agronomy

MDPI

Article

Suitability Maps of Bactrocera Oleae Presence by SDM Based on Pedo-Climatic and Topographic Predictors Data in Sicily

Giuseppe Antonio Catalano [1], Giovanni Pirrello [2], Provvidenza Rita D'Urso [2,*] and Claudia Arcidiacono [2]

[1] Department of Veterinary Sciences, University of Messina, Polo Universitario dell'Annunziata Viale Giovanni Palatucci n. 23, 98168 Messina, Italy
[2] Building and Land Engineering Section, Department of Agriculture, Food and Environment (Di3A), University of Catania, Via S. Sofia, 100, 95123 Catania, Italy; giovanni.pirrello@hotmail.com (G.P.); carcidi@unict.it (C.A.)
* Correspondence: provvidenza.durso@unict.it

Abstract

Climate change and increasingly restrictive pesticide regulations have created a growing need for new tools to support the integrated pest management (IPM) of the olive fruit fly, *Bactrocera oleae*, in cultivated areas of the Mediterranean. In this study, the environmental suitability for this phytophagous insect in eastern Sicily was mapped by using geographic information system (GIS) tools and species distribution models (i.e., Random Forest and MaxEnt). The models were trained on presence data of the fly, obtained from a network of pheromone traps and locations where olive trees were present, combined with climatic, topographic and soil predictors for both current conditions and the future climate scenario (2021–2040). Correlation analysis was utilised to select ten predictors from an initial set of 33 soil and climate variables. Model performance was evaluated by using 10-fold cross-validation based on accuracy measures Area Under the Curve (AUC), True Skill Statistic (TSS), and the difference between the training and testing AUC) to minimise overfitting. Both algorithms demonstrated excellent predictive performance, producing convergent suitability maps, with high values concentrated in the foothills and hills of the Iblean–Calatino area and low values along the coastal plains and at higher altitudes, where extreme temperatures and unfavourable soil textures reduce habitat suitability. Response curves highlighted the combined influence of moderate temperature and precipitation seasonality, balanced topsoil texture, and moderate slopes in defining the species' ecological niche. The proposed framework provides an operational basis for optimising monitoring networks and targeting IPM measures under current and near-future climate conditions.

Keywords: species distribution modelling; GIS; *Bactrocera oleae*; territorial analysis; IPM; climate projections

Academic Editors: Vjekoslav Tadić, Dorijan Radočaj, Mladen Jurišić and Francis Drummond

Received: 26 November 2025
Revised: 18 February 2026
Accepted: 21 February 2026
Published: 24 February 2026

1. Introduction

Global agricultural productivity is being compromised by many factors, primarily the rising atmospheric CO_2 levels driven by climate change [1]. Indeed, the number of insects has increased across a wider range of habitats due to higher average daily temperatures and more frequent extreme weather events, which have resulted in substantial crop losses [2].

At the same time, increasingly restrictive European regulations on pesticides (e.g., the ban on dimethoate under Regulation (EU) 2019/1090) are accelerating the ongoing transition to integrated pest management (IPM), which is based on preventive monitoring and predictive modelling (Directive 2009/128/EC) [3]. To respond to the need for new

decision-support tools, this study assesses the environmental suitability for the olive fruit fly (*Bactrocera oleae* (Rossi)) in cultivated areas of eastern Sicily, using species distribution models and subsequent GIS analyses.

In the olive production sector, the IPM strategy is applied to protect crops from attacks by the olive fruit fly (*Bactrocera oleae* (Rossi)) and the olive moth (*Prays oleae*). These pests have significantly impacted the olive sector (*Olea europaea* L.). In the 2024/25 crop year, Italian olive oil production reduced to 250,000 tons, the lowest point in a decade (−24%) [4]. This decline was due to water scarcity, abnormal temperatures during critical growth stages, and increased pest pressure. The *Bactrocera oleae* (named *B. oleae* hereafter) is one of the most prevalent and destructive pests in this global ecosystem. Its larvae feed on the fruit pulp, causing premature fruit drop and significantly reducing yield and oil quality [5]. Therefore, both demographic and phenological models have been developed to describe species population dynamics. Currently, olive farms and advisory services require high-resolution spatially accurate decision support tools to prioritise monitoring and optimise interventions under increasingly variable climatic conditions and reduced chemical control options. However, most current approaches focus only on broad-scale climatic suitability or local population trends. This means they fail to deliver maps that integrate key pedoclimatic constraints at a level of detail relevant to farms. Therefore, it is timely to produce high-resolution suitability maps under current and near-future climate conditions in order to support targeted control and risk-based IPM planning in Mediterranean olive systems.

In the literature, studies have mainly focused on two approaches. The first involves implementing models for large-scale analysis by correlating soil and climate predictors with insect demography or by conducting a phenological analysis to identify optimal environmental ranges for insect development. The second approach, conversely, analyses insect distribution at a smaller spatial scale and aims at studying population trends. However, these methods do not account for temporal interactions among pest populations and, therefore, lose spatial information, which can result in inaccurate results.

In a regional study conducted by [3], a modelling approach was central to understanding *B. oleae* infestation and predicting population dynamics. The study aimed at developing a monthly data-driven method to predict the presence of *B. oleae* infestations in olive groves at a regional level in Tuscany (Italy) by applying machine learning (ML) algorithms. To this end, a demographic model was constructed, counting the numbers of (i) eggs, (ii) live and dead first-, second- and third-instar larvae, (iii) live and dead pupae and (iv) exit holes. The combination of uncorrelated bioclimatic predictors was then used for the prediction of the infestation levels.

Furthermore, along the same lines, ref. [2] improved model calibration at a fine spatial scale in the Umbria region by using demographic data from 79 olive groves and climatic variables interpolated with the Inverse Distance Weighting (IDW) method from the Regional Hydrographic Service. The relationships between predictors and *B. oleae* presence suitability were analysed along the "Julian day", highlighting the seasonal patterns of the pest to support targeted protection measures.

Outside Europe, the olive fly poses a significant threat to olive production, prompting in-depth research into the pest in various parts of the world. In California, USA, Kumal [6] used phenological models to assess the spread of the pest in warmer inland valleys. In Australia, studies of areas with the most suitable climate have highlighted the risk of the pest becoming established in regions not previously affected. In North Africa and the Middle East, research has examined the potential interactions between irrigation practices and pest population dynamics in arid conditions. Finally, ref. [7] investigated the geographical relationships between the olive tree host and the olive fly parasite in Africa and the

Mediterranean region in relation to the impacts of climate change by using ecological niche modelling approaches.

Many studies in the scientific literature have used species distribution models (SDMs) to predict the potential distribution of species by integrating abiotic and biotic factors, as reported by [8]. This study aims at assessing the environmental suitability for the phytophagous olive fruit fly (*B. oleae*) in cultivated areas of eastern Sicily through the combination of GIS and SDMs. The high spatial detail of the proposed framework enables the identification of areas that are more susceptible to infestation under current conditions and in the near future, providing operational guidance for the planning of monitoring networks and the optimisation of plant protection measures in line with the current need for the sustainable precision-oriented management of Mediterranean olive groves. Additionally, response curves translate model outputs into interpretable pedoclimatic ranges and thresholds (e.g., temperature and precipitation seasonality, topsoil texture and topographic constraints), supporting decision-making at farm level. This can refine management in existing orchards (e.g., timing and intensity of monitoring and interventions) and inform site selection and risk screening for new or expanding olive plantations.

Unlike previous research, this study is novel in that it

i. Uses a high-resolution (1 ha) pedoclimatic model that integrates topsoil texture to characterise microenvironmental conditions relevant to pest development;
ii. Assesses the future suitability for *B. oleae* in global climate model simulations for near-term projections (2021–2040) compared with the current distribution of olive groves, identifying areas of current and potential risk;
iii. Uses response curves to define agronomically significant thresholds and predictor ranges, translating the ecological model's outputs into an operational tool that supports IPM planning and management in Mediterranean olive systems.

On this basis, the main research questions are the following:

1. Which SDM provides the most accurate prediction of the presence of *B. oleae* in eastern Sicily?
2. Which pedoclimatic factors mostly influence the environmental suitability for *B. oleae* presence under heterogeneous pedoclimatic conditions?
3. How do the response curves derived from SDMs clarify the relationship between environmental predictors and the ecological niche of *B. oleae*?
4. How would the spatial suitability for *B. oleae* change under future climate projections (2021–2040) according to global climate models?

2. Materials and Methods

In this study, a methodology based on SDM algorithms was defined to identify areas with the highest environmental suitability for B. oleae presence in Eastern Sicily. This is based on the known distribution of its host plant, *Olea europaea* L., and key environmental covariates for both current and future scenarios (Sections 2.1 and 2.2). Unlike previous research, this study integrates bioclimatic predictors generated through the 'dismo' package in Rstudio software 2022.07.0 (https://www.R-project.org, accessed on 14 October 2024) besides soil predictors within a regional geographical area (R Core Team, 2023) [9]. To this end, a collaboration was established with a local cooperative that supplied presence data for B. oleae gathered by using field monitoring traps.

Accuracy measures of the models have been considered for the assessment, and the contribution of each environmental variable to the environmental suitability for *B. oleae* presence was analysed by using response curves generated by Random Forest (RF) and MaxEnt models (Sections 2.3 and 2.4). Figure 1 shows the workflow for current and future

scenarios, from data acquisition and pre-processing to SDM calibration, validation and map-based analyses.

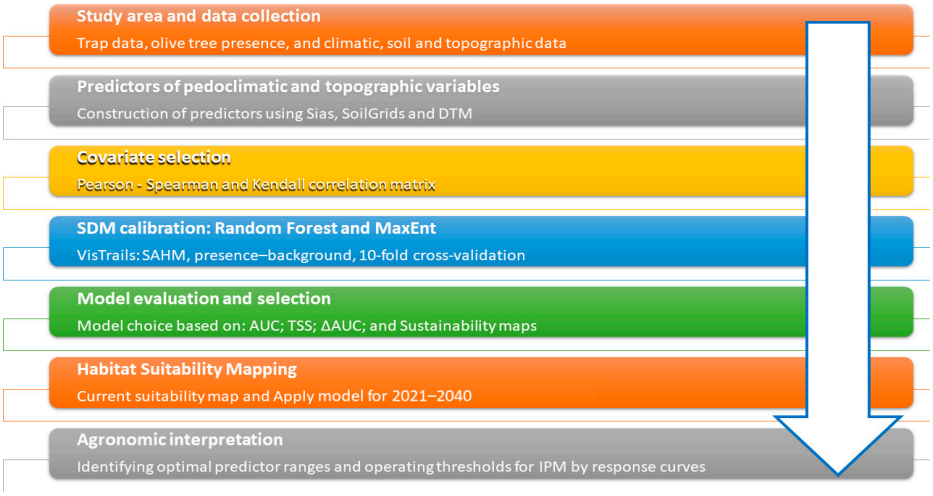


Figure 1. Methodological framework.

2.1. Study Area and B. oleae Occurrence Data

The study area included the provinces of Catania, Syracuse, and Ragusa (Figure 2), located in the south-eastern part of Sicily. This territory is characterised by different agrarian landscapes and a notable historical and cultural heritage, with a Mediterranean climate characterised by hot dry summers and mild wet winters. Crop selection and agronomic management are influenced by this climate and are calibrated to seasonal variations in temperature and water availability. Sicily produces high quality olive oils, i.e., *Monte ETNA* D.O.P. (Protected Designation of Origin) (Council Regulation (EEC) No. 2081/92—Official Journal of the European Union. Series L. 21 June 2022.) or *Monte Iblei* D.O.P. (Ministerial Decree of 4 December 2003), known for their unique flavour profiles due to the specific soil and climatic conditions. Olive groves are often located in arid or semi-arid environments, which requires deficit irrigation and efficient irrigation strategies, especially during the summer period when water resources are most limited. In Sicily, olive groves are often intercropped with other crops, such as citrus fruits, forming polyculture ecosystems. These ecosystems play a critical role in maintaining ecological balance and supporting the region's agricultural productivity.

