RESEARCH PAPERS

# Potential distribution of *Xylella fastidiosa* in Italy: a maximum entropy model

Luciano BOSSO[1], Danilo RUSSO[1,2], Mirko DI FEBBRARO[3], Gennaro CRISTINZIO[1] and Astolfo ZOINA[4]

[1] Dipartimento di Agraria, Università degli studi di Napoli Federico II, Via Università n. 100, 80055 Portici, Napoli, Italy
[2] School of Biological Sciences, University of Bristol, Woodland Road BS8 1UG, Bristol, United Kingdom
[3] EnvixLab, Dipartimento Bioscienze e Territorio, Università del Molise, Pesche, Italy
[4] CNR – Istituto per la Protezione Sostenibile delle Piante, UOD Portici, Via Università n. 133, 80055 Portici, Italy

**Summary.** Species distribution models may provide realistic scenarios to explain the influence of bioclimatic variables in the context of emerging plant pathogens. *Xylella fastidiosa* is a xylem-limited Gram-negative bacterium causing severe diseases in many plant species. We developed a maximum entropy model for *X. fastidiosa* in Italy. Our objectives were to carry out a preliminary analysis of the species' potential geographical distribution and determine which eco-geographical variables may favour its presence in other Italian regions besides Apulia. The analysis of single variable contribution showed that precipitation of the driest (40.3%) and wettest (30.4%) months were the main factors influencing model performance. Altitude, precipitation of warmest quarter, mean temperature of coldest quarter, and land cover provided a total contribution of 19.5%. Based on the model predictions, *X. fastidiosa* has a high probability (> 0.8) of colonizing areas characterized by: i) low altitude (0–150 m a.s.l.); ii) precipitations in the driest month < 10 mm, in the wettest month ranging between 80–110 mm and during the warmest quarter < 60 mm; iii) mean temperature of coldest quarter ≥ 8°C; iv) agricultural areas comprising intensive agriculture, complex cultivation patterns, olive groves, annual crops associated with permanent crops, orchards and vineyards; forest (essentially oak woodland); and Mediterranean shrubland. Species distribution models showed a high probability of *X. fastidiosa* occurrence in the regions of Apulia, Calabria, Basilicata, Sicily, Sardinia and coastal areas of Campania, Lazio and south Tuscany. Maxent models achieved excellent levels of predictive performance according to area under curve (AUC), true skill statistic (TSS) and minimum difference between training and testing AUC data (AUC_diff). Our study indicated that *X. fastidiosa* has the potential to overcome the current boundaries of distribution and affect areas of Italy outside Apulia.

**Key words:** emerging pest, EPPO, Maxent, olive quick decline syndrome, species distribution models.

## Introduction

Species distribution models can give understanding of the ecological and evolutionary factors that determine the spatial patterns exhibited by organisms (Guisan and Zimmermann, 2000). They have been used to study the relationships between environmental parameters and species richness (Guisan and Thuiller, 2005; Sinclair *et al.*, 2010; Wisz *et al.*, 2013), carry out analyses regarding species at risk (Doko *et*

*al.*, 2011; Bosso *et al.*, 2013; Roscioni *et al.*, 2013; 2014; Russo *et al.*, 2015), and make predictions about species distributions in past climatic conditions (Russo *et al.*, 2014) or future/projected situations (Rebelo *et al.*, 2010). Species distribution models also provide increasingly realistic scenarios to explain the influence of bioclimatic variables on plant diseases and epidemiology of pathogens, especially in the context of emerging plant diseases (Fabre *et al.*, 2011; Aguayo *et al.*, 2014).

Plant pathogens may expand their distributions or increase their pathogenic action due to modified environmental conditions favourable to disease de-

velopment in new regions (Garrett *et al.*, 2006). While numerous studies have considered the introduction of animal and plant species to new regions (e.g. Guisan and Zimmerman, 2000; Guisan and Thuiller, 2005), few have assessed habitat suitability for plant pathogens including fungi, viruses and bacteria. When plant pathogens reach new areas, they often cause dramatic effects on agricultural and native plant communities. Famous cases for Europe and the Mediterranean include *Erysiphe necator* (Glawe, 2008), *Ophiostoma ulmi* (Solla *et al.*, 2005), *Ceratocystis platani* (Anselmi *et al.*, 1994), and *Erwinia amylovora* (Jock *et al.*, 2002).

*Xylella fastidiosa* (Wells *et al.*, 1987) is a xylem-limited Gram-negative bacterium which is the causal agent of severe diseases in plants of prime importance for agriculture (alfalfa dwarf, almond leaf scorch, citrus variegated chlorosis, coffee leaf scorch, Pierce's disease of grapevine and phony peach disease) and forests (elm leaf scorch, oak leaf scorch, oleander leaf scorch, and sycamore leaf scorch) (Janse and Obradovic, 2010; Purcell, 2013). Many wild plants, such as grasses, sedges and forest trees may carry this pathogen showing symptoms, or, more often, asymptomatically (Janse and Obradovic, 2010). *Xylella fastidiosa* is native to the Americas, ranging from USA to South America, and mainly occurs in tropical/subtropical areas, although leaf scorch diseases also occur in much colder climates (Janse and Obradovic, 2010). The latest information about the worldwide distribution of this bacterium is available from the European and Mediterranean Plant Protection Organization (EPPO). *Xylella fastidiosa* has been assigned quarantine status in the EPPO A1 list. The regions affected are requested to adopt immediate measures including carrying out extensive surveys of susceptible hosts in order to identify the contaminated areas and determine the most appropriate disease management strategies (Janse and Obradovic, 2010; Carlucci *et al.*, 2013; Loconsole *et al.*, 2014). Once restricted to the Americas, the bacterium was first discovered in Europe near Gallipoli (Lecce province) of southern Italy in October 2013 (Loconsole *et al.*, 2014). Since the initial outbreak it has spread across the Apulia region affecting approximately 10,000 ha of olive trees (Saponari *et al.*, 2014).

