

Modelling the spatial distribution of mycetoma in Sudan

Rowa Hassan^{a,b,*}, Hope Simpson^c, Jorge Cano^c, Sahar Bakhiet^a, Eltayeb Ganawa^d, Daniel Argaw^e,
Melanie J. Newport^b, Kebede Deribe^{b,f}, and Ahmed Hassan Fahal^a

^aMycetoma Research Centre, Soba University Hospital, University of Khartoum, Khartoum, 11111, Sudan; ^bDepartment of Global Health and Infection, Brighton and Sussex Medical School, Falmer, Brighton, BN1 9PX, UK; ^cDepartment of Disease Control, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK; ^dDepartment of GIS and Cartography, Faculty of Geography and Environmental Sciences, University of Khartoum, Khartoum, 11111, Sudan; ^eWorld Health Organization, Control of Neglected Tropical Diseases, Innovative and Intensified Disease Management, Geneva, Switzerland; ^fSchool of Public Health, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia

*Corresponding author: Tel: +249129050542; E-mail: Roalbash2016@outlook.com; R.hassan@bsms.ac.uk

Received 17 March 2021; revised 18 April 2021; editorial decision 19 April 2021; accepted 26 April 2021

Background: Mycetoma is a neglected tropical disease that is reported worldwide and Sudan has the highest reported number of mycetoma infections across the globe. The incidence, prevalence and burden of mycetoma globally are not precisely known and its risk factors remain largely unelucidated.

Methods: This study aimed to identify the environmental predictors of fungal and bacterial mycetoma in Sudan and to identify areas of the country where these niche predictors are met. Demographic and clinical data from confirmed mycetoma patients seen at the Mycetoma Research Centre from 1991 to 2018 were included in this study. Regression and machine learning techniques were used to model the relationships between mycetoma occurrence in Sudan and environmental predictors.

Results: The strongest predictors of mycetoma occurrence were aridity, proximity to water, low soil calcium and sodium concentrations and the distribution of various species of thorny trees. The models predicted the occurrence of eumycetoma and actinomycetoma in the central and southeastern states of Sudan and along the Nile river valley and its tributaries.

Conclusion: Our results showed that the risk of mycetoma in Sudan varies geographically and is linked to identifiable environmental risk factors. Suitability maps are intended to guide health authorities, academic institutes and organisations involved in planning national scale surveys for early case detection and management, leading to better patient treatment, prevention and control of mycetoma.

Keywords: ensemble models, environmental modelling, machine learning, mycetoma, Sudan

Introduction

Mycetoma is a neglected tropical disease (NTD) that, in common with other NTDs, has numerous health and socioeconomic impacts on patients and communities.^{1–3} It is a chronic infectious disease that leads to destructive granulomatous inflammation, eventually causing deformity to the affected body part limb, impeding mobility.^{4,5} Mycetoma is reported worldwide, but it is endemic in tropical and subtropical regions.^{6,7} Mycetoma-endemic areas are characterised by moderate aridity, low humidity and a short rainy season.^{8–10} Sudan is considered one of the endemic regions for mycetoma among a list of other countries.^{11,12} In a recent review, mycetoma was noted to be present

in 102 countries around the world and Sudan had the highest burden of cases.¹³ More than 70 microorganisms have been incriminated in causing mycetoma and, according to the causative organism, it is classified into eumycetoma, caused by fungi, and actinomycetoma caused by bacteria.^{14,15} Eumycetoma is frequently reported in Africa and India, whereas actinomycetoma is more predominant in Asia, and in Central and South America.^{16,17} In Sudan, *Madurella mycetomatis* is the leading causative agent of eumycetoma, while *Streptomyces somaliensis*, *Actinomyces madurae* and *Actinomyces pelletierii* are the most frequently reported actinomycetoma causative organisms.^{18–21}

Mycetoma is a painless disease in the initial stages, of gradual onset and progress. Most affected patients are of low

socioeconomic status and have poor levels of health education.²² Furthermore, in mycetoma-endemic regions, the medical and health facilities are meager so most patients report late to medical care with advanced disease.^{19,20} For this reason, most of the reported patients are hospitalised rather than diagnosed at an earlier, more treatable stage of the disease.^{21,23} Late presentation of mycetoma patients leads to a reduced cure rate and a higher chance of recurrence, which in turn increases the rate of morbidity and disability.^{24,25} Despite the long history of mycetoma in Sudan and efforts to fight it, the incidence, prevalence and geographical distribution remain unclear.⁹ Mycetoma patients in Sudan are often poor and illiterate and contracting mycetoma adds to their socioeconomic burden.

Most of the epidemiological characteristics of mycetoma disease are uncertain. The route of infection, the incubation period and the factors that determine susceptibility and resistance to disease are largely unknown.^{22,26} Likewise, its incidence, prevalence and burden have not been properly estimated at global or national levels. This lack of data is one reason for the absence of control and prevention programmes targeting the disease.^{22,27}

Although the infection pathway is not completely known, bacteria and fungi causing mycetoma are thought to penetrate through skin lesions and wounds.²⁸ It is hypothesised that thorns from trees and shrubs of the *Acacia* genus are a major cause of wounds, through which mycetoma infectious agents can infect humans. In most mycetoma-endemic countries, thorny trees and shrubs such as *Acacia* species and other species from the pea family (Family: Fabaceae) are present. Many *Acacia* species bear sharp thorns that protect their leaves from herbivorous animals and, which on the ground, pose a threat to individuals walking barefoot or in open shoes such as sandals.²⁹ Other potential environmental risk factors that have been identified include hygiene practices and proximity to animals.²² Many individuals living in rural areas have poor access to water of adequate quality and sanitation facilities, which impedes good hygiene and could be one of the factors allowing the organism to establish infection, in the context of direct and sustained contact with soil, animals and their dung.³⁰ Many mycetoma-causative organisms have been found in dung-enriched soil, hence dung could also provide a natural habitat for mycetoma-causative agents.²⁶

The distribution of mycetoma has previously been modelled using environmental niche modelling techniques, which suggested central Sudan is suitable for mycetoma occurrence due to an evident overlap of *Acacia* distribution and mycetoma cases, but only 44 mycetoma confirmed cases were used and they were distributed between 12°S and 19°N.⁹ In this work, we aim to provide an update to these models by constructing models using a large occurrence dataset, categorised by type of mycetoma (actinomycetoma vs eumycetoma), using a wider set of environmental predictors and with a range of modelling algorithms. The present study was conducted at the Mycetoma Research Centre (MRC), University of Khartoum, the only WHO collaborating centre on mycetoma globally.³¹ The MRC was established in 1991 to tackle the burden of mycetoma in Sudan with well-trained dedicated staff for the diagnosis and management of mycetoma patients attending from all the states of Sudan. To date, more than 9500 patients have been seen and managed at the centre.

