Journal of Scientific & Industrial Research
Vol. 82, December 2023, pp. 1231-1249
DOI: 10.56042/jsir.v82i12.1973

Ecological Niche Modelling of an Industrially Important Mushroom - *Ganoderma lucidum* (Leys.) Karsten: A Machine Learning Global Appraisal

Manish Mathur[1]* & Preet Mathur[2]

[1]ICAR- Central Arid Zone Research Institute, Jodhpur 342 003, Rajasthan, India

[2]Jodhpur Institute of Engineering and Technology, Jodhpur 342 802, Rajasthan, India

Received 27 May 2023; revised 08 September 2023; accepted 10 September 2023

Species Distribution Modelling (SDM) involves utilizing observations of a given species and its surrounding environment to produce a sound approximation of the species' potential distribution. The intricate relationships between organisms and their surroundings, coupled with the profusion of data, have captured the attention of ecologists and statisticians alike. Consequently, they have directed their efforts towards exploring the potential of machine learning techniques. Our study employs an ensemble machine learning approach to simulate the global ecological niche modelling of *Ganoderma lucidum* fungus. This involves the utilization of various environmental predictors and the averaging of multiple algorithms to achieve a comprehensive analysis. 563 spatially thinned presence points of *G. lucidum* were projected with three bio-climatic time frames, namely current, 2050, and 2070, and four Representative Concentration Pathways (RCPs), namely 2.6, 4.5, 6.0, and 8.5, as well as non-climatic variables (surface soil features, land use, rooting depth and water storage capacity at rooting zone). We observed excellent model qualities as the Area Under the receiver operating Curve (AUC) approached 0.90. Random Forest was identified as the best individual algorithm, while the Maxent entropy was identified as the least effective for Ecological Niche Modelling (ENM) of *G. lucidum*. Globally, under the current bio-climatic and non-bioclimatic projection, optimum habitat for this fungus covers 12510876.3 km^2 area while, maximum area (13248546.9 Sq. km.) under this habitat class with future projections was recorded with RCP of 8.5 in 2070. The primary determinants of its current global distribution were ecosystem rooting depth, water storage capacity, and precipitation seasonality. While, with two future bioclimatic time frames and RCPs, Isothermality was identified as the most influential predictor. Based on our assessment, it has been determined that this particular fungus is exhibiting a persistent pattern of proliferation across the regions of Europe, America, and certain areas of India. The present investigation sought to underscore the importance of discerning the native habitats of this species, taking into account both current and anticipated climatic shifts. This knowledge is essential for effectively coordinating the artificial cultivation and natural harvesting of *G. lucidum*, which is necessary to meet the ever-increasing industrial demands.

Keywords: Bioclimatic variables, Ecosystem rooting depth, Ensemble machine learning, Random forest algorithm, Representative concentration pathways

Introduction

White-rot polypore fungi in the family Ganodermataceae produce double-walled, echinulate basidiospores. The roughness of pileus surfaces and basidiospore shape and size distinguish Ganodermataceae. Despite their diverse macro-morphologies, most *Ganoderma* species lack micro-morphological differentiation.[1] Ganoderma species' root and stem rot disease affects oil palm, rubber, and other trees worldwide.[2,3] *Ganoderma lucidum* is typical of subtropical and temperate regions.[4] *G. lucidum* is the main pathogen on Q*uercus, Prosopis, Acacian Albizia lebbeck*, and *Salvadora* species. Basal rot kills *Prosopis cineraria*, *Dalbergia sisso* and *Acacia tortaliis* trees in dry and semiarid India (Figs 1a, 1b, 1c and 1d). In Figs 1e and 1f, *G. lucidum* basidiocarps are shown. Though it rots wood, this mushroom has many health benefits. It is antibacterial, antifungal, antiviral (especially against herpes and HIV), anticancer, anti-inflammatory, antioxidant, and radical scavenging.[5,6] Polysaccharides like triterpenoids, steroids, fatty acids, and proteins/peptides were linked to these pharmacological properties in clinical studies.[7]

G. lucidum is prioritized for disease treatment and prevention in the Chinese Pharmacopeia. The Therapeutic Compendium and American Herbal Pharmacopoeia[8] classify it as a herb. *G. lucidum* supplements are sold as food and medication to boost immunity and metabolism. Coffee, powdered tea,

*Author for Correspondence
E-mail: eco5320@gmail.com

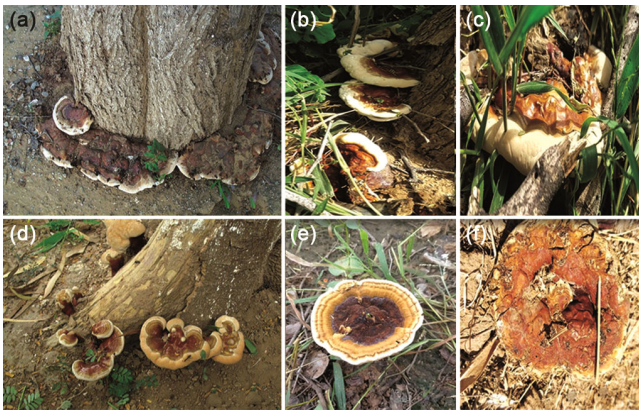

Fig. 1 — *Ganoderma lucidum* causing stem rot in *Prospois cineraria* (A) and in *Dalbergia sisso* (B and C) trees at planted at community grazing lands at Indian semi-arid areas in Sirohi district of Rajasthan; Fruiting bodies of fungus on stem base of *Acacia tortalis* (D*: photo courtesy by Dr. R.K. Bhansali); Fruiting bodies collected from filed (E and F)

supplements, beverages, syrups, toothpastes, soaps, and other products sell well.[9] *Ganoderma* extract is used in facial and cosmetic products because it inhibits the tyrosinase enzyme, which stops melanin production.[10] It improves male hair by reducing dihydrotestosterone and prostatic hyperplasia.[11] Over 1,000 Ganoderma products and 200 medications are available.[12] In China, Japan, and South Korea, *G. lucidum* products sell for over 2.5 billion USD annually.[13] *Ganoderma*-based nutraceuticals are also growing rapidly in India and are expected to reach $25 million USD in 2023.[14,15] Genuine and verified *Ganoderma*-based products are available on online consumer platforms such as Amazon, Flipkart, and others (www.vegamebeljepara.com; www.dazzlinggroup.com; www.dxnmalaysia.com).[16,17]

Due to its rarity and rising global demand for its raw material, this fungus is being artificially cultivated.(13) The collection and conservation of *G. lucidum*, which is vanishing in Indian arid and semi-arid environments, was stressed by Mawar *et al.*[18] in a field inventory. Studies have shown that sawdust, agricultural wastes (rice and wheat bran, sugarcane bagasse, peanut hulls, coconut fibre, etc.), and tea waste can be used for artificial growing on solid substrates.[19] Solid-state fermentation produces *G. lucidum*, which develops into a fruiting body over six months.[20] Fungi cultivation is laborious and uncontrollable, so alternatives are needed.[21] Given the pharmaceutical attributes and concomitant industrial requisites of *G. lucidum*, it is incumbent upon us to delve into its ecological niches and assess the influence of diverse bottom-up and top-down environmental determinants on its habitat suitability on a worldwide basis. In order to do this, we employed ensemble methods, which pooled the results of several individual algorithms into a single prediction.[22]

The speed and breadth of ecological analysis can be improved by using machine learning techniques to speed up the processing and analysis of large troves of raw data. For instance, tremendous progress has been made towards automating species identification from images in community science data as a result of collaborations between machine learning researchers and ecologists. While many computer scientists would be interested in making contributions to ecological problems, they may be put off by the need to learn new concepts and terminology. This is especially true for complex ecological issues that may not have a simple solution using current machine learning techniques. Species distribution modelling (SDM) is one such field, in which observations of a species and its environment are used to construct a best guess for the species' range. Given the abundance of data and the complexity of the interplay between species and their environments, this issue has garnered considerable interest from ecologists and statisticians, who have begun to turn their attention to machine learning approaches.

Our objective in employing Artificial Intelligence and Machine Learning was to approximate the quantity of suitable habitats that could viably sustain this particular species, taking into account the bio-climatic factors, soil composition, root depth, overall plant-accessible water retention capacity, and extant of cultivated and pastured land. The outcomes of the research would contribute to our comprehension of the accessibility of *G. lucidum* in their indigenous environs, especially in consideration of diverse climate change and greenhouse gas forecasts. Moreover, the identification of optimal habitats serves to enable the facilitation of extensive *in-vivo* collection, which may be correlated with the imperative to cultivate *G. lucidum in vitro* in order to satisfy industrial demands.

Material and methods

Data Collections

Presence records for *G. lucidum* were culled from various sources, including the Global Biodiversity Information Facility (https://doi.org/10.15468/dl.m3mh36),(23) the Indian Biodiversity Portal (https://indiabiodiversity.org/

species/show/33318), and scholarly works [6,13,24–29] and from our field-based inventories.[30,31] For this purpose, we used GIS ArcMap[32] and high-resolution Google Earth satellite images to pinpoint these locations in the WGS84 coordinate system. In cases where we didn't have access to occurrence data, we used Google Earth's (http://ditu.google.cn/) geocoding tools to pinpoint precise locations.

All of the aforementioned data was used to compile a CSV file of distributional hotspots. Using the Spatial Thin window of the R-based Graphical User Interface Wallace Software,[33] we filtered our data set to remove spatial autocorrelation and duplicate records with a thinning distance of 10 kilometres. These spatially thinned location points were utilized with *ConR* package to quantify its current Area of Occupancy (AOO) and Extent of Occurrence (EOO) and its current RET - IUCN (Rare, Endangered, Threatened – International Union for Conservation of Nature and Natural Resources) status.

Bio-Climatic (Bio) and Non-Bioclimatic Variables (Non-Bio)

Machine learning techniques can extrapolate future distributions of species from known locations.(34) The bioclimatic variables used to project distributions were drawn from WorldClim version 1.4 (https://worldclim.org/data/cmip6/cmip6clim30s.html - accessed on 21st December, 2022). Using DIVA-GIS version 7.5,(32) we downloaded and converted to ASCII (or ESRI ASCII) 19 bioclimatic variables for both the present and two future climate scenarios. Time frames 2050 and 2070 represent average values from the years 2041 to 2060 and 2061 to 2080, respectively.[35] The four Representative Concentration Pathways (RCPs) for which information was downloaded are the 2.6, 4.5, 6.0, and 8.5 pathways. Supplementary Table 1 gives information on each bioclimatic parameter, including its units and mathematical expressions. ISRIC World Soil Information (https://isric.org/soilgrids) was accessed in order to obtain seven different soil parameters (bulk density in kg/cm^3, Cation Exchange Capacity in $cmolkg^{-1}$, Soil pH in H_2O, Sand, Silt, Clay Percentages, Soil Organic Carbon Stock and nitrogen in $g\ kg^{-1}$ accessed on 15th August 2022). ArcMap was used to access these datasets via WMS servers and process them. (Complete instructions can be found at https://www.isric.org/instruction-wms.html).[36]

In order to quantify the relationships between climate, soil, and plants, the 95% rooting depth was identified as a crucial variable by the Global Climate Observation System's (GCOS) Terrestrial Observation Panel for Climate. International Land Surface Climatology Project (ISLSCP) provided the data on vertical root distribution that encompasses data points from various land covers (evergreen needleleaf, broadleaf, deciduous and mixed forests, grassland, croplands, protected and unprotected shrublands, and baren areas.[37] Data on this parameter was downloaded from http://daac.ornl.gov.

