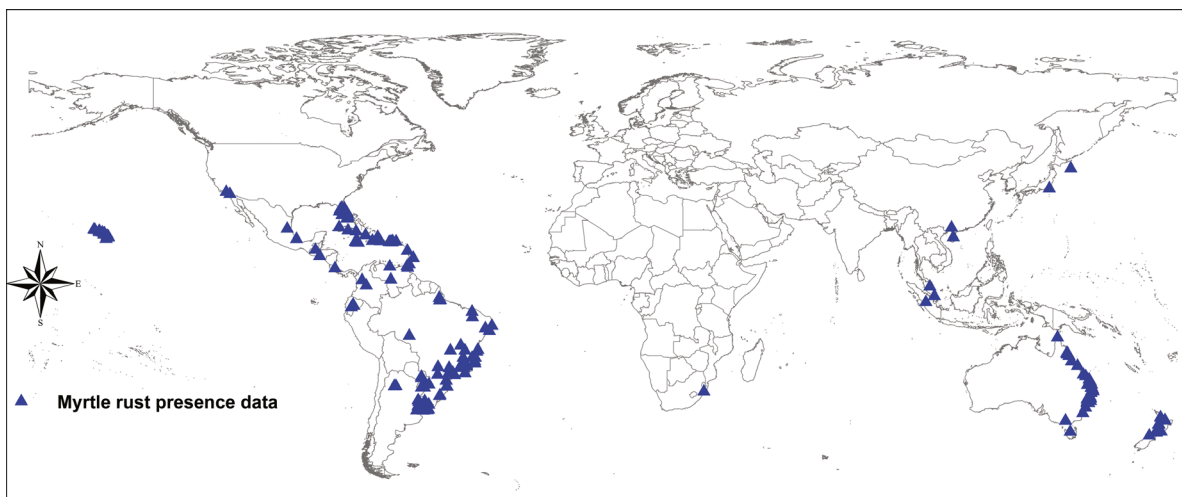


Figure 1 Current global distribution of myrtle rust caused by *Austropuccinia psidii*. The New Zealand presence points shown here were used as independent data to validate the output of the models developed in this study. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)].

*et al.*, 2018). More recently myrtle rust was reported from Singapore (du Plessis *et al.*, 2017) and New Zealand where it has not yet fully spread throughout the country ([www.mpi.govt.nz](http://www.mpi.govt.nz); accessed 25 April 2019).

Based on the wide host range and damage history of this pathogen in other regions of the world, myrtle rust is considered a serious risk to indigenous plants in the Myrtaceae family and to biodiversity in New Zealand (Kriticos & Leriche, 2008; Ramsfield *et al.*, 2010; Stewart *et al.*, 2017). Myrtle rust can potentially affect four main industries in New Zealand including nurseries growing myrtaceous plants, producers of mānuka honey and growers of feijoa and eucalypts (Clark, 2011). The rust could also alter the composition of New Zealand indigenous forest systems (Ramsfield *et al.*, 2010). In addition to economic and ecological values, the native species of the Myrtaceae family in New Zealand have significant sociocultural and spiritual significance. However, it is not known to what extent the climate in New Zealand is suitable for further spread of myrtle rust or how susceptible New Zealand myrtaceous species will be in different locations ([www.mpi.govt.nz](http://www.mpi.govt.nz), accessed 2 December 2017).

Because of the potential impact of myrtle rust and its rapid spread into new areas, several spatial models have been developed to predict further spread. Booth *et al.* (2000) used simple climatic mapping programs to identify high risk areas for myrtle rust in tropical ecoregions of the Americas and in Australia. While their model generally performed well, it did not predict southern parts of Queensland in Australia and Paraguay in South America as suitable, despite the fact that the presence of the rust is well documented in these areas. A more detailed climate-based model, named RISK, was developed based on climate data across Australia and relationships of temperature and relative humidity with disease development under controlled conditions (Booth & Jovanovic, 2012). Although, this model performed well for Australia, it did

not predict Tasmania, where myrtle rust was reported in 2015, as suitable (Tobias *et al.*, 2015). Alvares *et al.* (2017) developed a spatial regression model predicting myrtle rust in Brazil from calculated night temperatures and leaf wetness durations. Their model predicted a large area in central and northwest Brazil as highly suitable in most months of the year although only few reports of the rust are available from these regions to date. Elith *et al.* (2013) and Berthon *et al.* (2018) carried out another bioclimatic analysis of the risk of myrtle rust establishment in Australia using the MaxEnt model. Similarly, Stewart *et al.* (2017) used the MaxEnt model to project the global distribution of myrtle rust. Kriticos *et al.* (2008, 2013a) used CLIMEX Compare Locations and CLIMEX Match Regions to project the suitability of New Zealand for the establishment of myrtle rust. In addition, the NAPFAST model was used to project the potential global distribution of myrtle rust, including Australia and New Zealand, using 10 years of climate data (Magarey *et al.*, 2007). This model, which was based on early occurrences of myrtle rust (up to 2017), predicted extremely low probability of myrtle rust establishment in New Zealand and southern areas of New South Wales in Australia. These studies relied on single models and, although some of these models predicted the distribution of myrtle rust quite well, they were sometimes contradictory for some regions. A direct comparison of the performance of different models for the same region has not been carried out thus far.

It is critical to obtain a consensus model for the risk of myrtle rust establishment in New Zealand so that countermeasures could be designed to limit further spread. Therefore, the objectives of this study were to: (i) model the potential distribution of myrtle rust worldwide using different modelling approaches, with a focus on New Zealand; (ii) identify climatic variables important for the potential establishment of myrtle rust; (iii) compare the predictions of the likelihood of myrtle rust establishment by different models; and (iv) identify hotspots with a high

probability of establishment of myrtle rust according to a consensus model.

To address these objectives several species distribution models (SDMs) were used. In these models, the current distribution of the target species is related to historical environmental data to project the potential distribution. One of these models (CLIMEX) is considered semimechanistic, while two other modelling platforms used in this study (MaxEnt and Multi-Model Framework or MMF) are correlative models. They have been used in different studies to evaluate and project the potential distribution of various invasive organisms into new areas (Booth *et al.*, 2000; Kriticos *et al.*, 2013a; Narouei-Khandan *et al.*, 2016, 2017; Shimwela *et al.*, 2016).

## Materials and methods

### Data for occurrence of myrtle rust

The current geographical distribution of myrtle rust was gathered from a range of published studies (Graça *et al.*, 2011; Elith *et al.*, 2013; Stewart *et al.*, 2017) and specialists (see Acknowledgements). GOOGLE EARTH v. 5.1.35 and ARCGIS v. 10.2 were used to extract the presence location coordinates from some published studies. A total of 380 points were initially collected, which were reduced to 160 points after removing duplicate points. This refinement of presence data was done based on the resolution of the climate data to avoid environmental bias due to possible spatially autocorrelated presence points (Brown *et al.*, 2017; Fig. 1; Table S1). Unlike previous researchers, the presence data were not grouped according to different strains or genotypes/biotypes of the rust (Elith *et al.*, 2013; Stewart *et al.*, 2017). This was because, so far, there is no evidence that these strains have different environmental requirements. To investigate the possibility of dependence of *A. psidii* on geographic location, a principal component analysis on environmental similarity of the presence data was performed.