Materials and Methods

Data source

The data included in this study were extracted from the patient database at the MRC. The study included only patients from Sudan who were seen at the centre during 1991–2018. The variables of interest included the patients' demographic characteristics, details of their clinical presentation and diagnosis, place of birth, their current home address if different and their address at the time of onset of symptoms. For most people, the place of birth and current residence were within the same vicinity but if the person had moved outside of their birth area then the address where they were living at the onset of symptoms was used for the analysis. These locations were remotely georeferenced using ArcGIS 10.5 (Environmental Systems Research Institute [ESRI] Inc., Redlands CA, USA) to obtain geographical coordinates defining the mycetoma occurrence locations. The diagnosis of mycetoma and its classification into eumycetoma and actinomycetoma were based on a careful clinical interview and examination, grain culture, cytological examination of aspirates from mycetoma lesions and histopathological examination of surgical biopsy samples. Histopathological methods have been used for mycetoma diagnosis since the 1950s.³¹ Molecular diagnosis using conventional PCR with pan-fungal/bacterial primers and species-specific primers was introduced in 2017.

Environmental predictors and variable selection

Candidate predictors

Because some determinants of mycetoma occurrence remain unelucidated, to avoid exclusion of relevant predictors, we started with a wide range of candidate predictors and used correlation analysis followed by principal components analysis (PCA), a technique used to reduce the dimensionality of large datasets.³²

Continuous gridded datasets (rasters) of 52 environmental variables considered relevant to the mycetoma ecological niche were assembled. These raster datasets included precipitation and temperature, soil composition and pH, livestock distributions, proximity to water sources, elevation and related topographical variables, the predicted distributions of *Acacia* species and other thorny vegetation, measures of atmospheric moisture availability that determine potential vegetative growth and an index of vegetation coverage. Previous studies have indicated these variables to be associated with mycetoma occurrence in Sudan and elsewhere.^{9,33–35}

Climate variables such as minimum, maximum and mean temperature and precipitation levels were obtained from the WorldClim database (version 2.0), a repository of climatic indicators based on long-term data collection from weather stations.³⁶ In addition, from the Consortium for Spatial Information we obtained estimates of aridity (rainfall deficit) and potential evapo-transpiration (an indicator of atmospheric demand for moisture),³⁷ both modelled by using WorldClim datasets and a dataset of modelled elevation, based on data collected by the Shuttle Radar Topography Mission.³⁸

We compiled soil composition and pH datasets from the International Soil Reference and Information Centre-World Soil Information project.³⁹ Raster datasets of predicted livestock

Table 1. Distribution by state of mycetoma patients from Sudan seen at the Mycetoma Research Centre during 1991–2018

State	Eumycetoma	Actinomycetoma	Total
Al Jazirah	2163	234	2397 (34.4%)
Khartoum	730	280	1010 (14.5%)
White Nile	753	101	854 (12.2%)
North Kordofan	479	341	820 (11.7%)
Sennar	562	69	631 (9%)
River Nile	128	59	187 (2.7%)
South Darfur	118	62	180 (2.6%)
North Darfur	97	73	170 (2.4%)
Al-Qadafif	108	42	150 (2.1%)
Kassala	95	40	135 (1.9%)
South Kordofan	80	46	126 (1.8%)
West Darfur	44	36	80 (1.1%)
Northern	45	29	74 (1.1%)
Blue Nile	38	10	48 (0.7%)
East Darfur	34	15	49 (0.7%)
West Kordofan	19	12	31 (0.4%)
Red Sea	13	13	26 (0.4%)
Central Darfur	7	8	15 (0.2%)
Total	5513	1470	6983

distributions (cattle, chickens, sheep and goats), modelled by collaborators from the International Livestock Research Institute, the Food and Agriculture Organization of the United Nations and the Université Libre de Bruxelles, were obtained from the Gridded Livestock of the World version 2 database.⁴⁰ We downloaded spatial data on waterbodies and waterways from the OpenStreetMap project through the *Geofabrik* platform and produced continuous surfaces of straight-line (Euclidean) distance in km to each.⁴¹ We obtained the Enhanced Vegetation Index dataset, generated using satellite imagery data from the Moderate Resolution Imaging Spectroradiometer satellite, from the Earth Explorer NASA site.⁴² Full details of these data sources are shown in Supplementary File 1, Table 1.

Finally, we modelled the distributions of 11 species of thorny trees that are considered potential vehicles for the inoculation of mycetoma-causative organisms using occurrence records downloaded from the Global Biodiversity Information Facility⁴³ along with a suite of environmental variables. Full details of the modelling of these thorny plant species are given in Supplementary File 2.

Each of these gridded datasets was resampled to 1 km² resolution using bilinear interpolation and clipped and aligned to the outline of Sudan. Raster processing was performed using the *raster* package in R v. 3.3.2 and the final map layouts were created with ArcGIS 10.5 software (ESRI Inc., Redlands, CA, USA).

Variable selection

Initially, we calculated correlation coefficients between all candidate predictors across Sudan and identified groups of predictors that were correlated with a Pearson's correlation coefficient

>0.8. We then extracted values of each of the compiled raster datasets at each of the mycetoma-occurrence locations to produce a matrix of covariates (a table showing the x and y coordinates of the occurrences, and at each of these locations, the values of all candidate covariates). We then ran PCA separately for actinomycetoma and eumycetoma to identify the most relevant predictors of the distribution.

In this process, the original variables are reduced into a set of 'components' (or axes), which are linear combinations of variables and are uncorrelated with each other. The components are ordered by the fraction of the total dataset variance they contribute and the variance contribution of each component was summarised by a value termed an eigenvalue. From the minimum set of components that collectively contributed at least 80% of the total variation in occurrence locations, we identified the original variable with the highest influence (measured via the factor loading). We ensured that we did not select any variables that were found to be correlated with another included variable. This process resulted in a selection of uncorrelated variables that collectively contributed the greatest proportion of variance to the occurrence locations. We used the *prcomp* package in R to implement this analysis.

