

Robustness of local control strategies under modelling uncertainties: the 2004 avian influenza H5N1 outbreak in Thailand

Renata Retkute^{a,*}, Chris P. Jewell^b, Thomas P. Van Boeckel^c, Geli Zhang^d, Xiangming Xiao^d, Weerapong Thanapongtharm^e, Matt Keeling^a, Marius Gilbert^f, Michael J. Tildesley^b

^a*School of Life Sciences and Institute of Mathematics, University of Warwick, UK*

^b*Faculty of Health and Medicine, Furness College, Lancaster University, UK*

^c*ETH Zurich, Institute for Integrative Biology, Switzerland*

^d*Center for Spatial Analysis, University of Oklahoma, USA*

^e*Department of Livestock Development, Bangkok, Thailand*

^f*Biological Control and Spatial Ecology Universite Libre de Bruxelles, Belgium*

Abstract

The Highly Pathogenic Avian Influenza (HPAI) subtype H5N1 virus persists in many countries and has been circulating in poultry, wild birds. In addition, the virus has emerged in other species and frequent zoonotic spillover events indicate that there remains a significant risk to human health. It is crucial to understand the dynamics of the disease in the poultry industry to develop a more comprehensive knowledge of the risks of transmission and to establish a better distribution of resources when implementing control. In this paper, we develop a set of mathematical models that simulate the spread of HPAI H5N1 in the poultry industry in Thailand, utilising data from the 2004 epidemic. The model that incorporates the intensity of duck farming when assessing transmission risk provides the best fit to the spatiotemporal characteristics of the observed outbreak, implying that intensive duck farming drives transmission of HPAI in Thailand. We also extend our models using a sequential model fitting approach to explore the ability of the models to be used in "real time" during novel disease outbreaks. We conclude that, whilst predictions of epidemic size

*Corresponding author

Email address: r.retkute@warwick.ac.uk (Renata Retkute)

are estimated poorly in the early stages of disease outbreaks, the model accurately predicts the preferred control policy that should be deployed to minimise the impact of the disease.

Keywords: avian influenza, disease control, spatial heterogeneity, vaccination

1. Introduction

Since the emergence of highly pathogenic avian influenza (HPAI) H5N1 in the late 1990s, the virus has had significant impact on poultry industries around the world and posed a serious threat to public health. The majority of outbreaks occur in East and South Asia, a region that contains half of the population of the world. Whilst the probability of a human AI pandemic may appear low based on the poor ability of the virus to adapt to the upper human airway (Peiris et al. 2007), the death toll of such an event may be catastrophic. As a result of the HPAI H5N1 panzootic, there have been 860 confirmed human cases of avian influenza A (H5N1) in 16 countries and 454 deaths (a human case fatality of 53%) as of 30th October, 2017, based on World Health Organization (WHO) statistics. The majority of humans infected with the virus work in professions that involve close contact with potentially infected poultry (de Bruin et al. 2017) and therefore it is crucial to understand the dynamics of the disease in the poultry industry to develop a more comprehensive knowledge of the risks of transmission to humans.

In Thailand, H5N1 was first detected in the poultry industry in January 2004 (Tiensin et al. 2005). The first wave of infection took place from January to May and resulted in 193 reported AI outbreaks. A second infection wave started in July 2004 and culminated in March 2005, with 1,492 outbreaks notified during this period. Approximately 62 million birds were killed either through infection or through targeted culling as a control measure. In addition, during the 2004 H5N1 epidemic there were 17 human cases of infection and 12 deaths (Tiensin et al. 2005, Auewarakul 2008). In an attempt to improve case detection, the Thai government implemented an X-ray survey (Gilbert et al. 2006) in

September 2004. Control policies such as localised movement restrictions and 1km ring culling were introduced to reduce the risk of further spread of the disease; vaccination was not used.

