

Local and regional spread of banana xanthomonas wilt (BXW) in space and time in Kagera, Tanzania

M. M. Shimwela^{ab*}, J. K. Blackburn^c, J. B. Jones^d, J. Nkuba^a, H. A. Narouei-Khandan^b, R. C. Ploetz^e, F. Beed^f and A. H. C. van Bruggen^b

^aARI-Maruku, PO Box 127, Bukoba, Tanzania; ^bUniversity of Florida, Plant Pathology Department and Emerging Pathogens Institute, University of Florida, Gainesville FL23611; ^cSpatial Epidemiology and Ecology Research Laboratory, Department of Geography, University of Florida, and Emerging Pathogens Institute, University of Florida, Gainesville FL23611; ^dPlant Pathology Department, University of Florida, Gainesville FL23611; ^ePlant Pathology Department, University of Florida, TREC – Homestead FL33031, USA; and ^fResearch & Development Building, AVRDC – The World Vegetable Centre, Kasetsart University Bangkok, Bangkok 10900, Thailand

Banana is an important crop in the Kagera region of Tanzania. Banana xanthomonas wilt (BXW) was first reported in Kagera in 2006, and is now an important limiting factor in banana production, because all cultivars are susceptible and infected plants can fail to produce fruit. BXW is caused by *Xanthomonas campestris* pv. *musacearum* (Xcm), which is spread by farm tools, infected planting materials, and pollinating insects. Practices that address Xcm dissemination, such as mat removal, debudding and tool sterilization, have not prevented the spread of BXW in the region. Disease surveys were conducted in Kagera from 2007 to 2011 to assess BXW presence, monitor its intensity and evaluate its socioeconomic impacts. Spatiotemporal clusters of BXW were analysed with ARCGIS and SAS. The relationship between BXW clusters and environmental variables was examined using bivariate correlations in SPSS; two modelling approaches, MaxEnt (maximum entropy) and logistic regression, were used to predict the potential distribution of BXW in Tanzania. Disease progress over time was best described with the Gompertz model. Significant clustering of BXW was observed in all years and hotspots were located in the Muleba, Karagwe, Misenyi and Bukoba rural districts. These findings suggest that BXW spreads rapidly over short distances. BXW clusters were positively correlated with rainfall and negatively with temperature and altitude. According to MaxEnt, precipitation was the main factor associated with BXW development. MaxEnt and logistic regression predicted a wide potential distribution of BXW in Tanzania because the climate in all banana-growing regions is conducive for its establishment.

Keywords: climatic suitability, cluster analysis, hotspots, logistic regression, MaxEnt, *Xanthomonas campestris* pv. *musacearum*

Introduction

Banana xanthomonas wilt (BXW) is a serious emerging disease that affects all banana and plantain cultivars (*Musa acuminata* and hybrids of *M. acuminata* and *M. balbisiana*) that have been tested (Tripathi *et al.*, 2009). BXW was probably first described by Castellani (1939) on enset (*Ensete ventricosum*), a close relative of banana, in Ethiopia. Subsequently, Yirgou & Bradbury (1968) described the causal agent as a species of *Xanthomonas* (now known as *Xanthomonas campestris* pv. *musacearum* (Xcm)), and reported that banana was also susceptible (Yirgou & Bradbury, 1974). Although Yirgou & Bradbury (1974) recognized that BXW posed a threat to banana cultivation outside Ethiopia, the disease was limited to Ethiopia until it was reported in the Mukono district of Uganda in 2001 (Tushemereirwe *et al.*, 2004).

However, once BXW established in Uganda, it spread quickly and within 5 years the disease had spread to neighbouring countries, including Tanzania, Kenya, Rwanda, Burundi and the Democratic Republic of Congo (Carter *et al.*, 2010).

In Tanzania, BXW was identified for the first time in January 2006 in Kabare village in the Muleba district of the Kagera region of northwest Tanzania (Carter *et al.*, 2010; Fig. 1). Thereafter, the disease spread to all seven districts in Kagera and to the neighbouring regions of Mara and Kigoma. This raised concerns that BXW would spread further east to Kilimanjaro and Mbeya, which are major banana-producing regions in Tanzania.

The BXW pathogen is a Gram-negative, rod-shaped bacterium with a single polar flagellum (Yirgou & Bradbury, 1974). It invades the vascular system of banana, causing wilt and death of the plant. Its primary mode of natural transmission is by insect vectors, particularly stingless bees, which feed on ooze from fresh bract scars in male buds, especially those of cultivars with an ABB genome (cultivars that have one set of chromosomes donated by *M. acuminata* and two by *M. balbisiana*;

*E-mail: mshimwela@ufl.edu

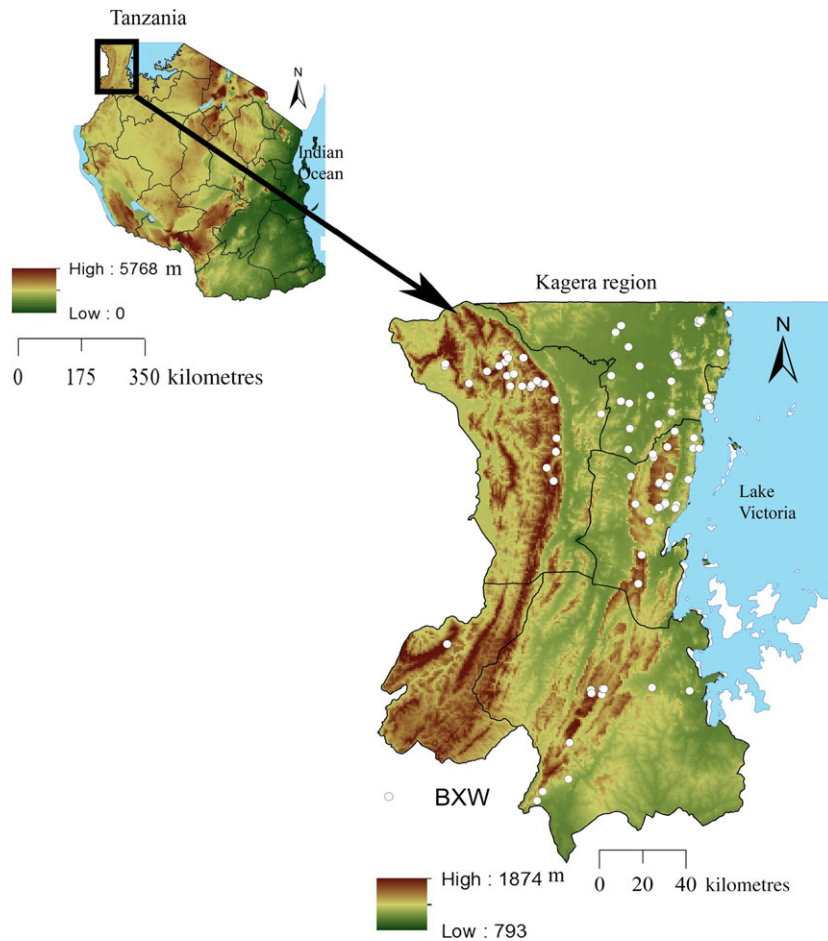


Figure 1 Locations of villages that had one or more fields affected with banana xanthomonas wilt during five surveys (2007–11) in Kagera region, Tanzania. [Colour figure can be viewed at wileyonlinelibrary.com].

Tripathi *et al.*, 2009). The pathogen is also disseminated by movement of infected planting material and via contaminated garden tools (Tripathi *et al.*, 2009). Nectar-collecting and fruit-eating birds and bats are also suspected of transmitting the bacterium. Symptomless banana bunches and leaves used to wrap bunches for transport to markets are another important source of Xcm inoculum that may be responsible for its long-distance spread (Nakato *et al.*, 2013).

The main symptoms of BXW are wilting and necrosis of leaves, wilting of the male flower, vascular discolouration, and premature and uneven ripening of fruits. Yellow ooze is excreted from cuts on any part of an infected plant. Initial symptoms on affected plants vary depending on the mode of infection. When Xcm transmission occurs via pollinating insects, male buds wilt and wither, whereas plants infected before flowering, via contaminated garden tools, display a progressive yellowing of leaves from the leaf tip toward the petioles. Most infected suckers die prematurely (Tripathi *et al.*, 2009).