Ensemble modelling

We built separate ensemble distribution models for actinomycetoma and eumycetoma, using the occurrence records and the suite of selected covariates for each. The ensembles were constructed from four algorithms from the BIOMOD (Centre d'Ecologie Fonctionnelle et Evolutive of the CNRS, Montpellier, France) package: generalized boosted regression model, random forest (RF), generalized linear models and generalized additive models.⁴⁴ For each model, we used the default parameters set by the *biomod2* R package. These algorithms are all classified as presence-absence models, meaning that they require both occurrence and absence points to determine environmental suitability, absence points being locations confirmed by surveys to be clear of the outcome under investigation. To account for the lack of absence points within the dataset, we generated 'pseudoabsence' points, representing areas presumably unsuitable for mycetoma. The use of pseudoabsence points to represent areas presumed to be unsuitable for a species is a well-established approach in species distribution modelling.^{45,46} We implemented pseudoabsence selection within the BIOMOD framework, using the 'surface range envelope' approach to define the area of assumed unsuitability. The envelope is estimated through a presence-only suitability model⁴⁷ that identifies the range of locations at which the values of the chosen environmental covariates are within a specified range (here between the 5th and 95th percentiles) of the covariate values at the occurrence locations.⁴⁴ A set of pseudoabsence samples (here a set of five samples), equal in size to the number of occurrence points, was randomly extracted from outside this envelope each time the model was run. Each algorithm was run 50 times with a random sample of 80% of the data points used for model training and the remaining 20% reserved for evaluation. The mean true skill statistic (TSS), the mean proportion correctly classified that is equivalent to kappa and the mean area under the curve (AUC)

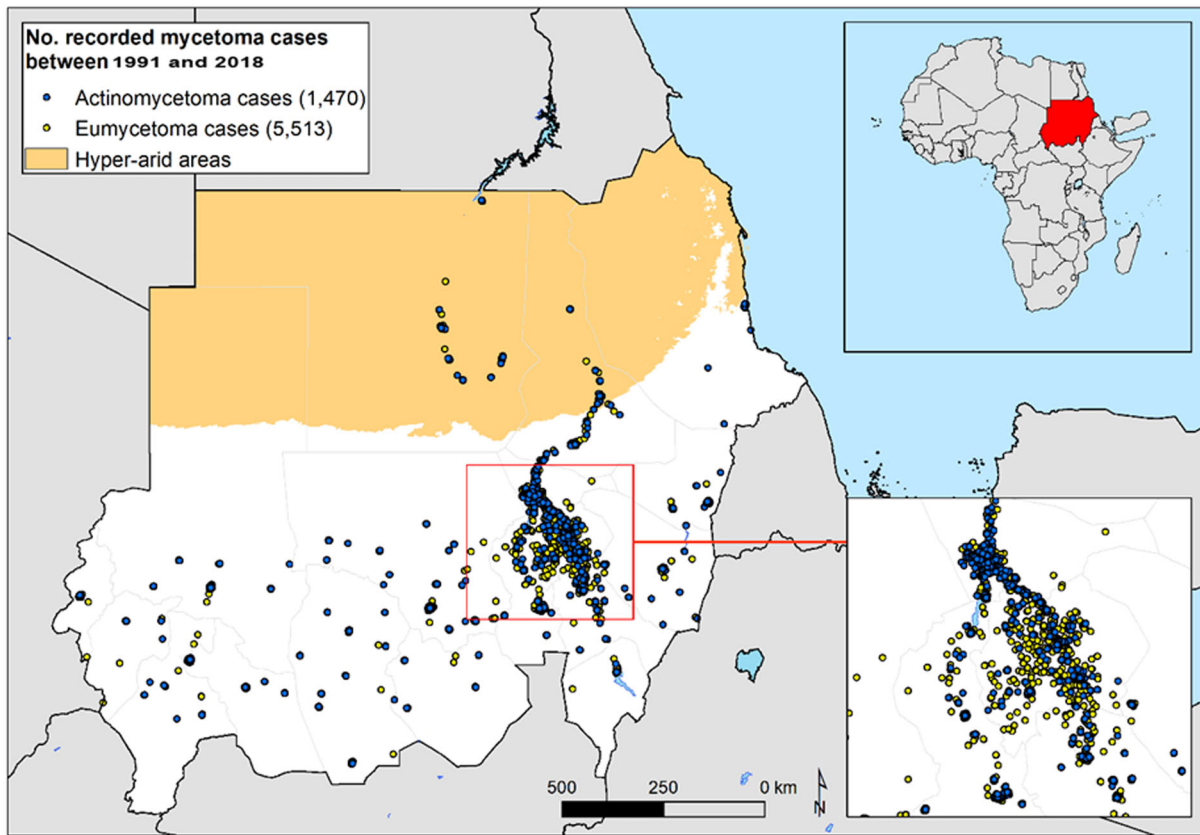


Figure 1. Records of mycetoma cases in Sudan (1991–2018); the upper right map shows where Sudan is situated in Africa; the blue dots represent Actinomycetoma and the yellow dots represent Eumycetoma. Mycetoma cases are distributed along the River Nile with a focus on central Sudan. Cases are distributed sporadically in the western part of Sudan.

of the receiver operating characteristic (ROC) were used to evaluate the performance of each algorithm.

Variables of final models

For eumycetoma, aridity index, soil calcium concentration, wetness index, mean diurnal temperature, distance to the nearest river, presence of cattle, goats and chickens, and predicted occurrence of *Acacia mellifera* and *Faidherbia albida* trees in the area, were included. For actinomycetoma, aridity index, distance to the nearest river, distance to the nearest water body (pond or lake), wetness index, soil sodium and iron concentrations, presence of cattle and sheep in the area, mean diurnal temperature and mean temperature in the coldest year quarter, were included.

All models with an AUC > 0.8 were compiled into the final ensemble distribution model. Within the ensemble, the predicted suitability value was the mean of the included models, weighted by their relative AUC. Upper and lower limits of suitability were estimated by calculating confidence intervals around the ensemble mean suitability value for each cell within the grid. Variable contribution plots were produced to show the relative contribution of each variable to the model, along with marginal effect plots to show the response of both mycetoma types to changes in each modelled covariate.