The concept of root zone soil water storage capacity offers valuable insights into the worldwide allocation of the total plant-accessible soil water storage capacity within the rooting zone, meticulously assessed across a 1.0 degree grid,(38) and in the present study, this information was downloaded from https://daac.ornl.gov/ISLSCP II/guides/root water storage 1deg.html. Assimilating NDVI-fPAR (fraction of Absorbed Photosynthetically Active Radiation) and atmospheric forcing data allowed us to calculate the amount of water stored in the rooting zone (in millimetres of H_2O). According to the method described by Ramankutty *et al.,*[39] we additionally downloaded a data-set on global croplands and pasture lands and used it.

Issue of multicollinearity

The analysis involved the examination of cross-correlation using the Pearson Correlation Coefficient (r), while also assessing multicollinearity to mitigate the risk of over-fitting. Furthermore, the Niche Tool Box (Ntbox)[40] was employed to eliminate data points with a cross-correlation coefficient of 0.85 or greater.[41] In this study, the allocation of model training and model validation was conducted with 70% and 30% of the data, respectively.[22]

Projection Assignment and their Transformations

Due to the disparate origins and resolutions of the Bio-Climatic (BIO) and Non-BIO variables, it is imperative to rectify their projections prior to data extraction and the subsequent prediction of the ensemble model. The achievement of this study was facilitated by employing a sequence of procedures in ArcMap utilizing ArcToolbox. Initially, the projection was established within the "projection and transformation" sub-window of Data Management Tools. The WGS 1984 EASE Geographic Coordinate System (GCS) was employed for this purpose. In order to accurately measure the extent of each habitat suitability class, it is necessary to utilize a Projected Coordinate System (PCS) within the "calculate

geometry" function of Arc Map. As a result, we have converted the projections of the habitat class raster file to WGS 1984 web Mercator (auxiliary sphere-3857). This step facilitates the computation of the area encompassed by a particular class, using a unit of measurement specified by the user (in this case, square kilometers were employed).

Species Distribution Modelling

The following modelling algorithms were utilized from the Dismo 1.1–4 package[42]: The utilization of a Generalized Linear Model (GLM) with a Gaussian distribution, the Generalized Additive Model-GAM,[43] Support Vector Machines (SVM), Random Forest-RF,[44] Multivariate Adaptive Spline- MARS, and Maximum Entropy- Maxent v. 3.4.1.[45] These are all prominent techniques employed in ecological research. These models and algorithms enable ecologists to analyze and understand complex ecological phenomena, such as species distribution patterns, community dynamics and habitat suitability. By employing these methods, ecologists can gain valuable insights into the intricate relationships between various ecological variables and make informed decisions regarding conservation and management strategies. The execution of Artificial Neural Network (ANN)[46] and Classification Tree Analysis (CTA)[47] was carried out using the biomod-2 package.[48] The default settings were employed for all algorithms under investigation. The SDM analysis was conducted in a two-step process: (a) determining the optimal model or algorithm based on predefined evaluation criteria (refer to the details below); and (b) employing the ensemble method to model using both climatic and non-climatic predictors.

K-fold cross-validation was employed to evaluate the models, wherein each algorithm underwent 10 folds and 10 replications.[49] Through this approach, the dataset is partitioned into 10 folds in a random manner for every iteration, and subsequently, the model is trained using 9 of these folds.[22] The assessment of the predictive algorithm's performance involved the consideration of several ecological metrics. These metrics included overall accuracy, sensitivity, specificity, Cohen's Kappa, True Skill Statistic (TSS), and Area Under the receiver operating Curve (AUC). Both TSS and KAPPA are metrics that rely on thresholds to assess the quality of a model's performance. Scores above 0.75 are indicative of favourable outcomes. The ecological performance of a model is expected to be below average if the value falls below the threshold of 0.50.[50] The utilization of the AUC metric was employed to assess the efficacy of the classifier in accurately labelling presence or absence data. The AUC statistic encompasses a range of grades that are indicative of the performance of a given system. These grades range from excellent (0.90–1.00), very good (0.8–0.9), adequate (0.7–0.8), acceptable (0.6–0.7), to poor (0.5–0.6).

In the realm of ecological analysis, specificity refers to the proportion of instances in which a model accurately detects the absence of a particular phenomenon. Sensitivity, in ecological terms, refers to the measure of the model's ability to correctly identify the presence of a particular species or phenomenon, expressed as a percentage. The identification of the contribution of each bio-climatic and non-bio-climatic variable was accomplished through the ranking of their respective variable importance.[51] The 'getResponseCurve' function of the SDM program in the R language[52] was employed to estimate the threshold limit of the most influential predictors on the habitat suitability of this species.

Post Modelling Analysis

Habitat Suitability

The ASCII raster outputs generated by the ensemble were imported into ArcMap, where the cell values ranging from 0 to 1 were utilized to classify and delineate various habitat types relevant to this particular species. Species habitat suitability is categorized into different levels, namely optimal, moderate, marginal, low, and inappropriate. This classification is determined by evaluating the constant point break criterion for each of these classes.[53] The raster calculator, a spatial analyst tool in ArcGIS, was employed to assess the overall extent, measured in square kilometres, of land falling within each of these ecological categories. The formula employed for the computation of the percentage alteration in average habitat suitability under ideal class circumstances was utilized to evaluate the potential ramifications of diverse climate scenarios on the projected suitability of habitat.[54,55]

$$\left[\left(\frac{(Future - Current)}{Current}\right) \times 100\right]$$

Non-Parametric Assessment of Internal Matrix of Habitat Classes

All of the examined projections were subjected to a regression analysis to determine the link between the number of polygons and the total area covered by the

four studied habitat suitability classes. We also performed a frequency distribution analysis to understand the internal matrix of polygon areas that fall under the optimal class, and the results were interpreted using the skewness, kurtosis, and Kolmogorov-Smirnov indices.

Network analysis is a graphical statistical method that facilitates rapid visualisation and interpretation of associations between numerous variables. In contrast to latent variable modelling, in which we try to reduce the complexity of the variables by extracting the information they have in common; we instead estimated the structure of all relationships between variables directly. Nodes, the circle-shaped components, that represent variables, make up networks which are graphic structures. Edges are the lines that tie nodes to one another. There are two types of networks: unweighted and weighted. Edges in weighted networks show the size of the relationships, whereas in unweighted networks they only show the relationship between nodes. In other words, the thicker the connection between them, the stronger is the relationship between nodes.

Networks can be categorised as directional or non-directional. In the former, arrows are present at one end of the edges to represent an influence or a path from one node to another. The edges of non-directional networks lack arrows, so they are simply lines that do not signify the direction in which relationships between variables is to be found.[39] In this study, we used JASP software's non-directional weighted network analysis[83] to understand the relationships between the four different habitat suitability classes and their associated predictors, namely the current Bio+Non-Bio, 2050, and 2070 with their respective RCPs ranging from 2.6 to 8.5.

Results

Data processing and Multicollinearity

A total of 1164 data points were gathered for this particular species, spanning various locations across the globe. To ensure data accuracy and avoid redundancy, we employed Wallace Software's Spatial Thin window technique, utilizing a thinning distance of 10 kilometres. This process effectively filtered out all but one occurrence of a record within each specific location. A total of 563 instances documenting the presence of *G. lucidum* were gathered to complete the ecological niche modelling (ENM) procedure. The ConR software was utilized to analyse the dispersed coordinates and ascertain the present Extent of Occurrence (EOO) and Area of Occupancy (AOO). The ecological findings of this exercise are visually represented in Fig. 2, showcasing a numerical value of 228,840,466.37 km^2 EOO, while the AOO covers 2252 km^2. These figures suggest that the species in question must be currently classified as having a status of 'least concern' and 'near threatened' on the red list, respectively. The distribution of this species encompasses various regions, including Spain, Austria, Hungary, the United Kingdom, Belgium, Denmark, the southern portion of Norway, Sweden, areas adjacent to New York, Washington, and Mexico in the USA, as well as the southern and western parts of India. Additionally, this species can be found to some extent in China, Taiwan, Thailand, Bangladesh, and near Cape Town in South Africa (Fig. 2).

The findings of the correlation analysis conducted on various bioclimatic variables can be found in Supplementary Table 2. In order to effectively tackle the problem of multicollinearity in species distribution modelling, we use the methodologies put forth by Pradhan *et al.*[41] Our observations indicate a

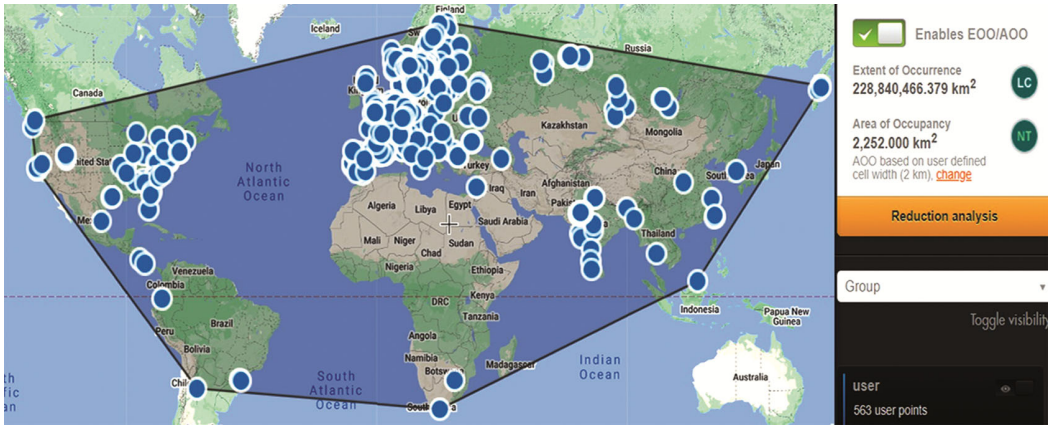

Fig. 2 — Graphical display of spatially thinned 563 presence point of *G. lucidum* convey Extent of Occurrence (EOO) and Area of Occupancy (AOO)

significant correlation between Bio-2 and Bio-19, as well as other bioclimatic variables, across various timeframes including the present and future scenarios projected by RCPs for 2050 and 2070. The correlation between the Warmest Quarter precipitation (Bio-18) and other biological variables was found to be the lowest. As a result, this particular variable was chosen for the ENM analysis of the species, considering all RCPs. Isothermality (Bio-3) and Mean Temperature of Wettest Quarter (Bio-8) exhibited similar patterns, with the exception of the years 2050 RCP 2.6 and 2070 RCP 8.5 respectively. In contrast to the aforementioned patterns, the analysis focused solely on the utilization of Annual Mean Temperature (Bio - 1) and Temperature Seasonality (Bio- 4) specifically for the year 2050 under the RCP 2.6 scenario. Among the soil variables considered, it is noteworthy that the cation exchange capacity (CEC), nitrogen content, and silt content display notable correlations with other variables. Consequently, these variables have been excluded from subsequent analysis.

