



OPEN

Predicting potential global distribution and risk regions for potato cyst nematodes (*Globodera rostochiensis* and *Globodera pallida*)

Yitong He^{1,2}✉, Rui Wang³, Honghai Zhao¹, Yonglin Ren², Manjree Agarwal², Dan Zheng⁴, Shan Gao⁵, Simon J. McKirdy²✉ & Dong Chu¹✉

Potato cyst nematodes (PCNs), golden (yellow) cyst nematode (*Globodera rostochiensis*, gPCN) and pale (white) cyst nematode (*G. pallida*, pPCN), are important invasive pests in many countries and regions where they can cause significant yield and economic loss for agriculture. Prediction and identification of habitats suitable for PCNs are critical for developing biosecurity strategies, both pre and post border, to maximise the potential for early elimination should an incursion occur. To date, the potential global distribution of PCNs has not been thoroughly studied. Therefore, this study conducted a species distribution model to illustrate the potential global distribution of PCNs and risk regions. In this study, the Maximum Entropy Model (Maxent) associated with the Geographic Information System (GIS) was employed to reveal the potential distribution of the gPCN and pPCN. In addition to bioclimate, soil quality was also included in the model. The global cultivated lands, whether the susceptible hosts were present or not, were used to assess the maximum potential risk regions. The limitation factors for PCNs distribution were also assessed. Results showed that 66% of the global land surface was suitable for gPCN or pPCN or both, and both species can colonise more than 75% of the global cultivated lands. The coldest quarter's mean temperature and precipitation were critical limitations in unsuitable regions. In summary, the global risk maps of PCNs contribute valuable additional information that complements previous national/regional distribution predictions. The results of this distribution research will contribute practical support for decision-makers and practitioners to implement biosecurity strategies from a global perspective, that incorporate prevention or promptly enforce control practices to limit the damage caused by future incursions.

Potato cyst nematodes (PCNs) are sedentary endoparasitic phytonematodes, including the golden (yellow) cyst nematode (*Globodera rostochiensis*, gPCN) and the pale (white) cyst nematode (*Globodera pallida*, pPCN). It is acknowledged that PCNs originated in South America and were introduced to Europe with potato breeding materials in the mid-nineteenth century^{1,2}; subsequently, Europe became the secondary spread centre³. Currently, PCNs are present on virtually all continents (except Antarctica), while gPCN and pPCN have been reported from 74 and 49 countries or regions, respectively^{4,5}. In Switzerland, a regular survey has been implemented annually since 1958, which has been continuously reporting the presence of PCNs⁶. Minnis et al. were the first to report a detailed field survey of PCNs in the United Kingdom, which identified severe infestation of both gPCN and pPCN⁷. In recent years, countries in most global regions have confirmed the presence of PCNs; these included Portugal, Turkey, Algeria, Morocco, India, Indonesia, Australia, Colombia, Kenya, Canada and USA^{8–17}.

¹Shandong Engineering Research Centre for Environment-Friendly Agricultural Pest Management, College of Plant Health and Medicine, Qingdao Agricultural University, Qingdao 266109, People's Republic of China. ²Harry Butler Institute, Murdoch University, Perth, WA 6150, Australia. ³State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, People's Republic of China. ⁴College of Economics, Qingdao Agricultural University, Qingdao 266109, People's Republic of China. ⁵Hebei Dahaituo National Nature Reserve Management Centre, Chicheng 075500, People's Republic of China. ✉email: hoyt.he@murdoch.edu.au; s.mckirdy@murdoch.edu.au; chinachudong@qau.edu.cn

In the infested field, PCNs can cause significant damage to agricultural production. For example, in Europe, approximately nine per cent of potato yield reduction is a consequence of PCNs infection, with total crop failure occurring in some heavily infested areas¹⁷. Despite potatoes being one of the most economically important hosts, PCNs can parasitise more than 170 species in the *Solanaceae* family, including tomato and aubergine, which are also of high economic value¹⁸. Economic assessment in Australia has estimated that gPCN was predicted to cause an economic loss of AU\$ 11.9 million to AU\$ 27.0 million per year¹⁹, while in Idaho (USA), pPCN caused approximately US\$26 million net loss in total agricultural output in 2006²⁰. Moreover, the severe impact of PCNs is also reflected in the difficulties of management. The PCN disease epidemic often takes a long time (> 30 years) to be recognised after its first infestation and is hard to eradicate once it is detected in fields²¹. The environmental adaptability of PCNs makes them even more challenging to control, given their ability to survive without hosts for periods up to 30 years^{4,22}, and their tolerance to a broad range of cold/warm temperatures^{4,5,23} (Supplementary Fig. S1).

Due to the widespread dispersal, severe damage, and the difficulties in management, a better understanding of PCNs potential distribution is important in order to facilitate biosecurity strategies, early detection and elimination. Some research focused on the regional or national spatial distribution patterns of PCNs have been published. Li predicted the potential distribution in China for both gPCN and pPCN with different habitat suitability models (HSMs)²⁴, and Contina et al. used spatial algorithms to illustrate the pPCN dispersal pattern in Idaho (USA)²⁵. However, few studies have elaborated on the potential global distribution of gPCN and pPCN. Thus, this study will focus on the PCNs distribution from a global perspective. The recognition of the global risk regions also provides a critical reference for the development of national/regional biosecurity strategies.

Model assessment is an efficient strategy to predict the potential distribution of a species²⁶. Species distribution models (SDMs), also known as habitat suitability models (HSMs), are commonly used to reflect the species' realised niche with survey data or historical records²⁶. Maximum entropy (Maxent) is one of the advanced SDMs, based on machine learning and mathematical statistics²⁷. The Maxent model possesses many advantages over its counterparts; for example, it does not require abundant ecological data of the targeted species and can use presence-only data to achieve high performance²⁸. Maxent can also support multiple data formats of environmental variables, which can be continuous or categorical²⁹. Given the uncertainty of absence records of PCNs and the various formats of environmental predictors, Maxent is an appropriate modelling method for this study.