Most BXW management practices were developed for moko disease, a bacterial wilt of banana caused by *Ralstonia solanacearum* phylotype II (Cellier *et al.*, 2012). Moko disease resembles BXW with respect to its epidemiology and the damage it causes (Thwaites *et al.*, 2000). Management practices recommended for BXW

include roguing of affected mats (whole mat removal), burying or burning of infected materials, removal of single stems from affected plants, disinfecting farm tools after every use, and male bud removal soon after the last hand of fruit is formed (Blomme *et al.*, 2014). Similar to moko disease, control of BXW is very challenging due to the absence of effective bactericides and acceptable resistant cultivars (i.e. those without male buds or persistent bracts, which are not infested by insect vectors). Even if effective bactericides were available, they would probably not be used in small-scale diversified banana fields that predominate in Eastern Africa.

Since BXW spread to Kagera, it has caused severe losses and affected the livelihoods of people who rely on banana as a staple crop. The numbers of poor households, vulnerable to food insecurity, have increased each year. By 2011, the combined loss of banana production and cost of buying alternative food crops was an estimated US\$10 million (Nkuba *et al.*, 2015).

Knowledge of where and when a disease occurs is crucial for identifying disease risk factors, improving the efficiency of surveillance methods and identifying control strategies (Madden *et al.*, 2007). A better understanding of the factors that are associated with BXW risk is needed in order to devise effective control and eradication strategies. Although epidemiological studies on

BXW have been conducted (Shimelash *et al.*, 2008; Adikini *et al.*, 2013; Nakato *et al.*, 2013), little is known about the distribution of BXW in space and time in relation to environmental factors (Bouwmeester *et al.*, 2010; Shimwela *et al.*, 2016).

Banana is a perennial crop that requires continuous management throughout the year. Altitude and weather conditions, particularly temperature and rainfall, influence growth of the crop as well as the spread of BXW (Shimelash *et al.*, 2008; Tripathi *et al.*, 2009). Transmission of Xcm occurs throughout the year, but varies according to altitude and season. For example, male bud infection that is mediated by insect vectors has not been observed at high altitudes in Ethiopia (Shimelash *et al.*, 2008), but is common at mid and low altitudes around Lake Victoria (Tripathi *et al.*, 2009). However, quantitative relationships between altitude, climatic factors and BXW incidence are not known, and the risk for its spread to other banana-growing regions in Tanzania has not been estimated.

Modelling and geographical information systems (GIS) allow spatiotemporal patterns of disease occurrence and risk levels to be computed over wide areas with limited data. Geospatial analytical techniques, including spatial autocorrelation, hotspot analysis, cluster analysis and temporal spatial analysis, are commonly used to determine patterns of disease spread, areas with high disease incidence, and the extent to which a disease could spread (Pfeiffer *et al.*, 2008).

One of the ways to assess the risk of spread is to evaluate climatic requirements for disease establishment (Shaw & Osborne, 2011; Yuen & Mila, 2015). How environmental variables relate to the spatial occurrence of a disease can be estimated with different modelling techniques and algorithms. Maximum entropy (MaxEnt) modelling and logistic regression have been employed widely to determine the potential areas with high risk of pathogen establishment (Narouei-Khandan *et al.*, 2016). MaxEnt is a presence-only ecological niche modelling technique that enables the efficient use of small sample sizes, and the use of both continuous and categorical variables (Phillips *et al.*, 2006). Logistic regression is frequently used to model plant disease distribution and risk (Shaw & Osborne, 2011), and can be implemented in a generalized linear model (GLM) framework for data with a binary distribution, such as species presence or absence (Hosmer & Lemeshow, 2000). The output from a logistic regression model can be used in GIS to map the predicted geographic distribution of areas most likely to sustain the species of interest.

The overall aim of this study was to describe the current spatial and spatiotemporal patterns of BXW spread, and relate current disease distribution to geophysical and climatological factors. To this end, modelling techniques and spatial and geostatistical analyses were used to study BXW and co-occurring environmental factors in the Kagera region of Tanzania. Five spatial datasets collected between 2007 and 2011 were used to: (i) investigate the local and regional distribution of BXW in Kagera in

space and time, (ii) examine relationships between BXW hotspots in Kagera and environmental variables, and (iii) develop a risk map for BXW development in Tanzania based on significant climatological and geophysical data.

Materials and methods

BXW data acquisition

Surveys were conducted between 2007 and 2011 to assess the presence of BXW disease, monitor its intensity, and to evaluate its socioeconomic impacts in Kagera (Fig. 1). The surveys were conducted by experienced ARI-Maruku research teams and well-trained agricultural extension officers from district-level agriculture and livestock development offices (DALDO), in collaboration with international institutes such as the International Institute of Tropical Agriculture (IITA), the USAID-funded regional C3P project, Bioversity International, and the Food and Agriculture Organization (FAO) of the United Nations. Datasets of BXW used in this study were obtained from scientists and extension officers who participated in surveys, and from published reports (Table S1). The BXW surveys used the same approaches to ensure comparable results and proposed actions (Blomme *et al.*, 2014). Detailed sampling methodologies and preliminary data analyses can be found elsewhere (Abele *et al.*, 2007; Rugalema & Mathieson, 2009; Nkuba *et al.*, 2015). All surveys employed a common structured household questionnaire through face-to-face interviews and BXW field inspections by research and extension personnel. The sample sizes (number of farms surveyed) were 204, 269, 261, 120 and 120 in 2007, 2008, 2009, 2010 and 2011, respectively. In all surveys, the locations were georeferenced and BXW was identified by the researchers and extension officers, based on diagnostic symptoms of the disease, including wilting of leaves, wilting of the male flower, premature and uneven ripening of fruits and excretion of yellow ooze from cut surfaces of affected plants. A follow-up survey was conducted in 2014 to acquire GPS coordinates for BXW-positive farms that had not been georeferenced in previous surveys. For geostatistical analysis, locations where the disease was observed were considered presence points and locations where BXW was not observed were considered absence points.

To assess the progress of BXW between 2006 and 2011, data were also obtained by DALDO officers during routine monitoring in 703 villages. In each village, 5–18 farms were visited, depending on the size of the village. The data obtained from these villages consisted of positive cases only, i.e. villages where BXW was observed, as confirmed by ward (administrative level 4, containing several villages) agricultural extension officers.

Geoprocessing of BXW data

Because datasets on the occurrence of BXW were acquired from different sources, in various formats and depths, geoprocessing of the data was necessary prior to further analysis; geoprocessing included projection and coordinate transformation, clipping, masking, feature selection, buffering, grid resampling and reclassification (Ormsby *et al.*, 2010). The datasets were processed within ArcGIS v. 10.2.1 using the SPATIAL ANALYST extension.

Temporal development of BXW

Disease incidence (y), measured as the proportion of 703 villages in Kagera with at least one BXW case throughout the year, was

used to generate disease progress curves as a function of time (x). The number of villages affected by BXW in each year was obtained from DALDO offices, based on the monthly reports of ward agricultural extension officers. Temporal development of cumulative BXW incidences were analysed by fitting monomolecular, logistic and Gompertz models to the observed data using PROC NLIN in SAS v. 9.3 (SAS Institute Inc.). The goodness of fit of the models was evaluated based on the regression of predicted versus observed values and visual inspection of residual plots (Madden *et al.*, 2007).

Spatial analysis of BXW cases in Kagera region

Various spatial analysis techniques, such as kernel density estimation, average nearest neighbours index (ANNI), spatial autocorrelation analysis, Getis and Ord statistics $G_i^*(d)$, and Ripley's K function are commonly used to detect spatial patterns of a disease in a given area (Pfeiffer *et al.*, 2008). In this study, two spatial analysis techniques, Ripley's K function and $G_i^*(d)$ were used to determine clustering of BXW cases in Kagera. Ripley's K function was used to estimate the size of the spatial scale of clustering of BXW cases in each year from 2007 to 2011. Further analysis with $G_i^*(d)$ was undertaken to determine where significant spatial clustering of BXW occurred in Kagera in each year.

