Climatic and spatial factors associated with citrus black spot. A Bayesian analysis of disease spread in South Africa

Joaquín Martínez-Minaya[1], *David Conesa*[2], *Antonio López-Quílez*[2], *Antonio Vicent*[1]

[1] Centro de Protección Vegetal y Biotecnología, Instituto Valenciano de Investigaciones Agrarias (IVIA) (Spain),

[2]Departament d'Estadística i Investigació Operativa, Universitat de València (Spain).

Abstract

Citrus black spot (CBS) is the main fungal disease of citrus worldwide. The Mediterranean Basin is still free of the disease. Bayesian spatial models were fitted in three stages of CBS epidemic in South Africa to improve the understanding of the factors influencing disease spread. All analyses were performed using INLA. Models showed better accuracy when climatic variables where introduced together with a dispersal kernel.

Keywords: species distribution models, INLA, *Phyllosticta citricarpa*, kernel.

1. Introduction

Citrus black spot (CBS) is a seriouos disease caused by the fungus *Phyllosticta citricarpa* (McAlpine) Van der Aa (syn. *Guignardia citricarpa* Kiely). The pathogen was first reported in Australia and is currently present in the main citrus-growing regions of southern and central Africa, South America and Asia. In 2010, CBS was reported in Florida (USA), the first detection in North America. The disease causes external blemishes on the rind which make the fruit unsuitable for the fresh market. Citrus-growing areas in the European Union (EU) are still free of CBS, thus phytosanitary measures are in place to avoid the entry of *P. citricarpa* (Directive 2000/29/CE).

A long-standing dispute is taking place about the appropiateness of the EU phytosanitary regulations. One of the key issues debated is the suitability of the climates in the EU citrus-growing areas for CBS stablishment and spread. Two studies conducted at global scale using the software CLIMEX indicated that the climates in the Mediterranean Basin were not conducive for CBS development [8] [11]. However, a recent CLIMEX study in the USA indicated that Mediterranean-type climate areas in California would be favourable for CBS [2]. Mechanistic (process-based) models were also used to estimate potential geographical range of CBS. One study considered the climates of the EU as suitable for the establishment of *P. citricarpa* [1], but another indicated that CBS was not expected to have an impact in areas with commercial citrus production in Europe [6].

The present study develops a spatial Bayesian analysis of CBS spread in South Africa across geographic regions in three stages of disease epidemic. Several environmental variables were considered to identify potencial asociations with disease distribution. South Africa was selected as a case study due to its climate diversity, with citrus regions covering up to ten different climate types [7].

2. Materials and methods

Data. Locations and dates where CBS was detected in South Africa from 1940 to 1950 were extracted from a previous work [10] and georeferenced. A raster layer (299 x 259 pixels) of CBS distribution in South Africa was generated from the original map [8] (Fig. 1). Environmental data from South Africa were acquired from the WorldClim database [5], which reports gridded mean values from the 1950-2000 period. A resolution of 5' (arc min) was used in all datasets.

Modeling approach. The presence and absence of CBS was modelled separately in 1945, 1950 and 2014 as a function of environmental variables and proximity to all known affected grid cells. Generalized linear models (GLM) with a link $logit$ were used to estimate the probability that a grid cell became infected. The predictor consisted of a linear function of the environmental variables and a spatial term. We tried to find out this spatial term using a geostatistical approach, but it did not performe adequately due to the strong spatial trend of CBS distribution data. The spatial term was then fitted using a nonlinear function and several dispersal kernels and bandwidths were evaluated [4]: In particular, if Y_i represented a variable indicating whether the site i was infected or not, it was modeled by means of:

$$Y_i \sim Ber(\pi_i), \quad \eta_i = logit(\pi_i) = \beta_0 + \sum_{j=1}^{m} \beta_j X_{ij} + \beta_{m+1} \sum_{k=1}^{N} K(d_{ik}, a) \ , \tag{1}$$

where β_0 was the intercept, β_j, $j = 1, \ldots, m$, were the scaling coefficients for the environmental factors, β_{m+1}, was the scaling coefficient for the spatial term, k refered to the known CBS-affected grid cells, and N was the total number of CBS-affected grid cells; d_{ik} was the euclidean distance between each CBS-affected grid cell k and target cell i. $K(d_{ik}, a)$, was the dispersal kernel, as the shape of the dispersal-by-distance function; and a, the bandwidth, was used to modify the form of the dispersal kernel, with low values of a indicating very limited long-distance dispersal and high values of a indicating a relatively high probability of long-distance dispersal.