### Environmental data

Historic climate data, including 19 bioclimatic variables, were acquired from WorldClim ([www.worldclim.org](http://www.worldclim.org)) and CliMond ([www.climond.org](http://www.climond.org)) for use in correlative models and CLIMEX, respectively. These variables were derived from long-term (1960–1990) monthly temperature and rain data that have widely been used in similar studies and provide 19 variables at global scale (Hijmans *et al.*, 2005); they are assumed to reflect the climate suitability for the growth and development of different organisms including plant pathogens (Elith *et al.*, 2013; Narouei-Khandan *et al.*, 2016; Brown *et al.*, 2017). The WorldClim dataset is one of few sources that also provide the same variables for future climate projections based on global climate models. This would be very beneficial if one is aiming to investigate and compare the effect of climate change on the species of interest. The resolution of both the WorldClim and CliMond data used in this study was 10 arc minutes (*c.* 18.6 km).

### Species distribution models

#### CLIMEX

CLIMEX is considered a semimechanistic model, with moderate complexity, that estimates the potential geographical

distribution of a species and its response to climate based on long-term climate data (Sutherst *et al.*, 2007; Khandan *et al.*, 2013; Kriticos *et al.*, 2013a). The process involves matching the species distribution and climate pattern using a calibration method rather than a statistical fitting process. Environmental needs of myrtle rust such as temperature (Elith *et al.*, 2013) and stress parameters were used in CLIMEX v. 4. The temperature parameter values of Kriticos *et al.* (2013a) were used as a starting point to calibrate potential pathogen growth. Stress indices that describe cold stress (CS), heat stress (HS), dry stress (DS), and combination of hot-dry (Kriticos *et al.*, 2013a) were used to investigate the species response to unfavourable conditions. In addition to parameters used by Kriticos *et al.* (2013a), hot-wet stress parameters were also applied.

Calibration of the CLIMEX model resulted in the final parameters described in Table 1, with an upper temperature threshold (DV<sub>3</sub>) of 30 °C. The upper temperature threshold (DV<sub>3</sub>) was adjusted to allow for the persistence of the pathogen in South American localities such as Quebracho and Tres Bocas in Uruguay, Paraguay and Espago Doeste in Brazil. The cold stress temperature threshold (TTCS) was set at 4 °C, which meant that the lower temperature threshold (DV<sub>0</sub> = 10 °C) did not constrain myrtle rust development. Soil moisture parameters were set to allow distribution of the pathogen in Maui, Hawaii and Santa Catarina and Espirito Santo in Brazil. Application of the dry-stress parameters resulted in the exclusion of the highlands around Petrolina, Brazil, northern Paraguay and southern Bolivia where the disease is not reported. The hot-wet stress parameters were also set to avoid over-prediction in hot and wet areas such as the Amazon region, some parts of central Africa and some parts of Florida. These adjustments resulted in unsuitability of parts of these areas for establishment of the pathogen. Detailed information about parameter fitting in the CLIMEX model can be found in Sutherst *et al.* (2007).

#### MaxEnt

MaxEnt is a correlative presence-only or presence-background model and has been claimed to outperform most of the existing correlative modelling approaches (Elith *et al.*, 2006; Merow *et al.*, 2013). This model compares presence points against

**Table 1** Parameter values used in developing CLIMEX model for myrtle rust.

Index	Parameter	Value	Unit
DV <sub>0</sub>	Lower temperature threshold	10	°C
DV <sub>1</sub>	Lower optimum temperature	15	°C
DV <sub>2</sub>	Upper optimum temperature	25	°C
DV <sub>3</sub>	Upper temperature threshold	30	°C
SM <sub>0</sub>	Lower soil moisture threshold	0.5	—
SM <sub>1</sub>	Lower optimum soil moisture	1	—
SM <sub>2</sub>	Upper optimum soil moisture	1.5	—
SM <sub>3</sub>	Upper soil moisture threshold	2	—
TTCS	Cold stress temperature threshold	4	°C
THCS	Cold stress temperature rate	−0.005	Week <sup>−1</sup>
DTCS	Cold stress degree-day threshold	15	°C
TTHS	Heat stress temperature threshold	35	°C
THHS	Heat stress temperature rate	0.002	Week <sup>−1</sup>
SMDS	Dry stress threshold	0.2	Week <sup>−1</sup>
HDS	Dry stress rate	−0.015	Week <sup>−1</sup>
TTHD	Hot-dry temperature threshold (0–50)	35	—
TTHW	Hot-wet temperature threshold (0–50)	33	°C
PHW	Hot-wet stress rate	0.04	Week <sup>−1</sup>

background sites, which are the locations where the presence or absence of the species is unknown. SDMTOOLBOX v. 2 (Brown *et al.*, 2017) was used to deal with sampling bias and to remove highly correlated variables after Pearson correlation tests. In addition, the random forest option in MMF was used to choose the most important variables among those that were not highly correlated. Because the number of presence locations comprised a large sample, linear, quadratic and hinge features were used to allow more complexity in the model. Additionally, response curves were produced to investigate the response of the pathogen to environmental factors and their effect on MaxEnt projection. Seventy-five percent of data were used to train/fit the model and 25% were set aside to test the model. If the extent of the background around the presence points is too large it can potentially result in failure of the model to predict the suitability of uncolonized locations that might be climatically suitable for the species (Brown *et al.*, 2017). To avoid this issue, the background optimization method suggested by Senay *et al.* (2013) was used for both MaxEnt and MMF. Thus, circles with a radius of 400 km around the presence points were found to be the most appropriate background extent (data not shown).

#### Multi-Model Framework (MMF)

The MMF models data of presence/absence of species using nine different species distribution models: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), logistic regression (LOG), naive Bayes (NB), classification and regression trees (CART), conditional trees (CTREE), *K*-nearest neighbour (KNN), support vector machines (SVM) and artificial neural networks (NNET) (Worner *et al.*, 2010). In MMF, a set of artificial absence points (pseudo-absences) are generated using one-class support vector machines (OCSVMs). Based on all the presence data, 100 ensemble models were fitted that proved to have the lowest prediction errors for absence of the pathogen. In this way, pseudo-absence points were generated that had a zero-probability of environmental suitability among all 100 ensemble models. This process resulted in many possible absence points, which were then clustered using *K*-mean clustering based on environmental similarity. From these clusters, the locations closest to the cluster centroid were selected as (pseudo-) absence points. The same bioclimatic variables selected by random forest for the MaxEnt model were also used in MMF.

#### Model validation

As stated earlier, the MaxEnt model was validated with 25% (randomly distributed) of the total dataset. Model validation in MMF was carried out by cross-validation (10-fold) and bootstrapping. Based on 10 different performance criteria (Table S2) calculated through bootstrapping and cross-validation, the models were ranked and the best model was selected accordingly (Table S2). The performance criteria were: accuracy, precision, recall (sensitivity), *F*-score, kappa index, specificity, true skill statistic (TSS), uncertainty,  $\chi^2$ , and area under curve (AUC). Further visual validation of all three models was carried out using the recent occurrence data of myrtle rust in New Zealand (New Zealand Ministry for Primary Industries, www.mpi.govt.nz; accessed 25 April 2019).