The spatial distribution of HPAI H5N1 has been studied in several countries
30 and regions using spatial statistical models (Gilbert et al. 2007 , Truscott et al.
2007 , Sharkey et al. 2008 , Gilbert et al. 2008 , Jewell et al. 2009 , Gilbert
et al. 2010 , Stegeman et al. 2011 , Ssematimba et al. 2012 , Paul et al. 2014
, Gilbert et al. 2014 , Artois et al. 2017) and mathematical models (Minh
et al. 2011 , Walker et al. 2012 , Hill et al. 2017) in East Asia. Evidence
35 from previous work suggests that different landscapes, production systems and
water-related variables are risk factors that will promote disease transmission.
In addition, mathematical models such as the one that we present here can
improve understanding about the transmission dynamics of infectious diseases
and help with an assessment of the effectiveness of control strategies applied
40 during outbreaks (Stegeman et al. 2011). A key challenge when implementing
control is the lack of clinical signs in many duck species - whilst chickens have a
very high mortality rate for H5N1 (Yu et al. 2007 , Jeong et al. 2009), ducks
are largely asymptomatic but can readily transmit the disease. In addition, in
Thailand, as in many other countries in South East Asia, the presence of free
45 grazing ducks, that feed year round on rice paddies, have been shown to be a
strong predictor for the presence of the disease in the landscape (Gilbert et al.
2007, 2008). It is therefore crucial to develop a more detailed understanding of
the factors leading to persistence of AI in poultry, the conditions that are most
suitable to disease transmission, and intervention strategies that will minimize
50 the future impact of the disease.

In this paper, we develop a set of mathematical models to simulate the
spread of HPAI H5N1 in the poultry industry in Thailand. We use detailed de-
mographic and epidemiological data from the 2004 epidemic and fit our models
to epidemiological surveillance (X-Ray) data. Our framework will allow us to
55 establish the risk factors that result in transmission of the disease that will in
turn enable us to effectively target intervention strategies. We will also consider

the ability of the models to be used in "real time", whereby model parameters are obtained by only using data that are observable at a given stage of the epidemic. This enables us to determine the predictive capacity of HPAI models
60 both in terms of establishing transmission risk and in determining the most appropriate intervention policy. It also allows us to quantify how progressive accumulation of information shapes the knowledge of outbreak dynamics and the robustness of control strategies.

2. Materials and methods

65 2.1. Data

In this paper, we use demographic data compiled by the Department of Livestock Development (DLD, Bangkok) that was constructed through the X-Ray surveys that were implemented in response to the H5N1 outbreak in 2004. Demographic data was recorded at the farm/owner level and consists of the
70 number of chicken and ducks in each flock, and a unique identifier of the sub-district where each farm is located.

Owing to imperfect reporting at the onset of the epidemic, data for the first wave of infection in Thailand are largely incomplete. Therefore, for parameter estimation and analysis, we use the outbreak data from the second wave of the
75 epidemic that took place from July 2004. In this second wave, more than 1400 outbreaks were reported and over 62 million birds died or were culled as part of the control policy (Tiensin et al. 2005). The outbreak data set includes the unique identifier of the subdistrict as well as the total number of sick, dead and destroyed animals.

80 In order to implement our mathematical model, we use geographical data for each subdistrict including the locations of subdistrict boundaries and the area of each subdistrict. We infer indices of neighbouring subdistricts by analysing intersection sets of subdistrict boundaries. We then calculate the length of all shared boundaries by fitting a linear spline through shared points on bound-
85 aries of neighbour subdistricts. The proportion of area used by rice paddy fields

within each subdistrict was derived from the moderate resolution imaging spectroradiometer (MODIS) sensor onboard the NASA Terra satellite (Xiao et al. 2006).

2.2. Mathematical models

90 We develop a meta-population framework for our model, following the approach of Buhnerkempe et al. 2014 . In our model, the basic unit of infection is the flock, and we assume that all birds (chicken and ducks) within a flock become infected such that an entire flock is classified as Susceptible, Exposed, Infectious or Removed.

95 In our modelling framework, we consider transmission via one of five mechanisms: (i) spatially independent transmission to any subdistrict in Thailand, (ii) transmission within an subdistrict, (iii) transmission across borders to neighbouring subdistricts, (iv) transmission driven by presence of rice paddies and (v) transmission driven by duck farm intensity. By varying the combination
100 of these factors, we have formulated four mathematical models with increasing complexity: a random process model (Model A), a spatial model (Model B), a spatial model incorporating rice density (Model C), and a spatial model incorporating duck intensity (Model D). Details of the equations governing each model are given in the Supplementary Information.