Model Performances

Excellent model qualities were recorded as AUC approached ~ 0.90 based on current climate evaluation criteria and RCPs corresponding to timeframes of 2050 and 2070. (Supplementary Tables 3 to 5). Random Forest was deemed the best algorithm to use because its AUC values were the highest across the board for climatic predictors. The Random Forest technique yielded Kappa and TSS values approaches 0.75, indicating high levels of model quality. The Maxent tool, when used with all climate predictors, produced low Kappa values, indicating that this widely used ENM technique was less effective for this species. The results of other individual algorithms, however, were more or less similar to one another, so we chose to use the ensemble model's results because it automatically averaged the results of the various algorithms and determined how much space was occupied under different classes of habitat suitability.

Variable Importance Percentage

In each analysis, the Variable Importance Percentage (VIP) was computed to assess the relative influence of the bioclimatic and non-bioclimatic predictors on the species' distribution pattern. With current bio-climatic, surface soil, rooting depth, water storage capacity and cropland and pasture land, our analysis revealed the influence of ecosystem rooting depth (VIP 17.5), croplands (VIP of 12.99) and water storage capacity (VIP = 10.36) on the distribution of this species (Fig. 3) in descending order, respectively. Among the bio-climatic and soil variables, precipitation seasonality (Bio-15; VIP = 7.77), clay percentage (VIP = 7.77) and SOC (VIP = 7.23) also showed their effects on this species. ROC curves of ecosystem rooting depth (Fig. 4) and water storage capacity (Fig. 5) revealed the trends by which these predictors affect the habitat suitability of this fungus.

VIP values for the bioclimatic years 2050 and 2070, along with the four RCPs are presented in Table 1. RCP 2.6 of 2050 time-frame exhibited the dominance of temperature seasonality (BIO-4) and annual mean temperature (Bio-1) with VIP values of 18.61 and 14.8 respectively. While with remaining RCPs with 2050 (RCP 4.5, 6.0 and 8.5) and all four RCPs with 2070 time-frames, Isothermality (Bio-3) was identified as most important factor that governs the suitability classes of this fungus (Table 1). We

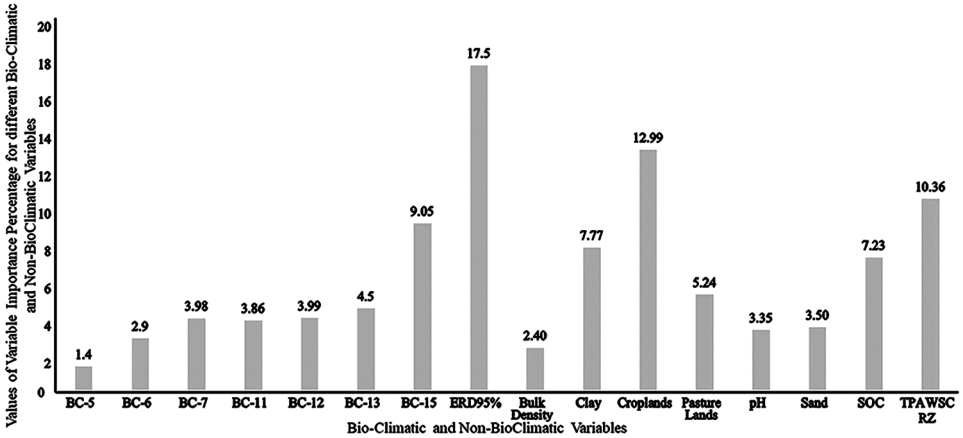

Fig. 3 — VIP of Current bio-climate, surface soil, ecosystem rooting depth (95%), total plant-available water storage capacity of the rooting zone (TPAWSC RZ) and area under crop and pasture lands variables

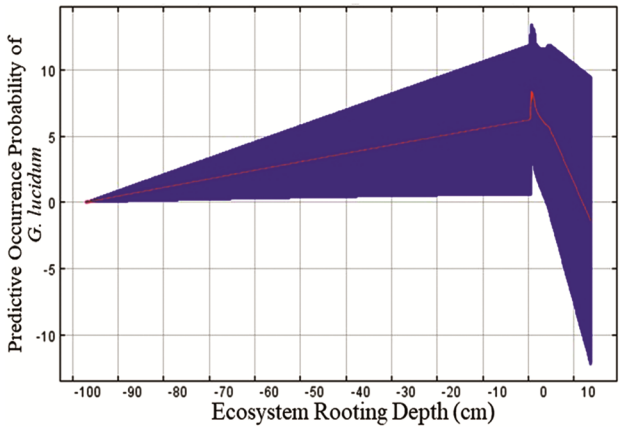

Fig. 4 — Response curve analysis depict impacts of ecosystem rooting depth on the habitat suitability of *Ganoderma lucidum*

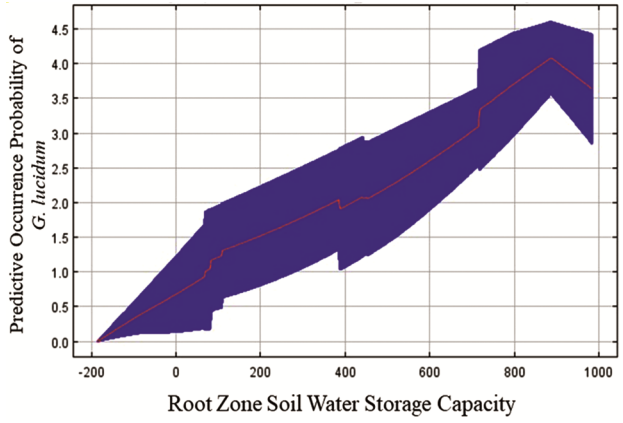

Fig. 5 — Response curve analysis depicts the impacts of total plant available water storage capacity on the habitat suitability of *Ganoderma lucidum*

measured an Isothermality value of less than 100 (~30), which indicated less extreme monthly temperature swings than annual ones. The ROC curves of these bio-climatic variables are presented in Supplementary figure 1.

Furthermore, we observed that the minimum temperature of the coldest month (Bio-6) closely matched the Bio-3 in terms of determining the habitat types of this fungus with the last three RCPs of 2050. Temperature annual range (Bio-7) indicated such tendencies with RCP 2.6 and RCP 4.5 of 2070. While the cooler quarter's mean temperature (Bio-11) and the rainiest quarter's precipitation (Bio-16) was the next most pertinent elements for the remaining RCPs of this climatic period.

Polygon Counts and Habitat Class Areas (Square Kilometres)

Number of polygons under different suitability classes with different projections is presented in Table 2. We recorded lowest number of all classes with current Bio + Non-Bio while the maximum number of polygons under optimum (381), moderate (800) and marginal (1048) were recorded with 2070 RCP 8.5. Interestingly the highest number of polygons (1875) under low suitability class was recorded with 2050 RCP 4.5. However, the largest drops in number of optimum polygons were recorded with 2050 RCP 8.5 (175) and 2070 RCP 4.5 (177).

Area under each habitat suitability class, calculated area (km^2) with different climatic and non-climatic variables is depicted in Table 3. Global spatial extents

Table 1 — Percentage contribution (or Variable Importance Percentage) of Bio-climatic variables of two-time frames with four RCPs. Bioclimatic variables with higher VIP values were used to make the link between species and their environment

Variables	2050				2070			
	RCP 2.6	RCP 4.5	RCP 6.0	RCP 8.5	RCP 2.6	RCP 4.5	RCP 6.0	RCP 8.5
Bio-1	14.8	x	x	x	x	x	x	x
Bio -3	x	23.63	25.6	23.12	22.37	29.52	22.55	23.19
Bio -4	18.61	x	x	x	x	x	x	x
Bio -5	x	8.52	10.61	5.04	10.07	12.73	x	14.43
Bio -6	x	19.75	18.86	22.57	x	x	5.54	x
Bio -7	7.26	x	x	x	17.4	23.76	x	x
Bio -8	13.72	7.95	6.24	4.12	5.63	5.18	6.4	x
Bio -9	x	4.99	6.11	5.4	7.23	x	8.34	X
Bio -10	x	x	x	6.83	x	x	6.97	X
Bio -11	X	x	x	x	x	x	14.98	X
Bio -12	11.54	8.64	x	x	x	x	-	X
Bio -13	X	x	x	x	x	9.15	10.59	X
Bio -14	4.87	4.35	4.95	x	11.08	x	9.48	14.89
Bio -15	X	x	x	7.71	10.22	13.92	9.07	11.45
Bio -16	8.67	6.64	12.96	10.68	10.31	x	x	18.13
Bio -17	9.05	8.18	7.42	7.93	5.64	X	x	X
Bio -18	11.4	7.27	7.93	6.45	x	5.73	6	17.88

x =eliminated variables due to their significant co-linearity with other variables

Table 2 — Numbers of polygons under four suitability classes with studied projections

Projections	Optimum	Moderate	Marginal	Low
Current Bio+Non-Bio	44	146	261	241
2050 RCP 2.6	164	473	846	1117
2050 RCP 4.5	316	688	866	806
2050 RCP 6.0	300	630	721	1541
2050 RCP 8.5	175	569	844	1875
2070 RCP 2.6	183	591	815	1334
2070 RCP 4.5	177	570	791	630
2070 RCP 6.0	253	624	779	1676
2070 RCP 8.5	381	800	1048	961

Table 3 — Area under different habitat suitability classes of *Ganoderma lucidum* with studied bio-climatic and non-bioclimatic variables

Environmental variables	Habitat suitability classes (Sq Km.)				
	Optimum	Moderate	Marginal	Low	Total Area
Current Bio & Non-Bio*	12510876.3	16325109.0	15407813.0	70899344.6	116767923.4
2050 RCP 2.6	8296264.5	15287614.2	17543945.2	75640099.3	126307117.4
2050 RCP 4.5	9911256.9	14608205.5	18858054.8	82929600.0	94168836.1
2050 RCP 6.0	10756475.8	8401995.5	17573263.6	57437101.0	125040657.4
2050 RCP 8.5	8900935.3	13272445.1	23524876.7	79342400.3	122198873.7
2070 RCP 2.6	8330954.8	15175554.8	24006584.3	74685779.6	126372954.0
2070 RCP 4.5	8902062.4	15511899.9	26135415.5	75823576.1	122312390.4
2070 RCP 6.0	9966043.9	11734253.5	22883141.1	77728951.7	130375926.4
2070 RCP 8.5	13248546.9	14039611.4	23187136.0	79900632.0	115143143.2

Non-Bio* = surface soil, ecosystem rooting depth (95%), total plant-available water storage capacity of the rooting zone (TPAWSC RZ) and area under crop and pasture lands variables

of different classes with predictors are presented in Fig. 6 (current bio-climatic and non-bioclimatic variables), fig. 7 (2050 RCPs 2.6 to 8.5) and fig. 8 (2070 RCPs 2.6 to 8.5). Maximum (13248546.94 sq. km) area under optimum class was recorded with 2070 RCP 8.5. Currently this class covering 128510876.39 sq. km area, while the least (8596264.55 sq. km) area for this class was recorded with 2050 RCP 2.6. With reference to current optimum area, we recorded per cent decline with all four studied RCPs with 2050 time-frame in an order of −33.69 (RCP 2.6), −20.78 (RCP 4.5), −14.02 (RCP 6.0) and −28.85 (RCP 8.5 Fig. 9A). However, with 2070 climatic projection using RCP 8.5 we recorded + 5.90 per cent increase in optimum area compared to current optimum area, while rest of the RCPs with this period showed a decline similar to 2050 in order of RCP 2.6>4.5>6.0 (Fig. 9A).