#### Model consensus

To identify the areas (or 'hotspots') that were projected as suitable for establishment of myrtle rust by all three modelling approaches, the output maps of CLIMEX, MaxEnt and MMF

were converted into binary maps in ArcMAP v. 10.2. To produce binary maps, the threshold of 10 percentile presence in the training data, occurrence threshold of 0.5 and an Ecoclimatic Index (EI) threshold of >1 were used for MaxEnt, MMF and CLIMEX, respectively (Jarnevich & Reynolds, 2011; Narouei-Khandan, 2014).

The performance of the consensus model was assessed by calculating the true positives and false negatives, estimating the sensitivity. Other performance data would not be possible to calculate as the consensus model is a combination of the individual models and independent absences/pseudo-absences are not available. A *z*-score test was performed between sensitivity values provided by the consensus model and the component models to investigate if the sensitivity of the consensus model was significantly better than those of the individual models.

#### Principal component analysis

To visualize the degree of similarity among the climates of global regions where myrtle rust occurs and New Zealand, principal component analysis (PCA) was performed on the 19 bioclimatic variables mentioned earlier. The bioclimatic data of the presence points were extracted in ArcGIS v. 10.2 and the data were transferred to R v. 3.4.3 where the PRINCOMP command was used to perform PCA analysis. The results (principal component 1 and 2) were transferred back and plotted in ArcGIS v. 10.2.

## Results

### Model performance, effect of environmental variables and potential distribution

#### Model performance

The CLIMEX model resulted in a very good fit to the current global distribution of myrtle rust. The calibrated model predicted areas as highly suitable ( $EI \geq 26$ ) that included locations where myrtle rust had been observed, with the exception of two points, one in Argentina and one in California, that were projected as being marginal for the establishment of the disease. The MaxEnt model gave an AUC of 0.74, which indicated a good model (better than a random model). In MMF, based on 10 performance criteria achieved by cross-validation and bootstrapping, the support vector machine (SVM) showed the best score, with an AUC of 0.95; therefore, MMF was ranked and selected as the best model (Table S2). All three models successfully predicted myrtle rust occurrences in New Zealand.

#### Environmental variables

The Pearson correlation test and random forest were used to remove highly correlated variables and to select the best candidate variables for both MaxEnt and MMF. This process resulted in the selection of 10 variables comprising: annual mean temperature (cbio01), mean of monthly temperature (max. temp. minus min. temp.) (cbio02), maximum temperature of the warmest month (cbio05), minimum temperature of the coldest month (cbio06), annual temperature range (cbio07), mean temperature of

the wettest month (cbio08), mean temperature of the driest quarter (cbio09), precipitation of the wettest month (cbio13), precipitation of the driest month (cbio14) and precipitation of the warmest quarter (cbio18). Analysis of variable contributions showed that temperature had the highest impact on myrtle rust distribution with the combination of annual temperature range (bio\_7) and annual mean temperature (bio\_1) contributing more than 66% to the prediction by the MaxEnt model (Table 2). The response curve of annual mean temperature (bio\_1) suggested that the probability of the presence of myrtle rust started to decrease above 25 °C (Fig. 2a). In addition, the response curve of precipitation in the warmest quarter (cbio18) showed that establishment of the pathogen decreased when precipitation of the warmest quarter exceeded 1800 mm (Fig. 2b). This confirmed the relevance of the hot-wet stress parameter used in the CLIMEX model (Table 1).

#### Potential distribution

All three models projected large areas in South America, Africa, East Asia, the Caribbean and southern USA as highly suitable for establishment of myrtle rust (Fig. 3a,b,c). However, there were apparent differences among the models in their projection, particularly in Europe and southern China. Europe was not projected suitable by SVM while there were suitable areas projected by CLIMEX and MaxEnt in the Balkans (Fig. 3a,b). Southern China was predicted as suitable by CLIMEX and MaxEnt (Fig. 3a,b) while SVM only projected the coastal area in the south as suitable (Fig. 3c). In South America, the projection of CLIMEX and MMF was in accordance with the current distribution of myrtle rust while the MaxEnt projection was generally low for South America with the exception of coastal eastern and southern areas in Brazil and Uruguay, which were projected to be highly suitable.

In Australia, all models projected coastal areas in eastern Australia as suitable, which matched the current distribution of the pathogen (Berthon *et al.*, 2018). In addition, small areas in western Australia and western China were projected as highly suitable by all three models where there are no reports of the pathogen yet

(Fig. 3a,b,c). The projection of all three models for New Zealand indicated that the Northland and Auckland areas in North Island (except the central highlands) were highly suitable for myrtle rust establishment. In addition, the rest of the North Island and small areas of the South Island around Nelson, Blenheim and Christchurch were projected as suitable by MaxEnt and CLIMEX (Fig. 3a,b,c). Compared to the Kriticos *et al.* (2013a) model, in the current CLIMEX model, the largest increases in EI values (up to 73 units) occurred in the North Island of New Zealand and in some areas in southern China. The most dramatic decreases in EI values (up to 53 units) were observed in South America, Central Africa, South-east Asia and the east coast of Australia. The model changes made some areas in South America and South-east Asia unsuitable or less suitable for myrtle rust (Figs S1 & S2).

Similarly, the differences between the predictions made by the MaxEnt model of Stewart *et al.* (2017) and the current MaxEnt model is apparent in areas of South America, China, South-east Asia, Australia and New Zealand (mostly in North Island).

#### Model consensus

The consensus model had an improved sensitivity (lower false negative rate) compared to its component models (Table 3). The results of the *z*-score test between the consensus model and the individual models indicated that the sensitivity of the consensus model was statistically significantly better than that of the SVM model, but it was not better than those of the CLIMEX and MaxEnt models (Table 3).

The consensus model highlighted the hotspots, or in other words the areas where at least two models projected climate suitability (Fig. 4). The total area of hotspots predicted by the consensus model was significantly smaller than the areas predicted as suitable by each individual model (Table S3). Almost all current presence data (156 out of 160) were located within hotspots projected as suitable by at least two models (Fig. 4). Areas where all three models agreed on suitability included Florida, the Caribbean, southern Brazil, southern Paraguay and Bolivia, eastern coastal areas in South Africa, eastern Madagascar, and eastern coastal areas in Australia, Papua New Guinea, southern China, southern Vietnam and Laos, Vanuatu and New Caledonia (Fig. 4). Other parts of the world were projected as being unsuitable or suitable by one or two models only.

In New Zealand, the models predicted that the most suitable areas for myrtle rust were in Northland, Auckland region and Waikato. All three models agreed on areas with suitable climates for establishment of myrtle rust. The whole of the North Island (except the central highlands) and the northern part of the South Island were predicted suitable by at least two models (Fig. 4).

#### Principal component analysis

The results of PCA showed that some of the myrtle rust presence points lie close to or within the New Zealand

**Table 2** Average contribution (out of 5000 iterations) of environmental variables to the prediction of the global distribution of myrtle rust caused by *Austropuccinia psidii* using MaxEnt.