Results

Dataset of mycetoma occurrence

The modelled data included 7812 unique points after the exclusion of patients (829) from outside Sudan and those lacking diagnostic or geographical information. The patients were seen at the MRC during 1991–2018 and they came from all the states of Sudan. The study included 5513 patients (79%) with confirmed eumycetoma and 1470 patients (21%) with actinomycetoma (Figure 1). Most of the mycetoma patients were from Al Jazirah State (34.4%) and Khartoum State (14.5%) (Table 1).

Environmental predictors of eumycetoma occurrence in Sudan

The PCA for eumycetoma occurrence in Sudan identified 10 variables that characterised the environmental conditions at the occurrence locations. These were aridity index, soil calcium concentration, wetness index, mean diurnal temperature, distance to the nearest river, presence of cattle, goats and chickens, as well as predicted occurrence of *A. mellifera* and *F. albida* trees in the area. Two variables—the predicted density of cattle and chickens—were excluded after the initial modelling step as each

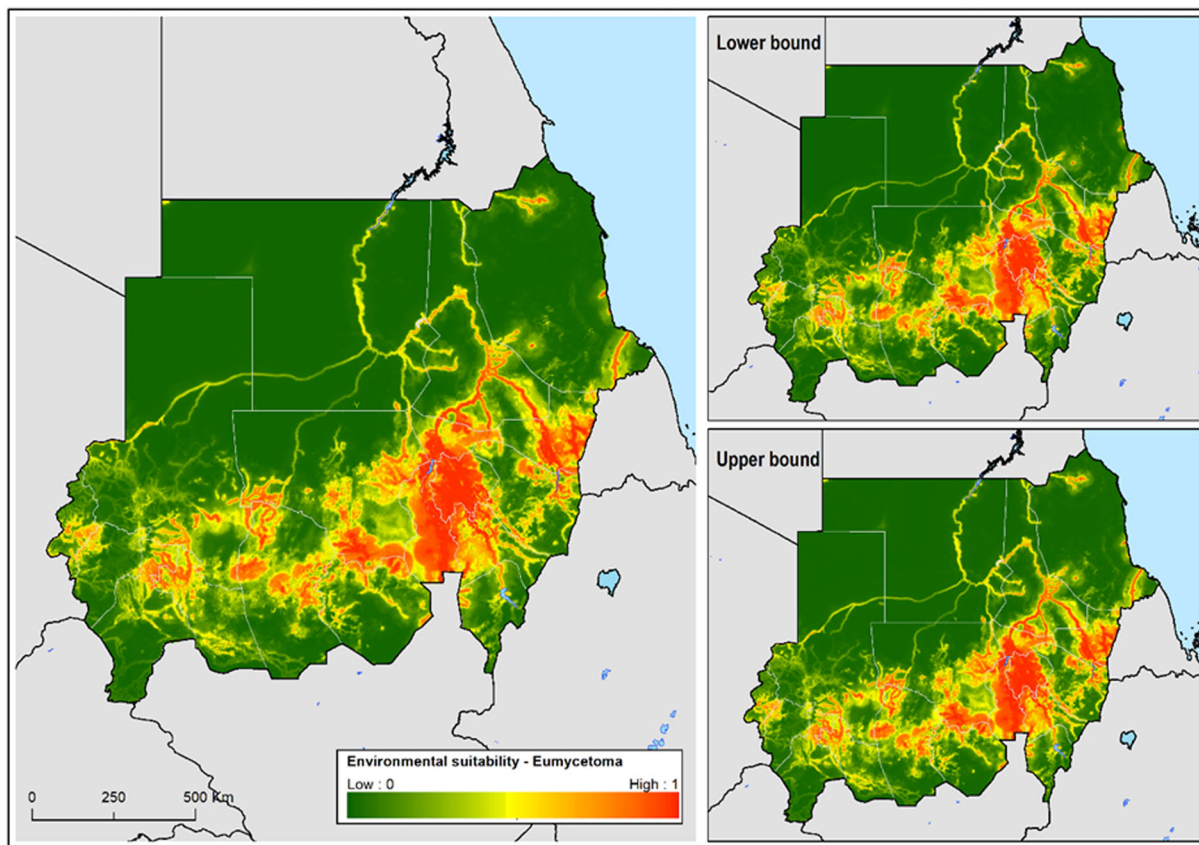


Figure 2. Predicted environmental suitability for eumycetoma in Sudan. Eumycetoma is predicted to occur along the River Nile and its tributaries with hotspots in the central and southeastern parts of Sudan.

Table 2. Validation metrics for ensemble models for eumycetoma and actinomycetoma suitability

		Weighted mean	Lower CI	Upper CI
Eumycetoma	TSS	0.909	0.909	0.911
	ROC	0.993	0.993	0.993
	kappa	0.898	0.897	0.898
Actinomycetoma	TSS	0.921	0.922	0.922
	ROC	0.995	0.995	0.995
	kappa	0.903	0.903	0.902

Abbreviations: ROC, receiver operating characteristic; TSS, true skill statistic.

contributed less than 1% of the variability in the response variable (eumycetoma occurrence).

All models for eumycetoma occurrence performed well with an ensemble ROC score of 0.993, sensitivity of 96.475% and specificity of 94.476%. The mean TSS was 0.909 and mean kappa score was 0.898 (Table 2).

Across the RF models, the distance to the nearest river was the strongest environmental predictor of suitability for eumycetoma occurrence and the suitability decreased with increased distance

to the nearest river. The diversity of thorny trees, represented by the number of species predicted to be present, was the second most important predictor of suitability for eumycetoma occurrence, with a greater diversity of thorny trees associated with a higher probability of occurrence of the disease. We found the suitability for eumycetoma to be higher in areas with a lower probability of occurrence of *A. mellifera* and arid areas. Soil calcium concentration suitable for eumycetoma occurrence was from 3000 to 11 000 mg/kg, while mean diurnal temperature range and the probability of occurrence contributed to a lesser extent; goat density contributed the least (Supplementary File 1, Figure 1).