In this study, the Maxent model associated with the Geographic Information System (GIS) was introduced to reveal the potential global distribution of gPCN and pPCN. Two assumptions were made to build up the foundation of this model. Firstly, this study assumed that either gPCN or pPCN would have an equal probability of presence and absence on the globe because both nematode species have adapted to almost all types of climates, including tropical, arid, temperate, continental, and tundra zones (Supplementary Fig. S1). The model, therefore, used the entire global land surface (except Antarctica) as background and predicted a maximum environmentally suitable range. This excluded the presence or absence of susceptible crop hosts. Secondly, this study assumed that all cultivated lands are potentially susceptible. Although potatoes, tomatoes or aubergines were not planted in all fields, they are possible to be planted on most of the cultivated lands due to their climate adaptation and cultivar diversity^{30–32}. It is even true when taking facility agriculture into account, with which the hosts can be planted in cultivated lands even without suitable environments; therefore, PCNs can potentially distribute there. Moreover, the broad host range (> 170 species) of PCNs also included several regional vegetables and breeding materials, such as *Lycopersicon pimpinellifolium*, *Physalis longifolia*, and *Solanum aethiopicum*^{4,5}, which increases the possibility of PCNs invasion in an expansive cultivated land. Thus, this study used the entire global cultivated land to predict the maximum potential risk regions. In conclusion, this paper aims to provide valuable information on the maximum potentially threatened region of PCNs invasion and a global reference for prevention and control strategies.

Materials and methods

Species presence data. The presence locations of gPCN and pPCN were gathered from web datasets, including Global Biodiversity Information Facility (<http://www.gbif.org>), Center for Agriculture and Bioscience International (<http://www.cabi.org/isc>) and European and Mediterranean Plant Protection Organization (<http://www.eppo.int>), as well as published government documents³³. Primarily, 170 and 79 distribution points were collected for gPCN and pPCN, respectively. Those presence records without geographical coordinates were searched by names on Google Earth (<https://earth.google.com/web/>) to add longitude and latitude. However, sample points need to be selected considering some points may disperse intensively in the same raster cell, which can cause sample bias and strongly affect the model's prediction³⁴. Thus, the PCNs occurrence points were trimmed by a grid map of 10 × 10 km cells (5 arcminutes which is the same resolution as environmental variables) so that only one point is present per cell, ensuring that every point has distinct environment values³⁵. Overall, 167 points for gPCN and 79 for pPCN were used for distribution prediction.

Environmental variables. Suitable environmental conditions are the main factor for species occupation and the foundation of species distribution modelling (SDM)²⁶. Bioclimate variables, such as temperature and precipitation are crucial and direct predictors for geographic distribution on large scales. Therefore, nineteen bioclimate variables with continuous data format were acquired from the WorldClim dataset (<https://worldclim.org>). Since PCNs are soil-dwelling species, soil quality variables were also introduced in this study to model the potential distribution of gPCN and pPCN. Thus, a soil quality variable with a categorical format was collected from the Harmonized World Soil Database (HWSD, <https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/>). All environmental layers had the same resolution of five arcminutes.

		Label	Type	Description
First set	Second set	Bio8	Bioclimate	Average temperature of wettest quarter
		Bio10	Bioclimate	Average temperature of warmest quarter
		Bio11	Bioclimate	Average temperature of coldest quarter
		Bio12	Bioclimate	Annual precipitation
		Bio17	Bioclimate	Precipitation of driest quarter
		Bio19	Bioclimate	Precipitation of coldest quarter
		Sq1	Soil quality	Nutrient availability (integrated soil texture, soil organic carbon, soil pH, total exchangeable bases)

Table 1. Two sets of environmental variables for model selection.

However, not all environmental variables are necessary for prediction because some correlated attributes may contain similar information and lead to model overfitting. Thus, the Pearson coefficient was calculated to eliminate multicollinearity ($r > |0.8|$) for a better response between environmental variables and geographical distribution (Supplementary Fig. S2a). Then, two sets of environmental variables were used to select the optimal model (Table 1). Variables in the first set were selected by Jackknife analysis for their importance (Supplementary Fig. S2b,c). The second set included six environmental variables based on biological and ecological characteristics of PCNs³⁶.

Maxent modelling. Maxent (Version 3.4.1)³⁷ was applied in this study to predict the environmentally suitable global distribution. Herein, different feature combinations and regularisation multipliers (β) were tested to optimise the model's performance^{38,39}. Maxent contains five features to illustrate the species-environment response (linear (l), quadratic (q), product (p), threshold (t) and hinge (h)). This study considered all possible feature combinations (31 in total). The regularisation multiplier varied between 0.5, 1, 2, and 3. In all, 248 candidate models have been evaluated for each species, with parameters reflecting all combinations of two sets of environmental variables, 31 feature class combinations, and four regularisation multiplier settings. Additionally, 80% of present localities were randomly selected for modal training, and the remaining 20% of localities were an independent test set to evaluate the model. The models were run with five replicates by cross-validation method, and the averaged results were automatically achieved via the programme.

The *kuenm* package⁴⁰ in R (version 4.1.2)⁴¹ was applied to write a batch run script for model calibration and evaluation. The Akaike information criterion corrected for small sample sizes (AICc) and the partial area under the receiver operating characteristic curve (pROC) were used as performance metrics^{42,43}. An omission rate of 5% was taken as the pROC calculation threshold. The candidate model with the lowest AICc value was taken as the best-performing final model. If there were candidate models with equal AICc values, the one with the highest pROC was selected as the final model.

Result visualisation and spatial analysis of the risk area. ArcMap (version 10.4.1)⁴⁴ was employed to visualise the prediction results. The results were reclassified to obtain environmentally suitable regions based on the Fixed Cumulative Value 5 (FCV5) because it is robust to abnormal values and prevalence, and balances the omission error at 5%⁴⁵. The logistic predictions lower than FCV5 were defined as unsuitable, while those higher than FCV5 were suitable. Then, the appropriate area (logistic prediction over FCV5) for gPCN and pPCN was overlapped to define the suitable habitat for both species. Furthermore, the environmentally suitable regions were overlapped with cultivated land to generate the risk map. The global land-use map was collected from the FAO database⁴⁶. The pixels with a cropland proportion above 5% were recognised as cultivated land. Then, the risk areas were classified into four levels (high risk, upper-medium risk, lower-medium risk, and low risk) based on the probability of observations²⁶. Finally, all maps were projected to the World Geodetic System-1984 (WGS84) Coordinate System.

Analyses of environmental response and limitations. First, the contribution and importance of each variable were analysed. The contribution of each variable was determined by the accumulated gain of each iteration in the training process. The variable importance was evaluated by calculating the area under the receiver operating characteristic curve (AUC)⁴⁷. Then the variables that limited potato cyst nematodes distribution in unsuitable areas were assessed. The environmental variable's response curves were used to determine each environmental variable's better-preformed interval (Cloglog value > 0.5)⁴⁷. The limitation factors were evaluated by the percentages of unsuitable areas dropped inside and outside the better-preformed interval.