The study areas were identified based on agronomic and operational criteria, favouring olive-growing districts where the pressure from *B. oleae* and the associated monitoring costs represent a critical element of crop management. In this context, suitability maps are intended as a decision-support tool to help rationalise trap placement and survey frequency, reduce monitoring costs and improve the spatial resolution of information on the pest's distribution. Specifically, the sampling area was defined by using basic data on *B. oleae* presence acquired from the pheromone-trap monitoring system of the Olive Producers Association (APO).

These data were collected from a monitoring network of pheromone traps distributed throughout the association's territory (Figure 1). To minimise spatial clustering, trap locations were selected so as to obtain, as far as possible, a single presence per 100 m grid cell (thinning threshold = 100 m). The network covered 41 farms and comprised 124 pheromone traps. The monitoring campaigns in the period 2019–2021 covered the summer months, i.e., from June to October. Points were selected based on farm location and size and semi-intensive farming systems, with the sampling design aiming at an approximate density of one trap per hectare. Throughout the monitoring season, the

traps were checked weekly to confirm the continuous presence of phytophagous insects during the considered period. The locations of occurrences were then verified in a GIS environment to ensure spatial consistency with the analysis grid and minimise duplication at pixel level. Residual spatial clustering was quantified by using a nearest-neighbour analysis (QGIS): the observed mean nearest-neighbour distance was 165.33 m (expected: 1974.96 m), yielding a nearest-neighbour index of 0.084, indicating a strongly clustered pattern in the occurrence dataset.

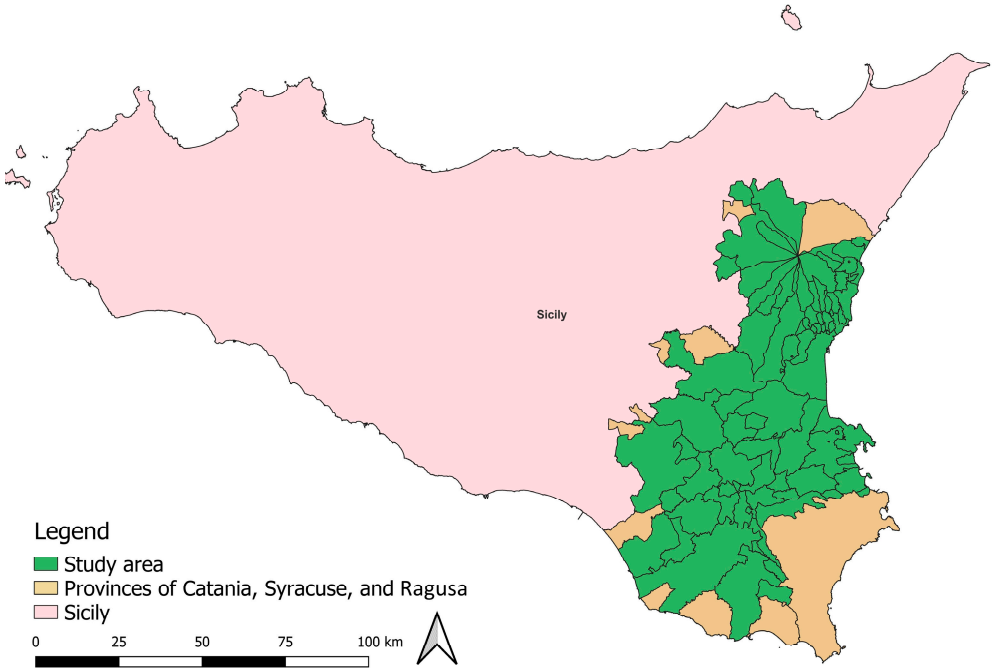


Figure 2. Location of the area under study in Sicily (Italy) and study area encompassing the municipalities of Catania, Syracuse, and Ragusa.

Data from APO were processed to produce an input map showing the presence of phytophagous organisms. Despite the fact that a limited number of presence points can reduce the predictive power of SDM [10–12], presence points were consolidated in a GIS (QGIS Desktop 3.22.1) by overlaying pheromone trap presence points, olive trees (*Olea europaea* L.) occurrence points, and a 250 m buffer around the traps. Specifically, a radius of 250 m was used to represent the local neighbourhood actually sampled by each trap, thereby avoiding unrealistic levels of precision at the point level in the occurrence dataset. Field monitoring and impact assessments usually link trap captures to management units within a radius of around 150 m, reflecting the short-range movements and spatial impact of traps in olive agroecosystems.

Therefore, a 250 m buffer provides a conservative representation of the pest's presence while remaining well below the reported dispersal distances of up to 5000 m for *B. oleae* [13]. After pre-processing and spatial aggregation, a total of 550 presence points was retained for modelling. In this regard, sensitivity runs using more restrictive presence sets were carried out. These analyses confirmed that when the effective number of presences is reduced, the models become more prone to overfitting and show reduced transferability. This is consistent with the SDM literature indicating that limited presence samples reduce predictive power [12].

As no reliable information was available to define true pseudo-absences, a presence–background approach was adopted, and background points were sampled across the study area. A total of 10,000 background points were sampled, corresponding to the minimum number recommended in the literature for robust SDM calibration [11]. Under

these conditions, this distribution of presence and background points may lead to a larger clustering of presence points compared to background points. Moreover, the overlap with the host plant presence map could introduce biases in the interpretation of environmental preferences. Therefore, the distribution of *B. olea* suitability was checked against the actual distribution of the host plant in the Results section, and the consistency of the results with the expected climatic conditions for the pest is discussed in Section 3.3.

2.2. Pedoclimatic Predictors

Climatic predictors for the study area were specifically generated by using data from the WorldClim global climate database at 30 arc-seconds spatial resolution (~1 km) (https://www.worldclim.org). The time span from 2003 to 2021 was chosen as the reference window for the current scenario, because it is broad enough to account for recent interannual variability in climate conditions, while also being representative of the operational climate in which field observations are located. According to [14], this timeframe meets the requirement of working with multidecadal periods in order to construct a stable 'recent' climate framework from which to derive the bioclimatic predictors considered in the models. The data were processed and analysed by using RStudio 2022.07.0 and QGIS 3.22.1, following well established methodologies [14,15]. Bioclimatic predictors were produced by using the "dismo" package in RStudio software [16], together with the function biovar utilised to transform monthly climatic data into the desired bioclimatic indices, i.e., 19 biovariables [17], included in the predictors' list.

The monthly climate data were averaged over the reference period of 2003–2021 to obtain 12 raster layers showing the multi-year average temperature and precipitation. These raster layers were then used as input for the 'biovars' function in the 'dismo' R package, which derives the 19 standard bioclimatic variables (Bio 1–Bio 19) according to WorldClim definitions [16].

In the case of the future scenario, the future climate conditions were obtained from the WorldClim database (SSP5-8.5 by CMCC-ESM2 global climate models, for near-term projections 2021–2040) (https://www.worldclim.org/data/cmip6/cmip6_clim30s.html, accessed 20 October 2024).

Since pedological factors are known to influence the vitality of the *B. oleae* pupas [18], the soil conditions in the study area were acquired from Soilgrid (https://soilgrids.org/) and SITR database (https://www.sitr.regione.sicilia.it/) and were considered in the simulations as input data. The following soil predictors and geomorphological information were included in the predictor list:

- Digital Terrain Model (DTM): The DTM, with a 20 m resolution, was acquired from the SITR geo-database through the GIS WFS (https://www.sitr.regione.sicilia.it/, accessed on 13 February 2022 and unavailable at present).
- Slope [%]: The slope of the terrain was computed, at a 20 m resolution, from DTM by using specific tools available in QGIS 3.22.1 (Sphinx, https://docs.qgis.org/3.40/it/docs/training_manual/index.html, accessed on 17 July 2025).
- Aspect [°]: The raster was produced from DTM by using specific tools available in QGIS 3.22.1 (Sphinx vectorization toolkit for QGIS, https://docs.qgis.org/3.40/it/docs/training_manual/index.html). This parameter contains values ranging from 0 to 360, which represent the direction of the slope, beginning with the north (0°) and continuing in a clockwise direction.
- Volume of water in soil at a −5 cm depth, in saturated soil of −10 kPa (named Vw-10 hereafter), at a field capacity of −33 kPa (named Vw-33 hereafter), and at a permanent wilting point of −1500 kPa (named Vw-1500 hereafter) was obtained

from Soilgrid at a 250 m spatial resolution (SoilGrids250) [19] (https://soilgrids.org, accessed on 2 October 2023).

- Sand Content, Silt Content, Clay Content [g/kg]: Raster images were extracted from SoilGrids at 250 m spatial resolution and used to characterise the topsoil texture. According to the established definitions, particles classified as sand have a diameter between 2 and 0.05 mm; particles designated as silt have a diameter between 0.05 and 0.002 mm; and particles identified as clay are the finest fraction, with a particle diameter less than 0.002 mm [19].

2.3. Implementation of Species Distribution Models (SDMs)

Vistrails:SAHM was the framework used for the entire process [20,21] (Figure 3); it was used to run the entire SDM workflow, covering all processes from handling and preprocessing the input data to calibrating the model and creating the map projection. The predictors were provided as raster layers (PredictorListFile) and the occurrences as a CSV dataset. All predictors were harmonised to a common template grid (TemplateLayer) (UTM WGS84 33N, EPSG:32633; spatial resolution of 100 m, resampling: Bilinear) and occurrence records were screened within the study area and aggregated at the level of the grid cell to reduce spatial duplication.

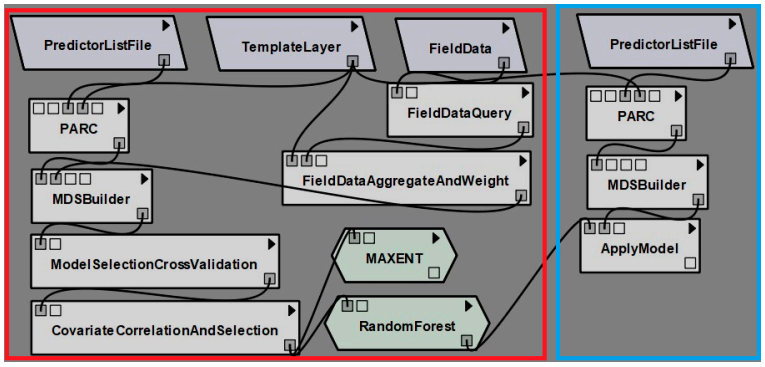


Figure 3. Workflow implemented in VisTrails-SAHM for SDM calibration and projection. The red panel summarises the steps used to build and validate the current suitability models, including data preparation, presence/background assembly, cross-validation, correlation screening, and model calibration using RF and MaxEnt. The blue panel illustrates the projection workflow utilised to generate suitability maps for the future scenario ('Apply Model').

In MDSBuilder, the background points were randomly placed at the centroids of pixels within the model extension. This was done without assigning more than one point to each pixel and was based on the guidelines provided by the VisTrails:Sahm software (version VisTrails 2.2.3 SAHM 2.0.1).

