PHC2018/05 - Using modelling to investigate the effectiveness of national surveillance monitoring aimed at detecting a *Xylella fastidiosa* outbreak in Scotland

Steven White (CEH), James Bullock (CEH), Stephen Cavers (CEH) and Daniel Chapman (University of Stirling)

Technical Report

Overview

Xylella fastidiosa is an important bacterial plant pathogen with a wide host range, causing significant economic impact in the agricultural and horticultural trades (Saponari, Giampetruzzi, Loconsole, Boscia, & Saldarelli, 2019; Tumber, Alston, & Fuller, 2014). Once restricted to the Americas, new severe outbreaks have recently been discovered in Italy, Spain and France, and to a lesser extent in other countries (EFSA Panel on Plant Health et al., 2019). Given the ever increasing global plant trade, the likelihood of this potentially devastating plant disease being introduced to novel locations, such as Scotland, is also increasing (Chapman, Purse, Roy, & Bullock, 2017). Therefore, understanding the potential spread in novel locations is important for accurate risk assessment and mitigation strategies. The Scottish Plant Health Centre requested a preliminary exploration of this potential threat in a Scottish context, with a view of informing contingency planning, which we address here.

A Scottish *X. fastidiosa* simulation model was developed based on knowledge of the epidemiology and pattern of spread of *X. fastidiosa* subsp. pauca in olive trees in Puglia, Italy, between 2013 and 2018. The Scottish model runs over a 200m gridded spatial landscape of Scotland, and within each grid cell the land cover varies amongst assumed host cover (i.e. the percentage of susceptible host plants within a grid cell), representing spatial variability in susceptible host plant distributions based on land cover. Within each grid cell we modelled the growth of disease at a yearly scale. This model is based on a modified compartmental SAIR model, where we explicitly model the density of uninfected susceptible host plants (S), asymptomatic infected host plants (A), symptomatic infected host plants (I) and removed (i.e. dead) host plants (R) over time. We modelled the dispersal of disease using a stochastic dispersal model which

represents local and long-range dispersal of the disease (e.g. vector movement). We then deployed national surveillance and risk-based strategies until the disease is detected, and reported on the success of detection. A flow diagram representing the model is shown in Figure 1. We also performed risk and sensitivity and surveillance analyses on the model, informing the effectiveness of surveillance monitoring for an *X. fastidiosa* outbreak in Scotland.

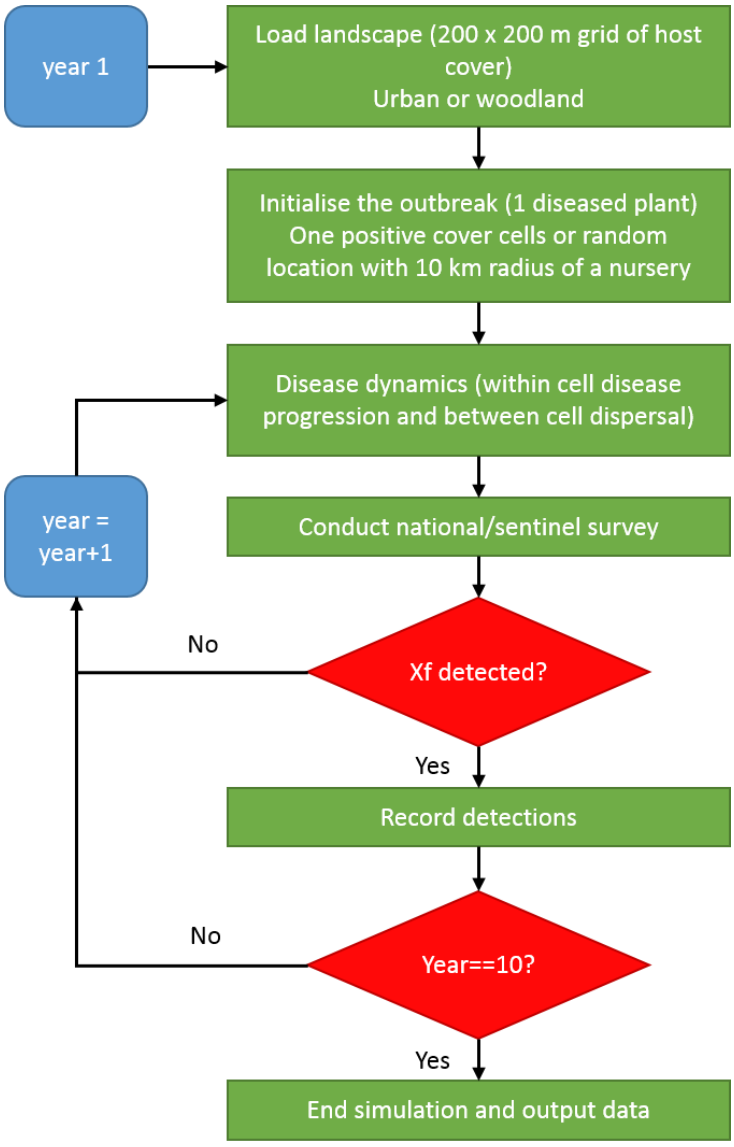

Figure 1: Flow diagram overview of the Scottish spread and surveillance model for testing surveillance measures on potential *Xylella fastidiosa* outbreaks.

Model Details

The potential spread of *Xylella fastidiosa* in Scotland and the effectiveness of surveillance measures was investigated using a spatially explicit, stochastic simulation model. The model presented here builds upon current work undertaken by the authors in the EU H2020 project XF-ACTORS (https://www.xfactorsproject.eu) and previous models (Chapman, White, Hooftman, & Bullock, 2015; White, Bullock, Hooftman, & Chapman, 2017). It should be noted that the work presented here is the first time that we have applied the Italian *Xylella* model in the UK, and the work undertaken in this study is a precursor to that being done in the UK-wide BRIGIT project (https://www.jic.ac.uk/brigit/), which will build on this study.

Previous models which are tuned for the *X. fastidiosa* outbreak in Puglia, Italy are unlikely to be directly applicable to contingency planning for *X. fastidiosa* entry into Scotland as the model is specific to one strain, and the climate and potential host plants differ from Italy, probably influencing the disease dynamics strongly. Nevertheless, these models are instructive to consider the scope of disease spread for informing contingency planning. To date, there are over 500 host plant species known (EFSA, 2018), with 4 known subspecies of *X. fastidiosa*, each with multiple sequence types (ST). The epidemiology of each ST for each host in a particular environment is currently not known. Therefore, predicting the potential spread for each of these combinations is impossible. In this work we examine two epidemiological scenarios: the worst-case scenario where the disease progresses similarly to that in Puglia, and a less severe scenario where the disease progresses at a much slower rate, modelled by reducing the transmission rate. The rationale for the slower rate being the well-known effect of temperature on the growth of *X. fastidiosa* (e.g. Feil and Purcell (2001). Since it is unknown which strain may be potentially introduced to Scotland, we do not know the host species that may be affected. In our approach we generalise the hosts into two broad classes, woodland hosts and host plants that may be found in urban areas. Work undertaken as part of the BRIGIT project (https://www.jic.ac.uk/brigit/) has shown that the highest levels of risk (denoted by the number of recorded host species in a given location) broadly overlaps with the urban and woodland cover maps (*unpublished*).