Variable		Contribution (%)
bio_7	Annual temperature range (bio_5–bio_6)	34.3
bio_1	Annual mean temperature	19.4
bio_14	Precipitation of driest month	13.6
bio_9	Mean temperature of driest quarter	13.5
bio_18	Precipitation of warmest quarter	5.5
bio_5	Maximum temperature of warmest month	5.2
bio_6	Minimum temperature of coldest month	3.6
bio_2	Mean diurnal range (mean of monthly (max. temp.–min. temp.))	2.3
bio_13	Precipitation seasonality (coefficient of variation)	1.5
bio_8	Mean temperature of wettest quarter	1.0

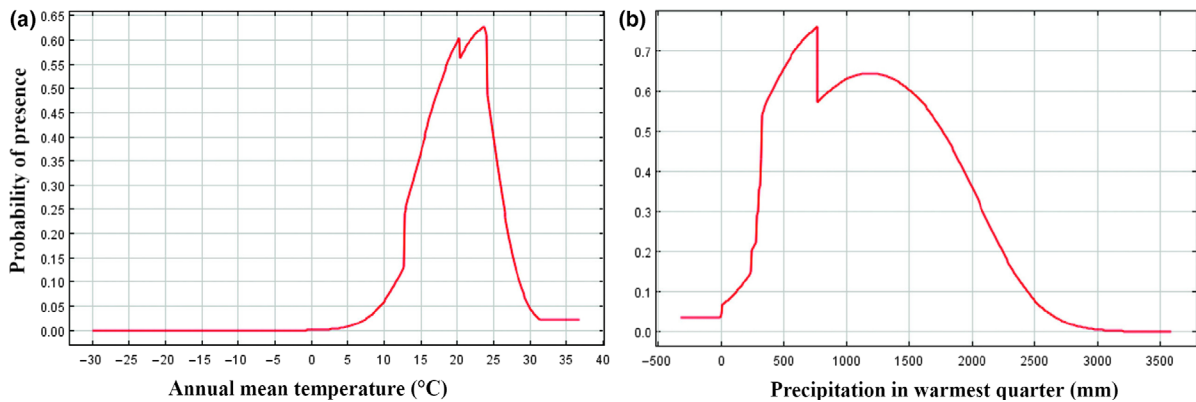


Figure 2 The probability of the establishment of myrtle rust in response to annual mean temperature (a) and precipitation of the warmest quarter (b), according to the model produced by MaxEnt. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)].

climate data space, which implies a close similarity in climate of global locations of myrtle rust and some areas in New Zealand (Fig. 5).

## Discussion

Several prediction models are already available for myrtle rust, but all of them rely on a single modelling approach (Magarey *et al.*, 2007; Kriticos & Leriche, 2008; Elith *et al.*, 2013; Kriticos *et al.*, 2013a; Alvares *et al.*, 2017; Stewart *et al.*, 2017; Berthon *et al.*, 2018). These individual models are associated with different sets of uncertainties in the predictions. The present study is the first to address the potential global and local establishment of myrtle rust using three different modelling methods (semimechanistic, presence-only and presence–(pseudo-) absence models). In addition, the modelling outcomes have been combined in a consensus model to account for uncertainties involved in individual models (Araújo *et al.*, 2005; Araújo & New, 2007).

All three models provided a good fit to the current distribution of myrtle rust worldwide. In addition, all three models successfully projected current areas of occurrence in New Zealand as suitable although these data were not used in the modelling process. There was a high degree of agreement among the models on the suitability of areas such as the eastern coastal parts of South Africa, Australia, the Caribbean and eastern coastal areas in Brazil. Nevertheless, MMF projected somewhat larger areas as suitable (in Brazil, China, central Africa and eastern Australia) than CLIMEX and MaxEnt. While MMF projected highly suitable areas in central and northwest Brazil, there are only a few reports of rust from these regions so far (Alvares *et al.*, 2017). This might be due to a lack of host plants or extensive surveys because the area is not easily accessible. For New Zealand, there was a good level of agreement among all three models, particularly in Northland and the Auckland region on the North Island. Both CLIMEX and MaxEnt projected most of the North Island as highly suitable (except central highlands) whereas the MMF (SVM model) projection was confined to coastal

areas of the North Island, which is in agreement with recent reports of myrtle rust in New Zealand ([www.mpi.govt.nz](http://www.mpi.govt.nz); accessed 25 April 2019).

The important role of temperature for myrtle rust establishment was highlighted by both CLIMEX and MaxEnt. In CLIMEX, none of the stress parameters used in calibration of the model had any effect on the projected establishment of myrtle rust in New Zealand and the model suggested that temperature has most effect on pathogen establishment in Northland where rain is not a limiting factor. The use of stress parameters in CLIMEX is one of the model's advantages as this allows characterization of areas with unfavourable conditions for species growth and establishment. For example, while all three models agreed that northern Mexico and southern California had unsuitable habitat, CLIMEX showed that the dry stress was the main limiting factor in those areas. The ability to identify specific limiting factors is not an option available in other models such as MaxEnt and MMF. In addition, CLIMEX was able to produce growth index charts for selected locations. For example, for Northland in New Zealand, CLIMEX indicated that the pathogen growth index was lower from May to early November than at other times of the year due to lower temperatures (Fig. 6).

The response curves produced by MaxEnt also provided useful information about the environmental factors that affect establishment of the pathogen. Both CLIMEX and MaxEnt indicated that the optimum temperature for establishment of myrtle rust is around 25 °C. When the temperature exceeded 25 °C, the probability of occurrence started to decrease, which was in agreement with available information on the temperature requirements of this pathogen (Dianese *et al.*, 1984; Tessmann *et al.*, 2001). Additionally, the response curve of precipitation in the warmest quarter implied that excessive precipitation combined with high temperature decreased the probability of myrtle rust establishment. The negative effect of rain on establishment of myrtle rust was mentioned in previous studies, which suggested excessive rain would wash out the urediniospores (Dianese *et al.*, 1984; Elith *et al.*, 2013; Zauza *et al.*, 2015).