Results

Model calibration and evaluation. In this study, 496 candidate models for gPCN and pPCN were calibrated (Supplementary Table S1 and Table S2). For gPCN candidate models, the AICc values varied from 4439.49 to 5287.05. However, two models equally obtained the lowest AICc value. The best-performed model (M_2_F_t_Set_1) was chosen by a higher pROC value (1.422), which used the multiplier *two*, feature *t*, and the first environmental variable set (Bio8, Bio10, Bio11, Bio12, Bio17, Bio19, and Sq1). The omission rate (OR) of gPCN's best-performed model was 21%. For pPCN candidate models, the AICc values varied from 2078.58 to 14,063.16. The best-performed model (M_3_F_lqh_Set_1) used the multiplier *three*, feature combination *lqh*,

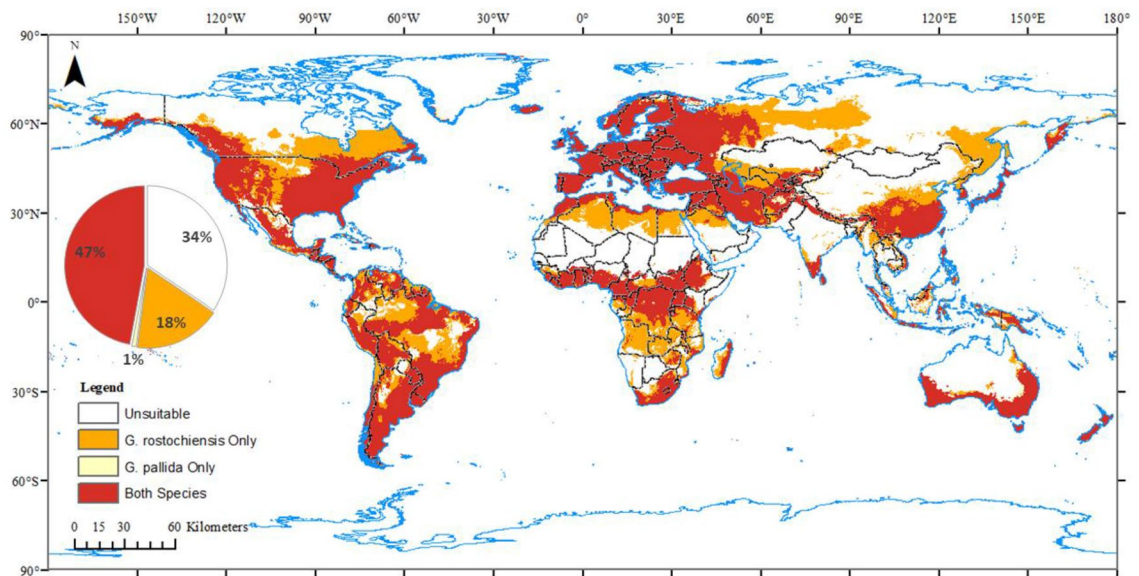


Figure 1. Environmentally suitable habitats for *Globodera rostochiensis* and *G. pallida* worldwide. White regions are unsuitable for both *G. rostochiensis* and *G. pallida*. Orange represents the areas only suitable for *G. rostochiensis*, while yellow regions are only suitable for *G. pallida*. Red areas are suitable for both *G. rostochiensis* and *G. pallida*. The pie chart shows the proportion of each suitable range on the entire global land surface. Global bioclimate data were acquired from the WorldClim open database (<https://worldclim.org>). The locations of species occurrence were collected from open databases: GBIF (<http://www.gbif.org>), CABI (<http://www.cabi.org/isc>) and EPPO (<http://www.eppo.int>). The species distribution model was conducted with Maxent (Version 3.4.1, http://biodiversityinformatics.amnh.org/open_source/maxent/), and results were modified with ArcMap (version 10.4.1, <https://www.arcgis.com/>).

and the first environmental variable set. This best-performed model's pROC and OR were 1.5 and 26%, respectively. Although some candidate models got relatively low OR (lower than 5%), they also gained higher AICc and lower pROC than the selected final models, which indicated a low goodness-of-fit.

Environmentally suitable habitats for potato cyst nematodes. The Maxent prediction indicated that the environmentally suitable habitats for both gPCN and pPCN are widely spread across the globe (Fig. 1). More than half (66%) of the land surface (excluding Antarctica) was suitable for at least one potato cyst nematode species. The areas suitable for both gPCN and pPCN were predicted to cover approximately 47% of the global land surface. The study identified that suitable habitats for both species could occupy the entire European continent, spreading from the British Isles to Iceland, extending to the Middle East and the Mediterranean coast of northern Africa. On the African continent, the suitable region also occupied the central and the southernmost points. The suitable area in North and South America expanded inland from the west and east coasts, taking approximately one-third of the continents. As for Asia, the suitable regions for both species were mainly distributed in southern China, the Korean Peninsula and Japan. Most suitable territories for the two nematode species in Oceania were distributed in the south and east coastal areas of Australia, and New Zealand.

The area only suitable for gPCN covered nearly 18% of the global land surface, which expanded along the boundaries of the area suitable for both species. These areas included central and eastern North America, central South America, the plateau of Eastern Europe and West Siberian, northern and central Africa, as well as central China and the region of Russia bordering northeast China. Moreover, the global land cover only suitable for pPCN was less than 1%, which indicated that over 97% of pPCN's suitable habitats might be overlapped by gPCN's habitats.

Risk regions for potato cyst nematodes. The risk map (Fig. 2) illustrated that the high-risk regions (the cultivated land with suitable environments for either gPCN or pPCN) covered 21% of the global land surface. These high-risk regions accounted for approximately 75% of global cultivated land. In contrast, the low-risk regions, which have neither suitable environment nor cultivated land, occupied 40% of the land cover.

The medium-risk areas were subdivided into two levels. Those environmentally suitable regions with only small areas of cultivated lands were defined as the upper-medium risk regions, while the cultivated land distributed in unsuitable environmental conditions was defined as lower-medium risk regions. Specifically, the upper-medium risk regions occupied 32% of the global land surface. Lower-medium risk regions occupied only 7% of the land surface.