*Xylella fastidiosa* is transmitted by various species of sap-sucking hopper insects. Each infection occurs after a vector has fed on an infected plant and then subsequently feeds on a healthy plant. In Apulia, *X. fastidiosa* is vectored to olive trees by the spittlebug, *Philaenus spumarius* (Hemiptera: Aphrophoridae) (Saponari *et al.*, 2014). However, other insect vectors are suspected to have also spread the bacterium in the region (Elbeaino *et al.*, 2014), including *Neophilaenus campestris* Fallén (Hemiptera: Aphrophoridae) and *Euscelis lineolatus* Brullé (Hemiptera: Cicadellidae), an aspect awaiting further investigations. Hoddle (2004) modelled the potential distribution of *X. fastidiosa* in California and projected the model to the rest of the world. The model was developed with CLIMEX (Sutherst and Maywald, 1985) using the climatic response of the pathogen from the native geographic range. The main prediction was that cold temperatures would not allow *X. fastidiosa* to colonize France and the northern and central areas of grape production in Spain and Italy.

In Italy *X. fastidiosa* has been studied through conventional field and laboratory approaches, as well as using serological and molecular assays (Carlucci *et al.*, 2013; Saponari *et al.*, 2013; Cariddi *et al.*, 2014; Loconsole *et al.*, 2014; Saponari *et al.*, 2014). However, no study has addressed the ecological requirements of the pathogen at a broad-scale. A fundamental reason to investigate the potential geographic distribution of *X. fastidiosa* is that the resulting information could be used to determine the suitability of an area at risk of invasion in order to adopt urgent and strict control measures and prevent further propagation and dissemination of the pathogen.

In this study, we developed a maximum entropy, presence-only distribution model (Maxent; Phillips *et al.*, 2006) for *X. fastidiosa*, to provide a preliminary analysis of its potential geographical distribution and determine which eco-geographical variables (EGVs) may favour its presence in other Italian regions outside Apulia.

## Materials and methods

### Study area

We considered the entire Italian territory between latitudes 45°N and 36°N and longitudes 6°E and 18°E (corresponding to ca. 301,000 km², elevation range = 0–4810 m a.s.l.) (Figure 1).

### Data collection

Records for *X. fastidiosa* were collected from the online database set up by the Apulia Region govern-

**Figure 1.** Study areas considered to model the potential distribution of *Xylella fastidiosa* in Italy: Apulia (training region, black) and the rest of Italy (projection area, black + dark grey).

ment: http://webapps.sit.puglia.it/freewebapps/ MonitoraggioXFSintesi. We used only observations referred to olive trees infected by *X. fastidiosa* in 2014–2015 (for further details on the records see: http://cartografia.sit.puglia.it/doc/Uso_del_WebGIS.pdf) (Table S.1). The resulting database included 1,258 records mostly concentrating in the southernmost part of Apulia. Records were screened in ArcGis (version 9.2) for spatial autocorrelation using average nearest neighbour analyses and Moran's I measure of spatial autocorrelation to remove spatially correlated data points (e.g. Russo *et al.*, 2014; 2015). The average nearest neighbour analysis calculates a nearest neighbour index based on the average distance from each feature to its nearest neighbouring feature. Moran's I measures spatial autocorrelation based on feature locations. After this selection, 50 fully independent presence records for *X. fastidiosa* were used to generate species distribution models (Table S.1).

**Selection of ecogeographical variables**

To carry out species distribution models for *X. fastidiosa*, we used a set of 21 Ecogeographical Vari-

ables (EGVs). We included altitude, 19 bioclimatic variables and land cover. Altitude and the 19 bioclimatic variables were obtained from the WorldClim database (www.worldclim.org/current) (Hijmans *et al.*, 2005). The data layers were generated through interpolation of average monthly climate data from weather stations on a 30 arc-second resolution grid (corresponding to a 1 km$^2$ resolution). Land cover was obtained from the Corine Land Cover IV level for the year 2012 (available at http://www.sinanet. isprambiente.it/it/sia-ispra/download-mais/co-rine-land-cover/corine-land-cover-2012/view). The complete list of all Corine Land Cover classes of the fourth level is available at http://www2.dmu.dk/1_ Viden/2_Miljoe-ilstand/3_natur/nordlam/nldocs/ wsOct01T1/jferanec.pdf. All variable formats were raster files (grid) with a 30-arc second resolution (0.93 × 0.93 km = 0.86 km$^2$ at the equator). To select the number of variables for the final distribution models, we first eliminated the highly correlated predictors by retaining those with a Pearson's |r| ≤ 0.80. From this first set of predictors, we considered those most relevant to the ecological requirements of *X. fastidiosa*, according to expert opinion and current knowledge (Hoddle, 2004; Janse and Obradovic, 2010; Carlucci *et al.*, 2013; Saponari *et al.*, 2013; Cariddi *et al.*, 2014; Loconsole *et al.*, 2014; Saponari *et al.*, 2014). This led to a final set of 11 variables (Table 1) used for model training.