Moderate suitability area was recorded maximum (16325109.05 sq. km) with current Bio and Non-Bio variables, while, lowest area under this class (8401995.54) was recorded with RCP 6.0 (2050). Results of per cent changes in this class with respect to current time-frames have been presented through fig. 9B. Results suggest that the area declined under this class with all RCPs belongs to both time frames (Fig. 9B). Highest decline (−48.53) was recorded with 2050 RCP 6.0 followed by 2070 RCP 6.0 (−28.12). Highest marginal area (26135415.59 sq. km) was recorded with 2070 RCP 4.5 and lowest (15407813.07 sq. km) was recorded with current BIO and non-BIO variables. Our per cent change calculation revealed overall gain under this suitability class which ranged from 13.86 (2050 RCP 2.6) to 69.62 (2070 RCP 4.5, Fig. 9C). Area under low suitability class was recorded maximum (82929600.08 sq. km) and minimum (57437101.07 sq. km) with 2050 RCP 4.5 and RCP 6.0, respectively. With reference to per cent change, we recorded gain under all scenarios except 2050 RCP 6.0 that exhibited decline (−18.99, Fig. 9D). Total calculated area under all the suitability classes was recorded maximum (130375926.47 sq. km) with 2070 RCP 8.5 and the lowest (94168836.15 sq. km) was recorded with 2050 RCP 6.0. Under total area, we recorded overall gain for the scenario with respect to current area (Fig. 9E) and as obvious exception was recorded with 2050 RCP 6.0.

Current spatial extent (12510876.39 sq. km.) of this fungus under the optimum suitability class is dominant in the European countries (10259449.39 sq.

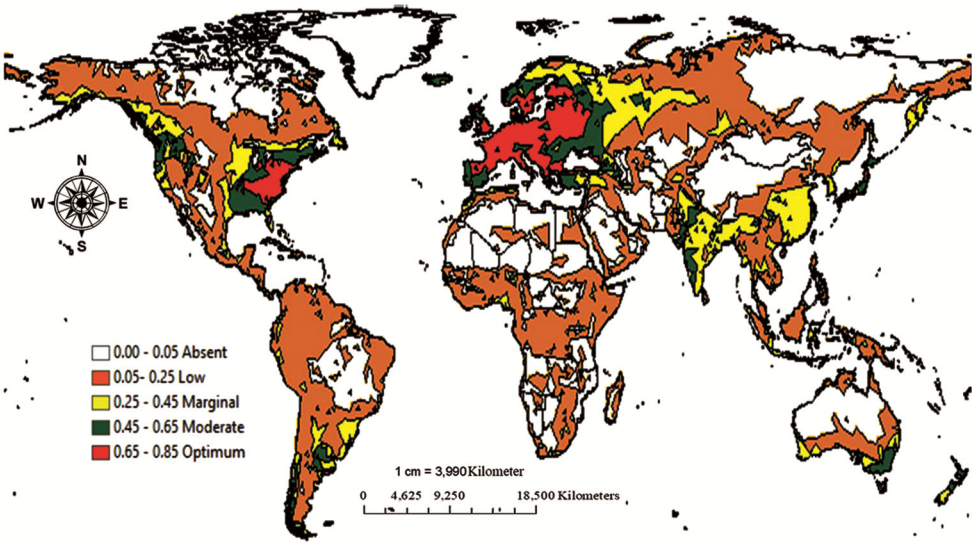

Fig. 6 — Current bio-climate, surface soil, ecosystem rooting depth (95%), total plant-available water storage capacity of the rooting zone (TPAWSC RZ) and area under crop and pasture lands variables

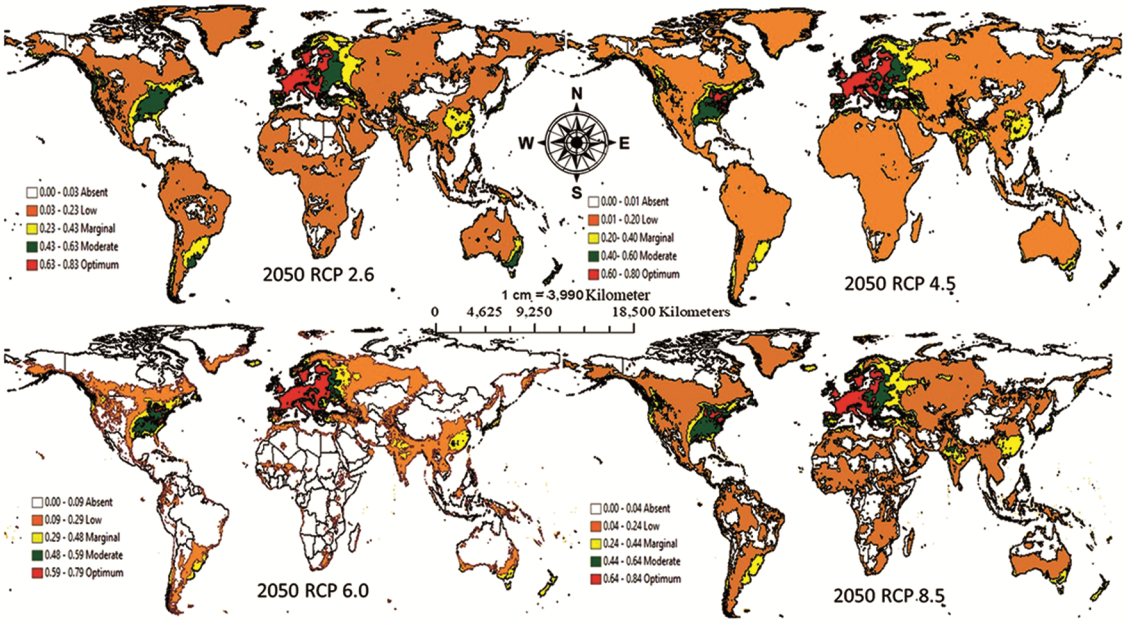

Fig. 7 — Habitat suitability of *G. lucidum* with 2050 bio-climatic predictors (a) RCP 2.6; (b) RCP 4.5; (c) RCP 6.0 and (d) RCP 8.5

km.) like France, Germany, Italy, Greece, Spain, Poland and Denmark. Our GIS analysis revealed such areas under 22 different polygons. Under the similar timeframe and suitability class, USA was identified as the second largest area where this fungus covers 2052618.3 sq. km area with 14 polygons. Philippines (93343.23 sq. km.), Argentina (36873.26 sq. km.) and India (17955.72 sq. km) have this species, categorized under the same class within two polygons each.

The findings of our regression analysis, which explore the relationship between the number of polygons and the corresponding area in km^2, are illustrated in Fig. 10. Analysis revealed significant linear relationships between these two variables, which can be interpreted as a correlation with respect to the area (Km^2.) = 43121 × Number of Polygon = 0.74^{**} ($P > 0.01$).

Distribution Analysis

Results of frequency distribution analysis for the optimum class calculated with all studied projections are presented in Table 4. A data set is symmetric if it looks the same when viewed from either the left or right side of the center. In a normal distribution, both

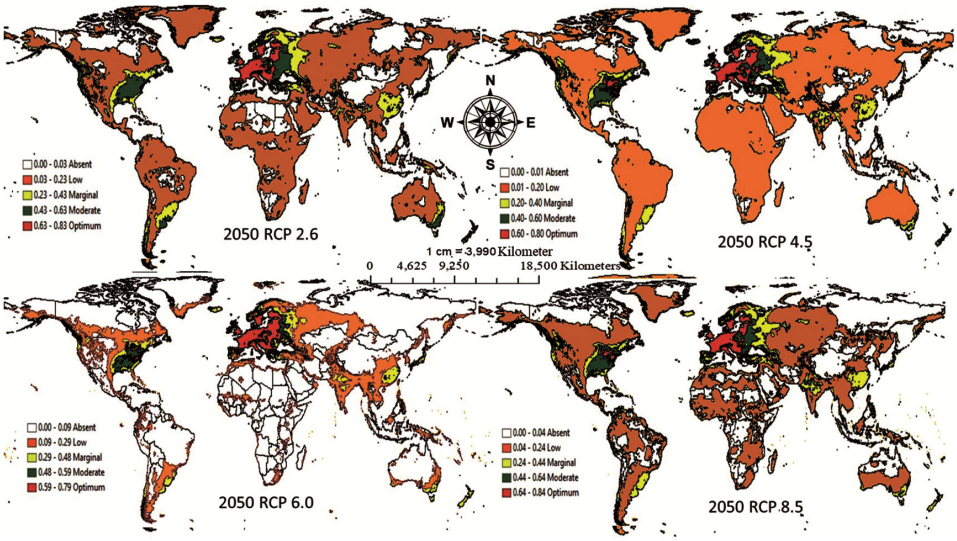

Fig. 8 — Habitat suitability of *G. lucidum* with 2070 bio-climatic predictors (a) RCP 2.6; (b) RCP 4.5; (c) RCP 6.0 and (d) RCP 8.5

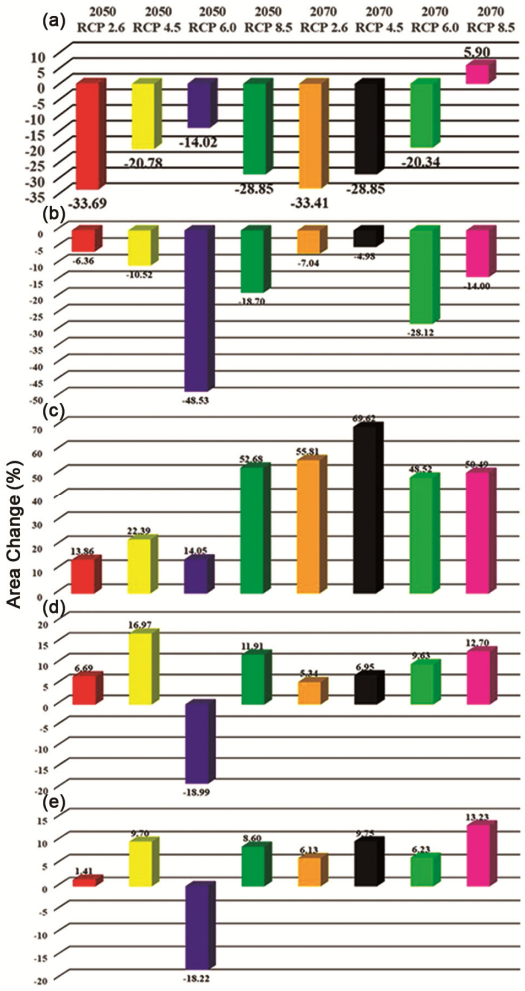

Fig. 9 — Percentage change area under different habitat suitability classes calculated with different future bioclimatic time-frames and RCPs with reference to current bio-climatic and non-bioclimatic variables

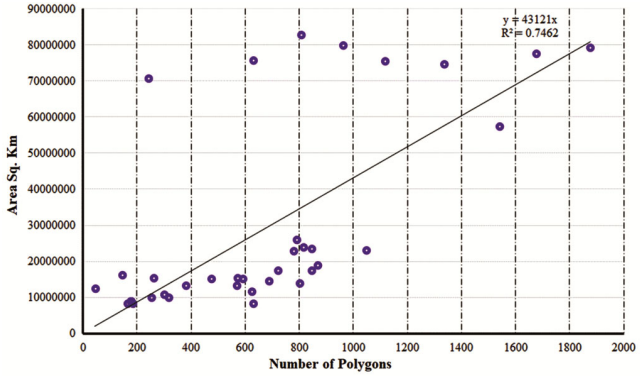

Fig. 10 — Regression analysis with number of polygons and area covered by them under all studied projection (no = 36; degree of freedom no-2 = 34 were 0.329 and 0.429 at 95 and 99 per cent, respectively)

skewness and kurtosis have values close to zero. If the skewness is less than −1 or greater than +1, we classify the distribution as highly skewed; if it's between −1/2 and +1/2, we call it moderately skewed; and if it's between −1/2 and +1, we call it approximately symmetric.[38] In the present study, the skewness recorded was more than 1 with all projection and it ranged from 5.89 (current Bio + Non-Bio) to 17.11 (2050 RCP 4.5) which indicates highly skewed distribution nature of areas covered under the category of optimum class.