105 2.2.1. Model A: Random process model

The simplest model for a HPAI outbreak is that of complete spatial randomness, whereby transmission events are distributed independently according to a uniform probability distribution over all susceptible flocks. Under this scenario, the infection pressure at time t in any susceptible flock is equal to the background term, δ , multiplied by the number of infectious flocks at time t , $n_I(t)$,
110 plus one. The biological meaning of the model is that the outbreak dynamics does not depend on any environmental factors or the structure of the poultry sector in Thailand. This naive model, whilst unlikely to be representative of

heterogenous transmission routes, provides a baseline for epidemic spread for
115 comparison with the spatially explicit models.

2.2.2. Model B: Spatial model

The Spatial model utilises a classical metapopulation approach, based upon
the model first used by Buhnerkempe et al. 2014 . This model is an extension of
the random process model, with the addition of within and between subdistrict
120 transmission. The model therefore incorporates local density-dependent spread
and contains three additive terms representing the different transmission scenar-
ios - the background term, δ , within-subdistrict transmission with transmission
rate β_W and local cross-border transmission with transmission rate β_B . The
framework assumes that transmission is dependent upon the poultry industry
125 features within a given subdistrict and within neighbouring subdistricts and
ignores the presence of other landscape features that may results in increased
transmission risk. Within and between subdistrict transmission depends on the
structure of poultry industry, and is parametrised by p_C and p_D (power laws for
a number of chicken and ducks) and ξ (multiplier for relative infectiousness and
130 transmisability of ducks to chicken).

2.2.3. Model C: Spatial (rice) model

Given that previous work suggests that the presence of free grazing ducks
can result in increased transmission risk in Thailand (Gilbert et al. 2007),
here we use data on rice fields in the country (Xiao et al. 2006) and extend
135 the meta-population model described in the spatial model above to consider the
likelihood of increased transmission risk in subdistricts with a high density of
rice paddies. In order to account for the role of rice paddies in the infection
dynamics, we use a sigmoidal function that operates as a scaling factor for the
infection pressure. Here a higher infection pressure is presumed to be present
140 for subdistricts with a higher proportion of rice paddy fields. This dependence
is described by the scaling, α_R , threshold, ϵ_R and power, n_R . Details of the
equation are given in the Supplementary Information.

2.2.4. Model D: Spatial (duck) model

Our final model builds upon the spatial model but additionally considers the
145 intensity of duck farming within a subdistrict as an additional factor influencing
transmission. We therefore categorise subdistricts based upon the intensity of
their duck farming industry defined as a \log_{10} of the size of largest duck flock
within a subdistrict plus one. Furthermore, in a similar way as for the spatial
(rice) model, we use a sigmoidal function such that a higher infection pressure
150 is assumed for subdistricts with more intense duck farming. This dependence is
described by the scaling, α_D , threshold, ϵ_D and power, n_D .

2.3. Fitting the models to outbreak data

Having defined the full modelling framework, we will consider this suite of
nested mathematical models (A-D) to determine the factors that contribute to
155 transmission of HPAI H5N1 within Thailand. Each model is fitted to outbreak
data using a Bayesian likelihood approach that determines the parameters that
best captures the observations (for details see Supplementary Material section
below).

For a sequential analysis, we re-estimated model parameters by incrementally
160 increasing the length of the outbreak data, evaluating how increasing amounts
of outbreak data altered and influenced the parameter estimation and model
projections.

2.4. Control options

Once fully parameterised, we will use the models to investigate the effec-
165 tiveness of a range of interventions that could be implemented to control the
outbreak. We will analyse outbreak projections under three competing strate-
gies: (i) culling of identified infected premises (IP) only, (ii) culling of IPs and
farms within a given radius of IPs (ring culling) and (iii) IP culling and reactive
vaccination with a given capacity within infected subdistricts.

170 The spatially aggregate nature of the meta-population model that we have
developed prevents us from applying a culling radius directly to each IP to

determine which flocks will be culled. In order to mimic ring culling in our metapopulation framework we therefore reduce the number of susceptible flocks within an infected subdistrict as a fraction of the subdistrict area to the area of a circle with a particular culling radius. We analyse projections based on the following culling radius values: 0.25km, 0.5km, 1km, 2km and 5km.

As an alternative to ring culling, we explore the impact of reactive vaccination within any subdistrict reporting infection. For this control policy, following the removal of infected flocks, we assume that a given percentage of the poultry within an infected subdistrict are vaccinated and become immune to infection. We make the assumption that any poultry that are targeted for vaccination will acquire immunity 7 days after the relevant infected premises is reported. We then explore the impact of vaccination upon the outbreak size and duration, given vaccine capacities of 50%, 70% and 90% of all poultry within each infected subdistrict. See Supplementary Information for more details. As optimal control actions for control of disease outbreaks may depend on objectives (Probert et al. 2016), we evaluated the effectiveness of local control by analysing projected number of infected flocks, number of culled flocks and duration of outbreak.