**Habitat suitability procedure**

We used Maxent ver. 3.3.3k (http://www. cs.princeton.edu/~schapire/maxent) (Phillips *et al.*, 2006) to model the potential distribution of *X. fastidiosa* in Italy. This approach relied on presence-only datasets and is especially useful when presence data are scarce. A target probability distribution is estimated by identifying the probability distribution of maximum entropy (most spread out, or closest to uniform), subject to a set of constraints that constitute the incomplete information about the target distribution. The pixels of the study area represent the space on which the Maxent probability distribution is defined: pixels with known species occurrence records constitute the sample points, and the environmental predictors comprise climatic variables, altitude, land use type or other environmental variables (Phillips *et al.*, 2006). This algorithm usually results in good predictive models compared with other presence-

**Table 1.** List of Ecogeographical Variables used in this study, their types and measurement units.

| Type | Ecogeographical Variable | Unit |
|---|---|---|
| Topographical | Altitude | m |
| Habitat | Land cover | - |
| Climatic | Annual Mean Temperature | °C |
| | Mean Temperature of Wettest Quarter | °C |
| | Mean Temperature of Driest Quarter | °C |
| | Mean Temperature of Warmest Quarter | °C |
| | Mean Temperature of Coldest Quarter | °C |
| | Precipitation of Wettest Month | mm |
| | Precipitation of Driest Month | mm |
| | Precipitation of Warmest Quarter | mm |
| | Precipitation of Coldest Quarter | mm |

only models (e.g. Elith *et al.*, 2006). It is based on a generative approach, rather than a discriminative one, which can give important advantages when the amount of training data is limited (Phillips *et al.*, 2006). Furthermore, it has good ability to predict new localities for poorly known species (Rebelo and Jones, 2010; Bosso *et al.*, 2013; Russo *et al.*, 2014; Russo *et al.*, 2015). To build the models, we used the presence records of *X. fastidiosa* selected as described above and the EGVs shown in Table 1. In the setting panel, we selected the following options: random seed, remove duplicate presence records, write plot data, regularization multiplier (fixed at 1), and 1000 maximum iterations. These settings are conservative enough to allow the algorithm to get close to convergence and the best performance (Phillips *et al.,* 2006); and, finally, 50 replicate effects with cross-validate replicated run type. For the latter procedure, 70% of records were randomly extracted to train the model and the remaining 30% to test it, and the procedure was repeated 50 times. In this way the training/testing datasets used by Maxent are independent of each other and selected at random. The 30% testing data subset is the most appropriate when dealing with datasets including a limited number of records

(Pearson *et al.*, 2007). A set of 10,000 background points was then randomly placed over the Apulia region. The average final map obtained had a logistic output format with suitability values from 0 (unsuitable habitat) to 1 (suitable habitat). We selected the 10th percentile (the value above which the model classifies correctly 90% of the training locations) as the threshold value to define the presence of *X. fastidiosa*. This is a conservative value commonly used in species distribution modelling studies especially when considering datasets gathered over a long time by different observers and methods of collection (e.g.: Rebelo and Jones, 2010; Bosso *et al.*, 2013; Russo *et al.*, 2015). We used this threshold to reclassify our model into binary presence/absence maps. Jacknife analysis was used to estimate the actual contribution that each variable provided to the geographic distribution models. During this process, Maxent generated three models. Firstly, each EGV was excluded in turn and a model was created with the remaining variables to check which one of these was the most informative. Secondly, a model was created by each individual EGV to detect which variable had the most information not featuring in the other variables. Thirdly, a model was generated based on all variables. Response curves derived from univariate models were plotted to assess how each EGV influenced presence probability. The model was projected to the whole Italian territory. In order to project the models calibrated in Apulia over the rest of Italy, the variables in the projection area must meet a condition of environmental similarity to the environmental data used to train the model. Therefore, we first ascertained that this condition was met by inspecting the Multivariate Environmental Similarity Surfaces (MESS) generated by Maxent (Elith *et al.*, 2010).

**Model validation**

We tested the predictive performance of the models with different methods: the receiver operated characteristics, analyzing the area under curve (AUC) (Fielding and Bell, 1997); the true skill statistic (TSS) (Allouche *et al.*, 2006); and the minimum difference between training and testing AUC data (AUC$_{diff}$) (Warren and Seifert, 2011). Such statistics were averaged across the 50 replicates run on the 70% (training) versus 30% (testing) dataset split. AUC assesses the discrimination ability of the models and its value ranges from 0 (equalling random distribution) to 1

(perfect prediction). AUC values > 0.75 correspond to high discrimination performances (Fielding and Bell, 1997). TSS compares the number of correct forecasts, minus those attributable to random guessing, to that of a hypothetical set of perfect forecasts. It considers both omission and commission errors, and success as a result of random guessing; its values range from -1 to +1, where +1 corresponds to perfect agreement and zero or less to a performance no better than random (Allouche *et al.,* 2006). The difference between training and test AUC data is minimized to reduce the risk that models are over-parameterized in such a way as to be overly specific to the training data (Warren and Seifert, 2011).

## Results

### Species distribution models for *Xylella fastidiosa* and model validation

Our model showed excellent levels of predictive performance, as indicated from AUC, TSS and AUC$_{diff}$ values (Table 2). The model predicted a high probability of presence at the Apulia's region tip (Figures 2a and 3a) and in Calabria, Basilicata, Sicily, Sardinia and coastal areas of Campania, Lazio and south Tuscany (Figures 2b and 3b). MESS analysis showed that the projection area had a medium to high environmental similarity with several regions in the training area (Table 3). The model achieved a regularized gain value of 1.8, indicating a good fit with occurrence data (Figure S.1). Six variables contributed to a total 90% of model prediction. The analysis of single variable contributions showed that precipitation during driest (40.3%) and wettest (30.4%) months were the main factors influencing model performance. Altitude, precipitation during warmest quarter, mean temperature for warmest and coldest quarter, and land cover provided a total contri-

bution of 19.5%. Based on the model's predictions, *X. fastidiosa* has a greater probability of occurring (> 0.8) at low altitudes (0–150 m a.s.l.) in areas characterized by low precipitation in the driest month (< 10 mm) and in the warmest quarter (< 60 mm), medium precipitation in the wettest month (80–110 mm), and by mean temperature > 8°C in the coldest quarter. Land cover analysis revealed that *X. fastidiosa* might essentially occur in: i) agricultural areas (75.5%) with intensive agriculture, complex cultivation patterns, olive groves, annual crops associated with permanent crops, orchards and vineyards; ii) forest (12.8%) mostly oak woodland; and iii) Mediterranean shrubland (11.8%) (Table 4).