Consequently, the dataset was partitioned by using a 10-fold cross-validation in the ModelSectionCrossValidation module. The approach involved the division of the occurrences into ten subsets, which were then utilised iteratively for model training and testing. The evaluation of the model was conducted through the utilisation of k-fold cross-validation (k = 10) within the SAHM workflow. Occurrence records were divided into ten partitions. For each iteration, models were fitted on n-1 partitions and evaluated on the remaining partition. This process resulted in fold-specific evaluation metrics. The results of the cross-validation process were summarised as the mean and standard deviation across folds. These were then used for the purpose of model comparison and selection. The final model was refitted by using the full dataset. Two SDMs, namely Random Forest (RF) and Maximum Entropy (MaxEnt), were used to generate environmental suitability maps of *B. oleae* presence. RF is an ensemble learning method composed of multiple decision trees

built on random subsets of data. MaxEnt is a machine-learning method optimised for presence-background data that estimates the most uniform distribution consistent with the observed data. These models have been widely applied for various objectives such as the estimation of potential biomass [17], optimisation of resource input [21], and relationship between fish and habitat [22]. Model parameterisation was optimised through a sensitivity analysis within the SAHM framework [11]. For Random Forest, we tested mtry = 1–2, ntrees = 500–5000 and nodesize = 1–5. For MaxEnt, we explored replicates = 1–20, maximum iterations = 5000–10000 and beta multiplier = 0.5–5. The final settings adopted, based on cross-validation performance, were RF: ntrees = 1000, mtry = 2, nodesize = 2 and MaxEnt: replicates = 1, beta multiplier = 1, maximum iterations = 5000.

The models were trained and tested by using 10-fold cross-validation [23]. To reduce information redundancy and limit the effects of multicollinearity on model calibration, all environmental variables were tested for predictor dependence. To this end, Pearson, Spearman and Kendall correlations were calculated between the covariates. Where two or more variables showed a high correlation ($r \geq 0.75$), only one variable from each group was retained in the final set. The choice between highly correlated variables was guided by two criteria [6,24]: (i) ecological coherence with the biology and developmental cycle of *B. oleae* and (ii) informative contribution. The latter was assessed via the proportion of deviance explained, favouring the most interpretable and useful variable for forecasting purposes. Accordingly, for some correlated pairs, the correlation filter was not applied when the retained predictor was considered to capture a biologically meaningful eco-physiological constraint of *B. oleae* and thus provided complementary information relevant to the species' developmental cycle and habitat suitability. Then, the Apply model module was also used to enable predictions of the future climate suitability for *B. oleae*. Finally, the continuous suitability maps (0–1) were reclassified to distinguish between unsuitable and suitable areas, based on the probabilistic threshold considered to calculate the TSS (Sensitivity = Specificity criterion). Values below the threshold were assigned to the unsuitable class. Above this threshold, the suitable areas were further categorised into three probability classes: low, medium, and high. This class representation is widely established in the SDM literature to enhance the operational readability of risk map (e.g., Ashraf et al. [7]).

2.4. Output Assessment

The results of the predicted presence maps and the performance of the models were evaluated by using accuracy measures derived from the confusion matrix, including the True Skill Statistic (TSS) and the Area Under the Curve (AUC), which are widely used metrics for evaluating the reliability of SDMs.

The AUC represents the area under the Receiver Operating Characteristic (ROC) curve. According to [25], the test is moderately accurate in the range 0.7–0.9 and highly accurate for values above 0.9. In these types of assessments, overfitting is generally considered and analysed because in such conditions the utility of the model is limited, indicating that the uncertainty regarding the true complexity of the environmental niche is high [26].

In the literature, the selection of the most suitable model is generally based on a comparative analysis of its performance. The most commonly used criteria are as follows:

(i). Maximising the AUC during the training phase by choosing the model with the highest AUC value calculated on the data used to build the model;
(ii). Maximising the testing AUC by choosing the model with the highest AUC value calculated on a randomly selected test dataset not used in the training phase.

In addition to these criteria, a joint analysis of training and testing performance allows the risk of overfitting to be assessed by computing the difference between the corresponding

AUC values (ΔAUC). According to [27], a ΔAUC difference larger than 0.05 between the training and testing phases suggests potential overfitting. This third approach was adopted in this study, with a threshold of $\Delta AUC \leq 0.05$ used to identify models that strike a better balance between predictive capacity and generalisation.

Regarding the True Skill Statistic (TSS), values range from −1, indicating a performance no better than random, to +1, indicating perfect agreement [28]. The TSS is computed as: (Sensitivity + Specificity − 1).

Model response curves were also analysed to evaluate the relationship between environmental predictors and habitat suitability, with the aim to understand the potential interactive adaptation of the pest. These curves describe the relationship between predictor values and the environmental suitability, supporting the identification of high-probability ranges and thresholds, and each predictor's contribution. Therefore, response curves were used to select the most suitable model for *B. oleae* distribution and to compare it with previous studies, supporting the verification of future predictions based on observed data [29].

In the simulation of future scenarios, the "Apply model" module was used to create future distribution forecasts of *B. oleae* and compare them with the current probability of olive presence.

3. Results

The Results section is organised into four main parts. Section 3.1 describes how non-collinear covariates were selected by using the PSK correlation matrix. Section 3.2 analyses the suitability patterns simulated by RF and MaxEnt algorithms. Section 3.2.1 then extends this assessment to the period from 2021 to 2040 under the SSP5-8.5 scenario by overlaying the predicted suitability with the presence of olive orchards in 2012. Finally, Section 3.3 presents and describes the response curves of the selected predictors, which are grouped into climatic, soil, and topographic variables.

3.1. Analysis of Covariates and Models' Metrics

Figure 4 summarises (i) the set of predictors retained after correlation analysis and (ii) the Pearson–Spearman–Kendall correlation matrix (PSK matrix) used for selection. The left column indicates the "Response/Importance" that reports the univariate contribution of each predictor to modelled suitability.

In accordance with the procedure delineated in Section 2.3, variable selection was conducted through a dual analysis of (i) pairwise correlations among predictors and (ii) the percentage of deviance explained (Figure 4). Predictors were retained based on an acceptable correlation structure and an explained deviance higher than 2%. However, it should be noted that two exceptions were applied during the process. Aspect was retained despite a lower explained deviance because it captures insolation-related microclimatic variability (Aspect), which can influence adult activity and habitat conditions at a local scale. Furthermore, BIO15 and BIO4 were both retained despite their correlation (r = 0.80), given the potential influence of both precipitation and temperature seasonality on the pest life cycle; this issue is discussed in Section 3.3.1.

Based on the reduction in collinearity between highly correlated predictors, ten covariates were selected for modelling: Bio 2 (Mean Diurnal Range) (Dev. Exp 2.8%), Bio 4 (Temperature Seasonality; SD × 100) (Dev. Exp 2.1%), Bio 15 (Precipitation Seasonality; coefficient of variation, %) (Dev. Exp 1.9%), Bio 16 (Precipitation of Wettest Quarter, mm) (Dev. Exp 2.8%), clay (0–5 cm) (Dev. Exp 3.5%), sand (0–5 cm) (Dev. Exp 3.4%), silt (0–5 cm) (Dev. Exp 2.4%), DTM (Dev. Exp 3.2%), slope (%) (Dev. Exp 3.8%), and aspect (°) (Dev. Exp 0.4%).

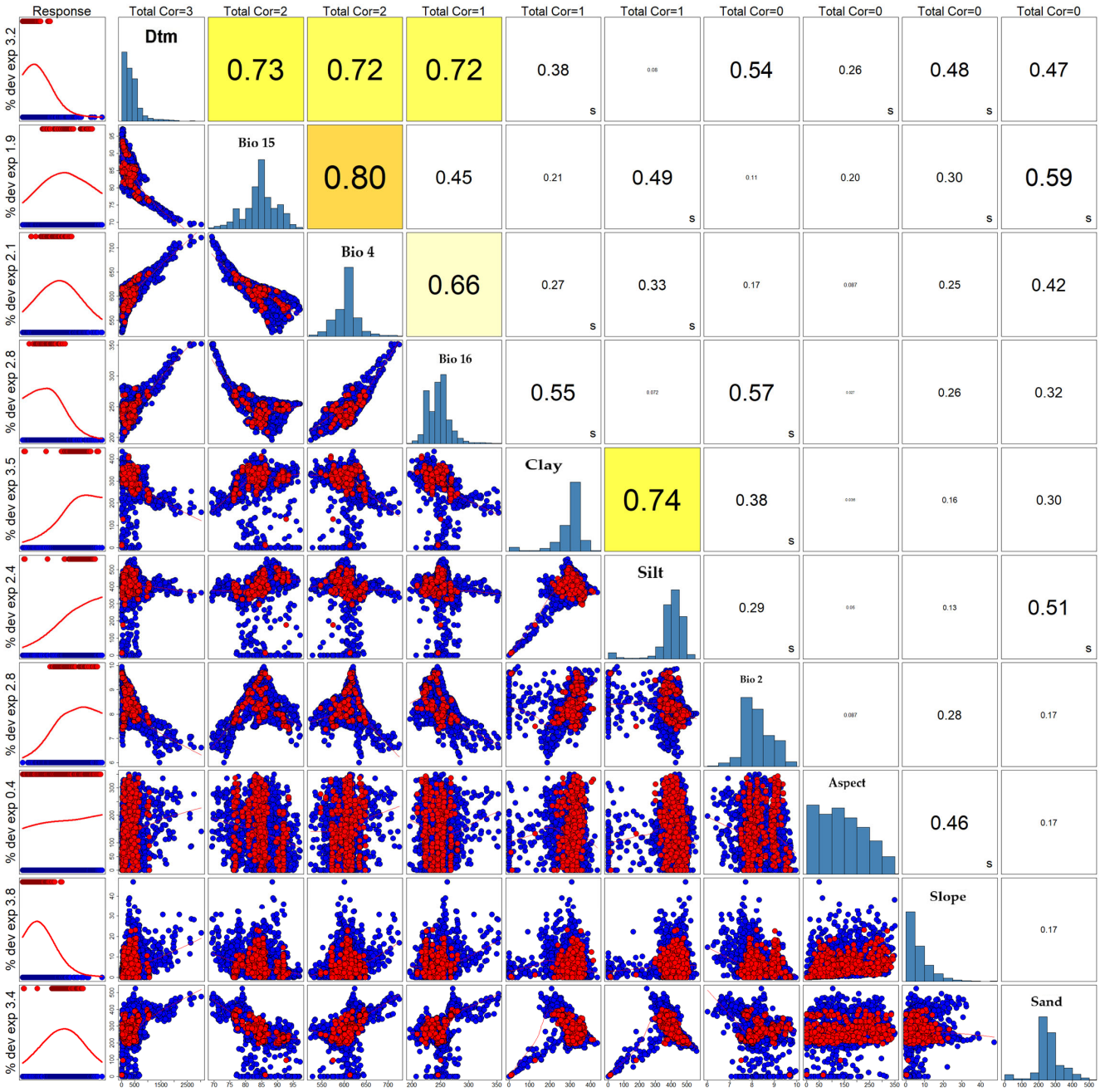


Figure 4. PSK matrix, where histograms for each variable are displayed along the diagonal. Scatter plots with smoothing curves are shown below the diagonal, with red points representing presence data and blue points representing absence data. Pearson correlation coefficients are reported in the upper panels, with correlation strength visually categorised as dark yellow for $r \geq 0.75$, yellow for $0.70 \leq r < 0.75$, and light yellow for $r < 0.70$.