At the time of developing this model, little was known about the distribution and density of vector species in Scotland and how they may vary in different habitats and potentially contribute to the epidemiology of *X. fastidiosa*. Therefore, we made the assumption that insect vectors are evenly distributed and are implicitly modelled. However, the PHC funded project led by Kirsty Park, PHC2018/06, is filling this knowledge gap by utilising samples collected in an existing biodiversity network to identify the presence

of potential insect vectors of *Xylella fastidiosa* in the UK and the results from this project could be used in future model refinement.

We assumed that an outbreak starts by the introduction of a single asymptomatic host plant into the landscape, the location of which is either confined to a 10km radius from an importing nursery (nursery locations supplied by the Scottish PHC – the names and exact locations of these nurseries have not been disclosed in this report) or at a random location, depending on the scenario being analysed (see Surveillance section below).

Landscapes

The urban and woodland landscapes for the modelled *X. fastidiosa* outbreaks were derived from Land Cover Map 2015 (Rowland et al., 2017). The woodland landscape is calculated from the 'Broadleaf Woodland' aggregate class and the urban landscape is calculated from the 'Built-up Areas and Gardens' aggregate class. The two 200 m resolution landscape maps give the distribution of cover types likely to hold a high density of susceptible host plants. To calculate percentage cover for each cover type we take the 25m LCM2015 raster and count the cells within the 200 x 200 m cell containing the cover type and divide by the total number of 25 x 25 m cells within the 200 x 200 m. The 200 m grid resolution was chosen since it approximates the 100 m buffer radius for felling around infected and host plants specified in Decision (EU) 2015/789. The two calculated landscapes are shown in Figure 2. To generalise our approach, and for added simplicity, we derived a single landscape that combined both urban and woodland landscapes (see Figure 3).

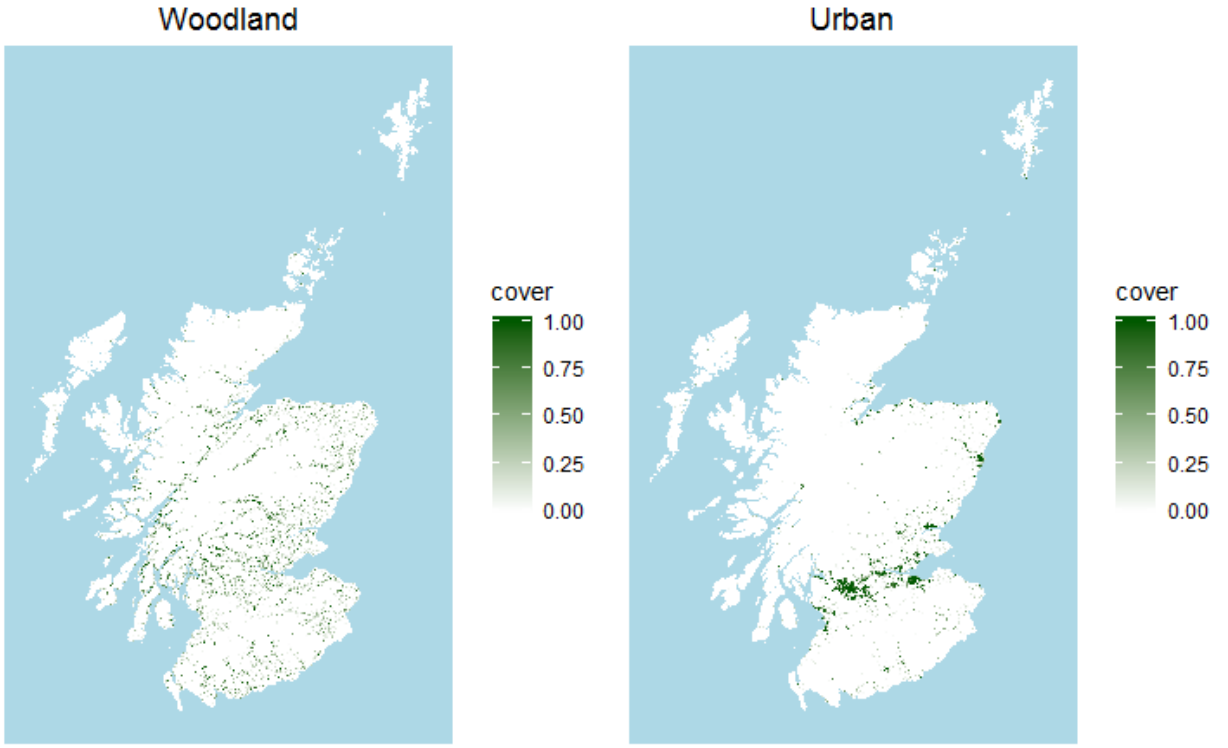

Figure 2: Land cover maps showing the distributions of susceptible host cover for broadleaf woodland and urban environments.