**Table 3.** Potential distribution of *Xylella fastidiosa* in the Italian regions (percent regional surface area potentially suitable for the bacterium) and environmental similarity between training and projection regions, according to the Maxent model and MESS analysis, respectively.

| Region | Potential distribution (%) | Environmental similarity (Category) |
|---|---|---|
| Sicily | 100 | Medium |
| Sardinia | 100 | Medium |
| Calabria | 93 | Medium |
| Puglia | 35 | High |
| Lazio | 35 | Medium |
| Campania | 32 | High |
| Basilicata | 30 | High |
| Tuscany | 15 | High |
| Other | 0 | Low |

**Table 2.** Predictive performances (mean and standard deviation values) achieved for species distribution models developed for *Xylella fastidiosa* in Apulia and the rest of Italy.

| Territory | AUC Training | SD | AUC Test | SD | AUCdiff | SD | TSS | SD |
|---|---|---|---|---|---|---|---|---|
| Apulia | 0.971 | 0.030 | 0.952 | 0.039 | 0.019 | 0.001 | 0.854 | 0.024 |
| Italy | 0.982 | 0.021 | 0.966 | 0.081 | 0.016 | 0.086 | 0.814 | 0.011 |

AUC = Area under curve; AUC$_{diff}$ = Minimum difference between training and testing AUC data; TSS = True skill statistic; SD = Standard Deviation.
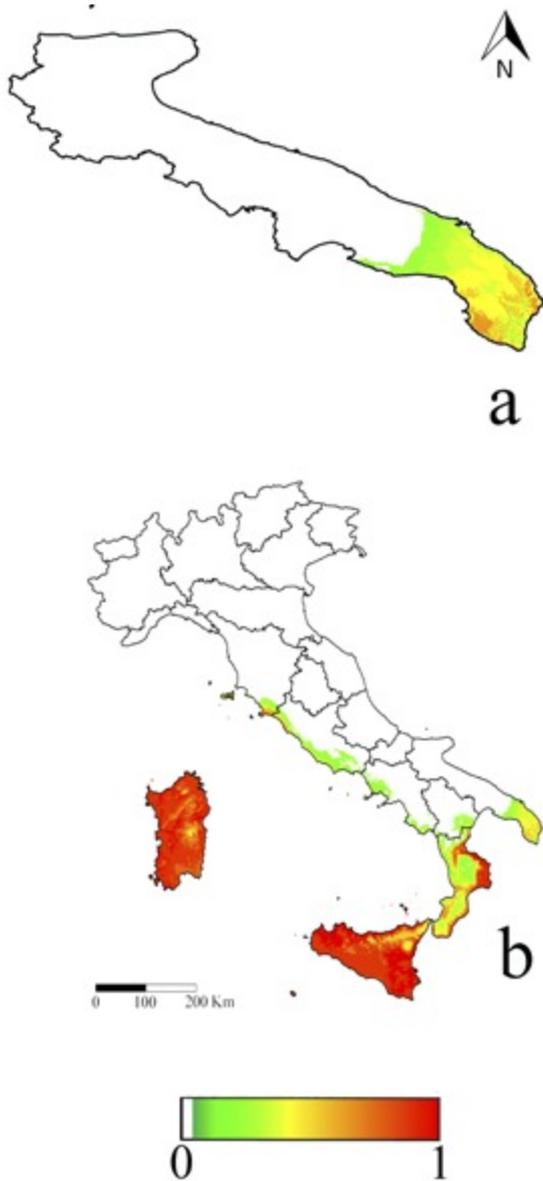
**Figure 2.** Species distribution models of *Xylella fastidiosa* in: Apulia (a) calibrated with Apulia records only; and Italy (b) calibrated with presence records from Apulia.

**Figure 3.** Binary map of *Xylella fastidiosa* in: (a) Apulia calibrated with Apulia records only; and (b) Italy calibrated with presence records from Apulia.

## Discussion

### Model performance

Our model detected a set of environmental variables useful for assessing the potential distribution of *X. fastidiosa* in Italy. The model's prediction power was firstly supported by the high gain value (1.8) achieved (e.g. Bosso *et al.*, 2013). In general, models whose AUC > 0.75 are regarded as reliable (Elith, 2002). In our case, AUC values > 0.9 demonstrate a very high predictive capacity – among the greatest obtained for published models (e.g. Rebelo and Jones, 2010; Domíguez-Vega *et al.*, 2012; Russo *et al.*, 2014). The model's high predictive capacity was also confirmed by the excellent values of $AUC_{diff}$ and TSS (Russo *et al.*, 2014; 2015). The projections of our mod-

**Table 4.** Land use classified after Corine Land Cover 2012 IV level, according to the MaxEnt model developed for *Xylella fastidiosa* in Italy.

| Macro category | Corine land cover IV level | Suitable habitat (%) |
|---|---|---|
| Agriculture | Intensive agriculture | 22.49 |
| | Olive groves | 20.54 |
| | Complex cultivation patterns | 18.15 |
| | Orchards | 5.77 |
| | Arable land with significant areas of natural vegetation | 4.49 |
| | Annual crops associated with permanent crops | 4.01 |
| Forest | Quercus ilex/Quercus suber and evergreen woods | 5.07 |
| | Deciduous mixed oaks woods | 3.82 |
| | Agroforestal areas | 2.22 |
| | Mixed woodland and shrubs | 1.67 |
| Mediterranean shrubland | Mediterranean shrubs (high and low shrubs) | 11.77 |

el to other regions of the Italian territory was legitimate, as shown by MESS analysis.