Ripley's K function analysis

The second-order spatial point-pattern analysis technique was used to determine the distance (d) at which clustering of the disease occurred in the Kagera landscape. Ripley's K function allows for testing of complete spatial randomness and is defined as the expected number of individuals within a distance (d) of randomly chosen individuals in a population (Gatrell *et al.*, 1996). Mathematically, the K function is defined as:

$$K(d) = \lambda^{-1}E \quad (1)$$

where λ is the intensity of the spatial process (in this study, the mean number of BXW cases reported per unit area) per year, and E refers to the edge-corrected expected value. The intensity, λ , is given by N/A where N is the number of BXW cases in the study area A . If the cases of BXW per year are randomly (Poisson) distributed, the expected value of $K(d)$ is πd^2 . $K(d)$ is linearized as $L(d)$ (Gatrell *et al.*, 1996):

$$L(d) = \sqrt{(K(d)/\pi)} \quad (2)$$

where $L(d)$ is a linear expression of the expected number of cases occurring within distance d in metres (m) of all cases. Output of the K function is represented graphically as plots of $L(d)$ versus distance d , which represents the expected cluster radius. More detailed information on the Ripley's K function method is provided in the supplementary text S1. The analyses of the K function were performed with the SPATIAL ANALYST extension in ARCGIS v. 10.2.1 and search distances of 1000 m with 20 steps (0–100) and 99 permutations, for a $P < 0.01$ confidence interval.

Getis and Ord statistic $G_i^*(d)$

The $G_i^*(d)$ statistic was used to determine the spatial clustering of events (i.e. disease cases) with respect to each other within a predefined distance threshold (Getis *et al.*, 2003). This is a

group-level statistic requiring data to be grouped. The grouping can be done within village or district boundaries (rayons) or by creating hexagons on the landscape. Therefore, a 2 km hexagonal grid surface (measured from the centre to a peripheral corner) was generated for the entire Kagera region using the GENPOINTINPOLY tool in geospatial modelling environment (GME) (Beyer, 2012). The 2 km grid was chosen to minimize ecological fallacy. This resulted in 1901 hexagons for the entire Kagera region. BXW cases per year were then spatially grouped onto the grid surface using ARCGIS v. 10.2.1. The $G_i^*(d)$ statistic was defined by Getis & Ord (1992) as:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^* \bar{x}}{s\{[(nS_i^* - W_i^{*2})/(n-1)]^{1/2}\}} \quad (3)$$

where W and w represent a weights matrix used to determine spatial structure and association among locations in a dataset, i is the cell of analysis in which all other cells (j) must fall within a distance (d) to be included, x is number of BXW cases, n is the total number of locations, S and s are the standard deviations, and \bar{x} is the mean of all BXW cases located within 2 km-hexagonal grid cells. The $G_i^*(d)$ statistics analysis was performed within ARCGIS v. 10.2.1 with fixed distance bands of 1, 5 and 10 km to examine the critical distances at which BXW cases were clustered. These distances were then mapped together as a 4-class choropleth map, with classes representing no significance (NS), 1, 5 and 10 km critical distances. Detailed information on the $G_i^*(d)$ statistics method is provided in supplementary text S1.

Environmental effects

The relationship between BXW clusters (hot spots identified by $G_i^*(d)$ statistics; Getis *et al.*, 2003) and environmental variables was examined using bivariate correlations. Environmental variables (Table S2) for Kagera, consisting of 19 bioclimatic variables (quarterly and annual temperatures and precipitation) and average altitude, were downloaded from the WorldClim-Global climate data website (www.worldclim.org; Hijmans *et al.*, 2005). Spatial resolution for all environmental rasters were approximately 1×1 km (30 arc-s). The corresponding raster values were extracted to points using the SPATIAL ANALYST tool in ARCGIS v. 10.2.1. The relationship between BXW clusters and environmental variables was examined using Pearson's correlation coefficient (2-tailed P value ≤ 0.05 significance) using SPSS v. 16.0 (SPSS Inc.).

Ecological niche modelling

MaxEnt

The MAXENT program v. 3.3.3 was used to predict the potential geographic distribution of BXW in Tanzania. The MaxEnt model requires point locations where species are known to occur (presence-only data) and environmental covariates (e.g. precipitation and temperature). MaxEnt randomly selects pseudo-absence values from a user-defined extent of the study background (those pixels in the study area without presence points). Briefly, MaxEnt uses the principle of maximum entropy to make predictions from incomplete information. MaxEnt estimates the unknown probability distribution by finding the most uniform distribution (maximum entropy) across the study area. Phillips *et al.* (2006) and Elith *et al.* (2011) provide detailed mathematical descriptions of MaxEnt.

To avoid model overfitting of spatially clustered presence points, and the inability to predict spatially independent data, the presence data were spatially rarified using the SDMTOOLBOX (Brown, 2014; Narouei-Khandan *et al.*, 2016) at a spatial resolution of 2 km (African Equidistant conic projection), which reduced the BXW presence points from 640 to 225. In MaxEnt, presence points are compared with background points, where the probability of species presence is unknown. Because the background extent can affect model predictions (Senay *et al.*, 2013), an exploratory approach was used to define the proper background extent. A 50-km buffer was defined as the appropriate distance at which the area under the receiver operating characteristic curve (AUROC) value was maximized (Narouei-Khandan *et al.*, 2016).

Highly correlated variables were determined and removed by performing pairwise Pearson's correlation tests in SDMTOOLBOX to avoid multicollinearity. Variables with a correlation >0.85 were removed, resulting in the selection of 10 bioclimatic variables (Table S2). To train the model, 75% of the presence data was used and 25% was used for model validation. The model was set to run with 5000 iterations and more than 10 000 background points with 15 replications. The model was also set to run the jackknife test, which evaluates variable importance, and to produce response curves showing the relationship between each environmental variable and the probability of disease occurrence (Phillips, 2006). The MaxEnt prediction performance was evaluated by area under the curve or AUROC (Phillips *et al.*, 2006). The AUROC plots true positive predictions (presence locations with correct predictions) versus false positive predictions (the absence locations that were incorrectly predicted as presence).

Regression modelling

Logistic regression and GIS have been widely used to predict geographic distribution of plant diseases (Shaw & Osborne, 2011; Yuen & Mila, 2015). ArcGIS is helpful in extracting the variables to run logistic regression and building maps based on the results from regression analysis. Logistic regression is a modelling technique that can be used to predict the probability of occurrence of an event as a function of the independent variables (Hosmer & Lemeshow, 2000). Logistic regression generates the model statistics and coefficients that predict a logit transformation of the probability that the dependent variable is 1 (probability of occurrence of a BXW event). Logistic regression involves fitting a dependent variable using the following equations:

$$Y = \text{logit}(p) = \ln[p/(1 - p)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n \tag{4}$$

$$p = e^y / (1 + e^y) \tag{5}$$

$$p = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n)}{1 + [\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n)]} \tag{6}$$

where *p* is the probability that the dependent variable (*Y*) is 1, β_0 = intercept, β_1 = coefficient, and *X* = predictors (environmental variables). The model was generated using a GLM logistic model through the RATTLE package in R (Williams, 2009). Predicted probabilities from the logistic regression were mapped using the raster calculator in SPATIAL ANALYST extension in ArcGIS v. 10.2.1.

The 225 spatially rarified BXW presence points were used and coded as 1. The absence points were selected randomly

across the entire landscape of Tanzania, including Kagera. The areas around presence points were removed prior to random selection of absence points, by drawing minimum convex polygons using the GENMCP tool in geospatial modelling environment (GME) resulting in five separate polygons. By using the GENRANDOMPTS tool in GME, 450 (1:2) absence points were randomly generated for the entire landscape of Tanzania, and coded those as 0. The results were the same when more absence points were used (data not shown). The presence (1) and absence (0) points of BXW were then combined using the 'append' function in the DATA MANAGEMENT tool (ArcMAP v. 10.2.1). Finally, the corresponding rasters of 10 selected environmental variables (Table S2) were extracted to points using EXTRACT MULTI-VALUES TO POINTS tool in SPATIAL ANALYST tools in ArcMAP v. 10.2.1.