Environmental predictors of actinomycetoma occurrence in Sudan

The PCA for actinomycetoma occurrence also revealed 10 variables as potential environmental predictors. The variables were aridity index, distance to the nearest river, distance to the nearest water body (pond or lake), wetness index, soil sodium and iron concentrations, presence of cattle and sheep in the area, mean diurnal temperature and mean temperature in the coldest year quarter. After the primary analysis, four variables—distance to the nearest river, distance to the nearest water body (pond or lake), mean temperature in the coldest year quarter and variety in thorny trees—were included.

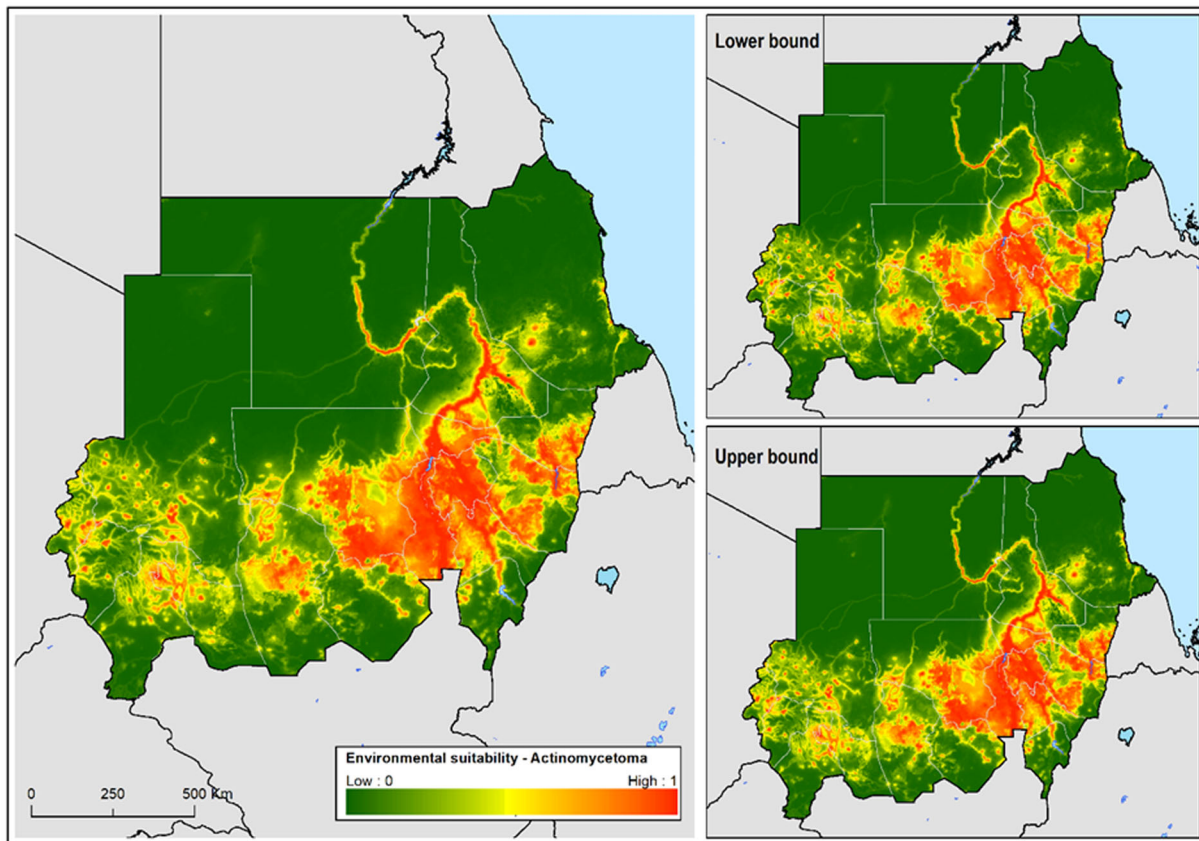


Figure 3. Predicted environmental suitability for actinomycetoma in Sudan. Actinomycetoma is predicted to occur along the River Nile and its tributaries with hotspots in the central and southeastern parts of Sudan with sporadic occurrence in the western part of Sudan.

All models for actinomycetoma occurrence performed well, with a ROC score of 0.995, sensitivity of 95.445% and specificity of 96.709%. The mean TSS was 0.921 and the mean kappa was 0.903 (Table 2).

Across the RF models, distance to the nearest water body was the major contributor to actinomycetoma occurrence, followed by the distance to the nearest river. Mean temperature during the coldest quarter over a range of 18–25°C contributed to the predicted suitability for actinomycetoma occurrence by 13%. Arid areas with a low concentration of soil sodium contributed by almost 11%. The diversity of thorny trees in an area showed no difference between areas with no diversity of thorny trees and areas with a diverse group of thorny trees (Supplementary File 1, Figure 3).

Predicted suitability for mycetoma occurrence in Sudan

Both eumycetoma and actinomycetoma were predicted to occur around the central and southeastern parts of Sudan and along the Nile river valley and its tributaries. The central states of Sudan, Al Jazirah, White Nile, Khartoum and Sennar State had the largest areas predicted as suitable for eumycetoma. The predicted distribution of actinomycetoma was slightly wider than that of eumycetoma, with more patchy suitability predicted in the western states of Sudan (Supplementary File 1, Figures 5 and 6). The

overlapping areas for actinomycetoma and eumycetoma risk are illustrated in Figure 4.

Discussion

We have applied ecological niche modelling techniques to identify environmental variables associated with the occurrence of mycetoma in Sudan. Models were run separately for eumycetoma and actinomycetoma and gave similar results in terms of the predicted distribution. Both types were predicted along the River Nile and its tributaries, mainly around the central and southeastern parts of Sudan covering Khartoum, Al Jazirah and White Nile States, the northern part of Sennar State and the eastern part of North Kordofan State, which are known to have a high burden of cases.⁴⁸ The models also predicted the occurrence of eumycetoma in the north and eastern parts of the country, where cases have not previously been recorded. This might indicate under-reporting or misdiagnosis of cases from these areas. Such areas must be considered as priorities for future surveillance activities and to offer targeted training of medical personnel for early detection and better diagnosis of mycetoma.