**Model limitations**

Although our model showed an excellent statistical validation and led to a robust map of the potential distribution of *X. fastidiosa*, caution is needed when interpreting its results for two main reasons.

First, some limitations may arise from the absence of "realized niche" predicting factors, such as biotic factors, including plant-pathogen-antagonist interactions. Our model did not take into account the potential distribution of the bacterium's vector. *Philaenus spumarius*, the known vector of *X. fastidiosa* in Italy (Saponari *et al.*, 2014), is widespread and ecologically flexible (Halkka *et al.*, 1967) so its presence is unlikely to act as a limiting factor for the expansion of *X. fastidiosa*. Besides, several other insect species are likely to carry the pathogen (Elbeaino *et al.*, 2014), reinforcing the *X. fastidiosa* invasion potential. Overall, we suggest that the reliability of our predictions is not significantly affected by the absence of the vector influence in the model.

Second, the WorldClim dataset we employed has its own limitations, because: 1) it is obtained by interpolating records from weather stations according to latitude, longitude and elevation; and 2) it is de-

veloped from records dating back to at least 15 years ago. However, these limitations apply to many other studies because WorldClim data are routinely used to model species potential distribution (Rebelo *et al.*, 2010; Doko *et al.*, 2011; Domíguez-Vega *et al.*, 2012; Bosso *et al.*, 2013; Aguayo *et al.*, 2014; Russo *et al.*, 2014; 2015).

We advise caution in interpreting our results and note that our model should not be interpreted as a risk map, which would have required the inclusion of further variables, such as natural (e.g. wind speed or direction) or artificial features (human-made infrastructures) potentially influencing invasion routes. As with most invasive species, infected vectors may also be potentially spread by accidental transportation by humans (Primack, 2014), making the resulting picture even more difficult to predict. Nonetheless, we have detected the most important environmental predictors occurring in the areas surrounding the likely introduction site and predicted the pathogen's potential distribution based on this analysis.

**Predicting the potential distribution of *Xylella fastidiosa* in Italy**

The two most important variables characterizing the potential distribution of *X. fastidiosa* were precipitation during the driest and wettest months, corresponding to a 70.4 % overall contribution. The

importance of low levels of precipitation in the driest month (< 10 mm), medium levels of precipitation in the wettest month (80–110 mm) and high mean temperature of the coldest quarter (> 8°C) may likewise also be explained in terms of climatic requirements of the bacterium and of its determined vector in Italy i.e. *P. spumarius* (Halkka *et al.,* 1967; Ejere and Okpara, 2010). The areas of greatest probability of presence corresponded to intensive agriculture – especially of olive trees – and to Mediterranean shrubland (low shrubs) situated at low elevations between ca. 0–150 m a.s.l (Table 4), particularly along the Tyrrhenian and Ionian coast (Figures 2 and 3). Calabria, Sicily and Sardinia have the greatest probability of offering suitable habitats for *X. fastidiosa* (Figure 2b; Table 3). Furthermore, the potential to attack vineyards, citrus and peach orchards by *X. fastidiosa* should not be underestimated, as this bacterium causes Pierce's and Phony peach diseases with extensive damage to the wine and peach industries in the USA (Purcell, 1997; Perring *et al.*, 2001). Furthermore, natural ecosystem such as Mediterranean shrubland (high shrubs) and oak forests (e.g. *Quercus ilex, Quercus suber, Quercus pubescens, Quercus robur* and *Quercus cerris*) could also be attacked by *X. fastidiosa* in internal areas of Calabria, Lazio, Sicily and Sardinia. As in the American continent, we have predicted a preference for internal and coastal areas with mild winters at low altitudes. For instance, in the USA, *X. fastidiosa* occurs with winter temperatures ≥ 1–4°C (Purcell, 1997). By contrast, the pathogen is probably less harmful where winter temperatures are colder, such as at higher altitudes, because freezing events can strongly decrease *X. fastidiosa* infections (Hopkins and Purcell, 2002; Hoddle, 2004).

A fundamental requirement for the establishment of any species outside its native range is that the recipient location must have a climate comparable to that of the invader's original range. Within agreement with findings obtained using CLIMEX by Hoddle (2004), our model confirms that Mediterranean climates can be particularly favourable for *X. fastidiosa*. Hoddle (2004) predicted that the bacterium could colonize the Italian central and southern grape-growing areas. Here the climate is especially favourable for its vector. By contrast, cold stress would not allow the bacterium to spread in the grape-growing provinces of Northern Italy. Cold stress represents the principal factor limiting the northern range of *X. fastidiosa* (Hoodle, 2004). Our model identified Tus-

cany as the latitudinal limit for the range expansion of *X. fastidiosa*. European, and especially Euro-Mediterranean areas, are characterized by a broad spectrum of climatic conditions and plant species, providing the bacterium with high chances to encounter hosts and environmental conditions favouring invasion and establishment.