Bayesian inference. Once the models were determined, parameters values were estimated by Bayesian inference. Parameters were treated as random variables, and prior distributions were assigned. Vague Gaussian distributions were used for the parameters involved in the fixed effects $\beta_j \sim N(0, 0.01)$, $j = 1, \ldots, m+1$, $\beta_0 \sim N(0, 0.00001)$, in order to allow empirically derived distributions, which permitted to fit the models using INLA [9]. Models were selected based on the Deviance Information Criterion (DIC) [9] for the goodness-of-fit and the logarithm of the Conditional Predictive Ordinate (LCPO)[3] for the predictive quality.

3. Results and discussion

A total of 411 locations were analyzed in 1945, 420 in 1950 and 2065 locations in 2014. Bayesian models showed that the relative contribution of climatic variables was different for each year, with very few similarities. In all three years analyzed, models including climatic variables and a dispersal kernel performed better than the models including only climatic variables. The DICs for the models with climatic variables were 46.043, 91.103 and 126.283 in 1945, 1950 and 2014, respectively. The DICs for the models including climatic variables and a dispersal kernel were 19.956, 80.925 and 59.535 for 1945, 1950 and 2014, respectively. LCPO values were also lower for the models with a dispersal kernel. (Table 1). The fact that the climatic variables associated with CBS distribution in South Africa varied depending on the stage of the epidemic illustrated the weakness of the models based on climate alone. The better overall performance of the models including a dispersal kernel indicated that spatial and dispersal factors are relevant in the distribution and spread of CBS. Therefore, previous studies where the risk of CBS expansion was estimated considering only climatic variables [2] [8] [11] should be interpreted with caution.

4. Bibliography

[1] EFSA, European Food Safety Authority. (2008). Scientific opinion of the panel on plant health (PLH) on a request from the European Commission on *Guignardia citricarpa* Kiely. *The EFSA Journal, 925*, 1-108.

[2] Er, H. L., Roberts, P. D., Marois, J. J., & van Bruggen, A. H. C. (2013). Potential distribution of citrus black spot in the United States based on climatic conditions. *European Journal of Plant Pathology, 137*, 635-647.

[3] Geisser, S. (1993). *Predictive inference. Monographs on statistics and applied probability.* CRC, London.

[4] Havel, J. E., Shurin, J. B. and & Jones, J. R. (2002). Estimating dispersal from patterns of spread: Spatial and local control of lake invasions. *Ecology, 83*, 3306-3318.

[5] Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*, 1965-1978.

[6] Magarey, R., Chanelli, S., & Holtz T. (2011). *Validation study and risk assessment: Guignardia citricarpa, (citrus black spot). USDA-APHIS-PPQ-CPHST-PERAL /NCSU.*

[7] Martínez-Minaya, J., Conesa, D., López-Quílez, A., & Vicent, A. (2015). Climatic distribution of citrus black spot caused by *Phyllosticta citricarpa*. A historical analysis of disease spread in South Africa. *European Journal of Plant Pathology*. In press.

[8] Paul, I., van Jaarsveld, A. S., Korsten, L., & Hattingh, V. (2005). The potential global geographical distribution of citrus black spot caused by *Guignardia citricarpa* Kiely: likelihood of disease establishment in the European Union. *Crop Protection, 24*, 297-308.

[9] Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society, 71*, 319-392.

[10] Wager, V. A. (1952). The black spot disease of citrus in South Africa. *Science Bulletin of the Department of Agriculture of the Union of South Africa, 303*, 1-52.

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

Table 1: Models for CBS distribution in South Africa including only climatic variables and combined with a dispersal kernel at three stages of disease epidemic.