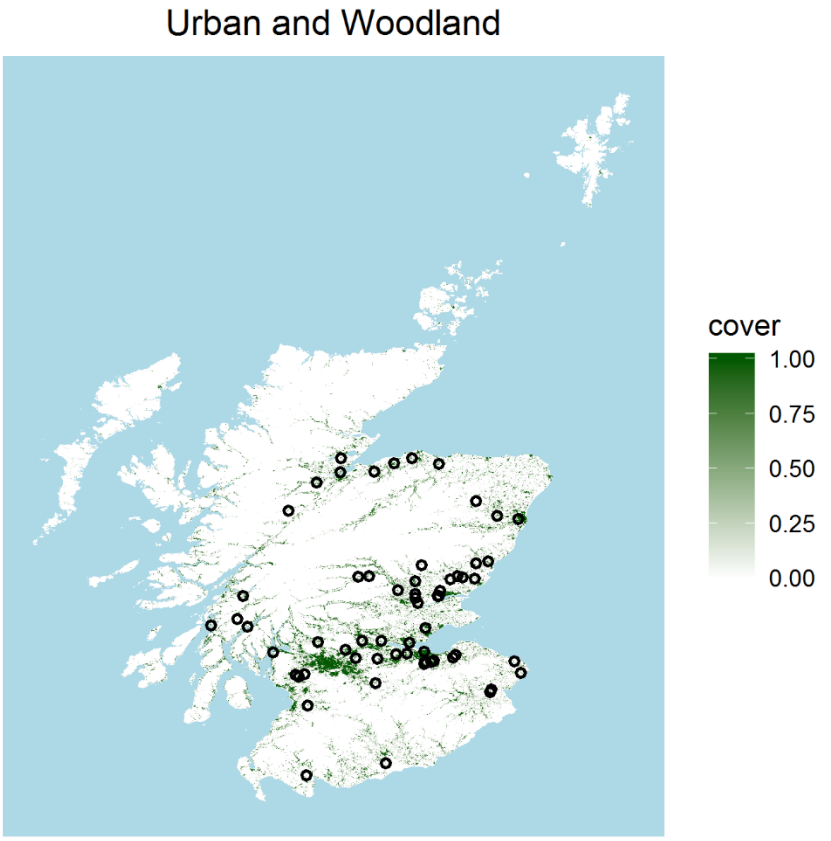

Figure 3: Combined urban and woodland cover. Black circles denote the approximate locations of the nurseries used in the analysis below (data supplied by Scottish PHC).

Temporal and Spatial Disease Dynamics

X. fastidiosa dynamics in infected grid cells were approximated with a standard deterministic compartmental model (Allen, Brauer, Van den Driessche, & Wu, 2008) in discrete annual time steps. The modelled host plant population was divided into four compartments – uninfected susceptible host plants (*S*), infected asymptomatic plants (*A*), infected symptomatic plants (*I*) and infected plants that have been killed or otherwise removed from the infective population (*R*) (Figure 4). The *S* proportion of the population moves into the *A* compartment at a rate determined by the current proportion of infective plants (*A* and *I*, with *A* assumed to be less infective than *I* due to lower bacterial concentrations). Once infected, plants then progress through the *A* and *S* compartments at a constant rate, finally moving into the *R* compartment. For the worst-case scenario, transmission and progression parameters were

estimated through a combination of expert knowledge about times for symptom development and canopy cover loss in olives, and infection growth observed in monitoring plot data from Puglia.

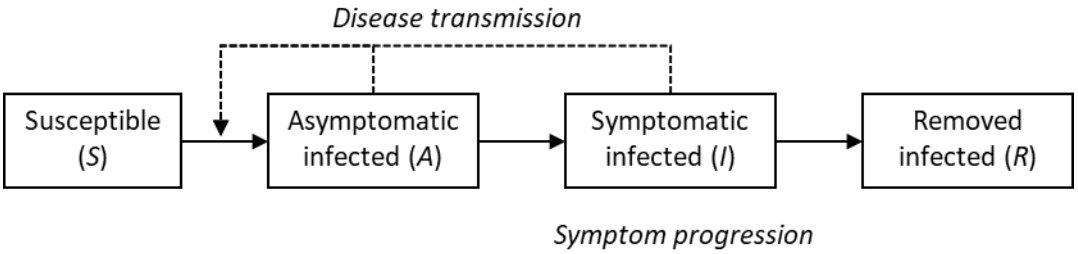

Figure 4: Schematic showing the compartmental epidemiology model for *Xylella fastidiosa* dynamics. Solid arrows show movements of individuals between compartments and the dashed arrows show how the infection transition depends on the density of infective compartments.

We defined the deterministic disease dynamics within *X. fastidiosa* infected grid cells by the following discrete-time versions of standard differential equations used in epidemiological compartmental modelling

$$\begin{aligned} S_{t+1} &= S_t - S_t\left(1 - e^{-\beta\upsilon(I_t+\tau A_t)/N}\right) \\ A_{t+1} &= A_t + S_t\left(1 - e^{-\beta\upsilon(I_t+\tau A_t)/N}\right) - A_t(1 - e^{-\alpha}) \\ I_{t+1} &= I_t + A_t(1 - e^{-\alpha}) - I_t(1 - e^{-\gamma}) \\ R_{t+1} &= R_t + I_t(1 - e^{-\gamma}) \end{aligned} \quad (1)$$

where the compartments *S*, *A*, *I* and *R* are defined as in Figure 4, *t* is the timestep (year) and $N = S + A + I + R$ is the total number of host plants in the grid cell. Ancient Olive groves in Puglia have a typical density ranging around of 10,000 trees per km^2, so we used this value to calculate *N*, as the worst-case scenario. Since grid cells were 200 x 200 m, the maximum *N* = 400 trees. This model was developed considering infection of olive trees, but for application to other settings, such as in this report, the compartmental model can be interpreted as modelling the fraction of the grid cell that is infected out of 400.

The transmission rate β represents the average number of transmissions from an infective plant. However, the transmission rate may be modified for outbreaks in novel locations, where the environment may have a significant effect on the epidemiology (parameter υ quantifies the reduction in transmission). We use this scaling parameter to explore the spread, surveillance and control at a lower transmission rate in Scotland (see Scenarios selection below). We assumed lower bacterial concentrations in asymptomatic host plants leading to lower infectivity than symptomatic plants (particularly true in olives). This is

modelled by parameter τ, which quantifies the relative infectivity of asymptomatic plants compared to symptomatic ones. The remaining parameters are the rates of symptom development (α) and removal (γ) where removal is defined as canopy loss or host plant death.

In data from monitoring plots in olive groves in Italy, we observed faster relative disease progression in smaller plots, which may reflect spatially limited disease transmission at very local scales. Assuming infection dynamics in grid cells with low susceptible cover are similar to the dynamics in small plots (which will be true if the susceptible cover is aggregated into monocultures such as olive groves), β was established as a function of the area of susceptible habitat (H) in the grid cell thus:

$$\beta = b_0 + b_1\sqrt{H}.$$

Scaling parameters b_0 and b_1 were estimated from Puglian plot disease progression data.

An example of worst-case scenario (Puglia-like) temporal dynamics of the within grid cell disease progression is shown in Figure 5.

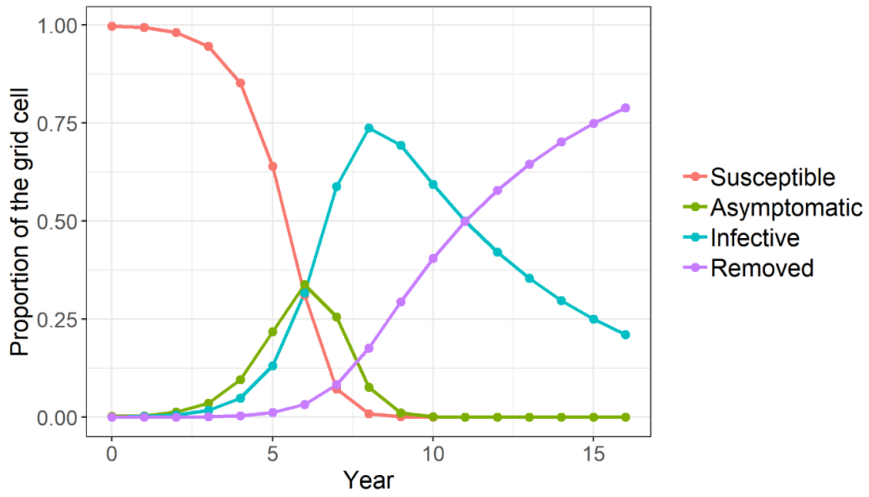

Figure 5: Example modelled *Xylella fastidiosa* Puglia-like temporal dynamics in a 200 x 200 m grid cell completely covered by susceptible habitat. Note that 'removed' refers to dead host plants killed by *X. fastidiosa*.