3. Results

3.1. Model Parameterisation

We used a Bayesian framework to parameterise each model to the 2004 epidemics of H5N1 in Thailand. In order to compare the four modelling frameworks and establish a goodness of fit to the observational data, we calculate the deviance information criterion (DIC). Model with a lower DIC value implies a closer fit to the observational data compared to the alternative models and therefore this model is preferred. For DIC calculations we use the parameter values generated from our MCMC runs for parameter estimation (Spiegelhalter et al. 2002). We calculated the following values for the four models: $DIC_A = 37478.2$, $DIC_B = 33763.5$, $DIC_C = 33596.7$ and $DIC_D = 33164.3$.

The DIC values for each model indicate that Model D, incorporating the intensity of the duck industry, is preferred over all other modelling frameworks, indicating that it is the intensive duck farming industry that is predominantly driving transmission in Thailand.

205 For all models, the fitted posterior mean and 95% credible intervals for estimated parameters are summarised in Table 1. As the complexity of the modelling framework increases, the value of the background term δ decreases, as local scale characteristics drive transmission for the more complex models. When considering rice paddy density as a risk factor (Model C) similar parameter estimates are found for the background term and within-subdistrict
210 transmission rate when compared with the purely spatial Model B. Interestingly Model D, that includes intensive duck farms as a risk factor, produces lower parameter value estimates for δ , β_W and β_B , suggesting that the presence of intensive duck farms may have a strong influence on transmission.

215 To establish the ability of each model to capture the spatiotemporal characteristics of the observed outbreak, we now simulate our models using the posterior distribution for parameters summarised in Table 1. We use these simulations to determine the ability of each model at capturing both the temporal and the spatial epidemic profiles from the 2004 outbreak. This will provide an
220 indication of whether the our modelling frameworks are able to mimic the fine-scale transmission dynamics between poultry farms in the country. The results are summarised in Fig 1.

Model D appears to most accurately predict the spatiotemporal dynamics of the HPAI outbreak in Thailand (Fig 1, second column). For this model, spatial
225 predictions of spread are much more representative of the true 2004 outbreak, with infected subdistricts predominantly in the centre and west of Thailand, in the region north of Bangkok. In addition, the temporal profile captures both the epidemic peak and the epidemic tail, providing supporting evidence of the significant role of the intensive duck farming industry in transmission of HPAI
230 H5N1 in Thailand.

The spatial model, Model B, proved to be a poor predictor of geographical

spread of the outbreak, with significant overestimates of infected farms in the east of the country and underestimates in the main epidemic hotspots. The model is also unable to accurately capture the temporal profile of the epidemic tail (Fig 1, third column). We see remarkably similar predictions for Model C (Fig 1, fourth column), though in this case overpredictions of spread in the east of Thailand are even greater, owing to the high densities of rice paddies in that part of the country. In addition, south eastern subdistricts of Thailand tend to have a higher average flock density, which will result in Models B and C overpredicting spread in these regions.

Random process model is unable to capture the spatial profile of the true outbreak, given that the model assumes random transmission characteristics. The model performs marginally better at predicting temporal behaviour, though slightly overestimates the epidemic peak and underestimates the number of farms infected in the epidemic tail (Fig 1, last column).

3.2. Sequential analysis

The model parameterisation that we have described in the section above determines parameter values by fitting the set of models to the entire observed spatiotemporal outbreak in Thailand. This gives us an understanding of the level of complexity required to capture HPAI epidemic dynamics. However, this approach only enables us to use our models retrospectively, to understand the model parameters and predictions of risk factors with the benefit of hindsight. However, it is also important to explore the predictive capacity of our models, so that we can assess whether the frameworks can be used to provide policy advice for ongoing epidemics.