The environmental limitations of potato cyst nematodes distribution. The analyses of variables' contribution and importance (Table 2) indicated that the mean temperature of the coldest quarter (Bio11) is

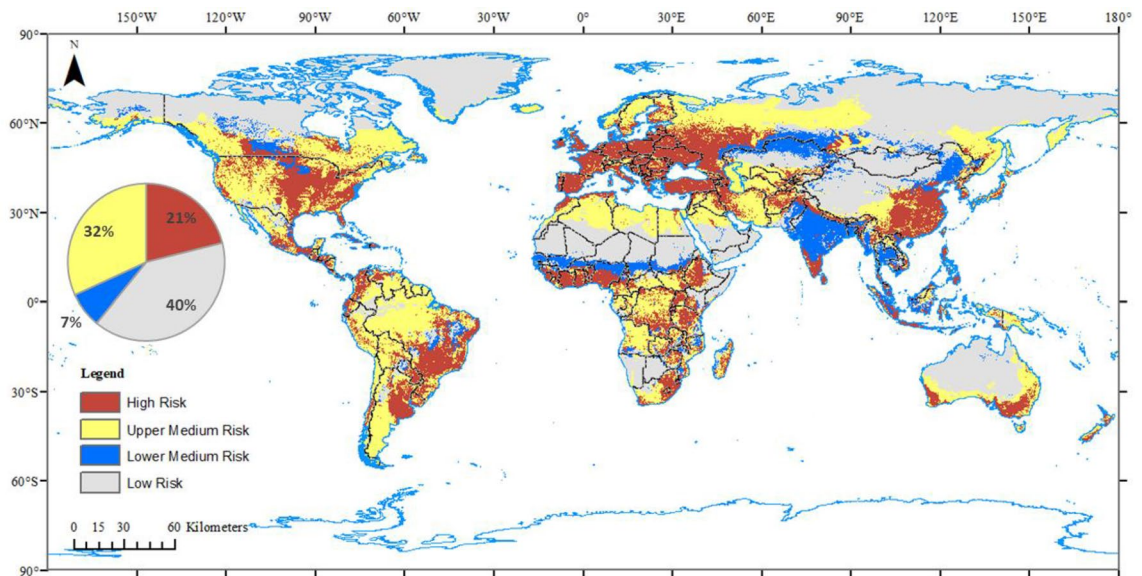


Figure 2. Global risk regions for *Globodera rostochiensis* and *G. pallida*. Grey areas have neither suitable environment nor cultivated land. Blue areas represent cultivated land with unsuitable environmental conditions. Yellow regions represent suitable environment with small areas of cultivated fields. Red areas are cultivated lands with suitable environments for at least one of *G. rostochiensis* and *G. pallida*. The pie chart shows the proportion of each risk level on the entire global land surface. Global bioclimate data were acquired from the WorldClim open database (<https://worldclim.org>). The locations of species occurrence were collected from open databases: GBIF (<http://www.gbif.org>), CABI (<http://www.cabi.org/isc>) and EPPO (<http://www.eppo.int>). The species distribution model was conducted with Maxent (Version 3.4.1, http://biodiversityinformatics.amnh.org/open_source/maxent/), and modified with ArcMap (version 10.4.1, <https://www.arcgis.com/>).

Variable	<i>G. rostochiensis</i>		<i>G. pallida</i>	
	Contribution (%)	Importance (%)	Contribution (%)	Importance (%)
Bio8	26.1	30.8	7.2	22.9
Bio10	5.8	4.2	0.5	1.5
Bio11	38.6	52.6	31.1	56.6
Bio12	13.3	7.1	2.1	4.4
Bio17	0.4	0.4	0.4	0.3
Bio19	15	4.5	51.2	13.9
Sq1	0.8	0.4	7.5	0.3

Table 2. The contribution and importance of each variable in *Globodera rostochiensis* and *G. pallida* models.

the most critical variable in either gPCN or pPCN distribution prediction. The mean temperature of the wettest quarter (Bio8) took second place, while the mean temperature of the warmest quarter (Bio10) was the least important predictor among the three temperature variables. As for precipitation predictors, precipitation of the coldest quarter (Bio19) contributed the most in both models, especially in pPCN distribution predictions (51.2% contribution). Annual precipitation (Bio12) was also an essential predictor, especially in the gPCN model (13.3% contribution). Although the soil quality variable (Sq1) showed a low proportion of importance, it was still an indispensable predictor, which gave contributions even higher than the precipitation of the driest quarter (Bio17). Particularly in the pPCN model, Sq1 contributed 7.5% of the gains.

The analyses of limiting factors in unsuitable areas (Fig. 3) illustrated that the mean temperature of the wettest quarter (Bio8) did not suppress the distribution of gPCN and pPCN. 46% of the unsuitable regions still had a favourable temperature (2–20 °C) for gPCN at the wettest quarter. Similarly, 55% of pPCN's unsuitable habitat had a favourable temperature (−4 to 20 °C) at the wettest quarter. Except for Bio8, all the other environmental variables restricted the distribution of gPCN. In the warmest quarter, the temperatures of the most unsuitable habitats were either too low (< 13 °C) or too high (> 25 °C) for gPCN. In the coldest quarter, the most unsuitable regions were too cold and dry, given that the temperature in 66% of the unsuitable land was lower than −5 °C, and 92% of the land had precipitation less than 98 mm. Moreover, the annual precipitation also limited gPCN's distribution. Over 80% of the unsuitable regions had precipitation less than 570 mm, and 91% even less than 70 mm in the driest quarter. Unlike gPCN, the mean temperature of the warmest quarter did not restrict pPCN's distribution. Over 50% of the unsuitable surface still had favourable temperatures (9–27 °C) for pPCN. The effects