The model also includes between-grid cell disease transmission as a stochastic process whereby infective plants in one grid cell transmit *X. fastidiosa* via implicitly modelled insect vectors to plants in other uninfected grid cells, spreading the disease to new locations. For this model, the transmission rate used for the within-grid cell model was scaled by a distance-decay function, termed a transmission kernel. This

ensures that new infections mostly appear in the neighbourhood of large infective populations. However, because the transmission kernel is specified with a long tail, the model also results in rare long-distance transmission events (e.g. vectors hitchhiking or wind dispersal). The kernel parameters were estimated by tuning the spread model to mimic the rate of spread and spatial pattern of disease observed in Puglia from 2013 to 2018.

Infection of new grid cells by dispersal was modelled stochastically using a mixture of short and long-distance exponential transmission kernels, to represent local diffusive-like movements of vectors as well as the rarer jumps across large distances that can spread the disease rapidly to new regions. The mixture kernel is described by

$$K(d_{ij}) = \frac{(1-L)e^{\frac{d_{ij}\ln 0.5}{d_{short}}} + Le^{\frac{d_{ij}\ln 0.5}{d_{long}}}}{2\pi d_{ij}}$$

where d_{ij} is the distance between two grid cells *i* and *j*, *L* is the proportion of long-distance dispersal and d_{short} and d_{long} are equivalent to the median short and long distance dispersal distances.

In the model, the transmission kernel scales a decline in transmission rate with increasing distance from sources of infection. We adapted the equation for infection of susceptible individuals in the compartmental model to formulate the probability that a susceptible host plant in a currently uninfected grid cell *i* becomes infected from any source of infection in the modelled landscape such that

$$\rho_{i,t} = 1 - e^{-\sum_j K(d_{ij})\beta_j(I_{t,j}+\tau A_{t,j})/N_j}$$

where *j* indexes over all other grid cells for summing the product of the dispersal kernel and the source transmission rate (see the Equation (1)).

$$S_{t+1} = S_t - S_t\left(1 - e^{-\beta\upsilon(I_t+\tau A_t)/N}\right)$$

$$A_{t+1} = A_t + S_t\left(1 - e^{-\beta\upsilon(I_t+\tau A_t)/N}\right) - A_t(1 - e^{-\alpha})$$

$$I_{t+1} = I_t + A_t(1 - e^{-\alpha}) - I_t(1 - e^{-\gamma})$$

$$R_{t+1} = R_t + I_t(1 - e^{-\gamma})$$

Potential infection of each host plant in the uninfected grid cell was simulated by drawing random numbers of new infections from the binomial distribution with probability $\rho_{i,t}$. If any plants became infected, they were moved to the asymptomatic (*A*) compartment triggering subsequent growth of a new disease focus in that grid cell.

An example of the spread model output is shown in Figure 6. Here the disease spreads out from the initial introduction in a non-symmetrical fashion locally, but new foci are also established some distance away from the introduction site, which also begin to spread locally, which is typical of the Italian outbreak (Bucci, 2019). It should be noted that the model is stochastic, and therefore repeated model runs with identical parameters will produce different patterns of spread each time. To analyse the behaviour of the stochastic model we ran each model 250 times, informed by trial and error, to ascertain the average behaviour.

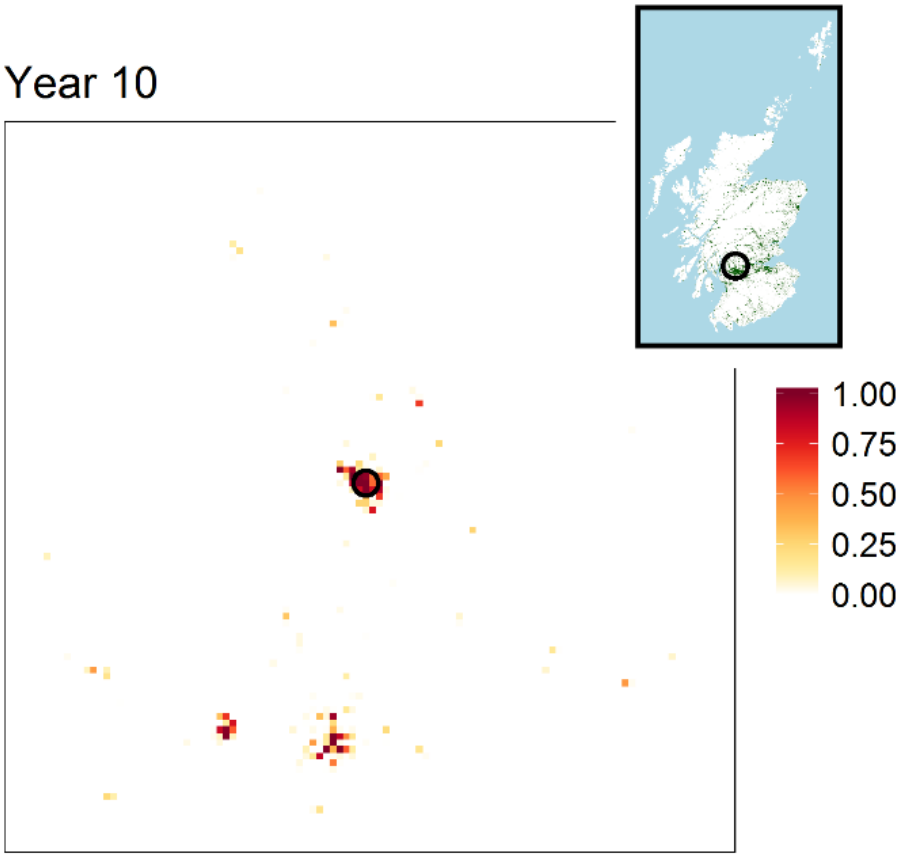

Figure 6: Example of *X. fastidiosa* spread over the woodland and urban landscape with Puglian worst-case scenario parameters. The black circle denotes the initial introduction of the outbreak, and the main plot is a magnification of the Scottish urban and woodland cover map (inlay). The simulation has run for 10 years since the infection was introduced, where the colour scale denotes the fraction of the grid cell infected (dark red indicated heavy infection). The simulation produces a patchy outbreak with several satellite foci distributed around the initial foci.