Year	Model [a]	DIC [b]	LCPO [c]
1945	$\eta_i = \beta_0 + \beta_1 b_{6_i} + \beta_2 b_{12_i} + \beta_3 b_{15_i} + \beta_4 b_{21_i}$	46.043	0.054
	$\eta_i = \beta_0 + \beta_1 b_{12_i} + \beta_2 b_{21_i} + \beta_3 \sum_{j=1}^{N} K(d_{ij}, 150)$	19.956	0.021
1950	$\eta_i = \beta_0 + \beta_1 b_{8_i} + \beta_2 b_{9_i} + \beta_3 b_{15_i} + \beta_4 b_{17_i} + \beta_5 b_{20_i}$	91.103	0.107
	$\eta_i = \beta_0 + \beta_1 b_{8_i} + \beta_2 b_{9_i} + \beta_3 b_{15_i} + \beta_4 b_{20_i} + \beta_5 \sum_{j=1}^{N} K(d_{ij}, 175)$	80.925	0.095
2014	$\eta_i = \beta_0 + \beta_1 b_{2_i} + \beta_2 b_{5_i} + \beta_3 b_{10_i} + \beta_4 b_{12_i} + \beta_5 b_{21_i}$	126.283	0.031
	$\eta_i = \beta_0 + \beta_1 b_{5_i} + \beta_2 b_{20_i} + \beta_3 \sum_{j=1}^{N} K(d_{ij}, 250)$	59.535	0.014

[a] b_2 = mean diurnal range (°C), b_5 = maximun temperature of warmest month (°C), b_6 = maximun temperature of coldest month (°C), b_8 = mean temperature of wettest quarter (°C), b_9 = mean temperature of driest quarter (°C), b_{10} = mean temperature of warmest quarter (°C), b_{12} = annual precipitation (mm), b_{15} = precipitation seasonality (mm), b_{17} = precipitation of driest quarter (mm), b_{20} = precipitation from October to January (mm), b_{21} = degree days base 10 from July to October (mm)

[b] Deviance Information Criterion

[c] Logarithm of the Conditional Predictive Ordinate

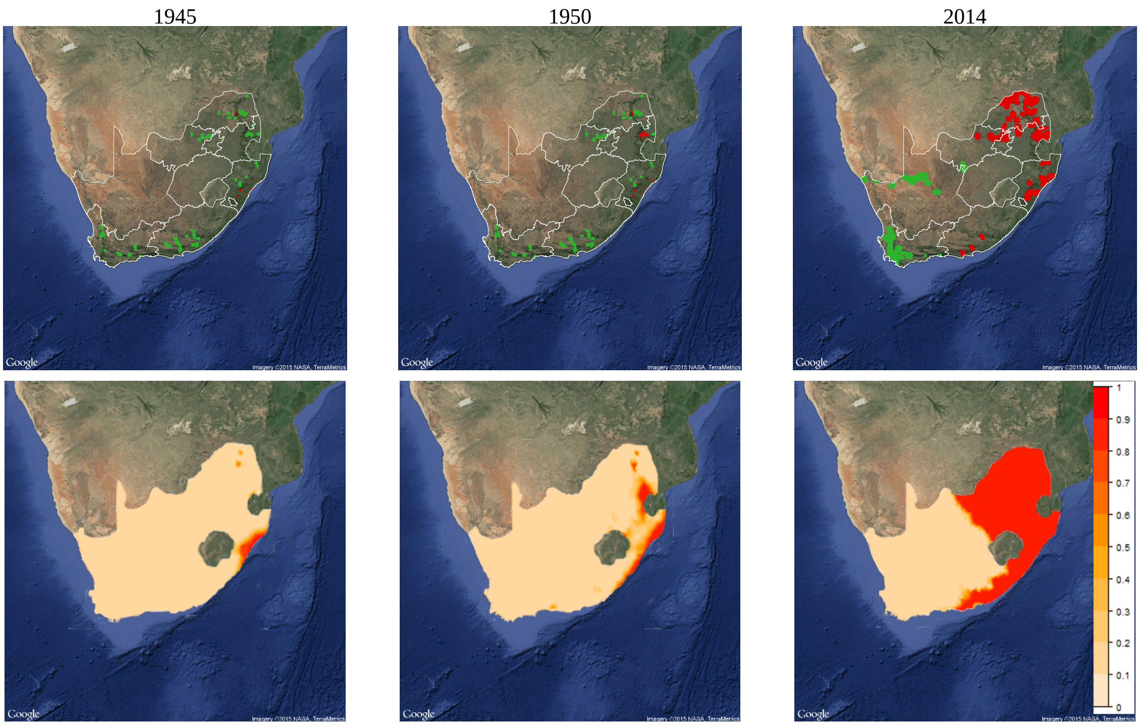

Figure 1: Top: Geographic distribution of citrus black spot (CBS) in South Africa. (●) CBS-affected citrus grid cells, (●) CBS-free citrus grid cells. Bottom: Predictions of the best models selected.