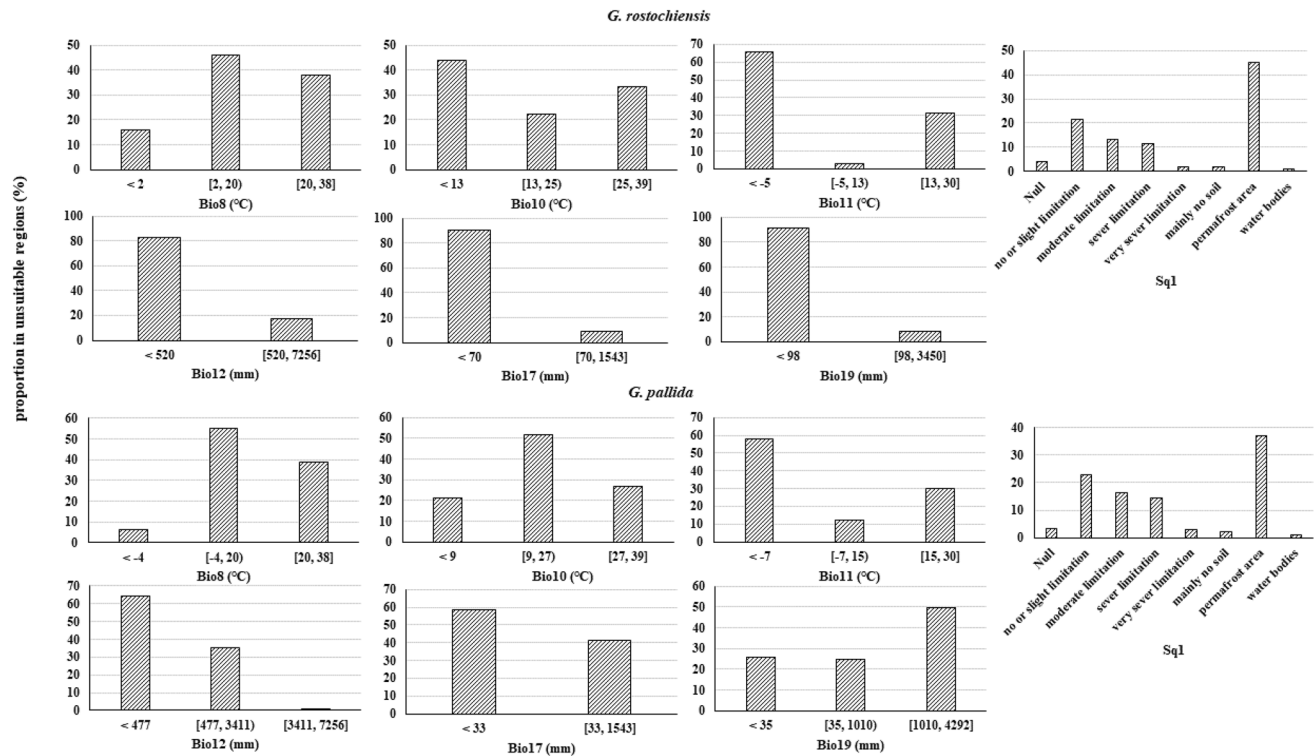


Figure 3. Histogram of environmental variables in unsuitable regions for both *Globodera rostochiensis* and *G. pallida*. *Bio8* average temperature of the wettest quarter, *Bio10* average temperature of the warmest quarter, *Bio11* average temperature of the coldest quarter, *Bio12* annual precipitation of driest quarter, *Bio17* precipitation of warmest quarter, *Bio19* precipitation of coldest quarter, *Sq1* soil quality (integrated soil texture, soil organic carbon, soil pH, total exchangeable bases).

of other bioclimate variables on pPCN are similar to those on gPCN. It is worth noting that half of the unsuitable areas had too much precipitation (> 1010 mm) in the coldest quarter for pPCN. Soil quality affected the distribution of gPCN and pPCN in the same way. Most of the unsuitable regions were permafrost land; besides, around 30% of unsuitable land had limitations (from moderate to no soil) on soil quality.

Discussion

Recent predictive modelling studies of species (plants, mammals, birds, insects) distribution have been based on climate variables or a combination of climate and topographic factors²⁶. Unlike the insects above ground, nematodes complete their life cycle underground. Therefore, soil quality characteristics, such as soil texture and pH, may significantly affect nematode survival and spread^{48,49}. Duyck et al. also acknowledged that soil type is an important variable in modelling nematode distribution⁵⁰. Stanton et al. demonstrated that including land-use-related variables can achieve a more robust model prediction⁵¹. Soil quality is such a variable related to land characteristics and parameters. As such, this research utilised both bioclimate variables and soil quality factors along with cultivated land data to model the predicted distribution and risk regions of PCNs globally. As bioclimate information is continuous data, while soil quality is categorical data the Maxent model was selected as the preferred SDM due to its advantage in applying both data formats⁵².

Maxent's tuning parameters are necessary to obtain higher confidence in species distribution prediction^{28,38}. In this study, parameter combinations were identified in response to nonlinear species-environment curves with appropriate complexity³⁸. The AICc is used as the selection criterion because it may ideally weigh the complexity of the estimated model against the goodness-of-fit⁵³. Besides, the mean AUC ratio with partial ROC was used as a reference for model selection, which overcame the limitation that conventional AUC equally weighs omission and commission error⁴². As a result, the predicting model in this study provided satisfactory performance.

The selected model predicted that the environmentally suitable habitats for gPCN were relatively larger than pPCN's and almost completely covered pPCN's suitable territories, which indicated the niche similarity between these species. In reality, gPCN has been reported to occur together with pPCN in many areas. For example, 25% of the infested fields in England and Wales were detected with both nematode species, while the proportion was 28% in Algeria^{7,8}. In Newfoundland (Canada), Portugal, and Kenya, gPCN was also accompanied by pPCN in some fields^{11,14,16}.

A key feature of this research was that it focused on the entire global cultivated lands instead of just focusing on areas with known susceptible hosts, in order to provide results that predict the maximum potential of PCNs global distribution. The prediction considered the global plantation structure as a changeable rather than fixed system because the major hosts of PCNs (potato, tomato, and aubergine) have demonstrated that they can be

planted in most cultivated lands, especially with anthropogenic management activities (e.g., breeding temperature and drought resistance cultivars, or building agricultural facilities such as greenhouse to improve planting conditions)^{30–32}. Moreover, the global challenges in recent years (such as climate change, the COVID-19 pandemic, and the Russia-Ukraine war) have been influencing the food supply and planting patterns globally^{54,55}.

This research provides a valuable contribution to future biosecurity management by predicting the distribution risk across cultivated and non-cultivated areas. The spread of gPCN and pPCN would be highly risky because PCNs' environmentally suitable habitats covered most global cultivated lands, which meet the essential requirement of species colonisation and survival. Given that potatoes, the primary host of PCNs, are planted intensively in Europe, China, the southern Himalayas and the Andes region⁵⁶, and many other regions of the globe, the invasion and spread of PCNs in these regions will initiate control/management strategies to address the risk. Furthermore, aubergine and tomatoes are also economically significant hosts¹⁸, which results in high risk in more extensive horticultural areas.

Global areas with a suitable environment but limited cultivated lands were considered upper-medium risk regions. Although high economic value hosts may not be present in these regions, PCNs could still colonise many *Solanaceae* plants¹⁸, which are present on all continents (except Antarctica)⁵⁷. Cultivated lands without suitable environments should also be treated with caution because colonisation may happen via high propagule pressure²⁶. Moreover, some hosts, such as tomatoes and aubergines, can be planted in greenhouses or other artificial conditions. Thus, even if the natural environment is unsuitable, these cultivated lands should be treated cautiously to prevent possible incursions. Also, anthropogenic effects (e.g., thriving trade of breeding materials) may exacerbate the risk of PCNs' dispersal, so strict biosecurity strategies should be executed particularly at the boundary of each risk region.

The results of the unsuitable regions are not definitive because our model used the presence-only data, while biotic factors, such as interspecific competition, are difficult to access³⁸. Therefore, our prediction of the suitable habitats is a comparable valuation of the ecological niche, and PCNs would establish in low suitability regions if they were to invade with large numbers and reproduce a substantial clutch size⁵⁸. Moreover, human cultivation activities that modify land areas probably increase this possibility and could cause significant uncertainties in the low risk regions⁵⁹. Despite these limitations in modelling the potential distribution, the results we present can provide fundamental and valuable information about different global regions with a higher possibility for PCNs to invade and establish. A greater understanding of the potential habitats achieved from these results can also help to assess the species' bilateral flow and invasion pathways.