Surveillance

The model was used to investigate surveillance strategies for detecting *X. fastidiosa* outbreaks. For this report we modelled two distinct first detection surveillance scenarios to compare and contrast. In our scenarios, infection is either introduced within a 10km radius of a nursery or at a random location in Scotland. We then modelled a mixture of national surveillance with concentrated risk-based surveillance around nursery location of varying radii and surveillance intensities (number of grid cells inspected) within the risk-based zones around the nurseries, controlled by the parameter σ, the proportion of survey near nurseries. If $\sigma = 1$, then only grid cells within the risk-based surveillance zone around the nurseries are picked for surveillance. Conversely, if $\sigma = 0$, then only positive cover cells are picked for surveillance outside the nursery risk-based surveillance zone. Intermediate values of σ denote a mixing of strategies. We repeated these scenarios for varying levels of surveillance effort (i.e. the total number of cells inspected).

We based our resolution for surveillance, (100 x 100 m cells), after the inspections as described in Decision (EU) 2015/789 and Decision (EU) 2017/2352 for eradication and containment strategies. We had chosen a 200 x 200 m grid cell resolution to coincide with the approximate felling radius of positively identified infected host plants. Therefore, we subdivided each grid cell into four 100 x 100 m squares to model the surveillance. Each surveillance square was subject to the national surveillance or concentrated risk-based surveillance, as described above.

If the modelled visual inspection of a 100 x 100 m square locates symptomatic hosts, then the model selects one random symptomatic host for simulated laboratory testing. If no symptomatic hosts are detected, the model assumes that a randomly selected asymptomatic host is tested. Test accuracy rates for infected symptomatic and asymptomatic host plants were estimated from the monitoring data from Puglia, suggesting a significant level of false negatives (approximately 5-10%) but negligible false positive rates. In the model, the visual inspection and laboratory testing combine into an overall probability of detecting *X. fastidiosa* in each grid cell during surveillance. These are converted stochastically into detection events through Bernoulli trials ('coin tosses') using the overall detection probabilities.

We model the above assumptions as follows. The probability of detecting symptomatic hosts during the visual inspection of a 100 x 100 m square is

$$P_1 = 1 - (1 - \Psi)^v$$

where *v* is the visual inspection rate (proportion of the square inspected) and Ψ is the proportion of the population with symptoms, calculated as

$$\Psi = \frac{\phi(S+A)+I+R}{N},$$

where ϕ is the (low) proportion of uninfected and asymptomatic plants displaying scorch like symptoms in the general population.

If symptoms are detected during the simulated visual inspection of a 100 x 100 m square, then the model assumes that one random symptomatic plant is taken for simulated laboratory testing. In the model, the probability that the sampled symptomatic plant is infected with *X. fastidiosa* is

$$P_2 = \frac{\phi A+I+R}{\phi(S+A)+I+R}.$$

If no symptoms are detected during the simulated visual inspection of a 100 x 100 m square, then the model assumes that a random asymptomatic plant is taken for simulated laboratory testing. The probability that the sampled asymptomatic plant is infected with *X. fastidiosa* is

$$P_3 = \frac{A}{S+A}.$$

During simulated laboratory testing, we assume no false positive results occur but some false negatives occur. This is based on Latent Class Analysis (Linzer & Lewis, 2011) of the agreement between symptoms, ELISA and PCR in the Puglian monitoring data and assessment of the rates at which negative ELISA results from symptomatic and asymptomatic trees are re-tested with the more accurate qPCR. Based on this analysis, false negatives occur in the model with rates f_A and f_S for asymptomatic and symptomatic samples, respectively.

The overall probability of getting a positive test result from an inspection of a 100 x 100 m square is therefore

$$P = P_1P_2(1 - f_S) + (1 - P_1)P_3(1 - f_A),$$

which represents the probability of getting a positive sample from a symptomatic or asymptomatic host plant.

This probability is then scaled up to the overall probability of detection in the 200 x 200 m grid cell, in which four 100 x 100 m inspections occur, as

$$P_{det} = 1 - (1 - P)^4.$$

Model parameters

Parameters used in the simulations are defined in Table 1. The model was designed for the spread of *X. fastidiosa* pauca in olive groves in Puglia, Italy so parameter ranges reflect our understanding of the epidemiology in that system, and its uncertainty. Parameter values used were 50 samples from posterior distributions obtained by fitting the spread model to the monitoring data from Puglia. Model fitting was done using Approximate Bayesian Computation (Beaumont, 2010) in which prior parameter distributions were estimated, then 500,000 simulations were performed with random draws from the prior distributions and the 50 best fitting parameters retained as an approximate posterior distribution. These posteriors are expressed as their median and 95% range. Some other model parameters were not estimated but varied in the simulation experiment (see Scenarios section) so are not included here.

Table 1: Estimated parameter values of the *X. fastidiosa* spread model used in the simulations.

Parameter	Meaning	Values	Justification
b_0	Transmission rate at low host cover (yr^{-1})	1.830 (1.625-1.959)	Prior estimated by fitting a simple transmission model to progression of symptoms in Italian monitoring plots of varying size.
b_1	Decline in transmission rate with increasing host cover (yr^{-1})	-0.0152 (-0.0198--0.0098)	Prior estimated by fitting a simple transmission model to progression of symptoms in Italian monitoring plots of varying size.
τ	Proportion of transmission achieved by asymptomatic hosts	0.050 (0.005-0.098)	*Xylella* concentrations are lower in asymptomatic hosts and bacterial acquisition rates by vectors on asymptomatic hosts is lower than on symptomatic hosts.
α	Symptom development rate (yr^{-1})	2.898 (2.317-4.111)	These values give 90-98% of infected hosts showing symptoms after one year, consistent with symptom development in olive trees (D. Boscia and M. Saponari, pers. comm.).