The same 10 variables used in MaxEnt were selected and used in the regression analysis. For each model run, a random 75% of the data was used for training and 25% for validation. For final model selection, predictors with the largest *P* values were removed in a stepwise fashion. AUROC was used to evaluate the predictive performance of the model using measures of specificity (absence of commission error) and sensitivity (absence of omission error). The overall error (based on a confusion matrix of actual versus predicted values) and Akaike information criterion (AIC) for all the test models were also compared. The model with lower AIC and overall error and higher AUROC was considered the best model.

Results

Temporal progress of BXW

In January 2006, the first cases of BXW were reported from Kabale village in the Muleba district of Kagera, and during that year the disease was also observed in

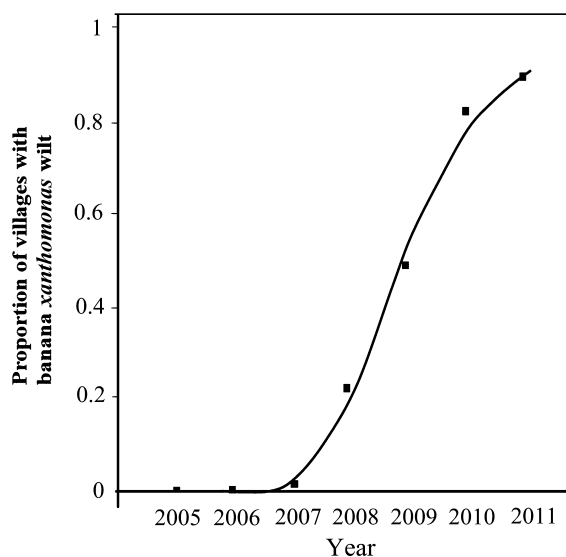


Figure 2 Disease progress curve for the proportion of surveyed villages with banana xanthomonas wilt (BXW) over the period 2005–11, in Kagera region, Tanzania. Squares and the continuous line indicate observed and estimated disease incidence, respectively, using the Gompertz model. BXW was first found in one field in one village in 2006.

Kabale B sub-village, and Bumilo and Magata villages. The Gompertz temporal model best described disease progress over time and had the highest R^2 (0.977), residual distribution close to normal and low MSE (0.0046). The Pearson correlation between the observed and predicted values based on the Gompertz model was 0.996. BXW spread rapidly from 2007 and, by 2011, disease incidence (positive villages out of all villages surveyed) approached 1 (Fig. 2).

BXW cluster analysis in Kagera region

The results of Ripley's K function analyses revealed that observed numbers of BXW cases were significantly clustered in Kagera in each year ($L(d) > d$; Fig. 3). Significant clustering of BXW cases in 2007 started at 4 km and increased with increasing distance, while in all other years clustering started below 2 km. Very strong clustering occurred between 5 and 10 km throughout the study period (Fig. 3). In 2008 and 2009, the significance of clustering dissipated at 25 km, while in 2007 and 2011 it continued to more than 30 km (Fig. 3).

Local clustering of BXW was recorded at critical distances (d_c) of 1, 5 and 10 km ($z \geq 1.96$ at $\alpha = 0.05$; Fig. 4). The total number of statistically significant clusters differed per year and per critical distance, with large numbers of clusters being recorded at 10 km and few at

1 km (Table 1; Figs S1 & S2). Localized clusters were denoted by small critical distances whereas dispersed clusters were indicated by large distances. Consistent clustering of BXW was recorded in 2008, 2009 and 2010, but by 2011, disease clusters were more spread out across the landscape, with an increased number at a smaller critical distance of 1 km compared to other years (Figs S3 & S4).

Environmental effects

Pearson's correlation analyses revealed significant relationships between environmental variables and BXW clusters (Table 2). The relationship between precipitation and BXW cluster occurrence was positive and highly significant ($P < 0.01$). Significant ($P = 0.05$) but weak negative correlations were obtained between temperature, precipitation seasonality, average altitude and BXW clusters.

Modelling potential geographic distribution of BXW in Tanzania

The training data of BXW was described well by the Max-Ent model. The AUROC value was 0.93, indicating that the model performed with high accuracy, and that the selected environmental variables explained the potential

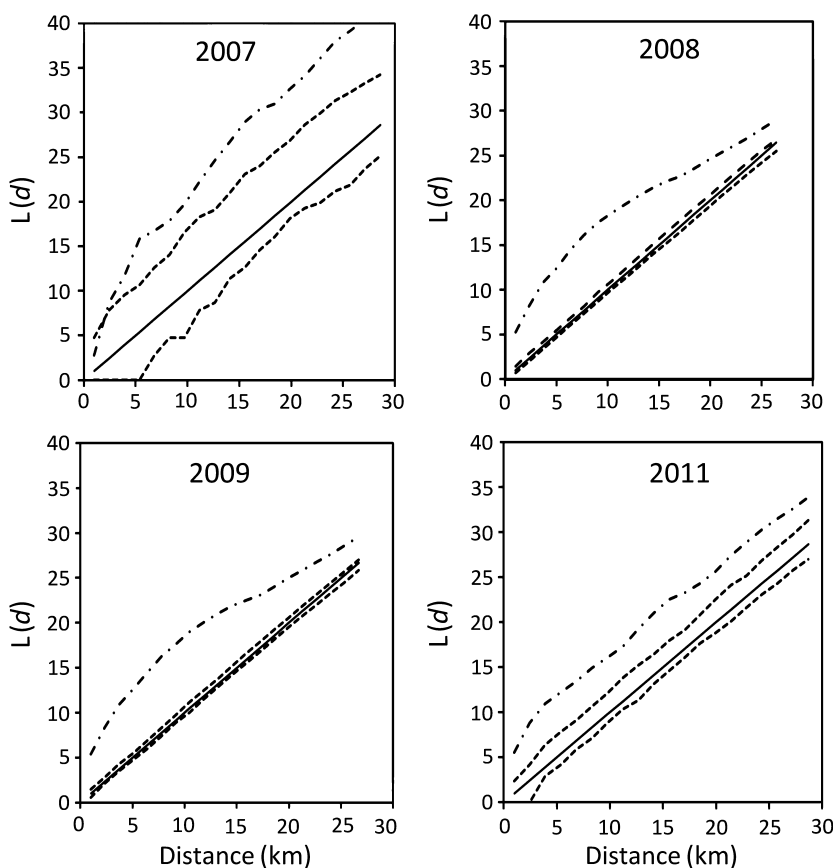


Figure 3 Ripley's K function analysis used to calculate the clustering of banana xanthomonas wilt (BXW) cases in Kagera region, Tanzania, from 2007 to 2011. The graph for 2010 (not shown) was very similar to that of 2009. The observed value of the test statistic $L(d)$ (linear expression of the expected number of cases occurring within distance d) at a given distance is marked by dash-dot lines, while expected distance is presented with a solid line. Observed $L(d)$ greater than expected ($L(d) > d$) indicates clustering. The boundaries of the 99% confidence interval on the null hypothesis of complete spatial randomness are denoted by plain dashed lines.

BXW distribution in Tanzania. The MaxEnt model predicted that all high and medium altitude banana production regions of Tanzania would be highly suitable for the establishment of BXW (Fig. 5). Logistic regression also resulted in high prediction accuracy (AUROC = 0.99), low AIC (123.5), high pseudo R^2 (0.93), and low overall error rate (0.0298), indicating good predictive performance as well as good discrimination between presence and absence points. The logistic regression model resulted in predictions very similar to those of the MaxEnt model (Fig. 6), except that the areas that would be suitable for BXW establishment were more extensive.