In this study, the environmental limitations showed that pPCN had a broader temperature adaptation than gPCN. The predicted suitable temperature range (from coldest to warmest quarter) for pPCN is -7 to 27 °C, while the suitable temperature range for gPCN is -5 to 25 °C. Jones et al. reported that the development and reproduction of female pPCN preferred 15 – 17.5 °C while gPCN favoured 17.5 – 22.5 °C⁶⁰, which supported the temperature range predicted in this study. The results from Kaczmarek et al. showed a broader temperature range with a similar trend (i.e., 13 – 25 °C for pPCN and 15 – 27 °C for gPCN), and they also presented that when the temperature was over 25 °C pPCN hatched more efficiently than gPCN⁶¹. Noteworthy is that pPCNs may out-compete gPCNs in the mixed-species soil due to their faster and more significant multiplication⁶². The broader suitable temperature range for pPCN utilised in this study can also support its greater competitiveness than gPCN. In reality, pPCN has shown a trend of gradually replacing gPCN in Britain²¹. Thus, the area only suitable for gPCN should also be treated discreetly with pPCN invasion.

The environmental limitations analysis also showed that the coldest quarter's mean temperature is the significant temperature-related bioclimate factor. However, with climate change, areas affected by this frigid climate may gradually shrink so that the suitable habitats for PCNs can be assumed to extend toward the poles and higher elevations. Some studies have pointed out that the increasing temperature can affect PCNs' survival^{60,63}, whereas whether the effect will be significant enough to shift their distribution still needs to be thoroughly studied. This study also indicated that precipitation significantly suppressed PCNs' distribution in unsuitable regions. An arid climate is also likely to adversely affect PCNs distribution. However, studies have focused on the effect of temperature increases rather than precipitation changes^{60,63}. Thus, the effect of precipitation on PCN's ecological adaptation and geographical distribution needs more thorough research, especially under the challenge of climate change.

Scientific precautionary management needs to be implemented to minimise the ecological, economic and social impact caused by pest invasion, in which the awareness of possible high-risk regions is a crucial element in biosecurity strategies⁶⁴. Incorporating climate suitability and host availability, as well as other specific suitability criteria such as soil characteristics, are the primary steps, from which policy decision-makers can focus on pathway analysis and detection regimes at borders⁶⁵. Since species invasion can significantly threaten global ecological systems and economies, knowledge of potentially suitable habitats is imperative to manage risk⁶⁶. It is well accepted that elimination and control practices need to be implemented as early as possible in order to achieve desired biosecurity outcomes⁶⁷, and predictive knowledge of potential distribution is a key element in this decision-making. This is particularly the case for nematode prevention as they are often overlooked due to symptoms on host plants not being readily obvious at the early stages of infestation⁶⁸, and once established and detected any management response will consume enormous resources and the probability of elimination decreases. For these reasons, the prediction of risk regions will provide significant support to decision-makers assisting in the accurate and early detection of PCNs to overcome this challenge. This highlights the valuable contribution the potential global risk maps for PCNs presented in this study can provide for future biosecurity strategies and early elimination practices.

In conclusion, this study presented the potential globally suitable distribution of PCNs based on both climate and soil factors. The risk regions were predicted based on suitable environments along with global cultivated land and achieved a maximum potential range from the perspective of potential variation. The prediction is an informative complement for regional/national studies, which can support further biosecurity decision-making

and strategies on PCNs. Additionally, soil quality played an indispensable role in PCNs distribution models and highlighted the need to include additional parameters relevant to the ecology of the species under consideration. The mean temperature in the coldest quarter and precipitation are the main environmental limitations for PCNs' distribution, and the authors identified that the response of these nematodes to low temperature and soil moisture needs further study to further progress potential distribution modelling.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Received: 6 July 2022; Accepted: 14 December 2022