Parameter	Meaning	Values	Justification
γ	Infection removal (canopy collapse or host plant death) rate (yr^{-1})	0.194 (0.176-0.230)	These values give median removal times of 3.0-3.9 years, consistent with expert opinion on time for canopy collapse in infected olive trees (D. Boscia and M. Saponari, pers. comm.).
d_{short}	Median of short distance dispersal kernel (km)	0.151 (0.043-0.431)	Fitted to the spread rate in Puglia and consistent with limited mark recapture data on vector movement range (D. Bosco, pers. comm.)
d_{long}	Median of long distance dispersal kernel (km)	10.7 (8.0-20.1)	Fitted to the spread rate in Puglia
L	Proportion of long distance dispersal	1.95×10^{-4} (1.18×10^{-5} -2.91×10^{-4})	Fitted to the spread rate in Puglia
ϕ	Proportion of uninfected host plants exhibiting scorch-like symptoms	7.72×10^{-4} (7.54×10^{-4}-7.91×10^{-4})	Prior estimated by latent class analysis on agreement between tests (inspection, ELISA and qPCR) on olive trees in Puglia
f_A	False negative rate of infected samples from asymptomatic host plants	0.109 (0.077-0.142)	Prior estimated by latent class analysis on agreement between tests (inspection, ELISA and qPCR) on olive trees in Puglia
f_S	False negative rate of infected samples from symptomatic host plants	0.087 (0.062-0.114)	Prior estimated by latent class analysis on agreement between tests (inspection, ELISA and qPCR) on olive trees in Puglia
v	Visual inspection rate, i.e. proportion of the host plants inspected during an inspection	0.975 (0.969-0.982)	Prior estimated by latent class analysis on agreement between tests (inspection, ELISA and qPCR) on olive trees in Puglia

Specification of scenarios

Simulation scenarios varied the parameters as shown in Table 2, resulting in 400 distinct scenarios. All combinations of the following scenarios were run 250 times, each with a different random draw from the model parameter posterior distributions, thus totalling 100,000 individual model runs.

To compare the various combinations of scenarios we plot the proportion of replicate simulations in which detection occurred in a 5 year period and compare this against a national surveillance strategy where there is an equal probability of surveying every grid cell across the landscape.

Table 2: Scenario combinations used in the simulation experiment.

Variable	**Scenarios**	**Justification**
Surveillance and disease discovery	National surveillance or a mixture of national surveillance and risk-based surveillance	Assess the effectiveness of surveillance
Infected plant introduction	Within 10 km of a nursery or a random location	Assess the effectiveness of surveillance, depending on introduction location
Risk-based surveillance width	Varying widths: 2.5, 5, 10 & 20 km	Varying how far to survey from the nursery
Proportion of survey near nurseries, σ	Varying proportions: 0, 0.2, 0.4, 0.6, 0.8 & 1	Varying risk-based surveillance intensities
Epidemiology	Puglia-like worst case scenario disease transmission or 90% reduced transmission	Worst case scenario from fitted olive disease progression; reduced transmission is likely in Scotland
Total surveillance effort	Varying efforts: 100, 500, 1000, 2000 grid cells	Vary the fixed number of grid cells inspected each year

Results

Figure 7 contains details of the proportion of infected plants detected after 5 years for simulations of the Puglia-like, worst case scenario disease transmission. In all scenarios, the proportion of infected plants detected increases with the number of grid cells inspected per year. When the risk-based surveillance is sufficiently wide (greater or equal to 10 km), and if the infected plant is introduced within 10 km of a nursery then the best surveillance strategy is to only have risk-based surveillance ($\sigma = 1$). However, if the risk-based surveillance is less wide, then the best surveillance strategy is to have a mixture of risk-based and national surveillance. In contrast, if the infected plant is introduced at a random location then risk-based surveillance works poorly in general, especially at low widths (at higher widths the risk-based surveillance may work well – see Summary & Discussion). In all scenarios, having a mixture of national and risk-based surveillance nearly always outperforms a strategy of national surveillance on its own.

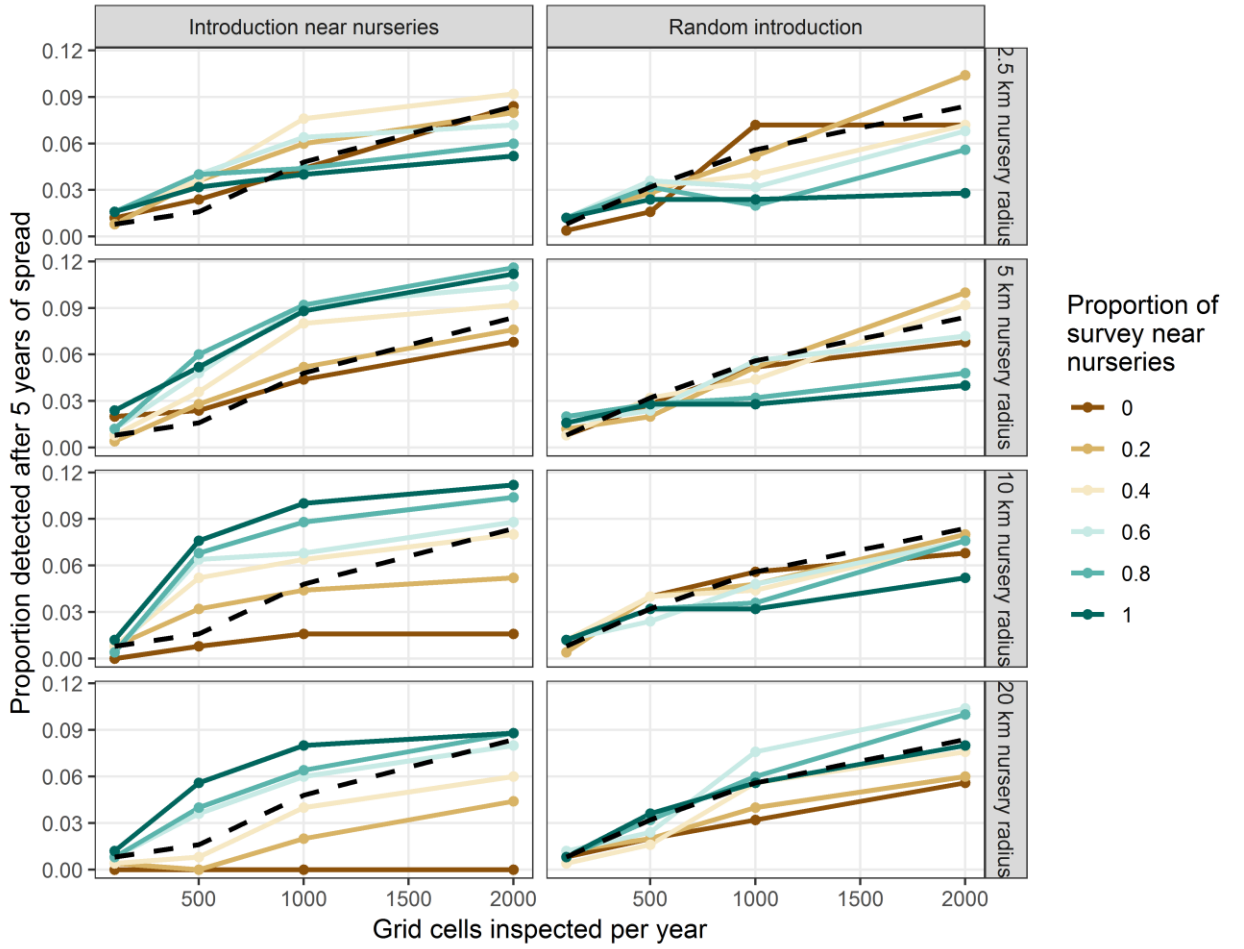

Figure 7: Proportion of simulations in which infected plants are detected by 5 years since introduction for the Puglia-like, worst case scenario epidemiology. The columns show the results for the two infected plant introduction scenarios: introduction within 10 km of a nursery (left); random introduction (right). The rows show the results for varying risk-based surveillance zone widths. The output is plotted against the total number of grid cells inspected per year. Colours indicate the proportion of survey near nurseries, σ. The black dashed line corresponds to the corresponding national surveillance (i.e. no risk-based surveillance).