Five environmental variables (mean temperature of warmest quarter (bio10), precipitation of the warmest quarter (bio18), annual precipitation (bio12), annual temperature range (bio7), and precipitation in the coldest quarter (bio19)) contributed significantly to prediction by the MaxEnt model, with contributions ranging from 7% to 36% (Table 3). Variable bio10 contributed most, followed by bio 18 and bio12. According to the jackknife

test, the variable that had the highest training gain when used alone was bio18 (Fig. S5). Model gain decreased most when bio19 was omitted, implying that bio19 had the greatest impact on model performance compared to other variables (Fig S5). MaxEnt identified distinct relationships between the probability of BXW occurrence and the top three variables that contributed most to disease prediction (Fig. 7). The probability of BXW presence decreased as bio10 increased, with the highest probability predicted at temperatures between 17 and 22 °C in the warmest quarter (Fig. 7a). The probability of BXW occurrence increased with increasing precipitation in the warmest quarter up to 700 mm, but declined with higher rainfall (Fig. 7b). This probability also increased with total annual rainfall up to 1400 mm and remained unchanged with a further increase in rainfall (Fig. 7c). As with MaxEnt, logistic regression indicated that bio18, bio19, bio7 and bio2 influenced BXW occurrence (Tables 3 & 4). Altitude and bio15 had very minimal contribution to the performance of both models.

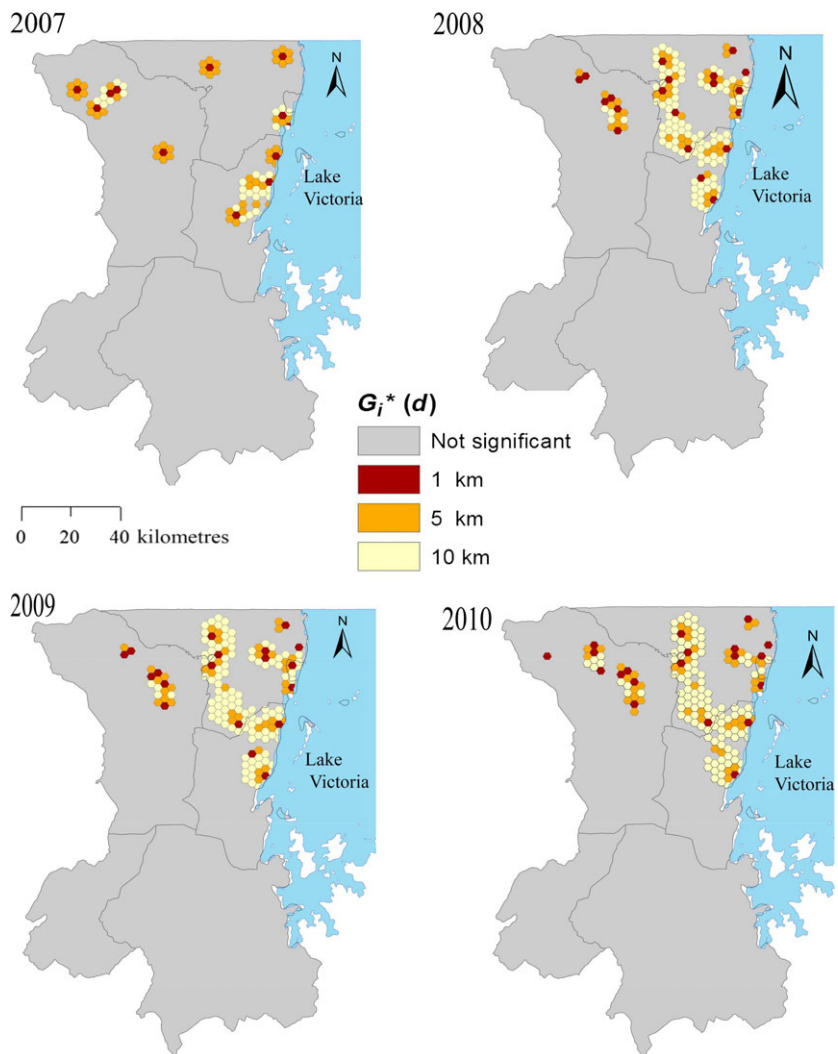


Figure 4 Spatial clusters of banana xanthomonas wilt in Kagera region using the local statistic $G_i^*(d)$ with three different threshold distances, in 2007–10. Significant clusters at distances of 1, 5 and 10 km are presented in red, gold and yellow. The $G_i^*(d)$ values that did not increase with these distances were designated as not significant. The cluster distribution in 2011 was not comparable to those in 2007–10 because the sample size was small and concentrated in only a few wards. [Colour figure can be viewed at wileyonlinelibrary.com].

Table 1 Values of critical distance (d_c) used in Getis and Ord $G_i^*(d)$ statistics, and number of significant local clusters of banana xanthomonas wilt detected for each distance over a period of 5 years

Year	Critical distance (d_c) (km)	No. of significant clusters ^a
2007	1	12
	5	52
	10	32
2008	1	20
	5	44
	10	100
2009	1	19
	5	42
	10	107
2010	1	20
	5	47
	10	110
2011	1	27
	5	53
	10	68

^a $z \geq 1.96$ at $\alpha = 0.05$.

Table 2 Bivariate correlations between banana xanthomonas wilt clusters and environmental variables in Kagera region, Tanzania

Code	Variable	Pearson correlation (r)
Bio18	Precipitation of warmest quarter	0.344**
Bio13	Precipitation of wettest month	0.327**
Bio12	Annual precipitation	0.325**
Bio16	Precipitation of wettest quarter	0.309**
Bio14	Precipitation of driest month	0.306**
Bio17	Precipitation of driest quarter	0.301**
Bio19	Precipitation of coldest quarter	0.045*
Bio15	Precipitation seasonality (coefficient of variation)	-0.119**
Bio4	Temperature seasonality (standard deviation \times 100)	-0.175**
Bio5	Maximum temperature of warmest month	-0.117**
Bio7	Temperature annual range	-0.116**
Bio3	Isothermality (bio2/bio7) (\times 100)	-0.076**
Bio6	Minimum temperature of coldest month	-0.072**
—	Average altitude	-0.064**
Bio9	Mean temperature of driest quarter	-0.043**
Bio10	Mean temperature of warmest quarter	-0.042**
Bio11	Mean temperature of coldest quarter	-0.036**
Bio1	Annual mean temperature	-0.025*

**Correlation significant at $\alpha = 0.01$ (2-tailed); *correlation significant at $\alpha = 0.05$ (2-tailed).

Discussion

Banana xanthomonas wilt spread rapidly in Kagera after it was first reported in 2006, and within 5 years its incidence approached 100%. Banana plants affected by BXW die and thus, the disease rapidly reduces yields and the area planted to the crop (Nkuba *et al.*, 2015). In an attempt to reduce the spread of BXW, the local government in Kagera

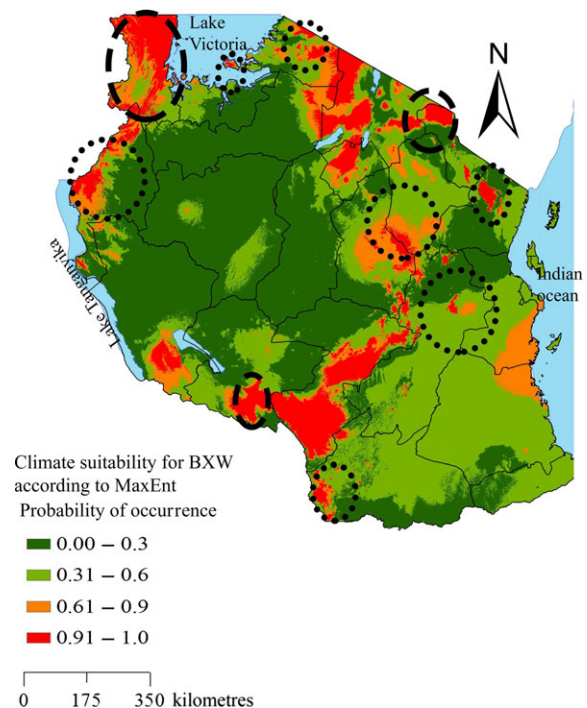


Figure 5 The potential distribution of banana xanthomonas wilt in Tanzania as predicted by the MaxEnt model (warmer colour indicates higher suitability). Bold dashed circles are major banana producing regions (Kagera in the northwest, Kilimanjaro in the northeast, and Mbeya in the southern highland of Tanzania); dotted circles are the areas with medium banana production.

initiated an intervention programme in 2013 that forced removal of affected plants (Shimwela *et al.*, 2016). Similar intervention measures were implemented in Uganda in 2005 (Tushemereirwe *et al.*, 2006), but the disease continued to spread due to a variety of factors (Shimwela *et al.*, 2016); nonetheless, epidemic development eventually slowed in some areas in Uganda (Tushemereirwe *et al.*, 2006; Kubiriba *et al.*, 2012).