The predictors identified by the PSK matrix were then used to train the MaxEnt and RF models. Table 1 summarises the performance of the models for training and 10-fold cross-validation, as described in Section 2.4. In addition to the area under the curve (AUC), the model performance was assessed by using the threshold-dependent statistic TSS, which combines sensitivity and specificity (TSS = sensitivity + specificity − 1). Under cross-validation, RF produced a sensitivity of 0.701 and a specificity of 0.979 (TSS = 0.681), whereas MaxEnt produced a sensitivity of 0.762 and a specificity of 0.850 (TSS = 0.612). This shows that RF produced fewer false positives, whereas MaxEnt detected a higher proportion of presences, though with more false positives.

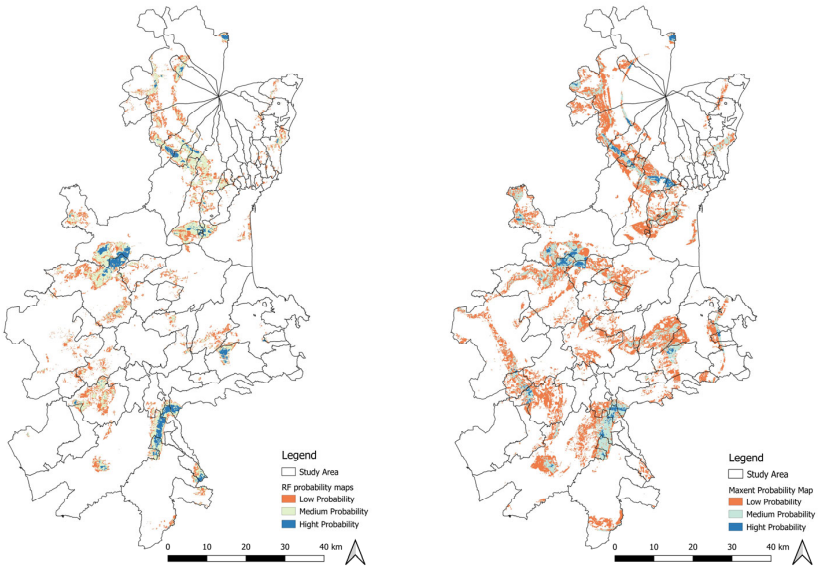
Table 1. Values of the training and cross-validation metrics for RF and MaxEnt.

	MaxEnt		Random Forest	
	AUC	TSS	AUC	TSS
Training	0.939	0.721	0.970	0.822
Cross Validation	0.892	0.612	0.967	0.681
ΔAUC	0.047		0.004	

However, as VisTrails:SAHM does not support spatial cross-validation (e.g., spatial blocking, checkerboard partitioning or geographically defined folds) within the workflow, the reported AUC and TSS values should be interpreted as internal performance estimates under random resampling and may be optimistic in the presence of spatial autocorrelation. The model reliability was also evaluated by examining the relevant areas of the environmental suitability maps and the ecological coherence of the response curves (see Section 3.3). These were used to determine whether the predicted suitability varied plausibly along the main environmental gradients.

3.2. Species Suitability Maps

Figure 5a,b show the suitability maps produced for the current scenario by the two algorithms, i.e., RF and MaxEnt.



(**a**) (**b**)

Figure 5. Suitability maps of *B. oleae*: (**a**) actual suitability by MaxEnt model; (**b**) actual suitability by RF model.

To calculate the threshold-dependent metrics (sensitivity, specificity and TSS) and their corresponding binary outputs, the probabilities were converted to presence/absence by using the Sensitivity = Specificity threshold implemented in SAHM (RF = 0.51; MaxEnt = 0.37). As the model evaluation relied on random 10-fold cross-validation, these accuracy estimates could be overly optimistic in the presence of spatial autocorrelation. Therefore, in addition to reporting the AUC and TSS, we interpreted the environmental suitability maps by using more restrictive suitability classes (see Section 2.3).

The RF produces a more selective surface. MaxEnt, on the other hand, is more inclusive in identifying areas of suitability, assigning moderate suitability values to regions with only a subset of ideal conditions. The differences at the boundaries of the identified areas

indicate "Low probability" ecological areas. The results of the two models show high spatial convergence (about 36,534 ha), delineating a specific distribution pattern. In both suitability maps, areas with higher levels of suitability are not evenly distributed across the territory due to heterogeneous soil and climate conditions. Environmental suitability for the olive fly tends to be very low along the coastal strips, as well as in the areas of the Catania Plain and at higher altitudes on Mount Etna. The most significant presence seems to be concentrated mainly in two inland areas: the first covers the Iblean Plateau, affecting the inland areas of the provinces of Ragusa and Syracuse; the second nucleus is located in the Calatino area, in the southwestern part of the province of Catania. The environmental suitability maps related to the presence of *B. oleae* show that the high suitability for its presence is restricted to lowland and foothill areas where olive cultivation has historically been practiced. Areas that are climatically suitable, but where olive trees are not cultivated, may not be occupied.

The area of overlap between the RF and MaxEnt models amounts to just 9.64% of the territory, equivalent to 52,848 hectares. The RF model delineates 52,320.0 ha as suitable, distributed across low (34,737.0 ha), medium (13,355.0 ha) and high (4228.0 ha) probability classes, with the suitability mainly concentrated in Mineo, Caltagirone, Palagonia, Adrano and Paternò. Conversely, the MaxEnt model defines 92,702.0 ha distributed across low (73,168.0 ha), medium (18,917.0 ha) and high (617.0 ha) probability classes, with environmental suitability mainly concentrated in the municipalities of Misterbianco, Caltagirone, Francofonte, Ragusa, and Mineo.

Based on the performance metrics reported in Table 1, RF outputs were used for the subsequent analyses, while MaxEnt results are reported for comparison.

3.2.1. Spatial Distribution of *B. oleae* in the Future Scenario 2021–2040

To evaluate the maximum risk scenario, the SSP5-8.5 scenario was selected, following the approach also adopted in the study by [7], which was used to identify the maximum potential expansion of climatically suitable areas for the pest.

As the models were trained by using climate data from 2003 to 2021, and the forecast covers the period from 2021 to 2040; the time needed for the crop to develop and enter in production was taken into account. For this reason, points indicating the presence of olive trees in 2012 were used.

In Figure 6, the areas classified as having high and medium pest suitability overlap in most cases with dense stands of olive trees arranged in multiple blocks within the Iblean–Calatino sector and along several inland valley systems. This pattern is particularly evident in the municipalities of Chiaramonte Gulfi and Militello in Val di Catania, where 68.57% and 82.42% of the municipal territories, respectively, are classified as having high or medium suitability, indicating high potential pest pressure at the landscape scale. Moreover, the comparison of Figure 5a with Figure 5b shows that the upper classes coincide with the current suitable areas. Some areas show medium to high suitability despite the absence of host presence points. In these cases, the soil and climate conditions are ideal, but the area is too far from crops. Therefore, if olive orchards were to be established in these areas in the future, they could be subjected to significant phytophagous pressure. A comparison between the RF actual and future environmental suitability maps revealed significant changes in the spatial distribution of potential risk for *B. oleae*. In the current scenario (Table 2), approximately 495,000 hectares (90.5%) of the territory are classified as risk-free, while only 9.5% present a certain level of suitability: ~35,000 ha (6.34%) in low, ~13,000 ha (2.44%) in medium and only 4224 ha (0.77%) in the highest class. In the SSP5-8.5 future scenario, the unsuitable area is reduced to around 393,000 hectares (71.66%). At the same time, areas with a relative probability of infestation increased significantly: low probabilities

increased fourfold to around 130,600 ha (corresponding to 23.84% of the study area), while medium probabilities expand to about two times the area in the present scenario, reaching around 22,000 hectares, or 4.08% of the study area. High suitability areas, on the other hand, slightly decreased to about 4300 hectares (0.77%) (Table 2). No significant increases have been observed in areas with a high probability of occurrence; these areas remain limited and selective. The projections of the SSP5-8.5 climate scenario are set to expand and connect areas potentially suitable for the spread of *B. oleae*, particularly in the hilly areas of Iblean–Calatino and along the inland valleys.

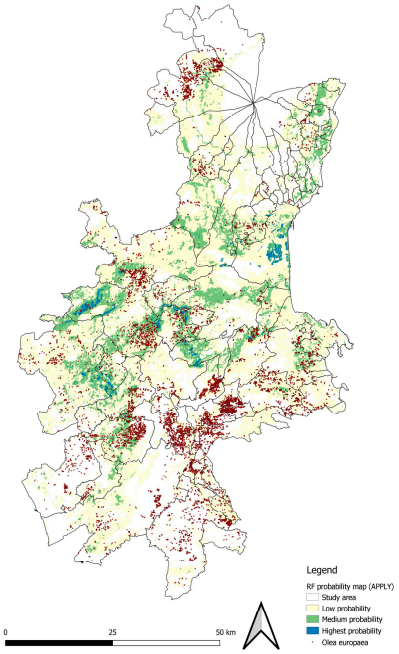


Figure 6. Overlay between olive tree presence points and the future scenario of *B. oleae* presence based on RF model.

Table 2. Change in the spatial extent of RF-derived suitability classes between the current scenario and the SSP5–8.5 projection (2021–2040) across the study area.

Suitability Class	Current [ha]	Current [%]	Future [ha]	Future [%]
Unsuitable	495,640	90.45	392,566	71.66
Low	34,737	6.34	130,629	23.84
Medium	13,355	2.44	22,340	4.08
High	4228	0.77	2330	0.43

3.2.2. Suitability Class Distribution in Olive-Associated Areas (Current and Future Scenarios)

The olive tree (Olea europea) layer was overlaid with the reclassified suitability maps to quantify the proportion of olive-associated area falling within each suitability class (Table 3). Suitability maps have been reclassified into four categories: unsuitable, low, medium and high by using the specific thresholds of the models (RF = 0.51; MaxEnt = 0.37). Values below the thresholds were assigned to the unsuitable class, whereas values equal to or above the threshold were divided in three equal intervals up to 1.00 (low, medium and high). Accordingly, for RF the class, we have the following: unsuitable <0.51; low 0.51–0.67; medium 0.67–0.84; high 0.84–1.00. For MaxEnt, we have the following: unsuitable <0.37; low 0.37–0.58; medium 0.58–0.79; high 0.79–1.00. Areas are expressed in hectares given the grid resolution of 100 m (1 grid cell = 1 ha).

Table 3. Change in the spatial extent of suitability classes between the current scenario and the SSP5–8.5 projection (2021–2040) in olive-associated areas.

Suitability Class	RF-Current	RF-Apply	MaxEnt-Current
Unsuitable	16,992 (80.5%)	15,457 (73.2%)	15,647 (74.1%)
Low	2096 (9.9%)	4561 (21.6%)	3345 (15.8%)
Medium	1117 (5.3%)	1004 (4.8%)	1600 (7.6%)
High	901 (4.3%)	84 (0.4%)	514 (2.4%)

Under current conditions, both RF and MaxEnt classified most olive-associated cells as unsuitable (80.5% and 74.1%, respectively; Table 3). The main differences between the models are observed in the intermediate classes: MaxEnt allocates a larger proportion of olive-associated area to low and medium suitability than RF (low: 15.8% vs. 9.9%; medium: 7.6% vs. 5.3%), whereas the high class remains comparatively limited in both cases (2.4% vs. 4.3%). In the RF future projection (Apply model), the class distribution shifts relative to the current RF map, with a lower proportion of unsuitable suitability cells (73.2%) and a higher proportion of low suitability (21.6%); the medium class remains close to current values, and the high class results are very limited (<1%) (Table 3).