Kurtosis measures how closely a distribution's shape resembles that of a Gaussian distribution; typically, it is zero for a Gaussian distribution and negative for a flatter distribution. A normal distribution with a three-point kurtosis serves as the reference standard. Mesokurtic refers to a normal

Table 4 — Examining the internal distribution matrix of the areas (sq. km.) classified as optimum habitat suitability enclosed in polygons under various projections

Projections	Polygon Numbers	Skewness	Kurtosis	Kolmogorov-Smirnov Test
Current BIO+NBIO	44	5.89	34.46	0.481
2050 RCP 2.6	164	11.42	135	0.487
2050 RCP 4.5	316	17.11	296.86	0.46
2050 RCP 6.0	300	17.04	290.53	0.472
2050 RCP 8.5	175	12.77	163.96	0.461
2070 RCP 2.6	183	12.88	169.97	0.471
2070 RCP 4.5	177	13.01	168.87	0.474
2070 RCP 6.0	253	15.44	240.06	0.481
2070 RCP 8.5	381	16.02	276.77	0.46

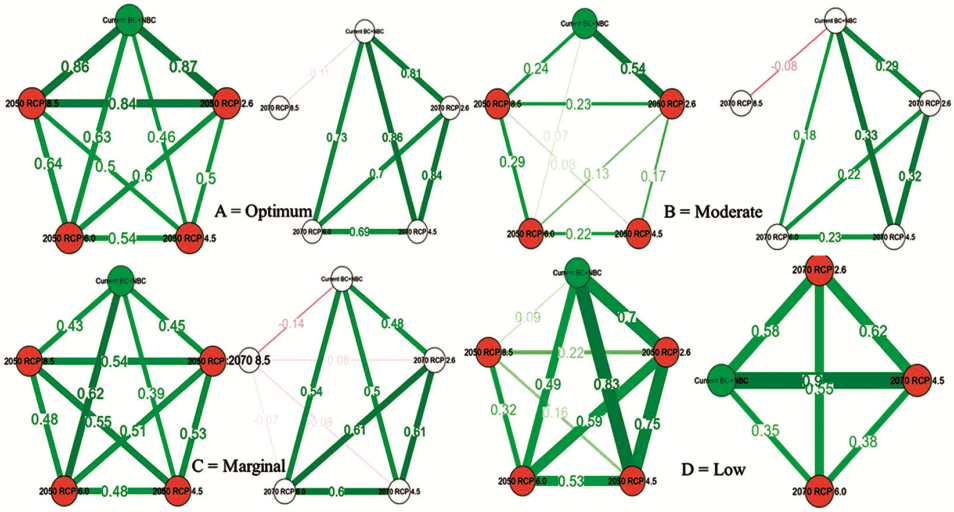

Fig. 11 — Network comparison between the areas (Sq. km.) classified under different suitability classes with different projections. Circle represent the projection; line or edge showed relationship with different nodes (edge width also corresponds to the strength of their relationship)

distribution with kurtosis exactly 3, platykurtic refers to one with kurtosis less than 3, and leptokurtic refers to one with kurtosis greater than 3. (central peak is higher and sharper). Our analysis revealed leptokurtic type distribution as the kurtosis value ranged from 34.46 (current Bio + Non-Bio) to 296 (2050 RCP 4.5).

The Kolmogorov-Smirnov (D) test relies on quantifying the most significant vertical divergence observed between the theoretical and empirical cumulative distribution functions. Additionally, the null and alternative hypotheses for this test posit that the data pertaining to a particular variable, obtained from various respondents, adhere to a normal distribution (H0), while the alternative hypothesis (HA) suggests that the data deviate from a normal distribution. If the determined P value for this test exceeds the alpha level (0.05), it is appropriate to accept the null hypothesis that the data adheres to a normal distribution. Conversely, if the P value is lower than the alpha level, it is appropriate to reject the null hypothesis and accept the alternative hypothesis.[57] Our K-S test values were ranged from 0.46 to 0.48 (Table 4) and as our calculated *P* values were lower than the alpha level, we reject the null hypothesis and accept the alternate hypothesis suggesting that the sample does not follow normal distribution.

Network Analysis

Results of network analysis are depicted in fig. 11. Within optimum category, current Bio +Non- Bio showed stronger relationships with 2050 RCP 2.6 (r^2 = 0.87) and 2050 RCP 8.5 (r^2 = 0.86). Both RCPs also exhibited higher relationships with each other (r^2 = 0.84 Fig. 11 A). With similar climatic time-frame, we did not have such stronger relationships for moderate and marginal suitability classes (Figs. 11 B and C). However, for low suitability class, our analysis revealed higher relationships of current Bio

+Non- Bio with 2050 RCP 2.6 (r^2 = 0.7) and 2050 RCP 4.5 (r^2 = 0.83) and RCP 2.6 and RCP 4.5 also had higher relationships with each other (r^2 = 0.75, Fig. 11 D). With 2070 climatic time-frame, our analysis revealed that RCP 8.5 had the weakest relationships with current BIO+NBIO (Figs. 11 A, B, C, D). With similar time-frame, optimum class showed stronger relationships between current BIO+NBIO and 2070 RCP 2.6 (r^2 = 0.81), RCP 4.5 (r^2 = 0.86), both these RCPs also significantly related with each other (r^2 = 0.86). Comparatively, moderate and marginal suitability classes exhibited weaker relationships. With low suitability class, we recorded strongest relationship between current Bio+Non-Bio and 2070 RCP 4.5 (Fig. 11 D).

Discussion

Plant fungi exhibit a global distribution and possess the ability to flourish in a wide range of climatic conditions, including tropical, subtropical, and temperate regions. In contrast to plants and animals, the advancement of ecoclimatic distribution or niche modeling in fungi is relatively limited. This could be attributed to the fact that conducting such analyses requires a comprehensive understanding of ecological principles.[58]

SDM encompasses a broad range of statistical methodologies employed to estimate the prospective geographical distribution and ecological preferences of a given species. [59] The utilization of fungal species distribution modeling (F-SDM) has witnessed a substantial surge in recent years across various sectors. F-SDMs have been developed for a diverse range of purposes, primarily classified into three overarching categories: to evaluate the environmental factors associated with the presence of organisms, predict their occurrence in specific regions, and utilize fungi as a representative species to explore both methodological and ecological aspects.[60] Research has demonstrated the significance of SDM and ENM in assessing plant fungi within various scientific surveys. These surveys encompass investigations aimed at determining the probability of disease transmission,[61–63] evaluating the vulnerability of plant hosts to the pathogen, and estimating associated costs.[64,65] The numerical output generated by SDM can manifest as a response curve, illustrating the influence of predictors on species distribution, or as an equation that links the anticipated expansion of a species to a set of environmental predictors. The application of SDM has been observed in various fields such as biogeography, global change biology, and conservation management.[66,67]

The utilization of SDM is accompanied by the inherent limitation of encompassing a vast array of alternatives, each exhibiting distinct outcomes contingent upon the prevailing circumstances. Consequently, this characteristic poses a heightened level of complexity when it comes to determining the optimal alternative for a given situation. According to researchers, instances such as this one, occur when it becomes crucial to anticipate the dispersal patterns of a species in response to different climate change scenarios. An additional concern arises when an abundance of anticipated ecological variables are incorporated, resulting in an overabundance of adjustments.[68] Over-adjustments often result in reduced applicability of models to novel datasets.[69] Utilizing ensemble techniques, which provide enhanced precision compared to individual methods, represents a viable approach to address this challenge. Ensemble modelling is proposed as a means to enhance the predictability of models and assess their similarity in output. When multiple models are integrated, each outcome signifies a distinct potential condition of the authentic distribution, as elucidated by ensemble (or occasionally consensus) modeling. The ecological approach employed in this methodology entails the utilization of averaging projections and amalgamating individual-model projections to generate a comprehensive final surface.[22]

The ensemble method is a valuable approach in ecological research as it allows for precise predictions of species-environment interactions and the subsequent changes in their distribution over time and space. This method offers advantages over using single models, particularly in the context of modeling future climate projections with a large number of species.[70] Additionally, it helps address the challenge of underfitting when predicting rare species. The validation of our decision to utilize the ensemble tool's output for the ENM of *G. lucidum* has been confirmed by these findings. This approach effectively combines the results of multiple algorithms, which exhibit minimal variations among them. Delgado-Baquerizo *et al.*[71] employed data derived from a comprehensive global survey and a long-term nine-year field experiment to demonstrate

the positive correlation between rising temperatures and the abundance of saprophytic fungi. Global predictions suggest that the species richness of plant fungi is expected to undergo an increase in response to various climate change and land-use scenarios. These findings have facilitated our understanding of the global dispersion patterns of *G. lucidum* and its adaptive responses to prevailing and anticipated environmental circumstances.

According to Větrovský *et al.,*[72] various environmental factors, play a significant role in shaping the global distribution of fungal species, with climate being particularly influential for the most prevalent ones. The researchers provided an assessment of the ecological significance of various factors, namely mean driest quarter temperature (Bio-9), precipitation seasonality (Bio-15), wettest quarter temperature (Bio-8), coldest quarter precipitation (Bio-19), mean diurnal range (Bio-2), gross primary production, bulk density, and pH. These factors were ranked in descending order based on their relative importance. By employing the Maxent algorithm, Wang *et al*.[73] conducted a predictive analysis to determine the potential distribution of *Pseudomonas syringae* pv. *actinidiae* (Psa) within China. Their research findings indicate that as per different emission projections for the future, the distribution and locations of suitable habitats vary from the current state. Furthermore, the primary factor driving the spread of Psa was predominantly attributed to the highest recorded temperature in April (19%), the average temperature during the coldest quarter (41%), the amount of precipitation in the month of May (11.5%), and the lowest temperature observed in October (10.8%). By employing Maxent machine learning technology, Bosso *et al*.,[74] conducted an investigation on the geographical distribution of *Xanthomonas fastidiosa*. They also identified specific features that enable the presence of this species within various regions of Italy. This species has been observed to have high probabilities (> 0.8) of occurrence in habitats characterized by specific conditions. These conditions include low altitude (150 m above sea level), low summer rainfall (10 mm), high mean temperature during the coldest quarter, low annual precipitation, and a combination of complex and agricultural patterns.