Figure 8 demonstrates the effect of the lower transmission rate scenario. In general, the emerging best surveillance strategies are similar to those in the Puglia-like transmission scenario. However, the proportion of outbreaks detected after 5 years is greatly reduced.

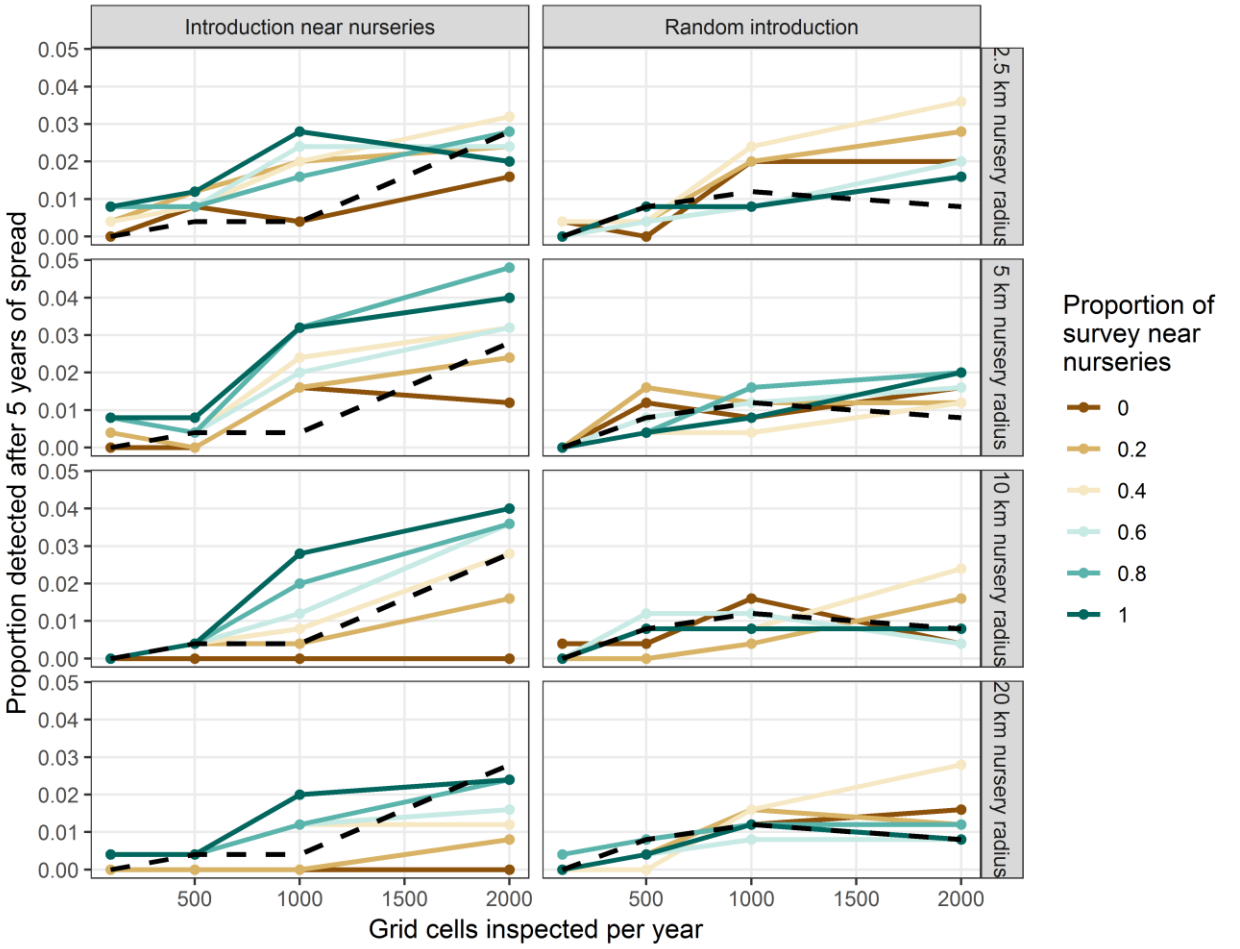

Figure 8: Proportion of infected plants detected after 5 years since introduction for the reduced disease transmission scenario. Individual panel descriptions are as in Figure 7 but note difference in y axis scale.

Similar qualitative results can be shown to the above when reflecting on the proportion of detections after longer periods (e.g. after 10 years).

Summary & Discussion

This preliminary study for predicting the potential spread of and surveillance strategies for detecting *X. fastidiosa* in Scotland has been made possible by utilising and adapting existing models and datasets, along with making plausible assumptions on the epidemiology from known outbreak locations (mainly Puglia). We explored various scenarios on the epidemiology, introduction location and the surveillance strategies (type and intensity), which were interpretable in the light of existing knowledge. The results, applications and potential model limitations/improvements are now discussed.

The form of the best surveillance strategy is sensitive to whether the introduction of the infected plant is near to a nursery (or location that is deemed to be a likely source for introduction). In the event that the introduction of an infected plant is close to a location that has been determined to be a potential risk, then a concentrated risk-based surveillance with low-level national surveillance elsewhere is predicted to be the best strategy. However, by contrast, if the introduction is not near to a location that has been determined to be a potential risk (e.g. due to dispersal via plant sales; or other modes of introduction) then a national surveillance strategy is predicted to work best in most cases. However, risk-based surveillance was predicted to work well in some scenarios, or a mixture of low intensity risk-based surveillance and national surveillance. This may be because the location of nurseries are located close to urban or woodland cover (see Figure *3*), which therefore suggests that risk-based surveillance will cover areas of a likely introduction, especially is the risk-based surveillance is sufficiently wide.

The predicted best strategy is largely independent of the transmission rate, even though spread and disease progression is greatly reduced when the estimated transmission rate for Puglia was reduced to 10% of its value. However, it should be noted that the probability of discovering infected plants is greatly reduced for the lower transmission scenario.

