The Hutton Criteria: a classification tool for identifying high risk periods for potato late blight disease development in Great Britain

SIOBHÁN R. DANCEY, PETER SKELSEY, DAVID E. L. COOKE

The James Hutton Institute, Invergowrie, Dundee DD2 5DA Scotland
Email: siobhan.dancey@hutton.ac.uk

SUMMARY

The Smith Period has been the national warning system for potato late blight in Great Britain (GB) for approximately 60 years, and it has not been assessed or revised since its inception. We assessed the performance of the Smith Period as a forecasting tool using Receiver Operator Characteristic (ROC) analysis on historical late blight outbreak- and weather-data from 2003-2014. The Smith Period was found to be a 'fair' diagnostic tool, with significant variation in performance across the different climatic regions of Great Britain. Based on these analyses and a series of controlled environment experiments, we developed a new forecasting system for late blight in GB – the Hutton Criteria. ROC analyses revealed the Hutton Criteria to be an 'excellent' diagnostic tool with a much improved uniformity in performance across GB. The Hutton Criteria have now replaced the Smith Period as the new national warning system for late blight in GB.

KEY WORDS

Phytophthora infestans, Smith Period, Hutton Criteria, Risk Criteria, Decision Support System

INTRODUCTION

Potato late blight remains a significant challenge to potato growers, but decision support systems can be used to inform growers of the risk of disease in order to optimise the timing of fungicide treatments. Each country generally has their own tailored DSSs for potato late blight, developed to suit the needs of their growers and varying in complexity, affordability, required inputs and outputs. The Smith Period is a set of temperature and relative humidity criteria developed in Great Britain (GB) to indicate high risk periods for potato late blight development. It is defined as two consecutive days where the minimum temperature is not below 10°C and there are at least 11 hours each day with a relative humidity ≥90% (Smith 1956-1). The aim of this study was to (1) evaluate how well the Smith Period has performed on recorded outbreaks from 2003 – 2014 across GB and (2) to test other models, guided by historical and experimental research, to determine whether alternative criteria offer significant improvements on the Smith Period.

MATERIALS & METHODS

The AHDB Potatoes funded Fight Against Blight (FAB) campaign has used a network of crop scouts across GB to sample and record potato late blight occurrences since 2003. The samples have been processed and recorded at The James Hutton Institute creating a FAB database of more than 2000 outbreaks recorded from 2003 -2014. Outbreak locations are reported by their post code district to provide anonymity to the grower. The Met Office provided the corresponding daily weather data, for April to September for each year, from a network of synoptic weather stations across GB, interpolated to 652 data points. The data comprised the daily minimum temperature and number of hours of relative humidity ≥90%, and was the data used by the AHDB Potatoes funded 'Blightwatch' system in that period to send risk alerts to growers based on the occurrence of Smith Periods in their post code district.

The potato late blight outbreak data was further subdivided based on nine defined climatic regions of GB (Figure 1).
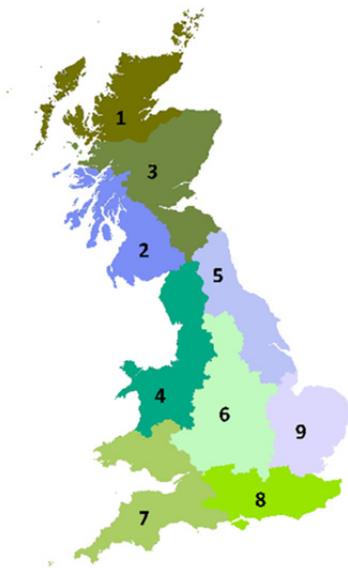

Figure 1. *Climatic regions of Great Britain used in this study; (1) Scotland North, (2) Scotland West, (3) Scotland East, (4) England North west & Wales North, (5) England North East, (6) Midlands, (7) England South West & South Wales, (8) South East England (9) East Anglia*

Five alternatives to the Smith Period (trial models) were defined based on conclusions from previous controlled environment experiments, historic literature and feedback from industry. Each consisted of two consecutive days of the temperature and relative humidity criteria being met: (1) Min temp 8°C & 11 hours RH ≥90%, (2) Min temp 10°C & 6 hours RH ≥90%, (3) Min Temp 6°C & 11 hours RH ≥90%, (4) Min Temp 8°C & 6 hours RH ≥90%, (5) Min Temp 6°C & 6 hours RH ≥90% (Smith 1956-2, Crosier 1934).