The ecological impact of *Macrophomina phaseolina* dry root rot and charcoal rot is amplified as temperature rises and moisture stress intensifies. As the temperature increases (reaching 35–40°C), the production of pathogenic microsclerotia occurs, thereby increasing the vulnerability of hosts to infection.[75] The aridity of the atmosphere and elevated temperatures serve as favorable conditions for the proliferation of microsclerotia, leading to heightened production of hydrolytic enzymes within these structures. This enzymatic activity plays a crucial role in promoting the onset of infection.[76–78] The composition of plant fungal communities (including Ascomycetes and Basidiomycetes) is influenced by abiotic environmental factors such as soil pH, nutrient levels, particle size distribution, and climate.[79–81] The temporal effects of these predictive factors, both top-down and bottom-up, are also observed. In the context of two Japanese mountain ecosystems, it was observed that variations in the fungal community composition were primarily attributed to disparities in climate conditions.[82] Conversely, in European ecosystems, the predominant factor influencing the composition and diversity of arbuscular mycorrhizal (AM) fungi was found to be soil characteristics.[81]

The ecological impact of *Macrophomina phaseolina* dry root rot and charcoal rot is amplified as temperature rises and moisture stress intensifies. As the temperature increases (reaching 35–40°C), the production of pathogenic microsclerotia occurs, thereby increasing the vulnerability of hosts to infection[75]. The aridity of the atmosphere and elevated temperatures serve as favorable conditions for the proliferation of microsclerotia, leading to heightened production of hydrolytic enzymes within these structures. This enzymatic activity plays a crucial role in promoting the onset of infection.[76–78] The composition of plant fungal communities (including Ascomycetes and Basidiomycetes) is influenced by abiotic environmental factors such as soil pH, nutrient levels, particle size distribution, and climate.[79–81] The temporal effects of these predictive factors, both top-down and bottom-up, are also observed. In the context of two Japanese mountain ecosystems, it was observed that variations in the fungal community composition were primarily attributed to disparities in climate conditions.[82,83] Conversely, in European ecosystems, the predominant factor influencing the composition and diversity of Arbuscular Mycorrhizal (AM) fungi was found to be soil characteristics.[81]

Drawing upon the insights provided by the Response Curve Function (ROC), we elucidate the ecological dynamics underlying the colonization

patterns of this fungus in relation to the ideal depth of roots (as illustrated in Fig. 12) and the availability of moisture. The projected outcome for optimal root depth is depicted in Fig. 4. Based on the analysis conducted, it has been determined that a depth of 4.5 cm is the most optimal for the growth of this particular fungus. Any depth exceeding this threshold significantly diminishes the suitability of the habitat for the fungus. The findings from our analysis of the second variable (Fig. 5) indicate a direct and positive correlation between moisture availability in the rooting zone and the suitability of the habitat for this particular fungus. This relationship remains consistent up to a moisture availability of 800 mm. Moreover, it is well-established that the propagation of this fungal infection occurs through direct transmission between roots, underscoring the significance of root lengths in this context. The positive correlation between root length and root availability is a widely recognized phenomenon in ecological studies. These plant structures, which play a crucial role in nutrient uptake and anchoring, are frequently found in abundance within landscapes characterized by high plant cover, such as forests and croplands. In terms of habitat suitability, it is worth noting that croplands hold greater significance for this particular species compared to pasture land due to the relatively shorter root length observed in the latter land use category.

Based on this finding, it can be inferred that the species in question is primarily influenced by energy variables or temperature-related variables, rather than water-related variables, within the bio-climatic context. Previous studies conducted on this particular species have yielded comparable findings, specifically pertaining to the fungus's development in relation to temperature.[13] In the context of various climatic timeframes and greenhouse gas scenarios, it has been observed that the species in question is significantly influenced by isothermality. Isothermality refers to the ratio between the temperature fluctuations experienced during day and night, as compared to the fluctuations occurring between summer and winter (over the course of a year). This factor holds utmost significance for the species under consideration. An isothermal value of 100 signifies that the diurnal temperature fluctuations are akin to the annual temperature fluctuations, whereas any value below 100 (around 30) suggests that the temperature variability within an average month surpasses that of the entire year. Lower isothermality values were observed, suggesting that the dispersal of this fungus may be more influenced by seasonal variations rather than monthly temperature fluctuations. The ROC curve of this bio-climatic predictor exhibited dissimilarities with the other ROC curves when considering the 2070 RPC 8.5 scenario. Based on the given temporal context and RCP, we have observed an increased magnitude in the spatial extent of this particular predictor.

Based on scientific research, it has been observed that the mycelial growth of this particular fungus is influenced by temperature variations. This environmental factor plays a significant role in facilitating the proliferation of bacteria on the fruiting

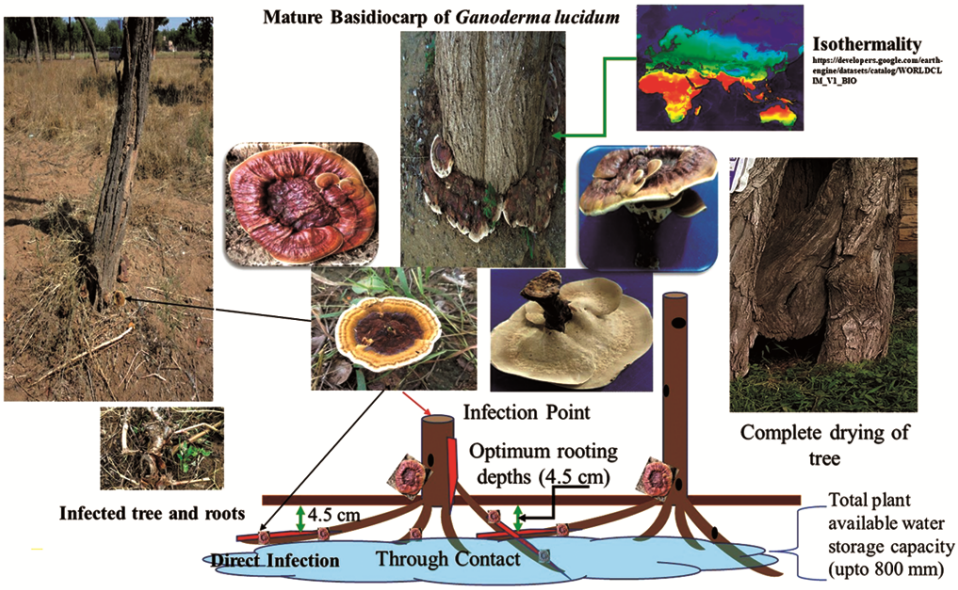

Fig. 12 — predisposing factors of *Ganoderma lucidum* under natural conditions

body of the fungus.[84,85] Through their enzymatic action, these bacteria play a crucial role in influencing the polysaccharide and other microelement composition within the fruiting bodies of *G. lucidum*. This support provided by the bacteria ultimately contributes to the growth and development of this fungus. Furthermore, *Sediminibacterium, Ochrobactrum*, and *Rhodococcus* demonstrate this facilitation behavior on their fruiting bodies in conditions where the temperature remains below 40°C.[52] As per Zhang *et al.,*[86] the colonization and development of *G. lucidum* are intricately linked to the moisture and nutrient levels present in the substrate. These factors play a crucial role in facilitating bacterial proliferation on the fruiting body. The density and growth of basidiocarp of *G. lucidum* are influenced by temperature. Specifically, the minimum and maximum temperatures for optimal density and growth range from 15 to 35℃. It is worth noting that the most favorable temperature range for growth falls between 30 and 35℃.[87]

Additionally, the worldwide distribution of this particular species has been verified by the Centre for Agriculture and Bioscience International. https://www.cabi.org/ISC/abstract/20056500566.[18,88–90] The spatial extent of this fungus, specific to its region, was thoroughly described. This field-based prediction allows us to estimate its production output in the wild, tailored to each specific region. Furthermore, this can be synchronized with its in-vitro propagation to fulfill industrial demands.[15] The area (sq. km) within each suitability class showed significant spatial variability according to our frequency distribution analysis for all examined projections. The single largest area for *G. lucidum* within the optimum class (9470055 sq. km.) was noticed with current Bio + Non-Bio, whereas the smallest area (90.04 sq. km.) for the same habitat type was found during 2070 RCP 6.0. Such trends for moderate (5988240 sq. km 2070 RCP 8.5 to 0.01 sq. km. during 2050 RCP 6.0), marginal (9535635 sq. km. 2070 RCP 2.6 to as low as 400 meters during 2050 RCP 4.5) and low (149656100 sq. km. 2070 RCP 8.5 to only 31 meters during 2070 RCP 4.5) suitability classes were also recorded. These trends revealed spatial heterogeneity for this fungus. With the aid of DIVA GIS, we identified smaller areas around the world, via creation of Keyhole Markup Language (KMZ) files that we provided as supplemental information with this manuscript. With the help of this tool, the smallest suitable area on earth can be located with pinpoint geocoordinates. With climate change projections, recognizing these areas not only makes it easier to comprehend its niche dynamics, but it can also help us to synchronise *G. lucidum* industrial demand and supply. The greatest potential for this species, for instance, is found in European nations, followed by a few U.S. states. To improve assessments from Asian, African, and Australian regions, field-based research in wild/remote locations is necessary. As a result, our current revelation can help future field-based explorations, as well as cross-check predictions from this study.

Conclusions

Ganoderma lucidum fungus is classed as near threatened, with a 2252 km^2 Area of Occupancy (AOO). Our ensemble machine learning technique produced excellent model characteristics (AUC 0.90) and Random Forest algorithm was comparatively more accurate than other individual algorithms. Its present global spread was mostly determined by ecosystem roots depth, water storage capacity, and precipitation seasonality. While bio-climatic-energy variables (temperature) will regulate its future worldwide distribution more efficiently than water factors (precipitation). This fungus, we may presume, is constantly expanding across Europe, North and South America, and parts of India.

Supplementary Data

Supplementary data associated with this article is available in the electronic form at https://nopr.niscpr.res.in/jinfo/jsir/JSIR_82(12)1231-1249_SupplData.pdf

List of abbreviations

Area of Occupancy (AOO)
Area Under the Receiver Operating Curve (AUC)
Artificial Neural Network (ANN)
Bio-Climatic (Bio) and Non-Bioclimatic Variables (Non-Bio)
Classification Trees (CART)
Ecological Niche Model (ENM)
Extent of Occurrence (EOO)
Generalized Additive Models (GAM)
Generalized Linear Models (GLM)
Genetic Algorithm (GARP)
Mahalanobis distances (MD)
Maximum Entropy (Maxent),

Multivariate Adaptive Regression Splines (MARS)
Random Forest (RF)
Receiver-Operating Characteristic (ROC)

Availability of data and materials

The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Author Contribution:

Manish Mathur: Conceptualization, resources, methodology, writing—review and editing

Preet Mathur software, validation, formal analysis, and visualization

Acknowledgements

The senior author expresses gratitude to the Director of ICAR-CAZRI for granting approval to participate in a training program on R-Programming. This training will enhance the author's ability to utilize ENM modeling techniques, thereby increasing their effectiveness in their work. Miss Preet Mathur, a student from Jodhpur Institute of Engineering and Technology in Jodhpur, India, expresses gratitude towards their Director for providing valuable academic assistance.

References

1 Bhansali R R, Ganoderma disease of woody plants of Indian arid zone and their biological control, in *Plant Defense: Biological Control, Progress in Biological Control*, edited by J M Merillon and K G Ramawat (*Springer Science*) 2012, 209–238, DOI 10.1007/978-94-007-1933-0_9.

2 Sankaran K V, Bridge P D & Gokulapalan C, Ganoderma diseases of perennial crops in India – an overview, *Mycopathologia* **159** (2005) 143–152.