Published online: 17 December 2022

References

- Evans, K., Franco, J. & De Scurrah, M. M. Distribution of species of potato cyst-nematodes in South America. *Nematologica* **21**, 365–369. <https://doi.org/10.1163/187529275x00103> (1975).
- Plantard, O. *et al.* Origin and genetic diversity of Western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci. *Mol. Ecol.* **17**, 2208–2218. <https://doi.org/10.1111/j.1365-294X.2008.03718.x> (2008).
- Price, J. A., Coyne, D., Blok, V. C. & Jones, J. T. Potato cyst nematodes *Globodera rostochiensis* and *G. pallida*. *Mol. Plant Pathol.* **22**, 495–507. <https://doi.org/10.1111/mpp.13047> (2021).
- CABI. *Globodera rostochiensis* (yellow potato cyst nematode). <https://www.cabi.org/isc/datasheet/27034> (2021).
- CABI. *Globodera pallida* (white potato cyst nematode). <https://www.cabi.org/isc/datasheet/27033> (2021).
- Ruthes, A. C. & Dahlin, P. The impact of management strategies on the development and status of potato cyst nematode populations in Switzerland: An overview from 1958 to present. *Plant Dis.* **106**, 1096–1104. <https://doi.org/10.1094/pdis-04-21-0800-sr> (2021).
- Minnis, S. T. *et al.* Potato cyst nematodes in England and Wales—Occurrence and distribution. *Ann. Appl. Biol.* **140**, 187–195. <https://doi.org/10.1111/j.1744-7348.2002.tb00172.x> (2002).
- Djebroune, A. *et al.* Integrative morphometric and molecular approach to update the impact and distribution of potato cyst nematodes *Globodera rostochiensis* and *Globodera pallida* (Tylenchida: Heteroderidae) in Algeria. *Pathogens* **10**, 216. <https://doi.org/10.3390/pathogens10020216> (2021).
- Vallejo, D. *et al.* Occurrence and molecular characterization of cyst nematode species (*Globodera* spp.) associated with potato crops in Colombia. *PLoS One* **16**, e0241256. <https://doi.org/10.1371/journal.pone.0241256> (2021).
- Hajjaji, A., Mhand, R. A., Rhallabi, N. & Mellouki, F. First report of morphological and molecular characterization of Moroccan populations of *Globodera pallida*. *J. Nematol.* **53**, e2021-07. <https://doi.org/10.21307/jofnem-2021-007> (2021).
- Camacho, M. J. *et al.* Potato cyst nematodes: Geographical distribution, phylogenetic relationships and integrated pest management outcomes in Portugal. *Front. Plant Sci.* **11**, 9. <https://doi.org/10.3389/fpls.2020.606178> (2020).
- Handayani, N. D. *et al.* Distribution, DNA barcoding and genetic diversity of potato cyst nematodes in Indonesia. *Eur. J. Plant Pathol.* **158**, 363–380. <https://doi.org/10.1007/s10658-020-02078-7> (2020).
- Bairwa, A. *et al.* Morphological and molecular characterization of potato cyst nematode populations from the Nilgiris. *Indian J. Agric. Sci.* **90**, 273–278 (2020).
- Mburu, H. *et al.* Potato cyst nematodes: A new threat to potato production in East Africa. *Front. Plant Sci.* **11**, 13. <https://doi.org/10.3389/fpls.2020.00670> (2020).
- Altas, A., Evlice, E., Ozer, G., Dababat, A. & Imren, M. Identification, distribution and genetic diversity of *Globodera rostochiensis* (Wollenweber, 1923) Skarbilovich, 1959 (Tylenchida: Heteroderidae) populations in Turkey. *Turk. Entomol. Derg. Turk. J. Entomol.* **44**, 385–397. <https://doi.org/10.16970/entoted.740223> (2020).
- Dandurand, L.-M., Zasada, I. A., Wang, X. & Mimee, B. Current status of potato cyst nematodes in North America. *Annu. Rev. Phytopathol.* **57**, 117–133. <https://doi.org/10.1146/annurev-phyto-082718-100254> (2019).
- Blacket, M. J. *et al.* Molecular assessment of the introduction and spread of potato cyst nematode, *Globodera rostochiensis*, in Victoria, Australia. *Phytopathology* **109**, 659–669. <https://doi.org/10.1094/phyto-06-18-0206-r> (2019).
- Sullivan, M. J., Inerra, R. N., Franco, J., Moreno-Leheude, I. & Greco, N. Potato cyst nematodes: Plant host status and their regulatory impact. *Nematropica* **37**, 193–201 (2007).
- Hodda, M. & Cook, D. C. Economic impact from unrestricted spread of potato cyst nematodes in Australia. *Phytopathology* **99**, 1387–1393. <https://doi.org/10.1094/phyto-99-12-1387> (2009).
- Koirala, S., Watson, P., McIntosh, C. S. & Dandurand, L. M. Economic impact of *Globodera pallida* on the Idaho economy. *Am. J. Potato Res.* **97**, 214–220. <https://doi.org/10.1007/s12230-020-09768-2> (2020).
- Trudgill, D. L., Elliott, M. J., Evans, K. & Phillips, M. S. The white potato cyst nematode (*Globodera pallida*)—A critical analysis of the threat in Britain. *Ann. Appl. Biol.* **143**, 73–80. <https://doi.org/10.1111/j.1744-7348.2003.tb00271.x> (2003).
- Duan, Y. X. *Plant Nematology* (Science Press, 2011).
- Peel, M. C., Finlayson, B. L. & McMahon, T. A. Updated world map of the Köppen–Geiger climate classification. *Hydrol. Earth Syst. Sci.* **11**, 1633–1644. <https://doi.org/10.5194/hess-11-1633-2007> (2007).
- Li, J. *Suitable Risk Assessment for Six Potential Invasive Nematodes in China*, Master thesis (Jilin Agriculture University, 2008).
- Contina, J. B., Dandurand, L. M. & Knudsen, G. R. A spatiotemporal analysis and dispersal patterns of the potato cyst nematode *Globodera pallida* in Idaho. *Phytopathology* **110**, 379–392. <https://doi.org/10.1094/phyto-04-19-0113-r> (2020).
- Guisan, A., Thuiller, W. & Zimmermann, N. E. *Habitat Suitability and Distribution Models with Applications in R* (Cambridge University Press, 2017). <https://doi.org/10.1017/9781139028271>.
- Phillips, S. J., Dudik, M. & Schapire, R. E. A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-First International Conference on Machine Learning* 83. <https://doi.org/10.1145/1015330.1015412> (Association for Computing Machinery, 2004).
- Phillips, S. J. & Dudik, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* **31**, 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x> (2008).
- Wan, J. *et al.* Predicting the potential geographic distribution of *Bactrocera bryoniae* and *Bactrocera neohumeralis* (Diptera: Tephritidae) in China using MaxEnt ecological niche modeling. *J. Integr. Agric.* **19**, 2072–2082. [https://doi.org/10.1016/S2095-3119\(19\)62840-6](https://doi.org/10.1016/S2095-3119(19)62840-6) (2020).
- Midmore, D. J. Potato production in the tropics. In *The Potato Crop: The Scientific Basis for Improvement* 728–793 (Springer Netherlands, 1992).
- Naika, S., de Jude, J. V. L., de Goffau, M., Hilmi, M. & van Dam, B. *Cultivation of Tomato: Production, Processing and Marketing* (Digrafi, 2005).