3.3. Assessment Through Analysis of Response Curves and Predictors

Figure 7 shows the contribution of each predictor to the suitability assigned by the model. The curves identify the ranges of predictors in which suitability increases, stabilises, or decreases, providing an interpretable link between the predictors and the environmental suitability maps generated by the models. In the following subsections, the analysis of the response curves and predictors is carried out by grouping the analyses by bioclimatic predictors, edaphic predictors, and topographic predictors.

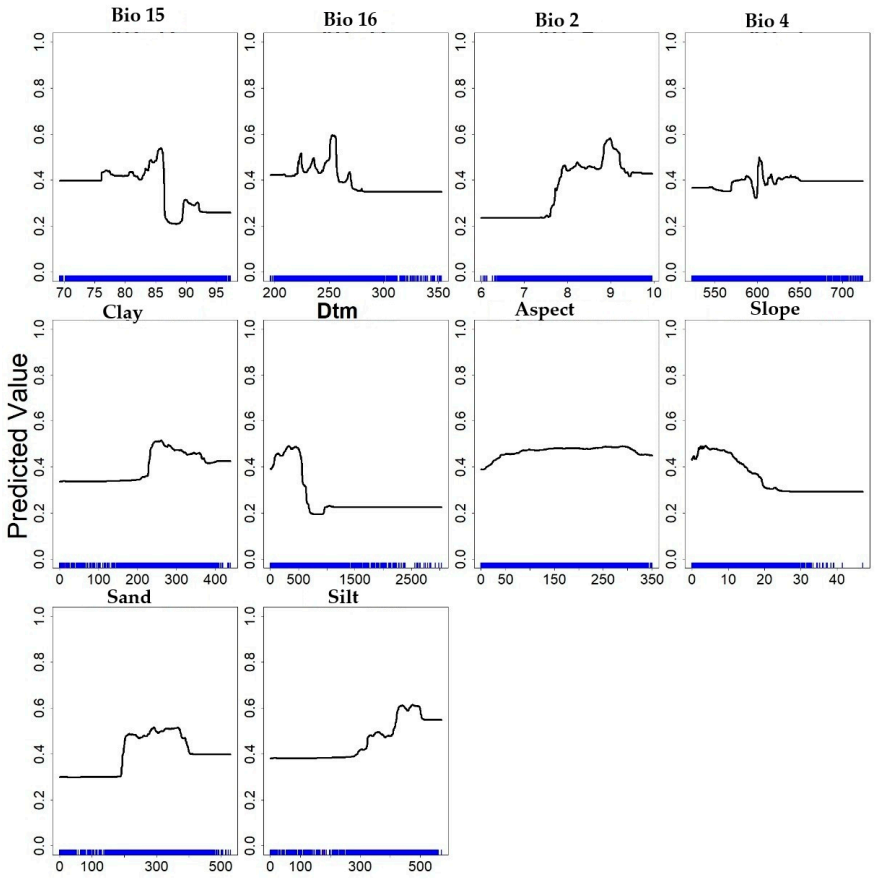


Figure 7. Response curves of the predictors for the RF model.

3.3.1. Climatic Predictors: A Focus on Bio 2, Bio 4, Bio 15, and Bio 16

The Bio 2 curve showed that intermediate values of diurnal temperature range are associated with a higher probability of occurrence. The probability of occurrence predicted by both models was mainly distributed between the third class (7.95–8.64) and the fourth class (8.64–9.1). These classes corresponded to areas located mainly in the foothills and hills of eastern Sicily (see Figure 8a), where the daily temperature fluctuations observed are more moderate than the values in inland areas. In these areas, both the MaxEnt and RF models assign high levels of environmental suitability, highlighting how this variable plays an essential role in determining potential distribution. Diurnal temperature ranges exceeding 10 °C cause a sharp reduction in the suitability for presence, highlighting a clear preference of *B. oleae* for climates with moderate diurnal temperature ranges.

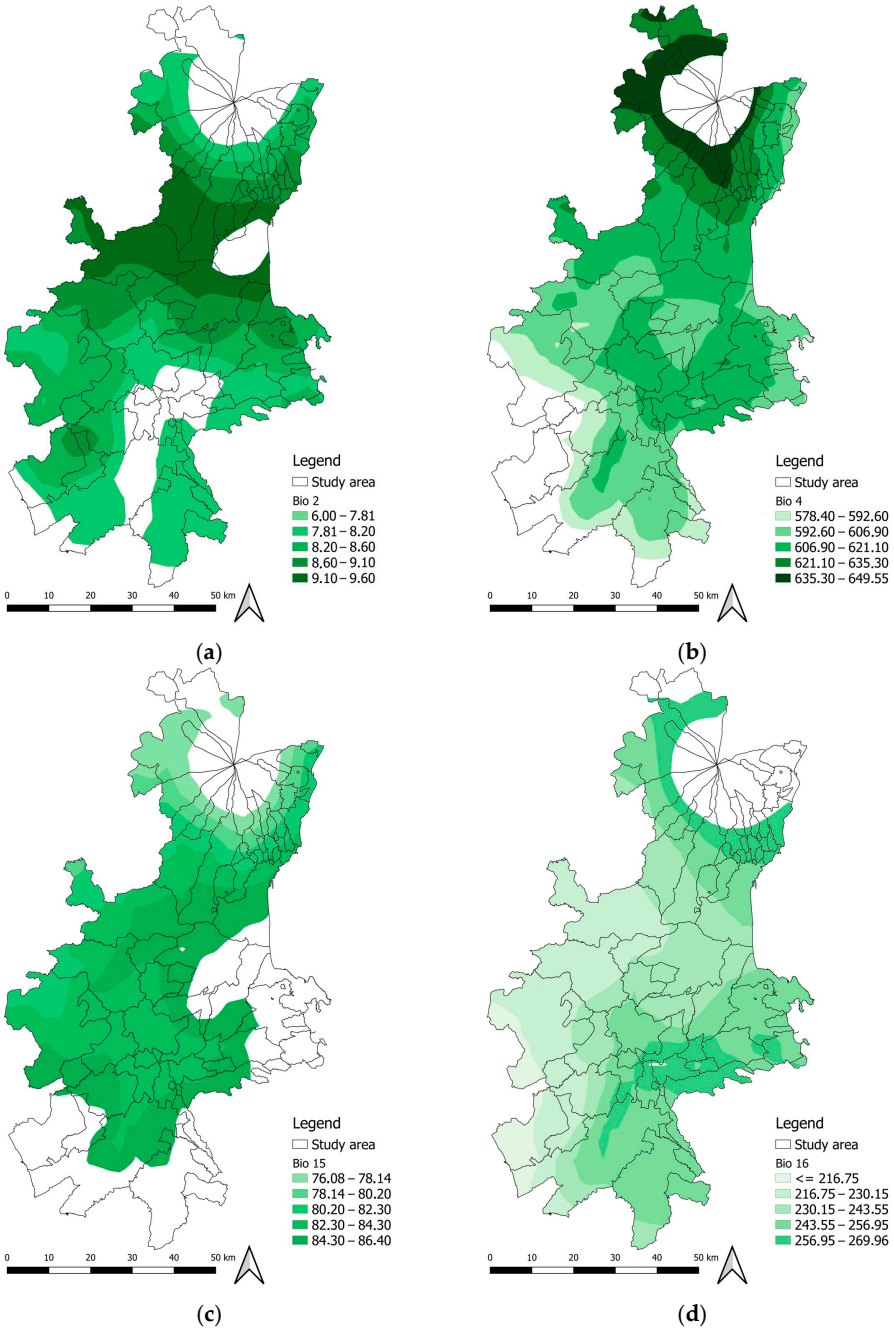


Figure 8. Climatic predictors' ranges in the study area: (**a**) Biovariable 2; (**b**) Biovariable 4; (**c**) Biovariable 15; and (**d**) Biovariable 16.

The Bio 4 predictor (Figure 8b) identifies the areas with the highest suitability as falling between 522 and 724.28. These values are found in both coastal and inland areas, where temperature variations during the year are harmful to *B. oleae*. These results demonstrate that spring acts as a second demographic filter after winter. The timing of emergence does not coincide with the plant's phenological phase, which causes pupal mortality through desiccation or the premature decline of adults; and it results in reduced seasonal infestation pressure. This phenomenon is more evident in the results of the RF model, which shows a more fragmented and selective potential area than the MaxEnt model. Internal areas with very hot summers are less suitable for *B. oleae* due to increased temperatures. High seasonality is penalised by the model as it leads to summer temperatures that are lethal for the insect.

The Bio 15 predictor (Figure 8c) shows optimal suitability at intermediate levels of rainfall variability. The medium (80.20–82.30%) and medium–high (82.30–84.30%) classes are widespread in the central-southern area, where precipitation is mainly concentrated in the autumn and winter months, with small contributions also in summer.

As for Bio 16 predictor (Figure 8d), the areas with the highest suitability are concentrated in zones with average (243.55–256.59 mm) and medium-high (243.55–256.59 mm) precipitation levels. These areas are arranged in two formations: an internal band running from west to east through the study area and a ring on the lower slopes of Etna. The Bio16 intervals (250–317 mm) maximise suitability by providing adequate autumn and winter rainfall. RF shows a more fragmented spatial pattern of suitability, whereas MaxEnt shows a more continuous pattern.

3.3.2. Soil Predictors (0–5 cm Texture: Sand, Silt, Clay)

The insect is sensitive to soil physical properties, such as the water retention capacity and aeration. The depth of pupation is significantly influenced by the soil type, with calcareous soils inducing deeper pupation, whereas alluvial soils with higher percentages of sand and silt result in an average depth of 1.16 cm.

Spatial results show that the highest probability is concentrated in areas where the clay content varies from 225 to 375 g/kg, clay in the medium–medium–high classes (206–357 g kg^{-1}; Figure 9a). The sand content in the soil (Figure 9b) has a negative correlation with suitability for presence. The suitability increases from low to intermediate values and decreases when sand falls into the high class (>364 g kg^{-1}; Figure 9b); these conditions are typical of the slopes of Mount Etna and some coastal areas, where higher porosity can promote surface drying in the absence of prolonged moisture [30]. As regards the silt content in the "Silt" soil, the medium and medium–high classes (458–571 g kg^{-1}; Figure 9c) coincide with maximum suitability. These percentages are recurrent in floodplains and inland basins, where water retention capacity combined with soil porosity provides a microenvironment favourable to pupal overwintering. The RF and MaxEnt models assigned a high probability of occurrence mainly in soils with a balanced texture between its components.

3.3.3. Topographic Predictors (DTM, Elevation, Slope, Aspect)

The Sicilian landscape is characterised by high levels of heterogeneity, which can affect the macroclimatic conditions. The DTM variable shows that the potential distribution of the suitability for *B. oleae* is strongly limited by altitude. Analysis of the response curves (Figure 7) shows a positive contribution between 32 m and 562 m above sea level, where the highest suitability values are concentrated, with a maximum weight of 0.48; at higher altitudes, the suitability decreases significantly. In the DTM variable map (Figure 10a), altitudes between 32 m and 562 m above sea level are those where olive cultivation is

most concentrated and correspond mainly to coastal plains, valleys and foothill areas. However, by overlaying the DTM variable on the environmental suitability maps of the RF and MaxEnt models (Figure 4), it can be seen that *B. oleae* is also present in suitable environments at relatively high altitudes, reaching about 1000 m above sea level. These habitats are found on some slopes of Mount Etna and in some parts of the Calatino area, where the local topography and climate create favourable microclimatic conditions.