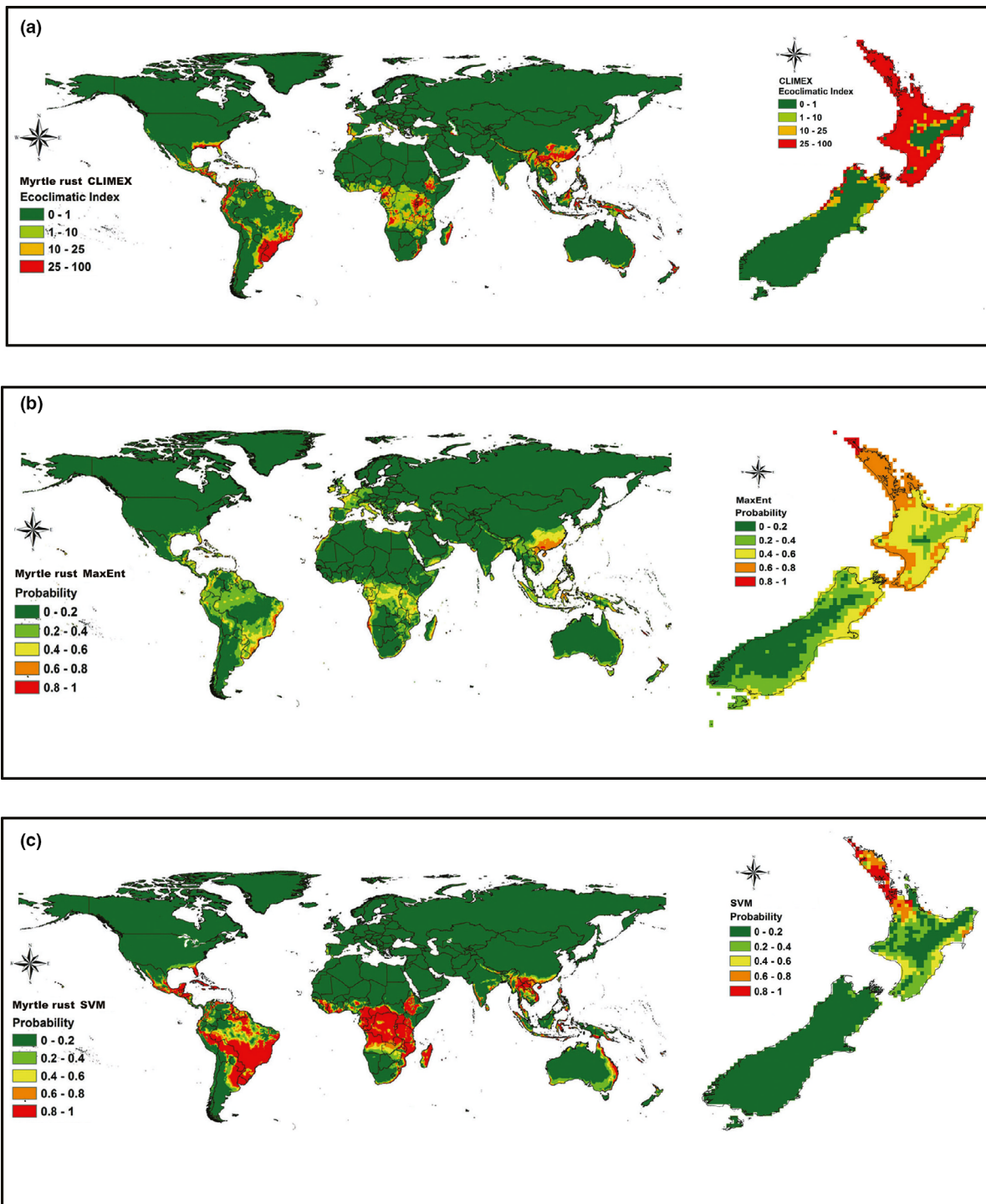


Figure 3 Potential suitability of the climate for myrtle rust at a global and New Zealand scale projected by the models developed using (a) CLIMEX, (b) MaxEnt, (c) support vector machine (SVM, the highest performing model in the multi-model framework). [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)].

The PCA showed that climate data of present locations of myrtle rust in Australia, Uruguay, Hawaii, Brazil and Japan are very similar to those of New Zealand, confirming the likelihood of pathogen establishment in large parts of New Zealand. The PCA did not show specific

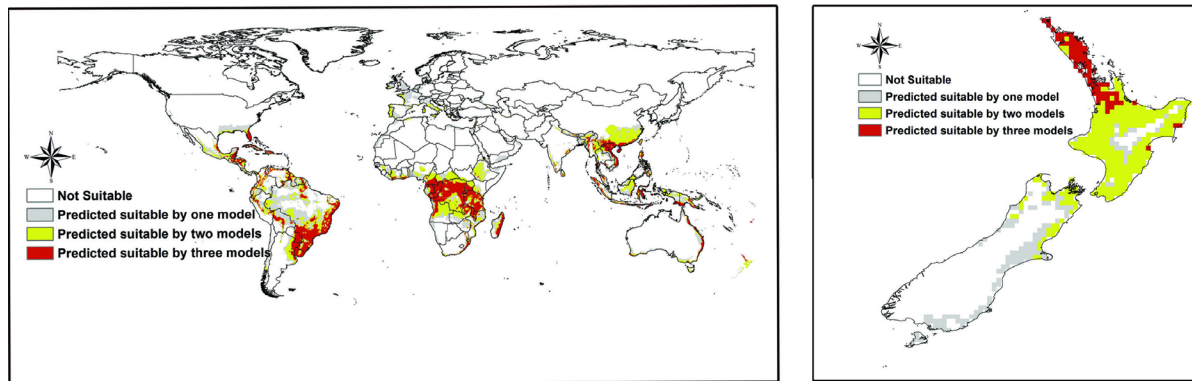
patterns and groupings (for example, according to potential genotype differences) among presence data in environmental space. More studies on the biology of myrtle rust are needed to investigate whether different strains have specific environmental responses. Modelling a

**Table 3** Comparison of accuracy measures of the consensus model and individual MaxEnt, CLIMEX and support vector machine (SVM) models for the projection of suitability for myrtle rust.

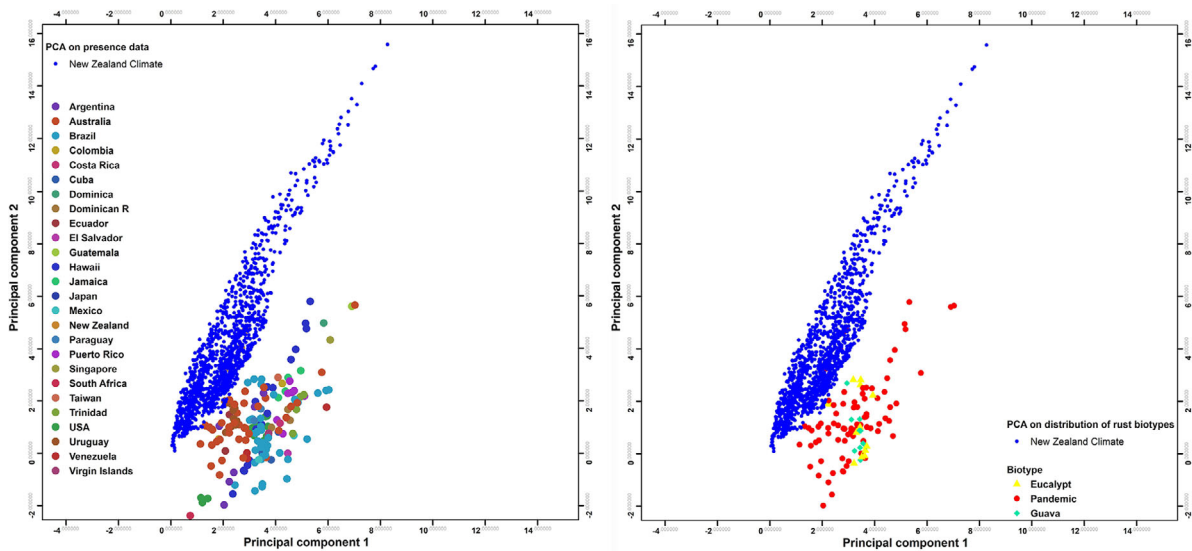
Model	Total (presence of myrtle rust)	True positive (TP)	False negative (FN)	Sensitivity (TPR) <sup>a</sup>	False negative rate (FNR) <sup>b</sup>
Consensus	160	156 (two models)	4	0.975	0.0250
MaxEnt		148	12	0.925	0.0750
CLIMEX		149	11	0.931	0.0688
SVM		139	21	0.868	0.1313

<sup>a</sup>TPR = TP/(TP + FN).