*Xylella fastidiosa* is an emerging threat for Italy and the entire Mediterranean basin, not only for the grape industries (Hoddle, 2004) but also for stonefruit and citrus cultivations, as well as Mediterranean shrub and forest species, as confirmed by our analysis. Besides Apulia, other Italian regions such as Basilicata, Campania, Tuscany, Lazio, Sicily and Sardinia are at risk, with potential highly detrimental consequences for the agro-forestry systems found there. The present research emphasises that these regions should be targeted with appropriate phytosanitary preventive measures to avoid the spread of the bacterium and the severe diseases it causes.

## Acknowledgments

## Literature cited

Aguayo J., F. Elegbede, C. Husson, F.X. Saintonge and B. Marçais, 2014. Modeling climate impact on an emerging disease, the *Phytophthora alni* - induced alder decline. *Global Change Biology* 20, 3209–3221.

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

Anselmi N., L. Cardin L. and G. Nicolotti, 1994. Plane decline in European and Mediterranean countries: associated pests and their interactions. *EPPO Bulletin* 24, 159-171.

Bosso L., H. Rebelo, A.P. Garonna and D. Russo, 2013. Modelling geographic distribution and detecting conservation gaps in Italy for the threatened beetle *Rosalia alpina*. *Journal for Nature Conservation* 21, 72–80.

Cariddi C., M. Saponari, D. Boscia, A. De Stradis, G. Loconsole, F. Nigro, F. Porcelli, O. Potere and G.P. Martelli, 2014. Isolation of a *Xylella fastidiosa* strain infecting olive and oleander in apulia, Italy. *Journal of Plant Pathology* 96, 1-5.

Carlucci A., F. Lops, G. Marchi, L. Mugnai and G. Surico, 2013. Has *Xylella fastidiosa* "chosen" olive trees to establish in the Mediterranean basin? *Phytopathologia Mediterranea* 52, 541–544.

Doko T., H. Fukui, A. Kooiman, A.G. Toxopeus, T. Ichinose, W. Chen and A.K. Skidmore, 2011. Identifying habitat patch-

es and potential ecological corridors for remnant Asiatic black bear (*Ursus thibetanus japonicus*) populations in Japan. *Ecological Modelling* 222, 748–761.

Domíguez-Vega H., O. Monroy-Vilchis, C.J. Balderas-Valdivia, C.M. Gienger and D. Ariano-Sánchez, 2012. Predicting the potential distribution of the beaded lizard and identification of priority areas for conservation. *Journal for Nature Conservation* 20, 247–253.

Ejere V.C. and H.T. Okpara, 2010. Aspects of the ecology of spittlebugs (Homoptera: Cercopidae) in Nsukka, south east, Nigeria. *Animal Research International* 7, 1242–1252.

Elbeaino T., Yaseen T., Valentini F., Ben Moussa I.E., Mazzoni V. and A.M. D'onghia, 2014. Identification of three potential insect vectors of *Xylella fastidiosa* in southern Italy. *Phytopathologia Mediterranea* 1, 328–332.

Elith J., 2002. Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. In S. Ferson, & M. Burgman (Eds.), Quantitative methods for conservation biology. New York: Springer-Verlag.

Elith J., C.H. Graham, R.P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, L. Jin, L.G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. Mcc Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz and N.E. Zimmermann, 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151.

Elith J., M. Kearney and S. Phillips, 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1, 330–342.

Fabre B., D. Piou, M.L. Desprez-Loustau and A. Marçais, 2011. Can the emergence of pine Diplodia shoot blight in France be explained by changes in pathogen pressure linked to climate change? *Global Change Biology* 17, 3218–3227.

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

Garrett K.A., S.P. Dendy, E.E. Frank, M.N. Rouse and S.E. Travers, 2006 Climate change effects on plant disease: genomes to ecosystems. *Annual Review of Phytopathology* 44, 489–509.

Glawe D.A., 2008. The powdery mildews: a review of the world's most familiar (yet poorly known) plant pathogens. *Annual Review of Phytopathology* 46, 27–51.

Guisan A. and W. Thuiller, 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8, 993–1009.

Guisan A. and N.E. Zimmermann, 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147–186.

Halkka O., M. Raatikainen, A. Vasarainen and A. Heinonen, 1967. Ecology and ecological genetics of *Philaenus spumarius* (L.) (Homoptera). *Annales Zoologici Fennici* 4, 1–18.

Hijmans R.J., L. Guarino, M. Cruz and E. Rojas, 2005. Computer tools for spatial analysis of plant genetic resources data. 1. DIVA-GIS. *Plant Genetic Resources Newsletter* 127, 15–19.

Hoddle M.S., 2004. The potential adventive geographic range of glassy-winged sharpshooter, *Homalodisca coagulata* and

the grape pathogen *Xylella fastidiosa*: implications for California and other grape growing regions of the world. *Crop Protection* 23, 691–699.

Hopkins D.L. and A.H. Purcell, 2002. *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. *Plant Disease* 86, 1056–1066.

Janse J.D. and A. Obradovic, 2010. *Xylella fastidiosa*: its biology, diagnosis, control and risks. *Journal of Plant Pathology* 92, S1.35–S1.48.

Jock S., V. Donat, M.M. López, C. Bazzi and K. Geider, 2002. Following spread of fire blight in Western, Central and Southern Europe by molecular differentiation of *Erwinia amylovora* strains with PFGE analysis. *Environmental Microbiology* 4, 106–114.

Loconsole G., O. Potere, D. Boscia, G. Altamura, K. Djelouah, T. Elbeaino, D. Frasheri, D. Lorusso, F. Palmisano, P. Pollastro, M.R. Silletti, N. Trisciuzzi, F. Valentini, V. Savino and M. Saponari, 2014. Detection of *Xylella fastidiosa* in olive trees by molecular and serological methods. *Journal of Plant Pathology* 96, 1–8.