In order to investigate the utility of our model predictions and control options in real time during outbreaks, we re-fit Model D only using outbreak data that would be available at particular points during the epidemic. The model is therefore parameterised using reported data up to day 10 and then sequentially on days 15, 20, 25, 30, 35, 50, 75, 100, 125, 150, 175 and 200 (the entire outbreak). We are interested in exploring how the prediction of the epidemiological

parameters change through time. Our results are summarised in figure 2. As the observation time, T , increases, the estimated value of the background term δ decreases. Meanwhile, within subdistrict and between subdistrict transmission rates are both observed to initially increase as T increases, before decreasing again after the epidemic peak.

When the model is only fitted to the first 10 days on the outbreak, we predict significant spread of the virus over large areas of the country (figure 3, second column). The model also dramatically overestimates outbreak size. This overestimate appears to be due to the parameterisation of the model on day 10 - the background term δ dominates (figure 2, fourth row) and the model predicts much lower within-subdistrict and cross border transmission (figure 2, fourth row). However, as the outbreak progresses, the model predicts that transmission is becoming more local, as seen by a reduction in the predictions of the value of δ . As the epidemic progresses, we see that the model performs much better at predicting the spatiotemporal profile of the outbreak. Indeed by day 30 (figure 3, third column), predictions of epidemic spread begin to be confined to regions where the true outbreak was observed. By day 50, the model is capable of producing accurate predictions of both the spatial and temporal profile (figure 3, fourth column), indicating that from this stage the ability of our modelling framework to generate robust predictions that could be used to inform policy in real time.

3.3. Evaluating the effectiveness of local control

We now investigate the effectiveness of interventions at reducing the impact of H5N1 in Thailand. We use the four models and explore the impact of ring culling at radii from 0.25km to 5km around each infected premises and vaccination within infected subdistricts with capacities of 50%, 70% and 90% of all birds within the subdistrict. The results are summarised in (Fig 4 A). We see a significant difference in predicted outbreak sizes and durations for the different control policies when we compare the predictions between the models. This suggests that, if we are interested in predicting the quantitative impact of a

given intervention, it is crucial to select the most appropriate model. However, regardless of the differences in the predictions of epidemic size, all four models predict that large radius ring culling (5km) is the most effective strategy at minimising the total number of infected flocks (Fig 4 A, top panel). This strategy
295 results in a significant number of culled flocks and in fact if we are interested in minimising the total number of culled poultry flocks, high capacity vaccination proves to be the most effective policy (Fig 4 A, middle panel). Models B, C and D predict that 5km ring culling results in the lowest outbreak duration across
300 all the control strategies investigated. However, we see somewhat different results for our preferred spatial (duck) model – in this case there is no significant difference between a 5km ring culling policy and a vaccination strategy with a 90% capacity (Fig 4 A, lower panel, Mann-Whitney test $p = 0.14$).

Finally, we investigate the predictive ability of control policies for the preferred Model D. We therefore simulate our model from the start of the outbreak,
305 using parameters that we have obtained on days 10, 15, 20, 25, 30, 35, 50, 100 and 200. We then explore the model predictions of the effectiveness of all competing ring culling and vaccination policies using model parameters obtained at each of these intervals. We observe that, even on day 10 when there is significant uncertainty in spatial spread parameters, Model D accurately predicts
310 that 5km ring culling is the policy that minimises the total number of infected flocks (Fig 4 B, top panel). Whilst the predicted epidemic sizes change for each control policy during the outbreak, the relative ordering of control policies remains consistent throughout. The same is true for total culled flocks, in that
315 high capacity ring culling is always preferred (Fig 4 B, middle panel).

However, if we are interested in minimising outbreak duration we see a somewhat different result. When the model uses only early outbreak data, predictions of epidemic duration are dramatically overestimated for IP only culling and all vaccination policies (Fig 4 B, lower panel). The dynamics of the simulated outbreak is mostly dominated by the background δ , so even under a vaccination
320 capacity of 90%, this gives a high enough infection pressure to sustain the outbreak for a long period of time. This overprediction is largely resolved by day 25

when vaccination is used, possibly owing to the more accurate predictions of the spatial extent of the disease at this point in the epidemic. We therefore conclude
325 that it is important to determine the objective of control when deciding upon an intervention policy and that model predictions of the spread of the disease and the impact of interventions should be considered with caution during the early stages of any influenza epidemic.