<sup>b</sup>FNR = FN/(FN + TP).



**Figure 4** Consensus model showing the hotspot areas, which two or three models both predicted as suitable for establishment of myrtle rust. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)].



**Figure 5** Principal component analysis showing similarity among climate of locations of myrtle rust (large spots) and New Zealand climate (small blue spots). The climate data were obtained from the WorldClim website ([www.worldclim.org](http://www.worldclim.org)). [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)].

species distribution based on presumed genotype/strain differences might be misleading as one may actually end up modelling host plant rather than pathogen distribution.

Hotspots with a high probability of establishment of myrtle rust were identified in the consensus model. The hotspot areas were mainly located in the

Caribbean, Central and South America, eastern coastal areas in Africa and Australia, as well as South-east Asia. These hotspot areas largely reflected the current presence data.

Similar to the present findings, the model of Booth *et al.* (2000) projected that coastal areas of northern New South Wales and Queensland in Australia would be

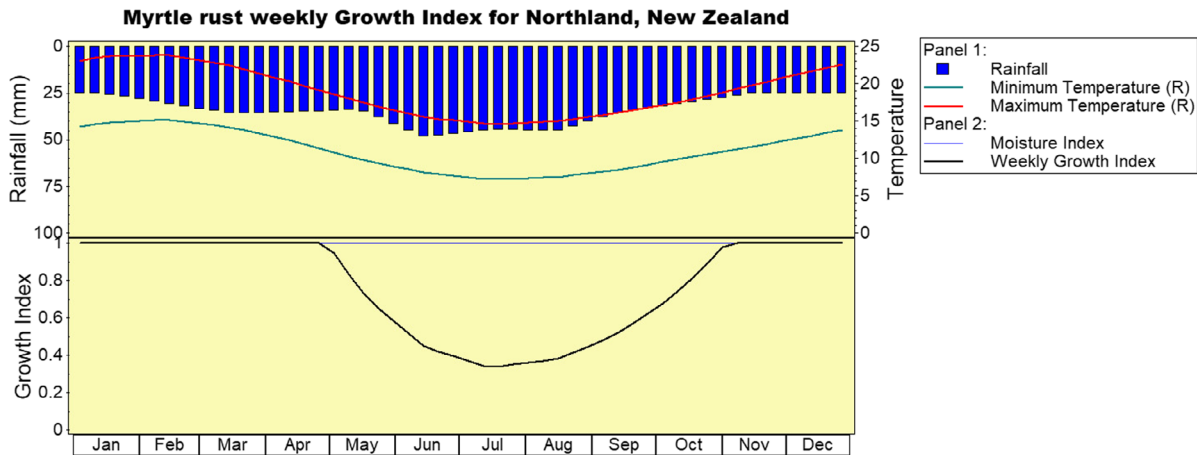


Figure 6 Weekly growth index of myrtle rust and climate variables around Auckland, New Zealand, produced by the CLIMEX model. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)].

suitable for establishment of myrtle rust. These regions, which have been affected by myrtle rust since 2010, were also projected as suitable by CLIMEX, MaxEnt and MMF in the current study. NAPPFAST (Magarey *et al.*, 2007) projected low suitability for the North Island of New Zealand and higher suitability for Central Otago on South Island, which is in contrast with the findings of the current and other existing models (Kriticos *et al.*, 2013a). Because the pathogen favours mild to warm temperatures, it might be expected that warmer parts of New Zealand would be more suitable for myrtle rust. Temperature parameter values, coarser resolution of climate data, fewer occurrence data and shorter data time-frame (10 year) used in NAPPFAST may have contributed to its improbable projection for New Zealand (Kriticos *et al.*, 2008). The model projections in the current study were in accordance with studies by Kriticos *et al.* (2008, 2013a) that predicted that most of the North Island is suitable for myrtle rust establishment using CLIMEX. While the CLIMEX parameter values used in the present study were similar to those used by Kriticos *et al.* (2013a), hot-wet stress parameters were also successfully integrated in CLIMEX, which were not used by Kriticos *et al.* (2013a). This integration of hot-wet stress parameters resulted in avoiding over-prediction in most parts of the Amazon and some parts of central Africa without compromising the projection for Florida. The role of hot-wet stress, caused by excessive rain and high temperatures, was confirmed by MaxEnt. Similar to the current findings, the Kriticos *et al.* (2013a) model, which combined the model results with distribution of the myrtle rust hosts in Australia, identified temperate temperature and moisture as the main factors in establishment of myrtle rust.

Using MaxEnt for their global distribution study, Stewart *et al.* (2017) projected lower probabilities for southern China, central Amazon, central Africa and East Asia than the model output presented here, probably because they used fewer presence data in those regions. Moreover, their presence data were not rarefied and the multicollinearity among predictor variables was not accounted

for, which could have affected model output (Miller, 2012). In addition to prediction of the general occurrence of myrtle rust, Stewart *et al.* (2017) and Berthon *et al.* (2018) developed models for different pathogen biotypes. While distinct biotypes of *A. psidii* that vary in virulence have been distinguished (Silva *et al.*, 2014; Stewart *et al.*, 2017), there is no evidence that these strains respond differently to environmental conditions.

Prediction models, particularly species distribution models (SDMs), could be valuable tools in risk assessment of serious biosecurity species (Araújo *et al.*, 2005; Kriticos *et al.*, 2013b). To properly interpret the results of SDMs, the uncertainties involved should be considered. The particular SDM used may actually contribute most to the uncertainty of the model output (Thibaud *et al.*, 2014; Watling *et al.*, 2015). However, there are other sources of uncertainty such as species presence data and their rarefaction and potential autocorrelations that can affect model prediction. There are also general limitations associated with each of the SDMs. For example, the parameter fitting process in the CLIMEX model is subjective; one may develop similar projection maps with various combinations of different growth and stress indices or different projection maps with similar climate data. Although the same climate data were used in the CLIMEX model by Kriticos *et al.* (2013a) and the CLIMEX model in the current study, differences in projected suitability of some areas between these studies can be attributed to the increased number of presence data used in the current model calibration. Kriticos *et al.* (2013a) used around 60 points at the global scale compared to the 160 points in the current model. Another source of uncertainty in the CLIMEX model output is the visual validation that is commonly used because CLIMEX models lack a standard statistical validation method.