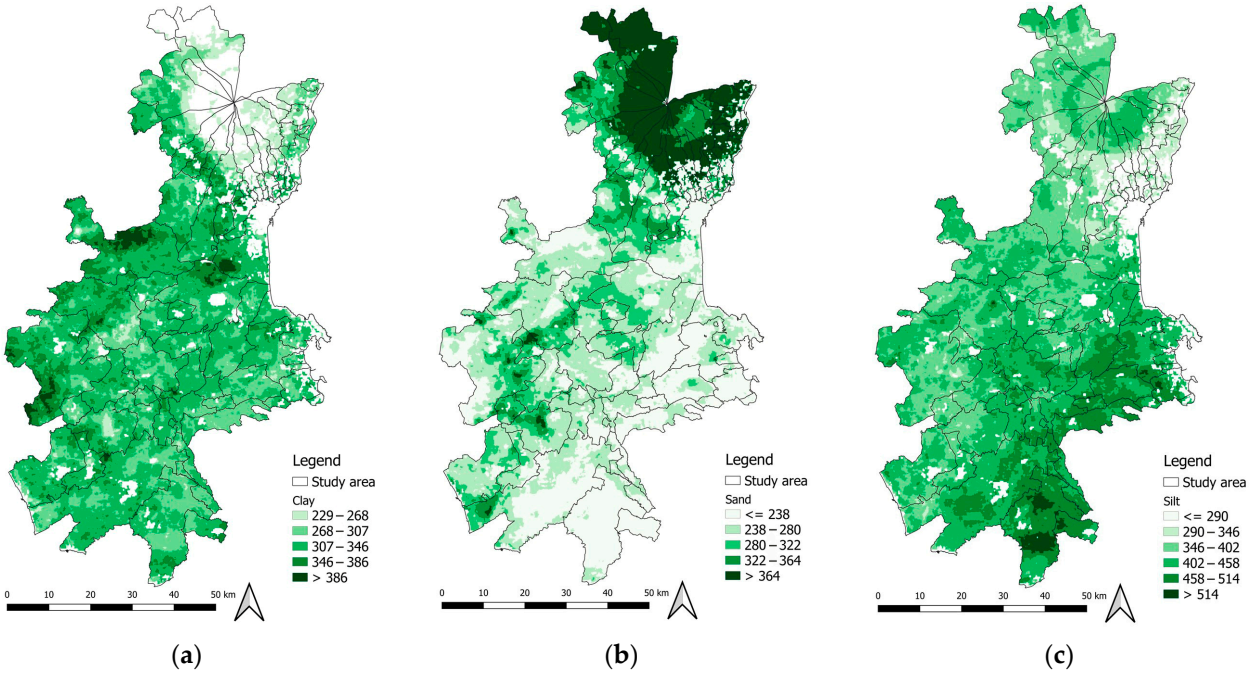


Figure 9. Components of soil texture ranges in the study area: (**a**) clay; (**b**) sand; (**c**) silt.

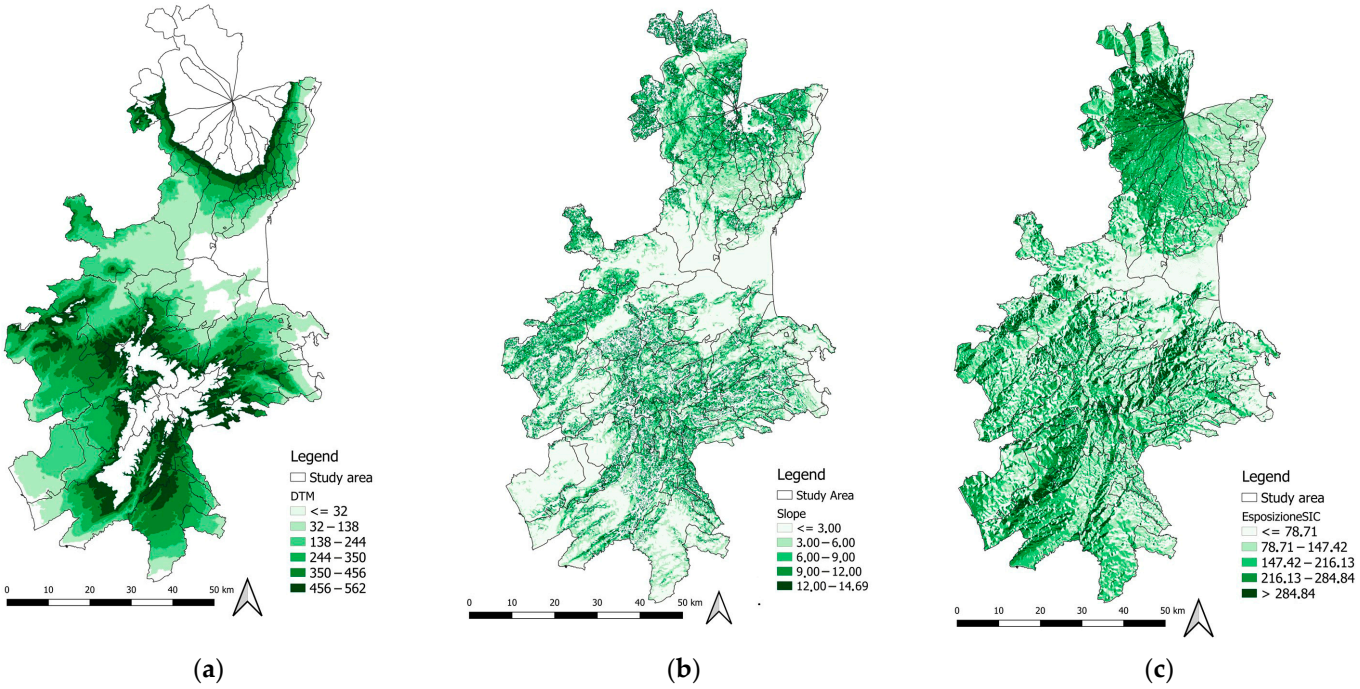


Figure 10. Topographic predictors' ranges in the study area: (**a**) DTM; (**b**) slope; (**c**) aspect.

The models indicate a clear preference for predominantly flat or sloping areas with values below 12%. This is reflected in the map in Figure 10b, where areas belonging to the

low–medium and medium classes (3.00–9) more often correspond to higher suitability than the medium–high/high classes (>9.00).

This relationship is likely to be an indirect effect linked to agricultural practices, as most olive groves are located on gently sloping land to facilitate cultivation and mechanisation. In contrast, the Aspect variable showed no clear preference, as the response curve indicated no appreciable increase in suitability across orientations. There is a weak correlation with the low suitability of north-facing slopes, which have microclimates that receive less radiation and are cooler and more humid (Figure 10c).

4. Discussion

4.1. Model Performance and Selection

Both algorithms achieved high predictive performance based on cross-validated metrics (AUC and TSS; Table 1), with negligible evidence of overfitting as indicated by low ΔAUC values. In addition to global accuracy, the spatial correspondence with monitoring observations supported RF as the most reliable representation of current suitability patterns in the study area; therefore, RF was selected as the reference model for the future projection, while the MaxEnt results were retained for comparison under current conditions.

According to D'Arrigo [25], models can be classified as "highly accurate" when the AUC values are higher than 0.90. In addition, the TSS values obtained in this study were higher than 0.60, further confirming the high predictive accuracy of the models [28]. Based on these criteria, the accuracy measures derived from 10 cross-validation replicates indicated high performance for both algorithms with values consistent with the ranges reported in regionally applied SDM of *B. oleae* [3,30]. Nevertheless, training AUC values can be a misleading performance measure in presence/background contexts [31]. For this reason, potential overfitting was assessed by examining the difference in AUC values between the training and testing phases (ΔAUC). Since the resulting (ΔAUC) remained below the commonly accepted threshold of 0.05 (Table 1), this suggested that overfitting was negligible.

Beyond the global metrics, the results of the algorithms were compared in terms of the spatial behaviour and correspondence with the observed data. The RF model showed superior performance in identifying areas with higher climatic suitability, as the areas with the highest probability of occurrence corresponded with the field observations (i.e., pheromone traps). On this basis, RF was chosen as the most accurate model for the case study, though accuracy measurement analyses were also verified by using the MaxEnt results.

4.2. Spatial Patterns and Ecological Constraints

Model-to-model differences are consistent with the distinct way RF and MaxEnt weight predictors and extrapolate in heterogeneous landscapes. The RF model produces a more selective surface, whereas MaxEnt is more inclusive, assigning moderate suitability values to regions with only a subset of ideal conditions. The differences at the boundaries of the identified areas indicate "low probability" ecological areas. These differences in suitability depend mainly on the weight given to environmental variables in different territorial contexts. At the same time, their strong overlap in areas of high suitability improves the accuracy of predictions [26], supporting confidence in the identification of the core suitable nuclei delineated by both algorithms.

The host availability further constrains the realised distribution. The suitability maps related to the presence of *B. oleae* show that the high suitability for its presence is restricted to lowland and foothill areas, where olive cultivation has historically been practiced. Areas that are climatically suitable, but where olive trees are not cultivated, may not be occupied.

As highlighted by [32], this is due to the fact that *B. oleae* is monophagous on *Olea europea*; therefore, the presence of the host limits its distribution even where the suitability for its presence is high.

The results of this study are also supported by the findings of [33], who verified that environmental suitability is not maximised in uniform landscapes, but at the interfaces between different environmental systems, such as foothills and hill margins, where local combinations of humidity, soil aeration and thermal stability create favourable micro-habitats.

The response curves provided an ecological interpretation of the predictor ranges associated with suitability patterns. Notably, one of the biologically significant results of this study is that winter temperatures influence the prediction mechanisms. Winter can alter the biotic potential and size of the population that will emerge in spring, defining the intensity of pest pressure for the following season. In fact, according to [2], milder winter temperatures are correlated with a more intense and earlier season of attacks.

For Bio 2, intermediate values of the diurnal temperature range are associated with higher suitability. In these areas, both the MaxEnt and RF models assign high levels of suitability, highlighting how this variable plays an essential role in determining potential distribution. According to [34], Bio 2 predictor is among the most influential in delineating the ecology of *B. oleae*. Diurnal temperature ranges exceeding 10 °C cause a sharp reduction in the suitability for its presence, highlighting a clear preference of *B. oleae* for climates with moderate diurnal temperature ranges.

For Bio 4, the areas with the highest suitability correspond to ranges that occur in both coastal and inland areas, where temperature variations during the year are harmful to *B. oleae*. As [2] observed, although high spring temperatures cause an acceleration of the first attacks, bringing forward their occurrence, this increase in temperature is negatively correlated with the overall probability of seasonal infestation. These results demonstrate that spring acts as a second demographic filter after winter. The timing of emergence does not coincide with the plant's phenological phase, which causes pupal mortality through desiccation or the premature decline of adults, and it results in reduced seasonal infestation pressure. This phenomenon is more evident in the results of the RF model, which shows a more fragmented and selective potential area than the MaxEnt model. Internal areas with very hot summers are less suitable for *B. oleae* due to increased temperatures. High seasonality is penalised by the model, as it leads to summer temperatures that are lethal for the insect. Experimental data from [12] showed that larval mortality reaches 100% above 32.5 °C.

For Bio 15, optimal suitability at intermediate levels of rainfall variability is consistent with areas where precipitation is mainly concentrated in the autumn and winter months, with small contributions also in summer. For Bio 16, the response suggests that precipitation intervals can maximise the suitability by providing adequate autumn and winter rainfall (e.g., Bio 16 intervals 250–317 mm).