Pearson R.G., C.J. Raxworthy, M. Nakamura and A.T. Peterson, 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34, 102–117.

Perring T.M., C.A. Farrar and M.J. Blua, 2001. Proximity to citrus influences Pierce's disease in Temecula Valley vineyards. *California agriculture* 55, 13–18.

Phillips S.J., R.P. Anderson and R.E. Schapire, 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190, 231–259.

Primack R.B., 2014. Essentials of Conservation Biology, 6th Edition, Sinauer, Sunderland, 603 pp.

Purcell A.H., 1997. *Xylella fastidiosa*, a regional problem or global threat? *Journal of Plant Pathology* 79, 99–105.

Purcell A.H., 2013. Paradigms: Examples from the bacterium *Xylella fastidiosa*. *Annual Review of Phytopathology* 51, 229–356.

Rebelo H. and G. Jones, 2010. Ground validation of presence-only modelling with rare species: A case study on *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology* 47, 410–420.

Rebelo H., P. Tarroso and G. Jones, 2010. Predicted impact of climate change on European bats in relation to their biogeographic patterns. *Global Change Biology* 16, 561–576.

Roscioni F., D. Russo, M. Di Febbraro , L. Frate, M.L. Carranza and A. Loy, 2013 Regional-scale modelling of the cumulative impact of wind farms on bats. *Biodiversity and Conservation* 22, 1821–1835.

Roscioni F., H. Rebelo, D. Russo, M.L. Carranza, M. Di Febbraro and A. Loy, 2014. A modelling approach to infer the effects of wind farms on landscape connectivity for bats. *Landscape Ecology* 29, 891–903.

Russo D., M. Di Febbraro, H. Rebelo, M. Mucedda, L. Cistrone, P. Agnelli, P. P. De Pasquale, A. Martinoli, D. Scaravelli, C. Spilinga and L. Bosso, 2014. What story does geographic separation of insular bats tell? A case study on Sardinian Rhinolophids. *PLoS ONE* 9: e110894. doi: 10.1371/journal.pone.0110894. pmid:25340737.

Russo D., M. Di Febbraro, L. Cistrone, G. Jones, S. Smeraldo, A.P. Garonna and L. Bosso, 2015. Protecting one, protect-

ing both? Scale-dependent ecological differences in two species using dead trees, the rosalia longicorn beetle and the barbastelle bat. *Journal of Zoology*, DOI: 10.1111/jzo.12271.

Saponari M., D. Boscia, F. Nigro and G.P. Martelli, 2013. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). *Journal of Plant Pathology* 95, 668.

Saponari M., G. Loconsole, D. Cornara, R.K. Yokomi, A. De Stradis, D. Boscia, D. Bosco, G.P. Martelli, R. Krugner and F. Porcelli, 2014. Infectivity and transmission of *Xylella fastidiosa* by *Philaenus spumarius* (Hemiptera: Aphrophoridae) in Apulia, Italy. *Journal of economic entomology* 107, 1316–1319.

Sinclair S.J., M.D. White and G.R. Newell, 2010. How useful are species distribution models for managing biodiversity under future climates? *Ecology and Society* 15, 8.

Solla A., J. Bohnens, E. Collin *et al.*, 2005. Screening European elms for resistance to *Ophiostoma novo-ulmi*. *Forest ScIence* 51, 134–141.

Sutherst R.W. and G.F. Maywald, 1985. A computerised system for matching climates in ecology. *Agriculture Ecosystems and Environment* 13, 281–299.

Warren D.L. and S.N. Seifert, 2011. Environmental niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Application* 21, 335–342.

Wells J.M., B.C. Raju, H.Y. Hung, W.G. Weisburg, L. Mandelco-Paul and D.J. Brenner, 1987. *Xylella fastidiosa* gen. nov., sp. nov.: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *International Journal of Systematic Bacteriolology* 37, 136–143.

Wisz M.S., J. Pottier, W.D. Kissling, L. Pellissier, J. Lenoir, C.F. Damgaard, C.F. Dormann, M.C. Forchhammer, J.A. Grytnes, A. Guisan, R.K. Heikkinen, T.T. Høye, I. Kühn, M. Luoto, L. Maiorano, M.C. Nilsson, S. Normand, E. Öckinger, N.M. Schmidt, M. Termansen, A. Timmermann, D.A.Wardle, P. Aastrup and J.C. Svenning, 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews of the Cambridge Philosophical Society* 88, 15–30.

**Table S.1.** Latitude, longitude e altitude of *X. fastidiosa* records used in Maxent model.

| No. | Latitude | Longitude | Altitude (m) | No. | Latitude | Longitude | Altitude (m) |
|---|---|---|---|---|---|---|---|
| 1 | 40.502 | 17.615 | 125 | 17 | 40.244 | 18.229 | 96 |
| 2 | 40.442 | 17.992 | 28 | 18 | 40.420 | 18.062 | 47 |
| 3 | 40.224 | 18.279 | 95 | 19 | 40.345 | 17.908 | 60 |
| 4 | 40.133 | 18.358 | 91 | 20 | 40.427 | 18.086 | 38 |
| 5 | 39.841 | 18.299 | 97 | 21 | 40.340 | 18.014 | 38 |
| 6 | 39.940 | 18.256 | 149 | 22 | 40.303 | 17.859 | 27 |
| 7 | 40.112 | 18.438 | 74 | 23 | 40.278 | 17.948 | 47 |
| 8 | 40.156 | 18.272 | 82 | 24 | 40.281 | 18.069 | 44 |
| 9 | 40.141 | 18.477 | 25 | 25 | 39.924 | 18.130 | 39 |
| 10 | 40.076 | 18.032 | 19 | 26 | 39.959 | 18.060 | 56 |
| 11 | 40.028 | 18.061 | 32 | 27 | 39.877 | 18.224 | 100 |
| 12 | 40.042 | 18.192 | 145 | 28 | 40.273 | 18.242 | 47 |
| 13 | 39.996 | 18.266 | 118 | 29 | 40.182 | 17.996 | 38 |
| 14 | 40.082 | 18.464 | 98 | 30 | 40.424 | 18.164 | 21 |
| 15 | 40.093 | 18.204 | 99 | 31 | 40.303 | 17.858 | 27 |
| 16 | 40.178 | 18.135 | 82 | 32 | 39.936 | 18.097 | 42 |