Other modelling methods have their own limitations. MaxEnt has been criticized for its poor transferability to new regions (Townsend Peterson *et al.*, 2007; Merow *et al.*, 2013). Other settings in MaxEnt such as feature selection, sample bias, spatial autocorrelation, and

background data extent, which can affect the model projection, have been discussed by Merow *et al.* (2013). Also, the presence–pseudo-absence models used in MMF do not provide information such as the importance of selected variables or response curves (Worner *et al.*, 2010; Narouei-Khandan, 2014).

Despite these caveats of individual models, the consensus model presented here, composed of three different types of models as recommended by Araújo & New (2007), minimized the uncertainties and resulted in predictions that reflected the observed occurrences and the predictions from previous models. The consensus model had the highest sensitivity and lowest false negative rate, with 156 out of 160 presence points lying within areas predicted as suitable. While consensus resulted in improved sensitivity compared to the component models, this improvement was not statistically significant. Models with high sensitivity generally minimize false negatives (omission error, under-prediction), but may result in increased false positives (commission error, over-prediction). However, minimizing false negatives is more important than false positives in the case of biosecurity threats from invasive species and models with low omission error should be preferred (Ward, 2006; Webber *et al.*, 2011). Although over-prediction may potentially result in increased biosecurity measures or management strategies where they are not needed, this issue can be managed by prioritizing high risk areas based on prior knowledge of species biology.

In conclusion, the most accurate projection of myrtle rust is attained by combining the output of several models and using good quality input data. It is difficult to fully control this disease in natural habitats because the pathogen mostly spreads by wind (Zauza *et al.*, 2015), but major efforts to limit its spread are warranted considering the potential economic, ecological and sociocultural impact of myrtle rust in New Zealand.

## Acknowledgements

The authors would like to thank the Bio-Protection Research Centre and Better Border Biosecurity ([www.b3nz.org](http://www.b3nz.org)) for funding this research. They also thank Drs Jane Elith, J. Stewart, Ned Klopfenstein, John Hanna, A. J. Carnegie, J. Y. Uchida and A. C. Alfenas for providing the distribution data. The authors would also like to thank Philippa Stevens, Mark Bullians and Dr Ursula Torres for their valuable comments on this manuscript.

## References

Alvares C, Sentelhas P, Mattos E *et al.*, 2017. Climatic favourability zones for eucalyptus rust in Brazil. *Forest Pathology* 47, e12301.  
 Araújo MB, New M, 2007. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* 22, 42–7.  
 Araújo MB, Whittaker RJ, Ladle RJ *et al.*, 2005. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* 14, 529–38.  
 Beenken L, 2017. *Austropuccinia*: a new genus name for the myrtle rust *Puccinia psidii* placed within the redefined family Sphaerophragmiaceae (Pucciniales). *Phytotaxa* 297, 53–61.

Berthon K, Esperon-Rodriguez M, Beaumont L *et al.*, 2018. Assessment and prioritisation of plant species at risk from myrtle rust (*Austropuccinia psidii*) under current and future climates in Australia. *Biological Conservation* 218, 154–62.  
 Booth TH, Jovanovic T, 2012. Assessing vulnerable areas for *Puccinia psidii* (eucalyptus rust) in Australia. *Australasian Plant Pathology* 41, 425–9.  
 Booth TH, Old KM, Jovanovic T, 2000. A preliminary assessment of high risk areas for *Puccinia psidii* (Eucalyptus rust) in the Neotropics and Australia. *Agriculture, Ecosystems & Environment* 82, 295–301.  
 Brown JL, Bennett JR, French CM, 2017. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* 5, e4095.  
 Carnegie A, Lidbetter J, Walker J *et al.*, 2010. *Uredo rangelii*, a taxon in the guava rust complex, newly recorded on Myrtaceae in Australia. *Australasian Plant Pathology* 39, 463–6.  
 Carnegie AJ, Kathuria A, Pegg GS *et al.*, 2016. Impact of the invasive rust *Puccinia psidii* (myrtle rust) on native Myrtaceae in natural ecosystems in Australia. *Biological Invasions* 18, 127–44.  
 Clark S. 2011. Risk analysis of the *Puccinia psidii* guava rust fungal complex (including *Uredo rangelii* myrtle rust) on nursery stock. Wellington, New Zealand: Ministry of Agriculture and Forestry. [<https://www.nzffa.org.nz/assets/452/puccinia-psidii-on-nursery-stock-ra.pdf>]. Accessed 22 October 2019.  
 Coutinho TA, Wingfield MJ, Alfenas AC *et al.*, 1998. Eucalyptus rust: a disease with the potential for serious international implications. *Plant Disease* 82, 819–25.  
 Dianese JC, Moraes TSDA, Silva AR, 1984. Response of *Eucalyptus* species to field infection by *Puccinia psidii*. *Plant Disease* 68, 314–6.  
 Elith J, Graham CH, Anderson RP *et al.*, 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–51.  
 Elith J, Simpson J, Hirsch M *et al.*, 2013. Taxonomic uncertainty and decision making for biosecurity: spatial models for myrtle/guava rust. *Australasian Plant Pathology* 42, 43–51.  
 Ferreira FA, 1981. Ferrugem do eucalipto ocorrência, temperatura para germinação deuredosporos, produção de teliosporos, hospedeiro alternativo e resistência. *Fitopatologia Brasileira* 6, 91–109.  
 Graça RN, Alfenas AC, Ross-Davis AL, *et al.*, 2011. Multilocus genotypes indicate differentiation among *Puccinia psidii* populations from South America and Hawaii. In: Fairweather ML, Palacios P, eds. *Proceedings of the 58th Annual Western International Forest Disease Work Conference, 2011*. Flagstaff, AZ, USA: Department of Agriculture, Forest Service, AZ Zone Forest Health, 131–4. [[https://www.fs.fed.us/rm/pubs\\_other/rmrs\\_2011\\_graca\\_r001.pdf](https://www.fs.fed.us/rm/pubs_other/rmrs_2011_graca_r001.pdf)]. Accessed 14 Oct 2019.  
 Hijmans RJ, Cameron SE, Parra JL *et al.*, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–78.  
 Jarnevich CS, Reynolds LV, 2011. Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. *Biological Invasions* 13, 153–63.  
 Khandan HN, Worner S, Jones E *et al.*, 2013. Predicting the potential global distribution of *Pseudomonas syringae* pv. *actinidiae* (Psa). *New Zealand Plant Protection* 66, 184–93.  
 Kriticos DJ, Leriche A. 2008. The current and future potential distribution of guava rust, *Puccinia psidii* in New Zealand. *MAF Biosecurity New Zealand Technical Paper*. Rotorua, New Zealand: Scion.  
 Kriticos DJ, Morin L, Leriche A *et al.*, 2013a. Combining a climatic niche model of an invasive fungus with its host species distributions to identify risks to natural assets: *Puccinia psidii* sensu lato in Australia. *PLoS ONE* 8, e64479.  
 Kriticos DJ, Maitre DC, Webber BL, 2013b. Essential elements of discourse for advancing the modelling of species' current and potential distributions. *Journal of Biogeography* 40, 608–11.  
 Loope L. 2010. A summary of information on the rust *Puccinia psidii* Winter (guava rust) with emphasis on means to prevent introduction