4. Discussion

330 We have developed a set of mathematical models to investigate the level of complexity required to make accurate predictions regarding the spread of HPAI H5N1 in Thailand. When developing models, there is always a trade off between including enough model complexity to be able to capture the, often very complex, epidemiological and demographic characteristics that can lead to
335 virus transmission, whilst at the same time keeping the model simple enough such that it is possible to parameterise. Here we explored that premise by considering four nested models of increasing complexity, where we accounted for random (non-distance based) spread, local spread of the disease, increased transmission owing to rice paddy density and increased transmission owing to
340 presence of intensive duck farming.

Our results indicate that regions with a highly intensive duck farming industry are most likely to be infected during the H5N1 outbreak - our model that includes this factor provides the closest fit to the observational data from the 2004 outbreak. Our analysis reveals locations where the probability of finding
345 infected animals is substantially higher, which can be used to identify specific areas where surveillance would be more beneficial than others in terms of increasing the probability of detecting the infection at early stages.

In addition, we have explored the ability of models of this nature to be used in real time during ongoing outbreaks. Many infectious disease models
350 are used retrospectively to determine the risks associated with spread and the impact of control for previous outbreaks. Whilst this has value, it is also crucial

to determine whether these models can be used in real time to advise policy makers during the course of an epidemic. Our results suggest that, during the early stages of HPAI outbreaks, there is significant uncertainty, owing to partial reporting of cases and potential for undetected infections over large areas. Models such as the one we describe here therefore only have limited predictive power in the very early stages of HPAI epidemics. However, as more data are accrued, the uncertainty in predictions decreases, such that after the first few weeks, we are able to accurately predict both the size and the spatial extent of the outbreak. This is highly beneficial in terms of being able to inform targeted surveillance policies to improve detection.

Despite this uncertainty, the models perform significantly better in determining the optimal control policies that should be deployed, even in the very early stages of the outbreak. This is promising, in that, despite over predictions of spread at epidemic onset, models are able to inform how to target interventions to reduce the risk of spread in the future. The caveat to this is the uncertainty in the impact of some control policies in the early stages of outbreaks - whilst the model accurately predicts that high radius ring culling is the preferred policy even when only using the first 10 days of the outbreak to parameterise the model, the duration of outbreaks is significantly overpredicted for IP only culling and vaccination policies.

For the future, it is important to explore mechanisms to improve the predictive power of infectious disease models during the early stages of disease outbreaks. It is simply not possible for policy makers to wait for uncertainty to resolve after the first few weeks before employing an intervention and therefore seeking methods to reduce uncertainty in model predictions at epidemic onset is vital. One area for future research would be to use data from previous outbreaks to provide more informed priors for epidemiological parameters. This would enable us to explore whether this can better establish the likelihood of spread in the first few days of a new epidemic when the outputs of models could potentially have the most significant impact upon reducing the spread of the disease.

5. Supplementary Material

5.1. Data

385 The demographic data describing poultry industry in Thailand during the 2004 outbreak has the following characteristics: number of subdistricts: 7,416; number of flocks: 3,303,160; number of chicken: 180,725,929; number of ducks: 19,745,049. This gives that the national poultry density during the outbreak was 395 *birds/km²*.

390 5.2. Mathematical models

We formulate four mathematical models with increasing complexity: a random process model (Model A), a spatial model (Model B), a spatial model incorporating rice (Model C), and a spatial model incorporating duck intensity (Model D).

395 Any district k is described by a number of flocks, F_k , number of susceptible flocks at time t , $S_k(t)$, and number of in infectious flocks at time t , $I_k(t)$. Then any flock i in a subdistrict k is described by a number of chicken, $NC_{k,i}$ and number of ducks, $ND_{k,i}$. For further calculations, we normalise number of chicken and ducks as follows:

$$\begin{aligned} nc_{k,i} &= \left(\frac{NC_{k,i}}{NC} \right)^{p_C}, \\ nd_{k,i} &= \left(\frac{ND_{k,i}}{ND} \right)^{p_D}, \end{aligned}$$

400 where power laws p_C and p_D account for virus transmission differences in the two species.