| No. | Latitude | Longitude | Altitude (m) | No. | Latitude | Longitude | Altitude (m) |
|-----|----------|-----------|--------------|-----|----------|-----------|--------------|
| 33 | 39.877 | 18.224 | 100 | 42 | 40.361 | 18.109 | 35 |
| 34 | 40.281 | 18.069 | 44 | 43 | 40.445 | 17.886 | 12 |
| 35 | 40.203 | 18.414 | 27 | 44 | 40.088 | 18.265 | 54 |
| 36 | 40.454 | 18.160 | 6 | 45 | 39.986 | 18.162 | 59 |
| 37 | 40.293 | 18.108 | 46 | 46 | 40.117 | 18.103 | 95 |
| 38 | 40.273 | 18.242 | 47 | 47 | 40.235 | 18.028 | 35 |
| 39 | 40.094 | 18.204 | 99 | 48 | 40.062 | 18.123 | 116 |
| 40 | 40.061 | 18.342 | 82 | 49 | 40.286 | 18.182 | 46 |
| 41 | 39.996 | 18.265 | 118 | 50 | 40.156 | 18.421 | 49 |
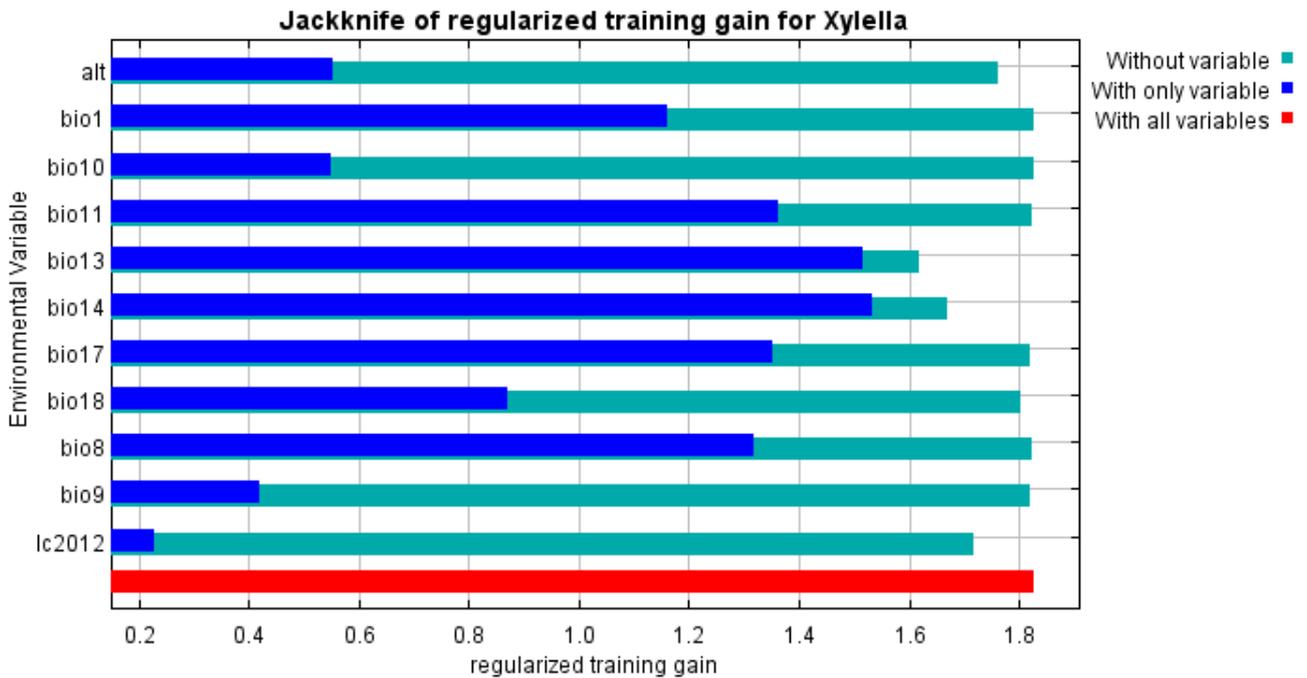


**Figure S.1.** Jackknife of regularized training gain for *X. fastidiosa*. The EGVs used for *X. fastidiosa* models were: Alt = Altitude; bio1 = Annual Mean Temperature; bio10 = Mean Temperature of Warmest Quarter; bio11 = Mean Temperature of Coldest Quarter; bio13 = Precipitation of Wettest Month; bio14 = Precipitation of Driest Month; bio17 = Precipitation of Driest Quarter; bio18 = Precipitation of Warmest Quarter; bio8 = Mean Temperature of Wettest Quarter; bio9 = Mean Temperature of Driest Quarter; lc2012 = Corine Land Cover 2012.