We conclude that having a mixture of risk-based and national surveillance is likely to be the best option for detecting potential *X. fastidiosa* outbreaks in Scotland.

Uncertainties Affecting the Assessment of the Model

The model used in this assessment was developed from a generic epidemiological model originally designed for *X. fastidiosa pauca* infecting olive trees in Puglia, Italy. The model necessarily simplifies the real-world epidemiology in Puglia and does not explicitly represent vector population dynamics or regrowth and infection of suckers from heavily infected olive trees.

Another important caveat is that epidemiological parameters may vary in different regions of Europe according to the bacterial subspecies, host species, host density within susceptible land cover types, vector ecology, climate and other factors not accounted for in the modelling. In this report we have made the simplifying assumption that transmission is lowered to try to understand the implications of an outbreak in a less favourable environment, such as Scotland. However, there may be other important differences that we have not considered. For example, asymptomatic periods and the infectivity of

asymptomatic plants may differ from the model for non-olive host plants in Scotland. Some of these aspects are currently being investigated in the BRIGIT project (https://www.jic.ac.uk/brigit/). We have had to assume, due to paucity of data, that all host plants and vectors are only found in urban and woodland areas. Furthermore, we assumed that the density of vectors does not vary between habitat types, and that the density of host plants in a grid cell is the percentage cover multiplied by a scaling factor (the approximate density of ancient olive trees in Puglia). We have had to make these assumptions in other studies (e.g. see EFSA Panel on Plant Health et al. (2019) for example). Improving the national mapped data on host and vector distributions and densities would help to identify potential high-risk areas when combining with epidemiological models.

Our model has shown that the best surveillance strategies depend on whether the introduction of the infected plant is close to a known high-risk location or not. In our model it is easy to prescribe this, but in reality, knowing the probable location of *X. fastidiosa* introduction is less well understood. Therefore, it would be beneficial to formally estimate the most likely locations of *X. fastidiosa* entry into Scotland and a PHC funded project, led by Samantha Broadmeadow (PHC2018/04), mapped the likelihood and impact of a Xylella outbreak in Scotland. Based on the results of that project, we would suggest that surveillance should focus on the central belt of Scotland targeting peri-urban environments and sites of special scientific interest.

Potential Future Model Developments

In this report we have answered questions on surveillance strategies for first detection. We have determined that combining risk-based surveillance with national surveillance is best for locating *X. fastidiosa* infected plants. However, the question remains of what to do once infected plants have been found. One approach to address this is to emulate the emergency measures described in Decision (EU) 2015/789, in which demarcated eradication buffer zones are established around detected *X. fastidiosa* occurrences, and surveillance and disease control measures are applied within these zones. Specifically, in each buffer zone all infected and uninfected hosts are felled within a 100m radius of the infected plant and an intense survey in a 10 km radius of the infected plant is undertaken. However, this may not be appropriate for a Scottish outbreak, where the transmission and spread may be reduced compared to a Puglia-like outbreak (say), as shown here. Furthermore, these strategies could be refined to reflect likely introduction locations and locations of susceptible host plants.

Surveillance and laboratory testing practices may vary from those implemented in the simulations which emulate those in Puglia. Potential Scottish protocols could be incorporated into the model to establish how these may affect detection.

Acknowledgements

This work was funded by the Scottish Government's Rural and Environment Science and Analytical Services (RESAS) Division through the Centre of Expertise for Plant Health. Steven White and Daniel Chapman are also funded by the Horizon 2020 XF-ACTORS project [SFS-09-2016] and the UKRI/Defra/Scot. Gov. BRIGIT project [BB/S016325/1], and have contributed to the EFSA Update of the Scientific Opinion on the risks to plant health posed by Xylella fastidiosa in the EU territory (EFSA Panel on Plant Health et al., 2019).

References

Allen, L. J., Brauer, F., Van den Driessche, P., & Wu, J. (2008). *Mathematical epidemiology* (Vol. 1945): Springer.

Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics, 41*(1), 379-406. doi:10.1146/annurev-ecolsys-102209-144621

Bucci, E. M. (2019). Effectiveness of the monitoring of X. fastidiosa subsp. pauca in the olive orchards of Southern Italy (Apulia). *Rendiconti Lincei. Scienze Fisiche e Naturali*. doi:10.1007/s12210-019-00832-6

Chapman, D. S., Purse, B. V., Roy, H. E., & Bullock, J. M. (2017). Global trade networks determine the distribution of invasive non-native species. *Global Ecology and Biogeography, 26*(8), 907-917. doi:10.1111/geb.12599

Chapman, D. S., White, S. M., Hooftman, D. A. P., & Bullock, J. M. (2015). Inventory and review of quantitative models for spread of plant pests for use in pest risk assessment for the EU territory. *EFSA Supporting Publications, 12*(4), 795E. doi:10.2903/sp.efsa.2015.EN-795

EFSA. (2018). Update of the Xylella spp. host plant database. *EFSA Journal, 16*(9), e05408. doi:doi:10.2903/j.efsa.2018.5408

EFSA Panel on Plant Health, Bragard, C., Dehnen-Schmutz, K., Di Serio, F., Gonthier, P., Jacques, M.-A., . . . Parnell, S. (2019). Update of the Scientific Opinion on the risks to plant health posed by Xylella fastidiosa in the EU territory. *EFSA Journal, 17*(5), e05665. doi:10.2903/j.efsa.2019.5665

Feil, H., & Purcell, A. H. (2001). Temperature-Dependent Growth and Survival of Xylella fastidiosa in Vitro and in Potted Grapevines. *Plant Disease, 85*(12), 1230-1234. doi:10.1094/PDIS.2001.85.12.1230

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *2011, 42*(10), 29. doi:10.18637/jss.v042.i10

Rowland, C. S., Morton, R. D., Carrasco, L., McShane, G., O'Neil, A. W., & Wood, C. M. (2017). *Land Cover Map 2015 (vector, GB)*. Retrieved from: https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73

Saponari, M., Giampetruzzi, A., Loconsole, G., Boscia, D., & Saldarelli, P. (2019). Xylella fastidiosa in Olive in Apulia: Where We Stand. *Phytopathology, 109*(2), 175-186. doi:10.1094/phyto-08-18-0319-fi

Tumber, K. P., Alston, J. M., & Fuller, K. (2014). Pierce's disease costs California $104 million per year. *California Agriculture, 68*(1), 20-29. doi:10.3733/ca.v068n01p20

White, S. M., Bullock, J. M., Hooftman, D. A. P., & Chapman, D. S. (2017). Modelling the spread and control of Xylella fastidiosa in the early stages of invasion in Apulia, Italy. *Biological Invasions, 19*(6), 1825-1837. doi:10.1007/s10530-017-1393-5