Receiver operator characteristic (ROC) curves are a means of assessing the success of binary classifiers, often used to evaluate medical diagnostic tests (Forman 2002, Fawcett 2004, Heagerty

2000). In our data sets we want to investigate the presence or absence of an 'alert' prior to an outbreak. The data is assembled to provide for each day (day = 1-28) prior to outbreaks the proportion of outbreaks (x) which have received an alert (1) and which have not received an alert (0). Using the series of unique x values for each data set as threshold values we determine a series of false positive and true positive rate pairs to plot from which the ROC curve is constructed. The further to the top left hand corner a point lands on an ROC chart the better as it has a high true positive rate and a low false positive rate, and thus the more and ROC curve pulls to the top left hand corner, the better it's performance. Points and curves falling along the diagonal line of an ROC plot are the equivalent of a diagnostic tool guessing, as the false positive rate is equal to the true positive rate. We calculated the area under the ROC curve (AUROC) and used this to quantify the performance of the alert systems (Fawcett 2004, Forman 2002). As the same outbreak data set is used for the historic analysis and the trial models it allows a comparison of the resultant AUROCs when the only factors to change are the alert criteria. The AUROC is an accepted measure for comparing ROC curves and testing for significant changes by using their values in ANOVA's (Hanley McNeil 1982, Bradley 1997). The curves were calculated in XLStat, MatLab and the resultant data was further visualized and spatially analysed in ArcGIS.

RESULTS

The Smith Period and models 1-5 have general patterns of occurrence across Great Britain which should be considered when interpreting results. The contrast between the Smith Period and model 2 is of specific note (Figure 2). Lowering the relative humidity criteria of the Smith Period produces an increase in the frequency of alerts across the central areas of GB.
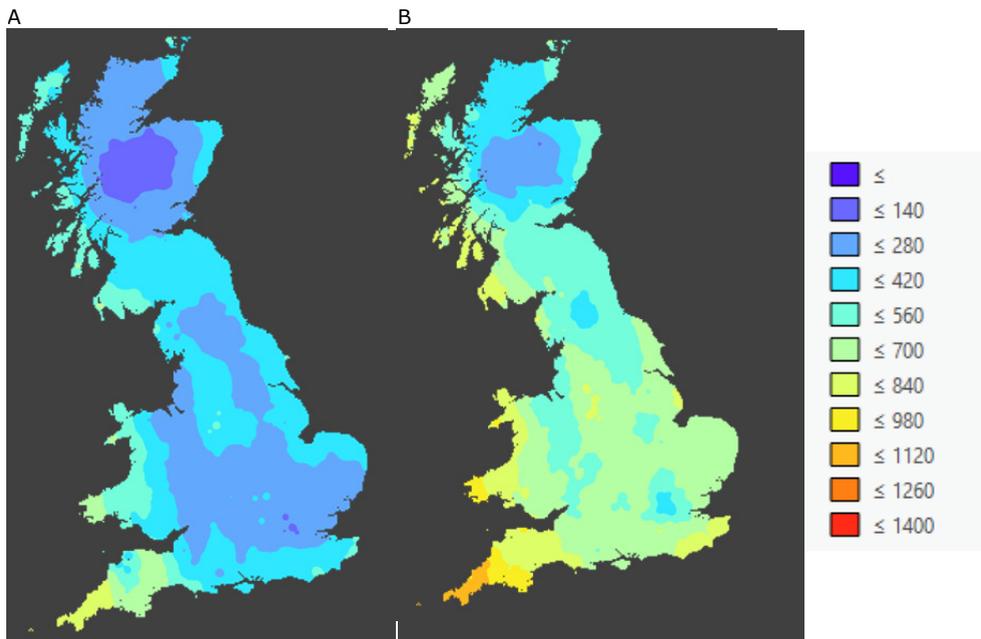

Figure 2. *Inverse distance weighted maps of the occurrence of (A) Smith Periods and (B) Model 2 across Great Britain from 2003 – 2014 from the 1st of April to the 30th of September*