- of additional strains to Hawaii. *US Geological Survey Open-File Report*, 1082.
- Magarey RD, Fowler GA, Borchert DM *et al.*, 2007. NAPPFAST: An internet system for the weather-based mapping of plant pathogens. *Plant Disease* **91**, 336–45.
- Marlatt R, Kimbrough J, 1979. *Puccinia psidii* on *Pimenta dioica* in south Florida. *Plant Disease Reporter* **63**, 510–2.
- Merow C, Smith MJ, Silander JA, 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**, 1058–69.
- Miller JA, 2012. Species distribution models: spatial autocorrelation and non-stationarity. *Progress in Physical Geography* **36**, 681–92.
- Narouei-Khandan HA, 2014. *Ensemble Models to Assess the Risk of Exotic Plant Pathogens in a Changing Climate*. Lincoln, New Zealand: Lincoln University, PhD thesis.
- Narouei-Khandan HA, Halbert SE, Worner SP *et al.*, 2016. Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling approaches, with emphasis on the USA. *European Journal of Plant Pathology* **144**, 655–70.
- Narouei-Khandan H, Harmon C, Harmon P *et al.*, 2017. Potential global and regional geographic distribution of *Phomopsis vaccinii* on *Vaccinium* species projected by two species distribution models. *European Journal of Plant Pathology* **148**, 919–30.
- Pegg G, Taylor T, Entwistle P *et al.*, 2017. Impact of *Austropuccinia psidii* (myrtle rust) on Myrtaceae-rich wet sclerophyll forests in south east Queensland. *PLoS ONE* **12**, e0188058.
- du Plessis E, McTaggart A, Granados G *et al.*, 2017. First report of myrtle rust caused by *Austropuccinia psidii* on *Rhodomyrtus tomentosa* (Myrtaceae) from Singapore. *Plant Disease* **101**, 1676.
- Ramsfield T, Dick M, Bulman L, Ganley R, 2010. *Briefing Document on Myrtle Rust, A Member of the Guava Rust Complex, and the Risk to New Zealand*. Rotorua, New Zealand: Scion. [<https://www.nzffa.org.nz/images/design/Briefing-paper-myrtle-rust.pdf>]. Accessed 14 Oct 2019.
- Roux J, Greyling I, Coutinho TA *et al.*, 2013. The myrtle rust pathogen, *Puccinia psidii*, discovered in Africa. *IMA Fungus* **4**, 155–9.
- Senay SD, Worner SP, Ikeda T, 2013. Novel three-step pseudo-absence selection technique for improved species distribution modelling. *PLoS ONE* **8**, e71218.
- Shimwela MM, Narouei-Khandan HA, Halbert SE *et al.*, 2016. First occurrence of *Diaphorina citri* in East Africa, characterization of the *Ca. Liberibacter* species causing huanglongbing (HLB) in Tanzania, and potential further spread of *D. citri* and HLB in Africa and Europe. *European Journal of Plant Pathology* **146**, 349–68.
- Silva ACD, Andrade PMTD, Alfenas AC *et al.*, 2014. Virulence and impact of Brazilian strains of *Puccinia psidii* on Hawaiian 'Ōhi 'a (*Metrosideros polymorpha*). *Pacific Science* **68**, 47–56.
- Soewarto J, Carriconde F, Hugot N *et al.*, 2018. Impact of *Austropuccinia psidii* in New Caledonia, a biodiversity hotspot. *Forest Pathology* **48**, e12402.
- Stewart J, Ross-Davis A, Graça R *et al.*, 2017. Genetic diversity of the myrtle rust pathogen (*Austropuccinia psidii*) in the Americas and Hawaii: global implications for invasive threat assessments. *Forest Pathology* **48**, e12378.
- Sutherst RW, Maywald GF, Kriticos DJ, 2007. *CLIMEX Version 3. User's Manual*. Melbourne, Australia: Hearne Scientific Software Pty Ltd.
- Tessmann D, Dianese J, Miranda A *et al.*, 2001. Epidemiology of a neotropical rust (*Puccinia psidii*): periodical analysis of the temporal progress in a perennial host (*Syzygium jambos*). *Plant Pathology* **50**, 725–31.
- Thibaud E, Petitpierre B, Broennimann O *et al.*, 2014. Measuring the relative effect of factors affecting species distribution model predictions. *Methods in Ecology and Evolution* **5**, 947–55.
- Tobias PA, Park RF, Külheim C *et al.*, 2015. Wild-sourced *Chamelaucium uncinatum* have no resistance to *Puccinia psidii* (myrtle rust). *Australasian Plant Disease Notes* **10**, 15.
- Townsend Peterson A, Papeš M, Eaton M, 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography* **30**, 550–60.
- Uchida J, Loope L, 2009. A recurrent epiphytotic of guava rust on rose apple, *Syzygium jambos*, in Hawaii. *Plant Disease* **93**, 429.
- Ward DF, 2006. Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions* **9**, 723–35.
- Watling JI, Brandt LA, Bucklin DN *et al.*, 2015. Performance metrics and variance partitioning reveal sources of uncertainty in species distribution models. *Ecological Modelling* **309**, 48–59.
- Webber BL, Yates CJ, Le Maitre DC *et al.*, 2011. Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions* **17**, 978–1000.
- Worner SP, Ikeda T, Leday G *et al.*, 2010. *Surveillance Tools for Freshwater Invertebrates*. MAF Biosecurity Technical Paper. New Zealand: Ministry of Agriculture and Forestry.
- Yong WTL, Ades PK, Bossinger G *et al.*, 2019. Geographical patterns of variation in susceptibility of *Eucalyptus globulus* and *Eucalyptus obliqua* to myrtle rust. *Tree Genetics & Genomes* **15**, 31.
- Zauza E, Lana V, Maffia L *et al.*, 2015. Wind dispersal of *Puccinia psidii* urediniospores and progress of eucalypt rust. *Forest Pathology* **45**, 102–10.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

**Table S1.** The geographic coordinates of locations of myrtle rust, of which the most recent (New Zealand data) were used to validate model outputs.

**Table S2.** The accuracy measures achieved by bootstrapping (a) and cross-validation (b) in the Multi-Model Framework developed for myrtle rust. Acronyms for models are: QDA, quadratic discriminant analysis; NB, naive Bayes; LDA, linear discriminant analysis; LOG, logistic regression; CART, classification and regression tree; CTREE, conditional tree; KNN, *K*-nearest neighbor; SVM, support vector machine; and NNET, neural network.

**Table S3.** The areas predicted as suitable for myrtle rust by the consensus model and component models. The average area of each pixel  $\approx 18.6 \times 18.6$  km.

**Figure S1.** Comparison of changes in value of the ecoclimatic index (EI) of the CLIMEX model developed in the current study and the CLIMEX model developed by Kriticos *et al.* (2013a). The warm colours show the most extreme increase in EI values of the current model compared to Kriticos *et al.* (2013a) and the cold colours show the most extreme decrease in EI values compared to Kriticos *et al.* (2013a).

**Figure S2.** A comparison between predictions of areas suitable for myrtle rust made by the Kriticos *et al.* (2013a) CLIMEX model (top) and the CLIMEX model developed in the current study (below).