3 Naidu Y, Siddiqui Y, Rafii M Y, Saud H M & Idris A S, Investigating the effect of basal stem rot infected oil palm wood blocks: biochemical and anatomical characterization, *Ind Crops Prod*, **108** (2017) 872–882.

4 Oke M A, Afolabi F J, Oyeleke O O, Kilani T A, Adeosun A R, Olanbiwoninu A A & Adebayo E A, *Ganoderma lucidum*: Unutilized natural medicine and promising future solution to emerging diseases in Africa, *Front Pharmacol*, (2022) 2876, doi: 10.3389/fphar.2022.952027\ (2002)

5 Ko H H, Hung C F, Wang J P & Lin C N, Antiinflammatory triterpenoids and steroids from Ganoderma lucidum and *G. tsugae*, *Phytochem*, **69 (1**) (2008) 234–239.

6 Basnet B B, Liu L, Bao L & Liu H, Current and future perspective on antimicrobial and anti-parasitic activities of Ganoderma sp.: an update, *Mycology,* **8** (2017) 111–124.

7 Bishop K S, Kao C H J, Xu Y, Glucina M P, Paterson R R M & Ferguson L R, From 2000 years of Ganoderma lucidum to recent developments in nutraceuticals, *Phytochem*, **114** (2015) 56–65.

8 Hapuarachchi K K, Elkhateeb W A & Karunarathna S C, Current status of global Ganoderma cultivation, products, industry and market, *Mycosphere*, **9(5)** (2018) 1025–1052.

9 Lai T, Gao Y & Zhou S F, Global marketing of medicinal Ling Zhi mushroom Ganoderma lucidum (W.Curt:Fr.) Lloyd (Aphyllophoromycetideae) products and safety concerns. *Int J Med Mushrooms*, **6(2)** (2004) 189–194

10 Chien C, Tsai M, Chen C, Chang S J &Tseng C H, Effects on tyrosinase activity by the extracts of ganoderma lucidum and related mushrooms, *Mycopathologia*, **166(2)** (2008) 117–120.

11 Meehan K, Composition to promote hair growth in humans, *U.S. Patent US9144542*, 29 September 2015.

12 Chen R Y, Kang J & Du G H, Construction of the quality control system of Ganoderma Products, *Edible Medicinal Mushrooms*, **24(6)** (2016) 339–344.

13 Bijalwan A, Bahuguna K, Vasishth A, Singh A, Chaudhary S, Dongariyal A, Thakur T K, Kaushik S, Ansari M J, Alfarraj S, Alharbi S A, Skalicky M & Brestic M, Growth performance of Gandoerma lucidum using billet method in Garhwal Himalaya India, *Saudi J Biol Sci*, **28** (2021) 2709–2717.

14 El Sheikha A F, Nutritional profile and health benefits of *Ganoderma lucidum* "Lingzhi, Reishi, or Mannentake" as functional foods: Current scenario and future perspectives, *Foods*, **11** (2022) 1030.

15 Bijalwan A, Kalpana B, Vasishth A, Singh A, Chaudhary S, Tyagi A, Thakur M P, Thakur T K, Dobriyal M, Kaushal R, Singh A, Maithani N, Kumar D, Kothari G & Chourasia P K, Insights of medicinal mushroom (Ganoderma lucidum): prospects and potential in India, *Biodiversity Int J*, **4(5)** (2020) 202–209.

16 Taofiq O, González-Paramás A M, Martins A, Barreiro M F & Ferreira, I C F R Mushrooms extracts and compounds in cosmetics, cosmeceuticals and nutricosmetics—A review, *Ind Crops Prod*, **90** (2016) 38–48.

17 Wu Y, Choi M H, Li J, Yang H & Shin H J, Mushroom cosmetics: The present and future, *Cosmetics*, **3** (2018) 22.

18 Mawar R, Ram L, Deepesh & Mathur T, Gandoderma in *Beneficial Microbes in Agro-Ecology Amaresan* edited by N, Kumar M S, K Annapurna, K Kumar & A Sankaranarayana (Academic Press, Elsevier, United Kingdom) 2020, ISBN: 978-0-12-823414-3, https://doi.org/10.1016/B978-0-12-823414-3.00031–9

19 Roy S M, Ara A, Jahan K, Das K, Munshi S K, Artificial cultivation of Ganoderma lucidum (Reishi medicinal mushroom) using different sawdusts as substrates, *Am J Biosci*, **3** (2015) 178–182.

20 Magday J, Bungihan M & Dulay R, Optimization of mycelial growth and cultivation of fruiting body of Philippine wild strain of *Ganoderma lucidum*, *Curr Res Environ Appl Mycol,* **4 (2)** (2017) 162–172.

21 Subedi K, Basnet B B, Pandey R, Neupane M & Tripathi G R, Optimizatin of growth condition and biological activities of Nepalese Ganoderma lucidum strain Philippine, *Adv Pharmaco Pharma Sci*, (2021), https://doi.org/ 10.1155/2021/4888979

22 Mathur P & Mathur M, Machine learning ensemble species distribution modeling of an endangered arid land tree *Tecomella undulata*: a global appraisal, *Arab J Geosci*, **16** (2023) 1–28.

23 GBIF.org GBIF Occurrence Download https://doi.org/10.15468/dl.m3mh36 (Accessed on 17August (2022).

24 Bhansali R R, Butt rots of Prosopis cineraria, in *Diversification of arid farming systems* edited by P Narain, A Kar, S Kathju & P Kumar (Arid Zone Research Association of India and Scientific Publishers, India) 2008

25 Bakshi R K, Reddy M A R & Singh S, Ganoderma lucidum root rot mortality in khair acacia catechu in re forested stands, *Eur J Forest Pathol*, **6(1)** (1976) 30–38.

26 Khara H S, Incidence of Ganoderma lucidum root rot on some tree species around Ludhiana, *Plant Dis Res*, **8(2)** (1993) 136–137.

27 Khara H S & Singh J, Diagnosis of Ganoderma lucidum root rot of trees, the cultural characteristics, spore germination and percentage of root-decay, *Plant Dis Res*, **12(2)** (1997) 108–112.

28 Pilotti C A, Stem rots of oil palm caused by Ganoderma boninense: Pathogen biology and epidemiology. *Mycopathologia*, **159(1)** (2005) 129–137, doi:10.1007/ s11046-004-4435-3

29 Shah K K, Tiwari I, Modi B, Pandey H P, Subedi S & Shrestha J, Shisham (Dalbergia sisso) decline by dieback disease, root pathogens and their management: a review, *J Agr Nat Res*, **4 (2)** (2021) 255–272.

30 Jindal S K, Singh D V, Moharana P C & Patel N, *Annual Report* (ICAR-Central Arid Zone Research Institute, Jodhpur, India) 2009, 156.

31 Jindal S K, Singh D V, Moharana P C & Patel N, *Annual Report* (ICAR-Central Arid Zone Research Institute, Jodhpur, India) 2010, 174.

32 Coban H O, Orucu O K & Arslan E S, MaxEnt modelling for predicting the current and future potential geographical distribution of Quercus libani Olivier, *Sustainability* 2671 (2020), doi:10.3390/su12072671.

33 Kass J M, Vilela B, Aiello-Lammens M E, Muscarella R, Merow C & Anderson R P, Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion, *Methods Ecol Evol*, **9** (2018) 1151–1156, https://doi.org/10.1111/2041-210X.12945

34 Wei B, Wang R L, Hou K, Wang X Y & Wu W, Predicting the current and future cultivation regions of Carthamus tinctorius L. using MaxEnt model under climate change in China, *Global Ecol Conser*, **16** (2018), https://doi.org/10.1016/j.gecco.2018.e00477

35 Hijmans R J, Guarino L, Cruz M & Rojas E, Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS, *Plant Genetic Resource and Newsletter,* **127** (2001). 15–19.

36 Cotrina Sánchez A, Rojas Briceño N B, Bandopadhyay S, Ghosh S, Torres G C, Oliva M, Guzman B K & Salas L R, Biogeographic Distribution of Cedrela spp. Genus in Peru Using MaxEnt Modelling: A Conservation and Restoration Approach*, Diversity*, **13** (2021) 261, https://doi.org/10.3390/ d13060261

37 Schenk H J & Jackson R B, ISLSCP II ecosystem rooting depths, in *ISLSCP Initiative II Collection Data set,* edited by F G Hall, G Collatz, B Meeson, S Los, E Brown de Colstoun, & D Landis (Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.) 2009, Available on-line [http://daac.ornl.gov/] from doi:10.3334/ORNLDAAC/929

38 Kleidon A, ISLSCP II Total plant-available soil water storage capacity of the rooting zone, in *ISLSCP Initiative II Collection Data set* edited by F G Hall, G Collatz, B Meeson, S Los, E Brown de Colstoun & D Landis (Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.) 2011, Available on-line [http://daac.ornl.gov/]. doi:10.3334/ORNLDAAC/1006

39 Ramankutty N, Evan A T, Monfreda C & Foley J A, *Global Agricultural Lands: Pastures, 2000* (Data distributed by the Socioeconomic Data and Applications Center) 2010. http://sedac.ciesin.columbia.edu/es/aglands.html.

40 Osorio-Olivera L, Lira-Noriega A, Soberon J, Townsend P A, Facon M, Contreas-Diaz R G, Martinez-Meyer E, Barve V & Barve N, Ntbox: an R package with graphical user interface for modelling and evaluating multidimensional ecological niches, *Methods Ecol Evol*, **11** (2020) 1199–1206, doi:10.1111/2041-210X.13452. https://github.com/luismurao/ntbox

41 Pradhan P, Strengthening Maxent modelling through screening of redundant explanatory Bioclimatic Variables with Variance inflation factor analysis, *Researcher*, **8(5)** (2016) 29–34

42 Hijmans R J, Phillips S, Leathwick J & Elith J, *dismo: Species Distribution Modelling* 2017, https://cran.r-project.org/web/packages/dismo/dismo.pdf

43 Wood S, mgcv: *Mixed GAM computation vehicle with automatic smoothness estimation* 2019, https://stat.ethz.ch/R-manual/R-devel/library/mgcv/html/00Index.html

44 Cutler F & Wiener R, *Random Forest: Breiman and Cutler's random forests for classification and regression*, 2018. https://doi.org/10.1023/A:1010933404324

45 Layola M R R, Semwal M, Rana T S & Nair N K, Predicting potential suitable habitat for Ensete glaucum (Roxb.) Cheesman using Maxent modeling, *Flora* **287** (2022), https://doi.org/10.1016/j.flora.2022.152007.

46 Ripley B & Venables W, *nnet: Feed-forward neural networks and multinomial log-linear models* 2020, https://cran.r-project.org/web/packages/nnet/nnet.pdf.

47 Therneau T, Atkinson B & Port B R, *rpart: recursive partitioning and regression trees* 2019, https://cran.r-project.org/web/packages/rpart/rpart.pdf.

48 Thuiller W, Lafourcade B, Engler R & Araujo M, BIOMOD—A platform for ensemble forecasting of species distributions, *Ecography,* **32** (2009) 369–373.