The Ripley's K plots indicated that BXW was spatially clustered in Kagera from 2007 to 2011. The $G_i^*(d)$ statistic identified localized clusters in the survey data, indicating that there was both local spread, possibly by insects, rain splash or contaminated planting materials and tools, and regional spread, possibly by further distribution of contaminated planting materials and banana bunches (Nakato *et al.*, 2013; Shimwela *et al.*, 2016). The clusters, also called hotspots, were concentrated in Muleba, Karagwe, Misenyi and Bukoba rural districts across all years, and seemed to disappear in some years and reappear in others. This could be due to successful removal of infected plants in some villages followed by a resurgence of the disease after a few months. Kubiriba *et al.* (2012) also reported a resurgence of BXW in Uganda, probably due to reintroduction of infected planting materials and the long incubation period up to 36 months (Ocimati *et al.*, 2013).

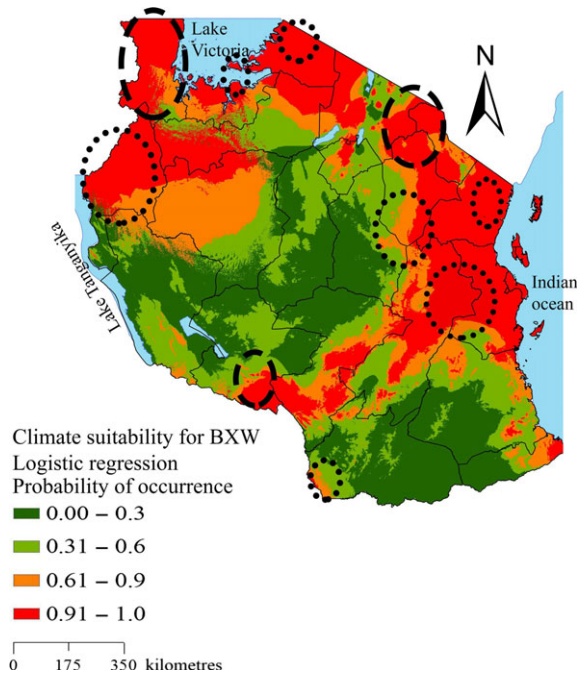


Figure 6 The potential distribution of banana xanthomonas wilt in Tanzania as predicted by the logistic model (warmer colour indicates higher suitability). Bold dashed circles are major banana producing regions (Kagera in the northwest, Kilimanjaro in the northeast, and Mbeya in the southern highland of Tanzania); dotted circles are the areas with medium banana production.

Another important finding of this study was the strong positive association of BXW occurrence with rainfall, especially in the warmest quarter (MaxEnt and logistic regression), but also in the coldest quarter (logistic regression). A positive correlation between BXW and rainfall had been reported earlier (Bouwmeester *et al.*, 2010; Tripathi *et al.*, 2009). Rain splash and wind-driven rain play an important role in the short- and long-distance dispersal of many plant-pathogenic bacteria such as *Xanthomonas axonopodis* pv. *citri*, the causal agent of citrus canker (Gottwald *et al.*, 2002). Thus, rainfall might also be responsible for the primary short-distance spread of BXW in Kagera, which would be consistent with the concentrated BXW clusters that were observed in this region.

In Kagera, BXW local spatial clusters were negatively correlated with temperature and average altitude of the BXW clusters. The negative correlation between temperature and BXW was confirmed in MaxEnt for the warmest quarter. Altitude was not selected as a contributing factor in MaxEnt although it was an important factor in logistic regression. Negative correlations between temperature and altitude and BXW were also found by Shimelesh *et al.* (2008) and Bouwmeester *et al.* (2010). The negative relationship between BXW and altitude was attributed to the reduced activity of insect vectors at higher altitudes (Shimelesh *et al.*, 2008; Tripathi *et al.*, 2009). In Kagera, the negative correlation between

Table 3 Percentage contribution (out of 5000 iterations) of selected independent environmental variables used in MaxEnt to model banana xanthomonas wilt distribution in Tanzania

Code	Variable	Unit	Contribution (%)
Bio10	Mean temperature of warmest quarter	°C	36.1
Bio18	Precipitation of warmest quarter	mm per 3 months	23.3
Bio12	Annual precipitation	mm per year	19.1
Bio7	Temperature annual range	°C	7.1
Bio19	Precipitation of coldest quarter	mm per 3 months	7.0
Bio4	Temperature seasonality	%	4.4
Bio2	Mean diurnal range	°C	1.9
Bio15	Precipitation seasonality	—	0.6
—	Average altitude	m	0.4
Bio14	Precipitation of driest month	mm per month	0.0

altitude and BXW incidence may be attributed to the higher rainfall at lower altitudes close to the shore of Lake Victoria. This explains the seemingly contradictory result that both temperature and altitude were negatively correlated with BXW occurrence, even though temperatures commonly decline with altitude. This may indicate that rainfall contributes more to the spread of BXW than insect vectors.

Banana xanthomonas wilt occurs throughout Kagera, and, according to recent informal reports, is also present in Kigoma and Mara. These locations were not included among the presence points, but were predicted to be areas conducive for BXW, and, thus, could be considered positive validation points. However, as far as the authors are aware, BXW is not present in other areas of Tanzania and this is the first study to predict the potential development of this disease elsewhere in Tanzania. According to MaxEnt and logistic regression, all areas with high and medium banana production (including Kagera, Kilimanjaro, Mbeya, Tanga, Kigoma, Mara, Morogoro and Arusha) are predicted to be highly suitable for BXW establishment, although one needs to realize that this prediction is based on extrapolation, a limitation of all correlative species distribution models. The probability of BXW establishment in the central part of Tanzania is predicted to be lower than that in the other areas of the country due to the dry climate in that area and the less extensive banana production than in other parts of Tanzania.

Václavík & Meentemeyer (2009) recommended using more than one modelling approach when predicting the spread of a species in a particular area because different models might predict different outcomes. For example, the BXW-conducive areas that were predicted in the present study by logistic regression were more extensive than those predicted by MaxEnt. MaxEnt is known to produce more conservative predictions than other niche models (Narouei-Khandan *et al.*, 2016).