This study showed the suitability for mycetoma occurrence was widespread in Khartoum State, and this can be attributed to the high influx of people from rural endemic areas to the state,

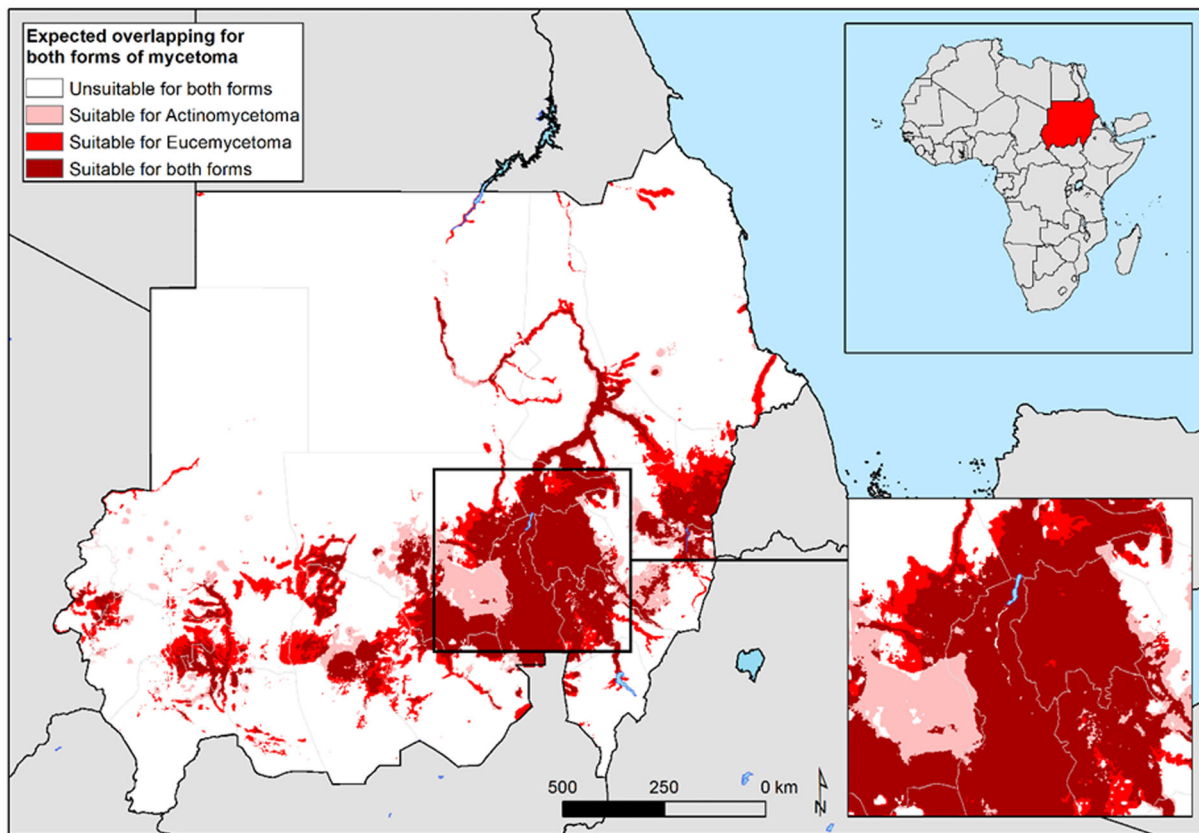


Figure 4. Overlaps of actinomycetoma and eumycetoma environmental suitability in Sudan. The map shows central Sudan to be the most suitable area for the occurrence of both types of mycetoma.

due to various socioeconomic difficulties, often secondary to having the disease. Also, Khartoum residents regularly visit their families in endemic regions and can contract the infection during such visits. Furthermore, much of the state beyond the city is undeveloped rural land that harbours the ecological and sociodemographic characteristics associated with mycetoma.

As previously mentioned, mycetoma is distributed globally in a belt form, and areas within the belt are dry with annual rainfall of 50–1000 mm and high temperatures.⁴⁹ Given the recognised influence of temperature on the survival of mycetoma-causative organisms,⁵⁰ it is not surprising that temperature-related variables had a large impact on the suitability for both types of mycetoma. Our results indicate that environments with mean temperatures in the coldest quarter of the year of 20–25°C are most suitable for actinomycetoma-causative organisms. Daily variation in temperature was a strong predictor of suitability for eumycetoma. It has been reported that fungal organisms that cause eumycetoma cannot survive in environments where the daily range in temperature is more than 15°C on average. This could be explained by the fact that the organisms do not live in regions that have extreme variations in temperature.⁵¹ According to our results, mycetoma seems to be predicted in arid areas; as mentioned previously in the literature, mycetoma prevails in reasonably arid areas with a short rainy season of 4–6 mo.⁵⁰

Analysis of the MRC data showed that most of the patients reside in locations close to water sources and the models predicted the same pattern. This observation is most likely explained by the fact that while mycetoma-causing species thrive in drier environments, people living in such environments settle near a water source to enable activities such as washing and growing crops. Thus distance to rivers correlates with mycetoma occurrence as a result of human behaviour, rather than environmental conditions near rivers supporting organism survival.

To our knowledge, the influence of the soil mineral components on the growth of mycetoma-causative organisms has not been studied thoroughly. Our findings suggested that soil calcium concentration had a moderate effect on suitability for eumycetoma and that sodium concentrations contribute to the disease occurrence suitability. In histopathological studies, calcium, among other elements, was found in abundance in the cell wall of *M. mycetomatis*. It is an important element in the formation of cement substance, which is a hard brown material containing melanin, heavy metal ions, proteins and lipids. Cement is a fungal protective mechanism, acting as a sink for harmful unpaired electrons and provides the cell walls with structural rigidity as well as storing water and ions to prevent desiccation.⁵² This may indicate that calcium could be an element in the organism's virulence by contributing to the myelinisation and sclerosis of the grain.⁵¹ In a recent cross-sectional study conducted in Mexico, cases were

prevalent among three soil types that had a high level of calcium.³⁵ This observation warrants further laboratory studies to explore the possible role of these soil elements in the pathogens' survival.