49 Rajamanickam V, Babel H, Montano-Herrera L, Ehsani A, Stiefel F, Haider S, Presser B & Knapp B, About Model Validation in Bioprocessing, *Processes,* **9** (2021) 961, https://doi.org/10.3390/ pr9060961

50 Ahmad R, Khuroo A A, Hamid M, Charles B & Rashid I Predicting invasion potential and niche dynamics of Parthenium hysterophorus (Congress grass) in India under projected climate changes, *Biod Conser*, **28** (2019) 2319–2344, https://doi.org/10.1007/s10531-019-01775-y

51 Irving K, Jahnig S C & Kuemmerlen M, Identifying and applying an optimum set of environmental variable in species

distribution models, *Inland Waters,* **10(1)** (2020) 11–28. (https://doi.org/10.1080/20442041.2019.1653111

52 Naimi B & Araujo M B, sdm: a reproducible and extensible R platform for species distribution modelling, *Ecography,* **39** (2016) 368–375, DOI: 10.1111/ecog.01881

53 Khan A M, Li Q, Saqib Z, Khan N, Habib T, Khalid N, Majeed M & Tariq A, MaxEnt modelling and impact of climate change on habitat suitability variations of economically important Chilgoza Pine (*Pinus gerardiana* Wall.) in South Asia, *Forests,* **13** (2022) 715 https://doi.org/10.3390/f130507150.

54 Mathur M, Spatio-Temporal Variability's in Distribution Patterns of *Tribulus terrestris*: Linking Patterns and Processes, *J Agr Sci Technol*, **16** (2014) 1187–1201.

55 Wright A N, Schwartz M W, Hijmans R J & Shaffer H B Advances in climate models from CMIP3 to CMIP5 do not change predictions of future habitat suitability for California reptiles and amphibians, *Clim Change*, **134** (2016) 579–591, https://doi.org/10.1007/s10584-015-1552-6

56 Epskamp S, Waldorp L J, Mõttus R & Borsboom D, The Gaussian graphical model in cross-sectional and time-series data, *Multi Beh Res*, **53(4)** (2019) 453–80. https://doi.org/10.1080/00273171.2018. 1454823

57 Love J, Selker R, Marsma M, Jamil T, Dropmannn D, Verhagen J, Ly A, Gronau, Q F, Smira M, Spskamp S, Matzke D, Wild A, Knight P, Rouder J N, Morey R D & Wagenmakers E J, JASP: Graphical statistical software for common statistical designs, *J Stat Softw,* **88(2)** (2019) 1–17, http://dx.doi.org/10.18637/jss.v088.i02

58 Ireland K B & Kiritcos D J Why are plant pathogens under-represented in eco-climatic nice modelling? *Int J Pest Manag*, **65(3)** (2019) 207–216, https://doi.org/10.1080/ 09670874.2018.1543910

59 Mathur M & Sundaramoorthy S, Woody Perennial Diversity at Various Land forms of the Five Agro-Climatic Zones of Rajasthan, India, in *Biodiversity and Chemotaxonomy Sustainable Development and Biodiversity* edited by K Ramawat (Springer, Cham) **24** 2019. https://doi.org/10.1007/978-3-030-30746-2_5. Print ISBN 978-3-030-30745-5. Online ISBN 978-3-030-30746-2.

60 Hao T, Guillera-Arroita G, May T W, Lahoz-Monfort J J Elith J, Using species distribution models for fungi, *Fungal Biol Rev,* **34(2)** (2020) 74–88.

61 Yonow T, Kriticos D J & Medd R W, The potential geographic range of *Pyrenophora semeniperda*. *Phytopathology*, **94(8)** (2004) 805–812.

62 Watt M S, Ganley R J, Kriticos D J & Manning L K, Dothistroma needle blight and pitch canker: the current and future potential distribution of two devastating diseases of Pinus species, *Can J forest Res*, **41(2)** (2011) 412–424.

63 Yonow T, Hattingh V & de Villiers M, CLIMEX modelling of the potential global distribution of the citrus black spot disease caused by *Guignardia citricarpa* and the risk posed to Europe, *Crop Prot*, **1(44)** (2013) 18–28.

64 Watt M S, Kriticos D J, Alcaraz S, Brown A V & Leriche A, The hosts and potential geographic range of Dothistroma needle blight, *Forest Ecol Manag,* **257(6)** (2009) 1505–1519

65 Ganley R J, Watt M S, Kriticos D J, Hopkins A J M & Manning L K, Increased risk of pitch canker to Australasia under climate change, *Australasian Plant Pathol*, **40(3)** (2011) 228–237.

66 Li Y, Li M, Li C & Liu Z, Optimized maxent model predictions of climate change impacts on the suitable distribution of cunninghamia lanceolata in China, *Forests,* **11** (2020) 302.

67 Pecchi M, Marchi M, Burton V, Giannetti F, Moriondo M, Bernetti I, Bindi M & Chirici G, Species distribution modelling to support forest management. A literature reviews, *Ecological Modell,* **411** (2019) 108817.

68 Breiner F, Guisan A, Bergamini A & Nobis M, Overcoming limitations of modelling rare species by using ensembles of small models, *Methods Ecol Evol*, **6** (2016) 1210–1218.

69 Merow C, Smith M, Guisan A, Mcmahon S, Normand S., Thuiller W, Rafael W, Zimmermann N & Elith J, What do we gain from simplicity versus complexity in species distribution models? *Ecography*, **37** (2014) 1267–1281.

70 Grenouillet G, Buisson L, Casajus N & Lek S, Ensemble modelling of species distribution: the effects of geographical and environmental ranges, *Echography*, **34** (2011) 9–17. https://doi.org/10.1111/j.1600-0587.2010.06152.x

71 Delgado-Baquerizo M, Guerra C A, Cano-Diaz C, Egidi E, Wang J T, Eisenhauer N, Singh B K & Maestre F T, The proportion of soil-borne pathogen increases with warming at the global scale, *Nat Clim Change,* **10(6)** (2020) 550–554 https://doi.org/10.1038/s41558-020-0759-3

72 Větrovský T, Kohout P, Kopecký M, Machac A, Man M, Bahnmann B D, Brabcová V, Choi J, Meszárošová L, Human Z R, Lepinay C, Lladó S, López-Mondéjar R, Martinović T, Mašínová T, Morais D, Navrátilová D, Odriozola I, Štursová M, Švec K, Tláskal V, Urbanová M, Wan J, Žifčáková L, Howe A, Ladau J, Peay K G, Storch D, Wild J & Baldrian P, A meta-analysis of global fungal distribution reveals climate-driven patterns, *Nature Comm,* **10** (2019) 5142 https://doi.org/10.1038/s41467-019-13164-8

73 Wang R, Li Q, He S, Liu Y, Wang M & Jiang G, Modelling and mapping the current and future distribution of *Pseudomonas syringae* pv. *actinidiae* under climate change in China. *PLoS ONE*, **13(2)** (2018) e0192153, https://doi.org/10.1371/ journal.pone.0192153

74 Bosso L, Russo D D, Febbraro M, Cristinzio G & Zonia A, Potential distribution of Xylella fstidiosa in Italy: a maximum entropy model, *Phytopathologia Mediterr*, **55(1)** (2016) 62–72.

75 Olaya G & Abawi G S, Effect of water potential on mycelial growth and on production and germination of sclerotia of *Macrophomina phaseolina*, *Plant Dis,* **80** (1996) 1347–1350.

76 Ratnoo R S, Jain K L & Bhatnagar M K, Variations in Macrophomina phaseolina isolates of Ash-gray stem blight of cowpea, *J Mycol Plant Pathol,* **27** (1997), 91–92.

77 Kaur S, Chauhan V B, Singh J P & Singh R B, Status of *Macrophomina* stem canker disease of pigeonpea in eastern Uttar Pradesh, *J Food Leg,* **25** (2012a) 76–78.

78 Kaur S, Dhillon G S, Brar S K, Vallad G E, Chand R & Chauhan V B, Emerging phyopathogen *Macrophomina* phaseolina: biology, economic importance and current diagnostic trends, *Crit Review Microbiol* (2012b) 1–16.

79 Zheng Y, Kim Y C, Tian X F, Chen L, Yang W, Gao C, Song M H, Xu X L & Guo Z L D, Differential responses of arbuscular mycorrhizal fungi to nitrogen addition in a near

pristine Tibetan alpine meadow, *FEMS Microbial Ecol*, **89** (2014) 594–605, doi: 10.1111/1574-6941.12361

80 Wang Y L, Gao C, Chen L, Ji N N, Wu B W, Li X C, Lu P P, Zheng Y & Guo L D, Host plant phylogeny and geographic distance strongly structure Betulaceae-associated ectomycorrhizal fungal communities in Chinese secondary forest ecosystems, *FEMS Microbial Ecol*, **95** (2019) fiz037, doi: 10.1093/femsec/fiz037

81 Van Geel M, Jacquemyn H, Plue J, Saar L, Kasari L, Peeters G, van Acker K, Honnay O & Ceulemans T, Abiotic rather than biotic filtering shapes the arbuscular mycorrhizal fungal communities of European seminatural grasslands. *New Phytol*, **220** (2018) 1262–1272. doi: 10.1111/nph.14947

82 Miyamoto Y, Sakai A, Hattori M & Nara K, Strong effect of climate on ectomycorrhizal fungal composition: evidence from range overlap between two mountains. *Multidisciplinary* J *Microbial Ecol*, **9** (2015) 1870–1879. doi: 10.1038/ismej.2015.8

83 Ren F, Zhang Y, Yu H & Zhang Y A Ganoderma lucidum cultivation affect microbial community structure of soil, wood segments and tree roots, *Scientific Rep*, **10** (2020) 3445, https://doi.org/10.1038/s41598-020-60362-2

84 Stajic M, Milenkovic I, Brceski I, Vukojevic J & Duletic L S, Mycelial growth of edible and medicinal oyster mushroom [Pleurotus ostreatus (Jacq.: Fr.) Kumm.] on Selenium-enriched media, *Int J Med Mushrooms,* **4(3)** (2002) DOI:10.1615/IntJMedMushr.v4.i3.70

85 Tanaka M, Knowles W, Brown R, Hondow N, Arakaki A, Baldwin S, Staniland S & Matsunaga T, Bio-magnetic recovery of selenium: Bioaccumulating of selenium granules in magnetotactic bacteria, *App Environ Microbiol*, **82** (2016) 3886, DOI: 10.1128/AEM.00508-16.

86 Zhang B, Yan L, Li Q, Zou J, Tan H, Tan W, Peng W, Li Z & Zhang X, Dynamic succession of substrate-associated bacterial composition and function during Ganoderma lucidum growth, *Peer J,* **6** (2018) e4975, https://doi.org/10.7717/peerj.4975

87 Naher L, Yusuf U K, Ismail A, Tan S G & Mondal M M A, Ecological status of Ganoderma and basal stem rot disease of oil palms (*Elaeis guineensis* Jacq.), *Aust J Crop Sci*, **7(11)** (2013) 1723–1727.

88 Tattar T A, *Diseases of Shade Trees* (Academic Press. Sand Diego, California) 1989, 391.

89 Wang X C, Xi R J, Li Y, Wang D M & Yao Y J, The Species Identity of the Widely Cultivated Ganoderma, 'G. lucidum' (Ling-zhi), in China, *PLoS ONE*, **7(7)** (2012). e40857, doi:10.1371/journal.pone.0040857

90 Hennicke F, Cheikh-Ali Z, Liebisch T, Macia-Vicente J G, Bode H B & Piepenbring M, Distinguishing commercially grown Ganoderma lucidum from Ganoderma lingzhi from Europe and East Asia on the basis of morphology, molecular phylogeny, and triterpenic acid profiles. *Phtyochem*, **127** (2016) 29–37, doi: 10.1016/j.phytochem. 2016.03.012.