ROC analysis ranks the Smith Period as a 'fair' classification tool with an AUROC of 0.686, and model 2 as an 'excellent' classification tool with an AUROC of 0.973 (Figure 3). AUROC results for the other trial models were 0.823, 0.849, 0.994, and 0.999 for models 1, 3, 4, and 5, respectively. It should be noted that the slight improvement in predictive accuracy for models 4 and 5 came at the expense of a much higher frequency of risk alerts: 7, 10, 16, 12, 24, and 30% of days prior to the reported outbreaks were classified as risk periods for the Smith Period and models 1 to 5, respectively. A comparison of the performance of the Smith Period and model 2 within the nine defined climatic regions showed a much greater spread in AUROC values for the Smith Period than for model 2 (Table 1). An ANOVA revealed highly significant differences in AUROC between regions [$F (8, 75) = 3.62$, $p = 0.001$] and between years for the Smith Period. The results for model 2 did not show highly significant differences for regions.
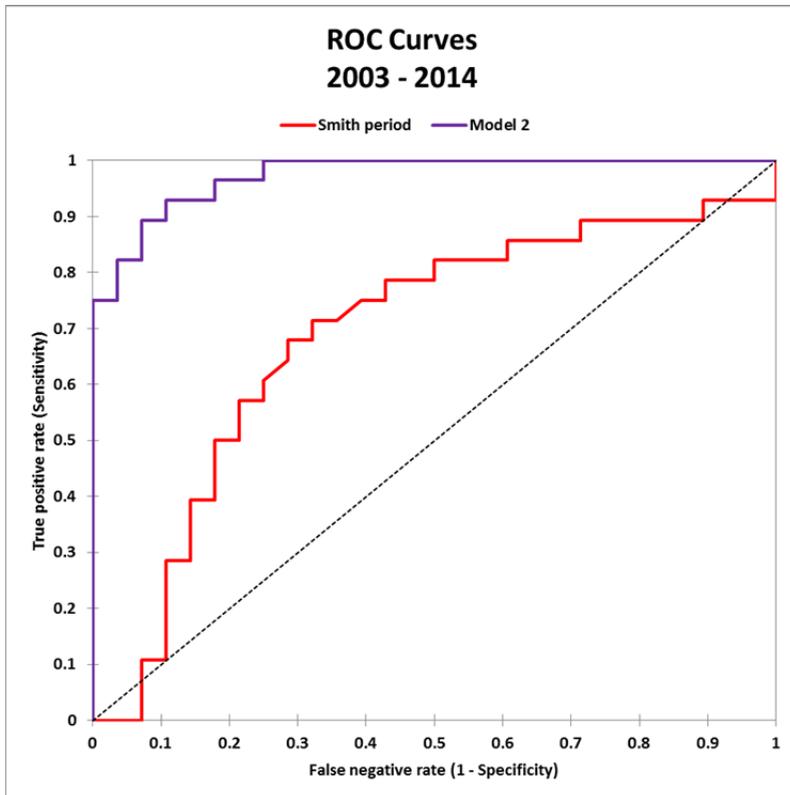

Figure 3. *Receiver operator characteristic curves for Smith Periods (red) and Model 2 (purple). These curves encompass the data for 28 days prior to >2000 reported potato late blight outbreaks from 2003 – 2014 across all of Great Britain. The Smith Period AUROC = 0.686 and Model 2 AUROC = 0.973*

Table 1. *Comparison of Area Under the ROC Curves for the Smith Period and Model 2 from 2003 – 2014 for Each Climatic Region*

	AUROC	
	Smith Period	**Model 2**
All of Great Britain	0.686	0.973
Climatic Region:		
Scotland West	0.225	0.915
Scotland East	0.740	0.946
England North East	0.684	0.976
England North West & Northern Wales	0.443	0.975
Midlands	0.628	0.987
England South West & Southern Wales	0.943	0.996
South East England	0.766	0.977
East Anglia	0.520	0.955

DISCUSSION

Our analyses revealed that the Smith Period has performed well across GB as a whole, but with a large degree of variation between climatic regions. A more uniform performance across GB would be desirable for a national warning system for late blight. It had been suspected previously from growers and evidenced in previous research (Chapman 2012), that the minimum temperature threshold was too high. Results from the trial models show an improvement of the low temperature models over the Smith Period, but not as great an improvement as was originally expected. Indeed there was no significant improvement in predictive performance between a minimum threshold of 8 or 6°C. Our experimental work provided evidence that lowering the relative humidity duration from 11 to 6 hours of ≥90% and maintaining the 10°C temperature threshold would lead to a marked improvement in predictive accuracy, and this was confirmed by the ROC analyses. Of the five trial replacement models, model 2 resulted in a significant improvement in overall predictive accuracy and uniformity in performance across the country, without a large increase in the frequency of risk alerts issued. These conclusions led us to select model 2 (the 'Hutton Criteria') as the replacement for the Smith Period. This growing season (2017) saw the Hutton Criteria implemented as the new national warning system for late blight on the AHDB Potatoes 'Blightwatch' website, and we will be assessing its performance using the above analyses together with the Fight Against Blight outbreak data for the 2017 season.

ACKNOWLEDGEMENTS

REFERENCES

Bradley, A.P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern recognition, 30(7), 1145-1159.

Chapman, A.C. (2012). The Changing *Phytophthora infestans* Population: implications for late blight epidemics and control. Dundee. The University of Dundee.

Crosier, W. (1934). Studies in the Biology of *Phytophthora infestans* (Mont) de Bar. Memoirs – Cornell University, Agricultural Experiment Station. 155(1) 3-40.

Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. Machine learning, 31(1), 1-38.

Forman, G. (2002). A method for discovering the insignificance of one's best classifier and the unlearnability of a classification task. In: Proceedings of the First International Workshop on Data Mining Lessons Learned (DMLL-2002). Available: http://www. hpl. hp. com/personal/Tom_Fawcett/DMLL-2002/Forman. pdf.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, 143(1), 29-36.

Heagerty, P.J., Lumley, T., & Pepe, M.S. (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics, 56(2), 337-344.

Smith, L.P. (1956,1).Potato Blight Forecasting By 90 Per Cent Humidity Criteria. Plant Pathology, (83 – 87).

Smith, L.P. (1956,2). Duration of High Relative Humidities. Meteorological Magazine. 85 (1), 229-232.