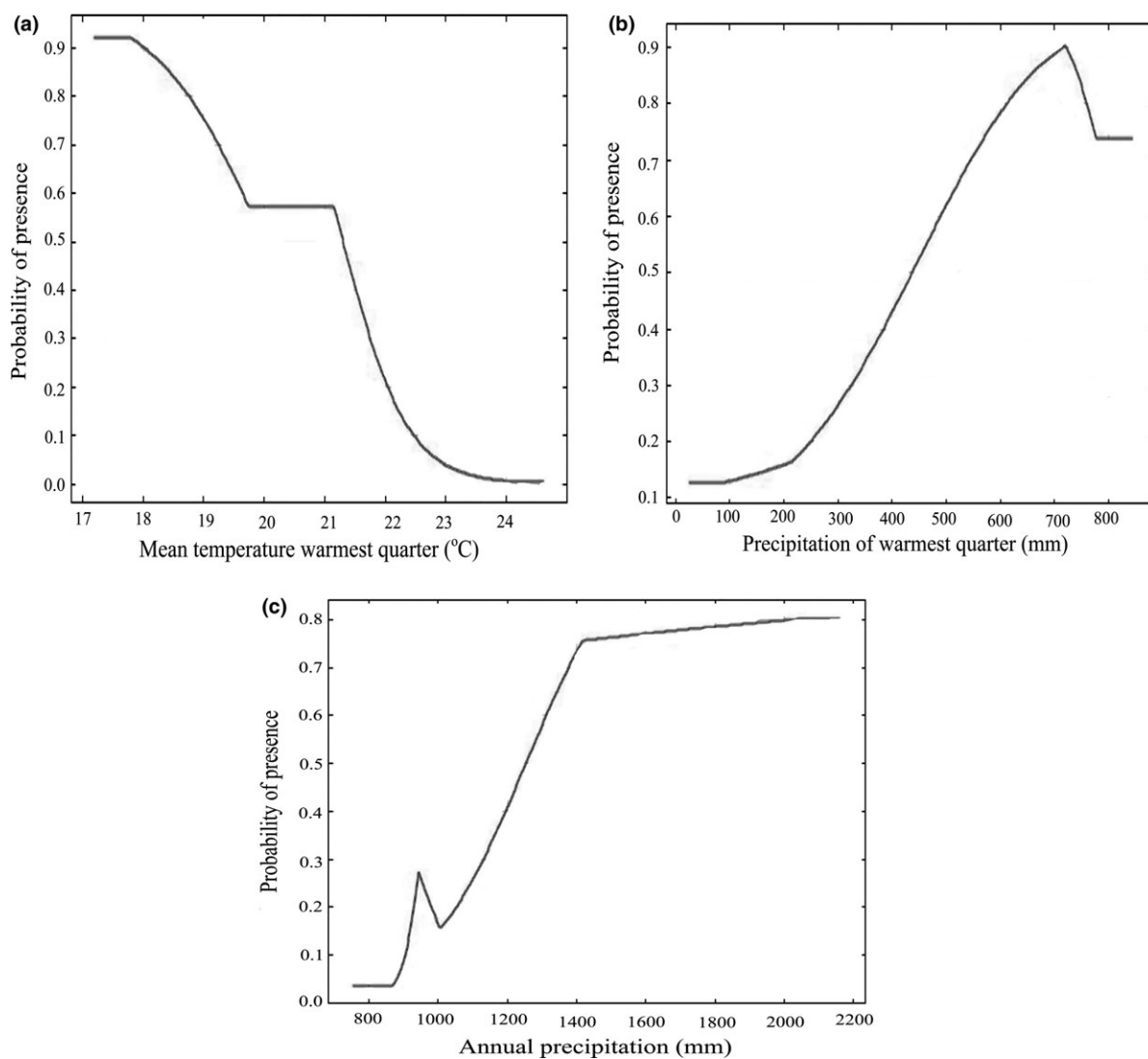


Figure 7 Response curves of three variables, (a) mean temperature in the warmest quarter (°C), (b) precipitation in the warmest quarter (mm), and (c) annual precipitation (mm), contributing most to the predictions of the occurrence of banana xanthomonas wilt (BXW) in Tanzania by the MaxEnt model.

The current absence of BXW in disease-conducive areas like Kilimanjaro, Mbeya, Tanga, Morogoro and Arusha is probably due to the dry climate and the long distance (approximately 1000 km) between these areas and the Kagera, Kigoma and Mara regions. Dispersal of Xcm by natural means (rain, wind and vectors) is limited to relatively short distances. However, long-distance spread is possible via the movement of infected planting materials, banana bunches, and possibly leaves used as packing material (Nakato *et al.*, 2013). As the climatic conditions are favourable for BXW establishment in the unaffected areas in Tanzania, it is very important to control the movement of banana materials into these areas. Recent improvements of the roads that connect Kagera with the rest of the country may increase the likelihood of spread of BXW from Kagera to other areas in Tanzania with suitable conditions for disease establishment.

Table 4 Model summary from stepwise logistic regression used to predict the presence/absence of banana xanthomonas wilt in Tanzania

Code	Variable	Parameter estimate		
		β -coefficient	Standard error	<i>P</i>
Bio18	Precipitation in warmest quarter	0.0151	0.003	<0.0001
Bio4	Temperature seasonality	-2.4957	0.583	<0.0001
—	Average altitude	0.0062	0.002	0.0004
Bio7	Temperature annual range	3.9678	1.206	0.0010
Bio2	Mean diurnal range	-4.3442	1.486	0.0034
Bio15	Precipitation seasonality	-0.1501	0.064	0.0188
Bio19	Precipitation in coldest quarter	0.0082	0.004	0.0218
	Constant	1.2745	6.155	0.8359

Disease maps that delineated the potential distribution of BXW in Tanzania in the present study were generated partially by extrapolation (by considering areas outside of Kagera) and thus, these maps may need to be viewed with some caution. Nevertheless, the maps could be used to develop better-directed surveillance systems for the disease in Tanzania. Furthermore, these maps could help policymakers establish regulatory and eradication policies for BXW in Tanzania, and assist the extension officers and national crop protection departments to decide where to enact control strategies. Based on extensive BXW surveys, Uganda was divided into three epidemic zones with different disease intensities (Tushemereirwe *et al.*, 2006). A strategy that targeted hotspot areas for intensive control measures was, at least temporarily, successful in that country (Tushemereirwe *et al.*, 2006).

Despite eradication efforts that were established in 2013, BXW spread rapidly over short distances, as indicated by cluster analysis (Shimwela *et al.*, 2016). This was probably due to the exchange of infected planting materials among neighbouring farmers and the use of farm tools without disinfestation, especially by traders during harvesting. It is also possible that Xcm cells in the ooze from cut stem surfaces were splash dispersed during the rainy season. Rainfall in the warm season was positively correlated with the probability of BXW spread, and it might be advisable to limit the cutting of diseased stems to dry and hot periods. Seasonality and management of the BXW epidemic in Kagera region were addressed in a parallel paper (Shimwela *et al.*, 2016).

In conclusion, BXW is strongly clustered over short distances in Kagera. Hence, hotspot-targeted interventions are recommended for disease management in the region. The authors propose that BXW management could be improved by implementing cultural practices during high temperatures and low precipitation conditions that do not promote disease development (Shimwela *et al.*, 2016). Finally, because climatic conditions are highly conducive for BXW establishment in other, unaffected banana-growing regions in Tanzania, strict quarantine measures and routine surveillance for the disease should occur in these areas. The BXW experience in Kagera highlights the importance of keeping these areas free of this disease.

Acknowledgements

This work was made possible by the generous support of the American people through the United States Agency for International Development (USAID)-funded Innovative Agricultural Research Initiative project (iAGRI; award no. CA-621-A-00-11-00009-00). The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of the US Agency for International Development or the United States Government. The authors thank Walter Bowen for facilitation of the funding and for his general guidance. They are grateful to the Esther B. O’Keeffe Foundation for contributing

additional funding to this research. The authors also acknowledge the contributions by Mgenzi Byabachwezi, Jerome Kubiriba, Justine Mchunguzi, Shaban Mkulila, Innocent Ndyetabura and Bulili Sayi for providing survey data and other valuable information from Tanzania.