32. Chapman, M. A. Eggplant breeding and improvement for future climates. In *Genomic Designing of Climate-Smart Vegetable Crops* 257–276 (Springer, 2020).
33. Management, D. o. C. *List of National Agricultural Plant Quarantine Pests Distribution Administrative Areas*. https://www.moa.gov.cn/nybg/2019/201906/201907/t20190701_6320036.htm (Ministry of Agriculture and Rural Affairs of the People's Republic of China, 2021).
34. Elith, J. *et al.* A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **17**, 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x> (2011).
35. Warren, D. L., Glor, R. E. & Turelli, M. ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611. <https://doi.org/10.1111/j.1600-0587.2009.06142.x> (2010).
36. Foot, M. A. *The Ecology of Globodera pallida (Stone) Mulvey & Stone (Nematoda, Heteroderidae) at Pukekohe, New Zealand, Doctoral thesis* (The University of Auckland, 1978).
37. Phillips, S. J., Dudík, M., & Schapire, R. E. Maxent software for modeling species niches and distributions (Version 3.4.1) [Internet]. http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed 24-10-2020.
38. 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* **36**, 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x> (2013).
39. Wan, J., Wang, R., Ren, Y. & McKirdy, S. Potential distribution and the risks of *Bactericera cockerelli* and its associated plant pathogen *Candidatus Liberibacter solanacearum* for global potato production. *Insects* **11**, 298. <https://doi.org/10.3390/insects11050298> (2020).
40. Cobos, M. E., Peterson, A. T., Barve, N. & Osorio-Olvera, L. kuenm: An R package for detailed development of ecological niche models using Maxent. *PeerJ* **7**, e6281. <https://doi.org/10.7717/peerj.6281> (2019).
41. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed 30-10-2020.
42. Peterson, A. T., Papeş, M. & Soberón, J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecol. Model.* **213**, 63–72. <https://doi.org/10.1016/j.ecolmodel.2007.11.008> (2008).
43. 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.* **21**, 335–342. <https://doi.org/10.1890/10-1171.1> (2011).
44. ESRI. ArcGIS Desktop: Release 10 (Version 10.4.1) (Environmental Systems Research Institute). Accessed 24-10-2020.
45. Kong, W., Li, X. & Zou, H. Optimizing MaxEnt model in the prediction of species distribution. *J. Appl. Ecol.* **30**, 2116–2128. <https://doi.org/10.13287/j.1001-9332.201906.029> (2019).
46. Fischer, G. *et al.* *Global Agro-ecological Zones (GAEZ v3.0)*. <http://pure.iiasa.ac.at/id/eprint/13290/> (2012).
47. Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E. & Blair, M. E. Opening the black box: An open-source release of Maxent. *Ecography* **40**, 887–893. <https://doi.org/10.1111/ecog.03049> (2017).
48. da Silva, J. C. P., de Medeiros, F. H. V. & Campos, V. P. Building soil suppressiveness against plant-parasitic nematodes. *Biocontrol Sci. Technol.* **28**, 423–445. <https://doi.org/10.1080/09583157.2018.1460316> (2018).
49. Kim, E., Seo, Y., Kim, Y. S., Park, Y. & Kim, Y. H. Effects of soil textures on infectivity of root-knot nematodes on carrot. *Plant Pathol. J.* **33**, 66–74. <https://doi.org/10.5423/ppj.Oa.07.2016.0155> (2017).
50. Duyck, P.-F. *et al.* Niche partitioning based on soil type and climate at the landscape scale in a community of plant-feeding nematodes. *Soil Biol. Biochem.* **44**, 49–55. <https://doi.org/10.1016/j.soilbio.2011.09.014> (2012).
51. Stanton, J. C., Pearson, R. G., Horning, N., Ersts, P. & ReşitAkçakaya, H. Combining static and dynamic variables in species distribution models under climate change. *Methods Ecol. Evol.* **3**, 349–357. <https://doi.org/10.1111/j.2041-210X.2011.00157.x> (2012).
52. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026> (2006).
53. Burnham, K. P. & Anderson, D. R. Multimodel inference understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304. <https://doi.org/10.1177/0049124104268644> (2004).
54. Din, A. U. *et al.* The impact of COVID-19 on the food supply chain and the role of e-commerce for food purchasing. *Sustainability* **14**, 3074. <https://doi.org/10.3390/su14053074> (2022).
55. Lang, T. & McKee, M. The reinvasion of Ukraine threatens global food supplies. *BMJ* <https://doi.org/10.1136/bmj.o676.o676.10.1136/bmj.o676> (2022).
56. Hijmans, R. J. Global distribution of the potato crop. *Am. J. Potato Res.* **78**, 403–412. <https://doi.org/10.1007/bf02896371> (2001).
57. Motti, R. The Solanaceae family: Botanical features and diversity. In *The Wild Solanums Genomes* 1–9 (Springer International Publishing, 2021).
58. Chytrý, M. *et al.* Projecting trends in plant invasions in Europe under different scenarios of future land-use change. *Glob. Ecol. Biogeogr.* **21**, 75–87. <https://doi.org/10.1111/j.1466-8238.2010.00573.x> (2012).
59. Lin, W., Cheng, X. & Xu, R. Impact of different economic factors on biological invasions on the global scale. *PLoS One* **6**, e18797. <https://doi.org/10.1371/journal.pone.0018797> (2011).
60. Jones, L. M. *et al.* Climate change is predicted to alter the current pest status of *Globodera pallida* and *G. rostochiensis* in the United Kingdom. *Glob. Change Biol.* **23**, 4497–4507. <https://doi.org/10.1111/gcb.13676> (2017).
61. Kaczmarek, A., MacKenzie, K., Kettle, H. & Blok, V. C. Influence of soil temperature on *Globodera rostochiensis* and *Globodera pallida*. *Phytopathol. Mediterr.* **53**, 396–405. https://doi.org/10.14601/Phytopathol_Mediterr-13512 (2014).
62. Hearne, R., Lettice, E. P. & Jones, P. W. Interspecific and intraspecific competition in the potato cyst nematodes *Globodera pallida* and *G. rostochiensis*. *Nematology* **19**, 463–475. <https://doi.org/10.1163/15685411-00003061> (2017).
63. Skelsey, P., Kettle, H., Mackenzie, K. & Blok, V. Potential impacts of climate change on the threat of potato cyst nematode species in Great Britain. *Plant. Pathol.* **67**, 909–919. <https://doi.org/10.1111/ppa.12807> (2018).
64. Carlton, J. & Ruiz, G. M. *Invasive Species: Vectors and Management Strategies* (Island Press, 2003).
65. Singh, S. K., Pains, D. R., Ash, G. J. & Hodda, M. Prioritising plant-parasitic nematode species biosecurity risks using self organising maps. *Biol. Invasions* **16**, 1515–1530. <https://doi.org/10.1007/s10530-013-0588-7> (2014).
66. Jiang, D., Chen, S., Hao, M. M., Fu, J. Y. & Ding, F. Y. Mapping the Potential global codling moth (*Cydia pomonella* L.) distribution based on a machine learning method. *Sci. Rep.* **8**, 8. <https://doi.org/10.1038/s41598-018-31478-3> (2018).
67. Simberloff, D. *et al.* Impacts of biological invasions: What's what and the way forward. *Trends Ecol. Evol.* **28**, 58–66. <https://doi.org/10.1016/j.tree.2012.07.013> (2013).
68. Ravichandra, N. Nematodes of quarantine importance. In *Horticultural Nematology* 369–385 (Springer, 2014).

Acknowledgements

The authors are grateful to friends who asked not to be named for constructive comments and generous help on this manuscript.

Author contributions

Conceptualisation: Y.H., Y.R., and R.W. Methodology: Y.H. and R.W. Resources: Y.H., H.Z., and S.G. Formal analysis: Y.H. Investigation: Y.H., R.W., H.Z., M.A., D.C., and Y.R. Writing-original draft preparation: Y.H.

Writing-review and editing: Y.H., R.W., M.A., Y.R., D.Z., S.G., D.C. and S.M. Supervision: D.C., Y.R., S.M. Project administration: D.C. and D.Z., Y.R. and S.M. Funding acquisition: D.C. All authors reviewed and granted the manuscript.

Funding

This research was funded by the National Key R&D Program of China, 2021YFD1400200; the Taishan Scholar Constructive Engineering Foundation of Shandong; and the First-class grassland science discipline program in Shandong Province, China.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26443-0>.

Correspondence and requests for materials should be addressed to Y.H., S.J.M. or D.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022