We found that environments with a greater number of species of Acacia trees were more suitable for eumycetoma, supporting previous evidence that suggested a role for skin penetration by thorns in the transmission of the disease.⁹ In our study, and in contrast to the study conducted by Samy et al. in 2014, not only was the distribution of Acacia trees tested, but also the role of diversity of these trees in mycetoma occurrence in both its types.⁹ However, areas with either no thorny trees or areas with more than five species of thorny trees were both suitable for actinomycetoma occurrence. In our study, eumycetoma, but not actinomycetoma, occurrence seems to be predicted by the diversity of thorny trees. This could be explained by the observation that eumycetoma-causative organisms live on the thorns. This study showed no strong association between the individual species types and mycetoma occurrence. These observations on the association between disease occurrence and individual tree species, their botanical properties and other parameters, need further investigation. Such studies are important to gain more in-depth knowledge about disease transmission and the life cycle of the causative microorganisms to develop ways to interrupt them.

In conclusion, the studied models have indicated that arid areas proximal to water sources, soil with low concentrations of calcium and sodium and with a variety of thorny tree species provide the most suitable environment for the occurrence of mycetoma. This work included environmental probable risk factors that have not been studied previously and which could be used to guide the implementation of preventive interventions and control strategies in the affected areas. Until there is robust evidence on disease transmission and the life cycle of the microorganisms involved, it may be advisable to limit the use of thorny trees in building animal enclosures and houses. However, such simple and effective interventions may not align with longstanding cultural practices and should be comprehensively discussed with all stakeholders and community members to ensure community acceptance. Protective measures such as the use of footwear should be encouraged in people who are in direct contact with these trees.

Our study also considered the suitable environmental indicators across the whole of Sudan with updated data from the MRC database records using a wide range of environmental risk factors. The results could help guide planning for national and international scale surveys for mycetoma identification and to design targeted training for medical staff on early detection and diagnosis. Moreover, they could aid in the design of national and global mycetoma advocacy and awareness programmes leading to early active case detection with early appropriate patient management in areas that are highly endemic for mycetoma.

Authors' contributions: RH, HS, JC, SB, EG, DA, MJN, KD and AHF conceived and designed the study; RH, HS, JC, SB, EG, DA, AHF, MJN and KD performed the study; RH, HS, JC, SB, EG, DA, MJN, KD and AHF analysed the data; RH and KD wrote the first draft of the manuscript. RH, HS, JC, SB, EG, DA, MJN, KD and AHF revised the manuscript. All authors read and approved the final manuscript.

Funding: This research was funded by the National Institute for Health Research (NIHR) Global Health Research Unit on NTDs at BSMS (16/136/29) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK government. KD is supported by the Wellcome Trust [grant number 201900/Z/16/Z] as part of his International Intermediate Fellowship.

Competing interests: The authors declare that they have no conflicts of interest.

Ethical approval: This study was approved by the Mycetoma Research Centre, Khartoum, Sudan (IRB, No. SUH 11/12/2018) and from the BSMS Research Governance and Ethics Committee (ER/BSMS435/1). Written informed consent was obtained from each adult patient, and verbal assent was obtained from minors (aged <18 y) in addition to signed written informed consent from the parent/guardian.

Data availability: Data are available upon request.

Supplementary data: Supplementary data are available at [Transactions](#) online.

Supporting information:

S1 File: Covariates, marginal effect plots and predicted occurrences

S2 Text: Modelling the environmental suitability for the species of thorny vegetation present in Sudan.

References

- Fahal A, Hassan AME, Abdel-Rahman ME, et al. A new model for management of mycetoma in the Sudan. *PLoS Negl Trop Dis*. 2014;8(10):e3271.
- Kwizera R, Bongomin F, Meya DB, et al. Mycetoma in Uganda: a neglected tropical disease. *PLoS Negl Trop Dis*. 2020;14(4):e0008240.
- Mohamed HT, Fahal A, van de Sande WW. Mycetoma: epidemiology, treatment challenges, and progress. *Res Rep Trop Med*. 2015;6:31–6.
- Siddig EE, Mhmoud NA, Bakhiet SM, et al. The accuracy of histopathological and cytopathological techniques in the identification of the mycetoma causative agents. *PLoS Negl Trop Dis*. 2019;13(8):e0007056.
- Zijlstra EE, van de Sande WW, Fahal AH. Mycetoma: a long journey from neglect. *PLoS Negl Trop Dis*. 2016;10(1):e0004244.
- van de Sande WW. Global burden of human mycetoma: a systematic review and meta-analysis. *PLoS Negl Trop Dis*. 2013;7(11):e2550.
- Reis CMS, Reis-Filho EGdM. Mycetomas: an epidemiological, etiological, clinical, laboratory and therapeutic review. *An Bras Dermatol*. 2018;93(1):8–18.
- Relhan V, Mahajan K, Agarwal P, et al. Mycetoma: an update. *Indian J Dermatol*. 2017;62(4):332.
- Samy AM, van de Sande WW, Fahal AH, et al. Mapping the potential risk of mycetoma infection in Sudan and South Sudan using ecological niche modeling. *PLoS Negl Trop Dis*. 2014;8(10):e3250.
- Bonifaz A, Tirado-Sánchez A, Calderón L, et al. Mycetoma: experience of 482 cases in a single center in Mexico. *PLoS Negl Trop Dis*. 2014;8(8):e3102.
- Castro LM, Piquero-Casals J. Clinical and mycologic findings and therapeutic outcome of 27 mycetoma patients from São Paulo, Brazil. *Int J Dermatol*. 2008;47(2):160–3.