References

- Abele S, Twine E, Legg C, 2007. *Food Security in Eastern Africa and the Greater Lake Region*. Ibadan, Nigeria: International Institute for Tropical Agriculture (IITA).
- Adikini S, Beed F, Tusiime G *et al.*, 2013. Spread of *Xanthomonas campestris* pv. *musacearum* in banana plants: implications for management of banana Xanthomonas wilt disease. *Canadian Journal of Plant Pathology* 35, 458–68.
- Beyer HL, 2012. Geospatial Modelling Environment (version 0.7.2.1). [<http://www.spataleecology.com/gme>]. Accessed 5 October 2016.
- Blomme G, Jacobsen K, Ocimati W *et al.*, 2014. Fine-tuning banana Xanthomonas wilt control options over the past decade in East and Central Africa. *European Journal of Plant Pathology* 139, 271–87.
- Bouwmeester H, Abele S, Manyong VM *et al.*, 2010. The potential benefits of GIS techniques in disease and pest control: an example based on a regional project in Central Africa. *Acta Horticulturae* 879, 333–40.
- Brown JL, 2014. SDMToolbox: a python-based GIS toolkit for landscape genetic, biogeography, and species distribution model analyses. *Methods in Ecology and Evolution* 5, 694–700.
- Carter BA, Reeder R, Mgenzi SR *et al.*, 2010. Identification of *Xanthomonas vasicola* (formerly *X. campestris* pv. *musacearum*), causative organism of banana xanthomonas wilt, in Tanzania, Kenya and Burundi. *Plant Pathology* 59, 403.
- Castellani E, 1939. Su un marciume dell’ ensete. *L’Agricoltura Coloniale, Italy* 33, 297–300.
- Cellier G, Remenant B, Chiroleu F, Lefeuvre P, Prior P, 2012. Phylogeny and population structure of brown rot- and Moko disease-causing strains of *Ralstonia solanacearum* phylotype II. *Applied and Environmental Microbiology* 78, 2367–75.
- Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ, 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17, 43–57.
- Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS, 1996. Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers* 21, 256–74.
- Getis A, Ord J, 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24, 189–206.
- Getis A, Morrison AC, Gray K, Scott TW, 2003. Characteristics of the spatial pattern of the dengue vector, *Aedes aegypti*, in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene* 69, 494–505.
- Gottwald TR, Sun X, Riley T, Graham JH, Ferrandino F, Taylor EL, 2002. Georeferenced spatiotemporal analysis of the urban citrus canker epidemic in Florida. *Phytopathology* 92, 361–77.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–78.
- Hosmer DW, Lemeshow S, 2000. *Applied Logistic Regression*. 2nd edn. New York, NY, USA: Wiley.
- Kubiriba J, Bagamba F, Rockfeller E, Tushemereirwe WK, 2012. The changing spread dynamics of banana Xanthomonas wilt (BXW) in Uganda. *Uganda Journal of Agricultural Sciences* 13, 53–60.
- Madden VL, Hughes G, van den Bosch F, 2007. *The Study of Plant Disease Epidemics*. St Paul, MN, USA: APS.
- Nakato VG, Beed FD, Ramathani I, Kubiriba J, Rwomushana I, Opio F, 2013. Risk of banana Xanthomonas wilt spread through trade. *Journal of Crop Protection* 2, 151–61.
- Narouei-Khandan HA, Halbert SE, Worner SP, VanBruggen AHC, 2016. Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling

- approaches, with emphasis on the USA. *European Journal of Plant Pathology* **144**, 655–70.
- Nkuba J, Tinzaara W, Night G *et al.*, 2015. Adverse impact of banana Xanthomonas wilt on farmers' livelihoods in Eastern and Central Africa. *African Journal of Plant Science* **9**, 279–86.
- Ocimati W, Ssekiwoko F, Karamura E, Tinzaara W, Eden-Green S, Blomme G, 2013. Systemicity of *Xanthomonas campestris* pv. *musacearum* and time to disease expression after inflorescence infection in East African highland and Pisang Awak bananas in Uganda. *Plant Pathology* **62**, 777–85.
- Ormsby T, Napoleon E, Burke R, Feaster L, Groessl C, Bowden L, 2010. *Getting to Know ArcGIS Desktop: Updated for ArcGIS 10*. 2nd edn. Redlands, CA, USA: ESRI Press.
- Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements ACA, 2008. *Spatial Analysis in Epidemiology*. Oxford, UK: Oxford University Press.
- Phillips S, 2006. A brief tutorial on MaxEnt. [http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc]. Accessed 5 October 2016.
- Phillips SJ, Anderson RP, Schapire RE, 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–59.
- Rugalema G, Mathieson K, 2009. *Disease, Vulnerability and Livelihoods on the Tanzania-Uganda Interface Ecosystem to the West of Lake Victoria*. Rome, Italy: FAO.
- Senay SD, Worner SP, Ikeda T, 2013. Novel three-step pseudo-absence selection technique for improved species distribution modelling. *PLoS ONE* **8**, e71218.
- Shaw MW, Osborne TM, 2011. Geographic distribution of plant pathogens in response to climate change. *Plant Pathology* **60**, 31–43.
- Shimelash D, Alemu T, Addis T, Turyagyenda FL, Blomme G, 2008. BXW in Ethiopia: occurrence and insect vector transmission. *African Crop Science Journal* **16**, 75–87.
- Shimwela MM, Ploetz RC, Beed FD *et al.*, 2016. Banana xanthomonas wilt continues to spread in Tanzania despite an intensive symptomatic plant removal campaign: an impending socio-economic and ecological disaster. *Food Security* **8**, 939–51.
- Thwaites R, Eden-Green S, Black R, 2000. Diseases caused by bacteria. In: Jones DR, ed. *Diseases of Banana, Abacá and Ensete*. Wallingford, UK: CABI Publishing, 213–39.
- Tripathi L, Mwangi M, Abele S, Aritua V, Tushemereirwe WK, Bandyopadhyay R, 2009. Xanthomonas wilt: a threat to banana production in East and Central Africa. *Plant Disease* **93**, 440–51.
- Tushemereirwe W, Kangire A, Ssekiwoko F *et al.*, 2004. First report of *Xanthomonas campestris* pv. *musacearum* on banana in Uganda. *Plant Pathology* **53**, 802.
- Tushemereirwe WK, Okaasai O, Kubiriba J *et al.*, 2006. Status of banana Xanthomonas wilt in Uganda. *African Crop Science Journal* **14**, 73–82.
- Václavík T, Meentemeyer RK, 2009. Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling* **220**, 3248–58.
- Williams GJ, 2009. RATTLE: a data mining GUI for R. *R Journal* **1**, 45–55.
- Yirgou D, Bradbury JF, 1968. Bacteria wilt of enset (*Ensete ventricosum*) incited by *Xanthomonas musacearum* sp. n. *Phytopathology* **58**, 111–2.
- Yirgou D, Bradbury JF, 1974. A note on wilt of banana cause by the enset wilt organism *Xanthomonas musacearum*. *East African Agriculture and Forest Journal* **40**, 111–4.
- Yuen J, Mila A, 2015. Landscape-scale disease risk quantification and prediction. *Annual Review of Phytopathology* **53**, 471–84.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Figure S1. Spatial distribution of hotspots of banana xanthomonas wilt at distances 1, 5 and 10 km for 2007 and 2008, as determined by Getis and Ord $G_i^*(d)$ statistics.

Figure S2. Spatial distribution of hotspots of banana xanthomonas wilt at distances 1, 5 and 10 km for 2009 and 2010, as determined by Getis and Ord $G_i^*(d)$ statistics.

Figure S3. Spatial distribution of hotspots of banana xanthomonas wilt at distances 1, 5 and 10 km for 2011, as determined by Getis and Ord $G_i^*(d)$ statistics.

Figure S4. Spatial clusters of banana xanthomonas wilt in Kagera region in 2011, using the local statistic $G_i^*(d)$ with three different threshold distances. Significant clusters at distances of 1, 5 and 10 km are presented in red, gold and yellow, respectively. The $G_i^*(d)$ values that did not increase with these distances were designated as not significant. The cluster distribution in 2011 was not comparable to those in 2007–10 because the sample size was small and concentrated in only a few wards.

Figure S5. The results of the jackknife test of variable importance. The environmental variable with highest gain when used alone was bio18, which therefore appeared to have the most useful information by itself. The environmental variable that decreased the gain most when it was omitted was bio19, which, therefore, appears to have the most information that is not present in the others variables.

Table S1. Summary of the sources of data of the banana xanthomonas wilt cases used in this study.

Table S2. Variables used for georeferenced logistic regression and MaxEnt distribution modelling, including average altitude and 19 climatic variables downloaded from the WorldClim-Global climate data website (www.worldclim.org).

Text S1. Methodology for Ripley's K function and Getis and Ord $G_i^*(d)$.