The algorithm differences are also reflected in the curve-based interpretation: RF shows a more fragmented and selective pattern, whereas MaxEnt outlines a more continuous potential distribution extending into transitional zones that approach the species' limits.

Soil predictors provide additional constraints linked to the portion of the life cycle occurring in the soil. The presence of *B. oleae* is influenced by soil characteristics, particularly texture, as the pest completes part of its life cycle in the soil during the winter season [35]. In addition, the presence of *B. oleae*, as demonstrated in the work of [17], is mainly found in the first few centimetres of the topsoil, with a maximum frequency between 0 and 3 cm. The insect is sensitive to its physical properties, such as the water retention capacity and aeration. The depth of pupation is significantly influenced by the soil type, with calcareous soils inducing deeper pupation, whereas alluvial soils with higher percentages of sand and

silt result in an average depth of 1.16 cm. In sandy contexts, higher porosity can promote surface drying in the absence of prolonged moisture [30]. Conversely, floodplains and inland basins combine water retention capacity and soil porosity, providing a microenvironment favourable to pupal overwintering. It can, therefore, be concluded that clay, silt and sand are compositional fractions of soil texture that must be interpreted jointly. The RF and MaxEnt models assigned a high relative probability of occurrence mainly in soils with a balanced texture between its components, according to Miller's triangle, with moderate clay and silt, and low sand values; this result is in line with the findings of [17].

Topographic predictors modulate suitability through terrain-driven heterogeneity. The Sicilian landscape is characterised by high levels of heterogeneity, which can affect the macroclimatic conditions. Local topography and climate can create favourable microclimatic conditions and allow suitable environments also at relatively high altitudes. Data from [17] confirm this result: altitude is a mitigating predictor of pest suitability, but not as constraining as other climate predictors such as the mean temperature or temperature range. High-altitude olive groves can be considered climatic refuges where the development of *B. oleae* is intrinsically slower, and the overall infestation pressure is significantly lower. Slope-related patterns are likely to include an indirect component linked to agricultural practices, as most olive groves are located on gently sloping land to facilitate cultivation and mechanisation. Aspect effects are comparatively weaker, with cooler and more humid microclimates on north-facing slopes potentially contributing to lower suitability under some conditions.

The suitability maps (Figure 4) should not be interpreted as maps of the absolute probability of occurrence, but rather as indices of the relative environmental suitability for the species. Although the resulting map shows the most suitable abiotic conditions for the species, it does not describe its actual niche, which is further limited by other factors such as predators or competitors, host plant distribution and dispersal barriers [36]. Furthermore, management practices such as irrigation and canopy management were not considered, although they can alter the local microclimate and topsoil moisture, thereby influencing adult activity and pupal survival. Consequently, local pest pressure may differ from the suitability mapped within the same pedoclimatic setting. Based on this, an area considered highly suitable by the model may not necessarily be occupied by the phytophagous species; rather, the results should be interpreted as an index ranking the relative suitability of different areas [37,38]. According to [39], biodiversity research should integrate the combined effects of climate change and land-use change. In this context, the modelling framework adopted here aims to capture interactions among multiple pedoclimatic predictors at the local scale.

4.3. Future Scenario Implications and Management

According to the 2021–2040 SSP5–8.5 projection, the RF model does not predict the emergence of extensive new high-suitability hotspots. Instead, changes in suitability are primarily expressed as a redistribution towards low-to-medium classes across broader areas, with high-suitability zones remaining limited. The most visible changes are concentrated in the hilly Iblean–Calatino region and along the valleys inland. Forecasting future pest trends also enables evaluation of their potential impact on olive production. Therefore, changes in pest suitability were compared with the presence of the plant, within the landscape. This comparison made it possible to verify the spatial consistency between areas that are potentially suitable for *B. oleae* and areas that are actually cultivated, thus providing initial validation of the model. To evaluate the maximum risk scenario, the SSP5-8.5 scenario was selected, an approach also adopted in the study by [7] which was used to identify the maximum potential expansion of climatically suitable areas for the pest. Since the models

were trained by using climate data from 2003 to 2021, and the forecast covers the period from 2021 to 2040, the time needed for the crop to develop and enter production was taken into account. As a consequence, points indicating the presence of olive trees in 2012 were considered. Some areas showed medium to high suitability despite the absence of host presence points. In these cases, the soil and climate conditions are ideal, but the area is too far from crops. Therefore, if new olive crops are to be planted, a soil and climate analysis must be carried out to assess the suitability of the area for phytophagous insects. This is due to mild winters favouring the survival of *B. oleae* in the pupal stage, which increases the risk of infestation [40]. Factors related to the host (e.g., cultivar, and phenology) were not included, although they may affect attack times locally. Therefore, the projections indicate the environmental suitability for B.Olea rather than the expected intensity of infestation.

From an ecological point of view, the future climate scenario does not predict new intense hotspots; on the contrary, it displays large-scale spread, promoting many areas currently classified as unsuitable to areas with low and medium probability pedoclimatic conditions. No significant increases have been observed in areas with a high probability of occurrence; these areas remain limited and selective. The projections of the SSP5-8.5 climate scenario were carried out to assess the potential expansion of areas suitable for the spread of *B. oleae*, particularly in the hilly Iblean–Calatino region and along the inland valleys. The results of this analysis have direct implications for management. According to [38], timely and targeted integrated pest management (IPM) strategies are needed for areas that are already cultivated with olive trees and overlap with areas of high value, as well as for areas where crops will be cultivated in the future. According to [41], the agricultural phase constitutes a significant contributor to the environmental impact of the olive oil life cycle.

Information generated by species distribution models has the potential to inform agronomic decisions and form the basis of new circular economy strategies. However, future projections should be interpreted as climate-driven changes in potential suitability for B.Olea overlaid on the current pedological and land use template. This is because the study does not account for potential shifts in olive cultivation, such as upslope or poleward movements, nor for changes in soil conditions or landscape reconfiguration. Future studies could integrate these maps with olive tree climate suitability maps. Furthermore, integrating them with maps detailing the utilisation of inputs such as chemical or organic fertilisers, including digestate, could be a valuable addition.

The strategic utilisation of biomass as organic fertilisers depends on the geographical proximity between olive mill sites and agricultural areas. Therefore, GIS platforms emerge as valuable tools for supporting territorial planning. Furthermore, integrating suitability maps with LCA analyses facilitates the identification of optimal locations and appropriate scale for new industrial facilities. This supports the promotion of plant establishment in areas with stable biomass flows, where effective pest management ensures the long-term sustainability of the incoming raw material. Based on the results, the methodology developed in this study enabled the optimisation of the monitoring by creating a climate suitability map for the pest. In addition, the comparison of response curves across the territory was suitable to improve pest management, providing farmers and local authorities with useful information for management and planning purposes while containing pest monitoring costs.

4.4. Operational Integration into IPM and Transferability

The resulting maps and response curves offer an integrated approach that can be incorporated into IPM planning as part of a spatial decision support system. In the current scenario, the maps enable areas of varying pest pressure to be identified within the territory. This information can be used to optimise trap placement and monitoring efforts

by focusing field surveys in areas of medium-to-high suitability and reducing efforts in areas of persistently low suitability. In this context, the results of the suitability analysis do not replace field-based IPM thresholds but provide spatial prioritisation to help allocate resources before damage becomes evident. One practical development would be to match suitability classes to local damage thresholds (e.g., trap capture levels and/or fruit infestation rates used by cooperative advisory services), translating the spatial risk layer into action thresholds for monitoring intensity and intervention timing. The same logic can support long-term planning for the 2021–2040 projections by identifying areas where suitability is increasing. This can guide the expansion of monitoring networks and the prioritisation of high-risk olive-growing areas in advance of potential outbreaks. Although this workflow was developed for *B. oleae* in eastern Sicily, it can be easily transferred to other perennial-parasitic cropping systems in the context of climate change, provided georeferenced monitoring data are available and predictors are harmonised within a common spatial model. Integrating the suitability results with crop phenology, management practices and local intervention thresholds would further improve the interpretability and relevance of decisions in such contexts.

5. Conclusions

This research study demonstrated how a high-resolution modelling approach can be effective in delineating the ecological niche of *B. oleae* at a local scale, particularly in the study area of eastern Sicily.

The high predictive power and strong spatial concordance between the two algorithms used in the study (RF and MaxEnt) confirmed the effectiveness of this method. The complex interactions between the selected predictors have made it possible to produce sustainability maps showing the potential distribution of *B. oleae*. The climatic predictors associate the suitability for the species with a moderate temperature regime, defined by Bio 2 and Bio 4, and a seasonal balance of precipitation, defined by Bio15 and Bio16. Soil texture was a critical factor in regulating soil water availability, determining survival in the pupal stage and future infestation intensity. Topography acted as a filter, limiting suitability to areas with altitudes and slopes compatible with olive cultivation practices and, consequently, with the microclimatic requirements of *B. oleae*.

The environmental suitability maps showed a non-homogeneous distribution of the pest in the territory, with environmental suitability increasing in areas where combinations of humidity, soil aeration and thermal stability create microhabitats favourable to the development and proliferation of the species.

These maps can directly support IPM by guiding trap placement, survey frequency, and the prioritisation of interventions in areas predicted as persistently suitable. Overlaying the future suitability projections for the SSP5–8.5 scenario for the period 2021–2040 with the 2012 olive presence map indicated that most existing olive-growing areas will remain at high potential risk, while additional zones, with high suitability but currently without olive groves, are identified as potential future hotspots for infestation. Other climate scenarios could be further investigated to explore future suitability projections based on updated land use and soil evolution scenarios, when available.

The approach developed in this study provides a valid and practical foundation for analysing the potential risks and planning the use of land in agricultural ecosystems. The limitation of this study is connected to the focus on pedoclimatic and topographic data, excluding anthropogenic factors.

In order to improve the functionality of the proposed approach as a decision-support instrument, subsequent developments should assess the incorporation of anthropogenic and crop-related drivers (e.g., irrigation, cultivar/phenology, pruning intensity, and local

management practices) and evaluate whether these variables enhance spatial predictions in marginal areas. Moreover, the integration of SDM outputs with vegetation indices and surface moisture/temperature proxies, as remotely sensed indicators, has the potential to facilitate the updating of possible risk patterns throughout the season and support adaptive monitoring. Further research could include anthropogenic variables, such as management practices, irrigation regimes, and land-use changes, to refine the model accuracy and better capture the complexity of agricultural systems.

Author Contributions: Conceptualisation, G.A.C. and C.A.; methodology, G.A.C., P.R.D. and C.A.; software, G.A.C. and G.P.; validation, P.R.D.; formal analysis, G.A.C., P.R.D. and C.A.; investigation, G.A.C. and G.P.; resources, C.A.; data curation, G.A.C. and G.P.; writing—original draft preparation, G.A.C., G.P. and P.R.D.; writing—review and editing, G.A.C., P.R.D. and C.A.; visualisation, G.A.C. and P.R.D.; supervision, C.A.; project administration, C.A.; funding acquisition, C.A. All authors have read and agreed to the published version of the manuscript.

Funding: The work of prof. Claudia Arcidiacono was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1032 17 June 2022, CN00000022—CUP: E63C22000960006). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them. This research is also in line with the 'Piano incentivi per la ricerca di Ateneo 2024–2026 (DR 2306 del 3 June 2024)—Pia.ce.ri, Linea 2′ project on 'Innovative solutions and strategies for sustainability of Cities, Land and Society' (SIA3) by the University of Catania, coordinated by Claudia Arcidiacono.