- 12 Darré T, Saka B, Mouhari-Toure A, et al. Mycetoma in the Togolese: an update from a single-center experience. *Mycopathologia*. 2018;183(6):961–5.
- 13 Emery D, Denning DW. The global distribution of actinomycetoma and eumycetoma. *PLoS Negl Trop Dis*. 2020;14(9):e0008397.
- 14 Fahal A, Mahgoub ES, Hassan AME, et al. Mycetoma in the Sudan: an update from the Mycetoma Research Centre, University of Khartoum, Sudan. *PLoS Negl Trop Dis*. 2015;9(3):e0003679.
- 15 Hay RJ, Fahal AH. Mycetoma: an old and still neglected tropical disease. *Trans R Soc Trop Med Hyg*. 2015;109(3):169–70.
- 16 Dieng M, Niang S, Diop B, et al. Actinomycétomes au Sénégal. Étude de 90 cas. *Bull Soc Pathol Exot*. 2005;98(1):14–7.
- 17 Lim W, Eadie K, Konings M, et al. *Madurella mycetomatis*, the main causative agent of eumycetoma, is highly susceptible to olorofim. *J Antimicrob Chemother*. 2020;75(4):936–41.
- 18 Schibli A, Goldenberger D, Krieg A, et al. Painless swelling of the forefoot and recurrent subcutaneous abscesses of the lower leg—Two distinct presentations illustrating the spectrum of eumycetoma in a nonendemic country. San Francisco, CA: Public Library of Science; 2017.
- 19 Sow D, Ndiaye M, Sarr L, et al. Mycetoma epidemiology, diagnosis management, and outcome in three hospital centres in Senegal from 2008 to 2018. *PLoS One*. 2020;15(4):e0231871.
- 20 Welsh O, Al-Abdely HM, Salinas-Carmona MC, et al. Mycetoma medical therapy. *PLoS Negl Trop Dis*. 2014;8(10):e3218.
- 21 Fahal A, Mahgoub ES, Hassan AME, et al. Head and neck mycetoma: the mycetoma research centre experience. *PLoS Negl Trop Dis*. 2015;9(3):e0003587.
- 22 Bakhiet SM, Fahal AH, Musa AM, et al. A holistic approach to the mycetoma management. 2018;12(5):e0006391.
- 23 Wadal A, Elhassan TA, Zein HA, et al. Predictors of post-operative mycetoma recurrence using machine-learning algorithms: the Mycetoma Research Center experience. *PLoS Negl Trop Dis*. 2016;10(10):e0005007.
- 24 Abbas M, Scolding PS, Yosif AA, et al. The disabling consequences of Mycetoma. *PLoS Negl Trop Dis*. 2018;12(12):e0007019.
- 25 Suleiman SH, Wadaella ES, Fahal AH. The surgical treatment of mycetoma. *PLoS Negl Trop Dis*. 2016;10(6):e0004690.
- 26 van de Sande WW, Fahal AH, Goodfellow M, et al. The mycetoma knowledge gap: identification of research priorities. *PLoS Negl Trop Dis*. 2014;8(3):e2667.
- 27 Fahal AH, Suliman SH, Hay R. Mycetoma: the spectrum of clinical presentation. *Trop Med Infect Dis*. 2018;3(3):97.
- 28 Welsh O, Vera-Cabrera L, Salinas-Carmona MC. Mycetoma. *Clin Dermatol*. 2007;25(2):195–202.
- 29 Charles-Dominique T, Barczy JF, Le Roux E, et al. The architectural design of trees protects them against large herbivores. *Funct Ecol*. 2017;31(9):1710–7.
- 30 Azrag R, Bakhiet S, Almalik A, et al. Is mycetoma a vector-borne disease? The first report on the detection of *Madurella mycetomatis* in ticks. *bioRxiv*. 2019:775767.
- 31 Mycetoma Research Centre. Available at: <http://www.mycetoma.edu.sd/> [accessed July 10, 2020].
- 32 Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374(2065):20150202.
- 33 Ahmed AO, van Leeuwen W, Fahal A, et al. Mycetoma caused by *Madurella mycetomatis*: a neglected infectious burden. *Lancet Infect Dis*. 2004;4(9):566–74.
- 34 Zijlstra EE, van de Sande WWJ, Welsh O, et al. Mycetoma: a unique neglected tropical disease. *Lancet Infect Dis*. 2016;16(1):100–12.
- 35 Cardenas-de la Garza JA, Welsh O, Cuellar-Barboza A, et al. Climate, soil type, and geographic distribution of actinomycetoma cases in Northeast Mexico: a cross-sectional study. *PLoS One*. 2020;15(5):e0232556.
- 36 WorldClim. Global Climate data. Available from: <https://www.worldclim.org/data/index.html> [accessed April 10, 2020].
- 37 CGIAR-CSI. Consortium for Spatial Information. Available at: <http://www.cgiar-csi.org> [accessed April 10, 2020].
- 38 Shuttle Radar Topography Mission (SRTM). Elevation data. Available at: <https://earthdata.nasa.gov/learn/articles/nasa-shuttle-radar-topography-mission-srtm-version-3-0> [accessed April 10, 2020].
- 39 ISRIC. World Soil Information Soil property maps of Africa at 1 km. Available at: <https://www.isric.org/> [accessed April 10, 2020].
- 40 Livestock Geo-Wiki. Livestock. Available at: <https://livestock.geo-wiki.org/home-2/> [accessed April 10, 2020].
- 41 OpenStreetMap. OpenStreetMap contributors. Available at: <https://www.openstreetmap.org/> [accessed April 10, 2020].
- 42 usgs. Enhanced Vegetation Index. Available at: <https://earthexplorer.usgs.gov/> [accessed April 10, 2020].
- 43 Global Biodiversity Information Facility Thorny trees species. Available at: <https://www.gbif.org/> [accessed April 10, 2020].
- 44 Thuiller W, Lafourcade B, Engler R, et al. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*. 2009;32(3):369–73.
- 45 Bhatt S, Gething PW, Brady OJ, et al. The global distribution and burden of dengue. *Nature*. 2013;496(7446):504–7.
- 46 Pigott DM, Bhatt S, Golding N, et al. Global distribution maps of the leishmaniasis. *Elife*. 2014;3:e02851.
- 47 Sui Z. Modeling tree species distribution and dynamics under a changing climate, natural disturbances, and harvest alternatives in the southern United States. Starkville, MS: Mississippi State University; 2015.
- 48 Fahal A, Sabaa AA. Mycetoma in children in Sudan. *Trans R Soc Trop Med Hyg*. 2010;104(2):117–21.
- 49 Serrano J, Dorado A, Chacon Y, et al. Mycetomas in the endemic area of Argentina. *Curr Fungal Infect Rep*. 2020:1–5.
- 50 Ahmed AO, van Leeuwen W, Fahal A, et al. Mycetoma caused by *Madurella mycetomatis*: a neglected infectious burden. *Lancet Infect Dis*. 2004;4(9):566–74.
- 51 Ibrahim AI, El Hassan AM, Fahal A, et al. A histopathological exploration of the *Madurella mycetomatis* grain. *PLoS One*. 2013;8(3):e57774.
- 52 Ahmed AA, van de Sande W, Fahal AH. Mycetoma laboratory diagnosis. *PLoS Negl Trop Dis*. 2017;11(8):e0005638.