Data Availability Statement: The data presented in this study are available at the Sicilian Region for Land Information System (SITR) (https://www.sitr.regione.sicilia.it, accessed on 19 September 2025), the global database WorldClim (https://www.worldclim.org, accessed on 16 July 2020), and Soilgrid250 (https://soilgrids.org, accessed on 2 October 2023). The presence points of *B. Olea* are property of APO Catania; therefore, this information is available upon request by contacting the corresponding author.

Acknowledgments: The authors acknowledge the Sicilian Regional Land Information System (SITR) for data provision (https://www.sitr.regione.sicilia.it/). The authors are grateful to the reviewers for the valuable and constructive comments that significantly improved the manuscript.

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

Abbreviations

The following abbreviations are used throughout this manuscript:

B. oleae	*Bactrocera oleae* (Rossi)
Bio 2	Mean Diurnal Range
Bio 4	Temperature Seasonality
Bio 15	Precipitation Seasonality
Bio 16	Precipitation of Wettest Quarter
CMCC-ESM2	Centro Euro-Mediterraneo sui Cambiamenti Climatici-Earth System Model 2
CV	Cross-Validation
DOP	Protected Designation of Origin
DTM	Digital Terrain Model
GIS	Geographic Information System
IDW	Inverse Distance Weighting
IPCC	Intergovernmental Panel on Climate Change
IPM	Integrated Pest Management
ISMEA	Istituto di Servizi per il Mercato Agricolo Alimentare
MaxEnt	Maximum Entropy

RF	Random Forest
ROC	Receiver Operating Characteristic
SDM	Species Distribution Model
TSS	True Skill Statistic
AUC	Area Under the Curve

References

1. IPCC. *Climate Change 2021: The Physical Science Basis*; Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press: Cambridge, UK, 2021.
2. Rondoni, G.; Mattioli, E.; Giannuzzi, V.A.; Chierici, E.; Betti, A.; Natale, G.; Petacchi, R.; Famiani, F.; Natale, A.; Conti, E. Evaluation of the effect of agroclimatic variables on the probability and timing of olive fruit fly attack. *Front. Plant Sci.* **2024**, *15*, 1401669. [CrossRef]
3. Volpi, I.; Guidotti, D.; Mammini, M.; Petacchi, R.; Marchi, S. Managing complex datasets to predict *Bactrocera oleae* infestation at the regional scale. *Comput. Electron. Agric.* **2020**, *179*, 105867. [CrossRef]
4. Istituto di Servizi per il Mercato Agricolo Alimentare (ISMEA). Tendenze e Dinamiche Recenti Olio D'oliva—Luglio 2025. Available online: https://www.ismeamercati.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/13559 (accessed on 18 September 2025).
5. Benelli, G.; Daane, K.M.; Canale, A.; Niu, C.Y.; Messing, R.H.; Vargas, R.I. Sexual communication and related behaviours in Tephritidae: Current knowledge and potential applications for Integrated Pest Management. *J. Pest Sci.* **2014**, *87*, 385–405. [CrossRef]
6. Kumar, S.; Neven, L.G.; Yee, W.L. Evaluating correlative and mechanistic niche models for assessing the risk of pest establishment. *Ecosphere* **2014**, *5*, art85. [CrossRef]
7. Ashraf, U.; Chaudhry, M.N.; Peterson, A.T. Ecological niche models of biotic interactions predict increasing pest risk to olive cultivars with changing climate. *Ecosphere* **2021**, *12*, e03714. [CrossRef]
8. Jin, Z.; Yu, W.; Zhao, H.; Xian, X.; Jing, K.; Yang, N.; Lu, X.; Liu, W. Potential Global Distribution of Invasive Alien Species, *Anthonomus grandis* Boheman, Under Current and Future Climate Using Optimal MaxEnt Model. *Agriculture* **2022**, *12*, 1759. [CrossRef]
9. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing; R Core Team: Vienna, Austria, 2023.
10. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
11. Catalano, G.A.; D'Urso, P.R.; Maci, F.; Arcidiacono, C. Influence of Parameters in SDM Application on Citrus Presence in Mediterranean Area. *Sustainability* **2023**, *15*, 7656. [CrossRef]
12. Guisan, A.; Graham, C.H.; Elith, J.; Huettmann, F.; NCEAS Species Distribution Modelling Group. Sensitivity of predictive species distribution models to change in grain size. *Divers. Distrib.* **2007**, *13*, 332–340. [CrossRef]
13. Preu, M.; Frieß, J.L.; Breckling, B.; Schröder, W. Case Study 1: Olive Fruit Fly (*Bactrocera oleae*). In *Gene Drives at Tipping Points*; Springer: Cham, Switzerland, 2020; pp. 79–101. [CrossRef]
14. Perez-Navarro, M.A.; Broennimann, O.; Esteve, M.A.; Bagaria, G.; Guisan, A.; Lloret, F. Comparing climatic suitability and niche distances to explain populations responses to extreme climatic events. *Ecography* **2022**, *2022*, e06263. [CrossRef]
15. Nietupski, T.C.; Kim, J.B.; Tortorelli, C.M.; Lemons, R.; Kerns, B.K. *Ventenata dubia* projected to expand in the western United States despite future novel conditions. *Ecosphere* **2024**, *15*, e4979. [CrossRef]
16. Hijmans, R.J.; Phillips, S.; Leathwick, J.; Elith, J. *dismo: Species Distribution Modeling*, R package version 1.1-4; CRAN: Windhoek, Namibia, 2017.
17. Catalano, G.A.; D'Urso, P.R.; Arcidiacono, C. Predicting Potential Biomass Production by Geospatial Modelling: The Case Study of Citrus in a Mediterranean Area. *Ecol. Inform.* **2024**, *83*, 102848. [CrossRef]
18. Dimou, I.; Lambropoulos, P.A.; Kapatos, E.T. Soil factors affecting pupal distribution and mortality of the olive fruit fly *Bactrocera oleae* (Gmelin) (Diptera: Tephritidae). *J. Appl. Entomol.* **2003**, *127*, 469–473. [CrossRef]
19. Poggio, L.; De Sousa, L.M.; Batjes, N.H.; Heuvelink, G.B.; Kempen, B.; Ribeiro, E.; Rossiter, D. SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil* **2021**, *7*, 217–240. [CrossRef]
20. Da Silva, R.S.; Kumar, L.; Shabani, F.; Picanço, M.C. Potential risk levels of invasive *Neoleucinodes elegantalis* (small tomato borer) in areas optimal for open-field *Solanum lycopersicum* (tomato) cultivation in the present and under predicted climate change. *Pest Manag. Sci.* **2017**, *73*, 616–627. [CrossRef]
21. Catalano, G.A.; Maci, F.; D'Urso, P.R.; Arcidiacono, C. GIS and SDM-Based Methodology for Resource Optimisation: Feasibility Study for Citrus in Mediterranean Area. *Agronomy* **2023**, *13*, 549. [CrossRef]

22. See, K.E.; Ackerman, M.W.; Carmichael, R.A.; Hoffmann, S.L.; Beasley, C. Estimating carrying capacity for juvenile salmon using quantile random forest models. *Ecosphere* **2021**, *12*, e03404. [CrossRef]
23. Hao, T.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography* **2020**, *43*, 549–558. [CrossRef]
24. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]
25. D'Arrigo, G.; Provenzano, F.; Torino, C.; Zoccali, C.; Tripepi, G. I test diagnostici e l'analisi della curva ROC. *G. Ital. Nefrol.* **2011**, *28*, 642–647. [PubMed]
26. Warren, D.L.; Seifert, S.N. Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* **2011**, *21*, 335–342. [CrossRef] [PubMed]
27. Mukherjee, T.; Sharma, V.; Sharma, L.K.; Thakur, M.; Joshi, B.D.; Sharief, A.; Thapa, A.; Dutta, R.; Dolker, S.; Tripathy, B.; et al. Landscape-level habitat management plan through geometric reserve design for critically endangered hangul (*Cervus hanglu hanglu*). *Sci. Total Environ.* **2021**, *777*, 146031. [CrossRef] [PubMed]
28. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232. [CrossRef]
29. Van Beest, F.M.; McLoughlin, P.D.; Mysterud, A.; Brook, R.K. Functional responses in habitat selection are density dependent in a large herbivore. *Ecography* **2016**, *39*, 515–523. [CrossRef]
30. Ortega, M.; Sánchez-Ramos, I.; González-Núñez, M.; Pascual, S. Time course study of Bactrocera oleae (Diptera: Tephritidae) pupae predation in soil: The effect of landscape structure and soil condition: Predation of *Bactrocera oleae*. *Agric. For. Entomol.* **2018**, *20*, 201–207. [CrossRef]
31. Lobo, J.M.; Jiménez-Valverde, A.; Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **2008**, *17*, 145–151. [CrossRef]
32. Caselli, A.; Favaro, R.; Petacchi, R.; Angeli, S. Infestation of the gall midge *Dasineura oleae* provides first evidence of induced plant volatiles in olive leaves. *Bull. Entomol. Res.* **2022**, *112*, 481–493. [CrossRef]
33. Lantero, E.; Matallanas, B.; Callejas, C. Current Status of the Main Olive Pests: Useful Integrated Pest Management Strategies and Genetic Tools. *Appl. Sci.* **2023**, *13*, 12078. [CrossRef]
34. Benhadi-Marín, J.; Santos, S.A.P.; Baptista, P.; Pereira, J.A. Distribution of *Bactrocera oleae* (Rossi, 1790) throughout the Iberian Peninsula based on a maximum entropy modelling approach. *Ann. Appl. Biol.* **2020**, *177*, 112–120. [CrossRef]
35. Neuenschwander, P.; Michelakis, S.; Bigler, F. Abiotic factors affecting mortality of Dacus oleae larvae and pupae in the soil. *Entomol. Exp. Appl.* **1981**, *30*, 1–9. [CrossRef]
36. Peterson, A.T.; Soberón, J.; Pearson, R.G.; Anderson, R.P.; Martínez-Meyer, E.; Nakamura, M.; Araújo, M.B. *Ecological Niches and Geographic Distributions*; Princeton University Press: Princeton, NJ, USA, 2011.
37. Merow, C.; Smith, M.J.; Silander, J.A., Jr. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **2013**, *36*, 1058–1069. [CrossRef]
38. Titeux, N.; Henle, K.; Mihoub, J.-B.; Regos, A.; Geijzendorffer, I.R.; Cramer, W.; Verburg, P.H.; Brotons, L. Global scenarios for biodiversity need to better integrate climate and land use change. *Divers. Distrib.* **2017**, *23*, 1231–1234. [CrossRef]
39. Wang, X.G.; Levy, K.; Nadel, H.; Johnson, M.W.; Blanchet, A.; Argov, Y.; Pickett, C.H.; Daane, K.M. Overwintering survival of olive fruit fly (*Diptera: Tephritidae*) and two introduced parasitoids in California. *Environ. Entomol.* **2013**, *42*, 467–476. [CrossRef]
40. Turner, M.G.; Gardner, R.H.; O'Neill, R.V. *Landscape Ecology in Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2001.
41. Cinardi, G.; D'Urso, P.R.; Arcidiacono, C.; Ingrao, C. Accounting for circular economy principles in Life Cycle Assessments of extra-virgin olive oil supply chains—Findings from a systematic literature review. *Sci. Total Environ.* **2024**, *945*, 173977. [CrossRef] [PubMed]