We can write a general expression for the force of infection to a susceptible flock, k :

$$\begin{aligned} \lambda_k(t) &= (1 + \alpha_R H(rc_k, \epsilon_{rc}, n_{rc}) + \alpha_D H(dc_k, \epsilon_{dc}, n_{dc})) \\ &\quad \times \left(\delta(1 + n_I(t)) + S_k(t) \left(\beta_W W_k I_k(t) + \beta_B \sum_{i \in \mathcal{NB}_k} B_{k,i} I_i(t) \right) \right) \quad (1) \\ H(x, \epsilon, n) &= \frac{1}{(x/\epsilon)^n + 1}, \quad (2) \end{aligned}$$

Here rc_k is the fraction of the area of the subdistrict occupied by rice paddies;
 405 the intensity of duck industry is formalised as $dc_k = \log_{10}(\max_{i=1..F_k}(ND_{k,i}) + 1)$;
 $n_I(t)$ is the number of infected flocks at time t , whilst β_W and β_B are the rates
 of spatial transmission within and between subdistricts, whilst W_k and $B_{k,i}$ are
 the rates of within-subdistrict transmission and local cross-border transmission.
 The term δ is spatially independent and allows transmission between any pair
 410 of subdistricts. We refer to δ from here onwards as the "background" term.

The rate of within-subdistrict transmission for subdistricts with a number
 of flocks $F_k > 1$ is given by

$$W_k = \frac{d_k}{F_k^2(F_k - 1)} \sum_{i=1}^{F_k} \sum_{j=1, j \neq i}^{F_k} \left((nc_{k,i} + \xi nd_{k,i}) (nc_{k,i} + \xi nd_{k,i}) \right), \quad (3)$$

where d_k is the density of flocks (i.e. number of flocks divided per subdistrict
 area) and F_k is a number of flocks in a subdistrict k .

415 The rate of local cross-border transmission between subdistricts with a number
 of flocks $F_k > 0$ and neighbouring subdistricts $\{i | i \in \mathcal{NB}_k, F_i > 0\}$ is given
 by

$$B_{k,i} = \frac{d_i L_{k,i} \Omega_{k,i}}{A_k F_k^2 F_i} \sum_{kk=1}^{F_k} \sum_{ii=1}^{F_i} \left((nc_{k,kk} + \xi nd_{k,kk}) (nc_{i,ii} + \xi nd_{i,ii}) \right), \quad (4)$$

where $L_{k,i}$ is the length of a shared boundary between neighbour districts k and
 i and A_k is the area of subdistrict k .

420 The term $\Omega_{k,i}$ is a scaling factor for between-subdistrict transmission in-
 formed by the mean distance between neighbouring subdistricts taking into
 account the size of each subdistrict, following the approach of Buhnerkempe
 et al. 2014 such that:

$$\Omega_{ki,i} = \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \kappa(\sqrt{x^2 + (y_i + y_k)^2}) dy_k dy_i dx,$$

$$\kappa(r) = \frac{\zeta}{1 + \left(\frac{r}{\theta}\right)^\alpha},$$

where α and θ are the shape and scale parameters of the distance-dependent
 425 local spread kernel respectively (Buhnerkempe et al. 2014). The normalisation

constant ζ is defined such that

$$\int_0^{\infty} 2\pi\kappa(r)dr = 1.$$

5.3. Bayesian approach for parameter estimation

We use a Bayesian approach to fit the meta-population model to the outbreak data (Jewell et al. 2009). The likelihood function has the following expression:

$$L = \prod_{k \in \mathbf{I}_T} \prod_{i=1, t_{ki} \neq t_0}^{n_k} (1 - \exp(-\lambda_k(t_{ki})))^{\hat{I}_k(t_{ki})} \left(\exp\left(-\sum_{t_0+1}^{t_{ki}-1} \lambda_k(t)\right) \right)^{\hat{I}_k(t_{ki})} \\ \times \prod_{k \in \mathbf{I}_T} \left(\exp\left(-\sum_{t_0+1}^T \lambda_k(t)\right) \right)^{S_k(T)} \prod_{k \in \mathbf{S}_T} \left(\exp\left(-\sum_{t_0+1}^T \lambda_k(t)\right) \right)^{F_k}$$

430 where n_k is the total number of infection events in a subdistrict k and t_{ki} is the time of the i^{th} infection event in a subdistrict k . We introduce $\hat{I}_k(t)$ to be the number of flocks that have been infected at time t , and note that this is different from $I_k(t)$ (i.e. the number of infectious flocks at time t). For each of the parameters, we update the parameter value using a random walk
435 Metropolis algorithm with a Gaussian proposal: $\theta^* \sim N(\theta, \sigma)$, where θ and θ^* are parameter vectors, and σ is a covariance matrix.

The proposed set of updated parameters θ^* is then accepted with probability

$$\min \left\{ 1, \frac{L(\theta^*) q(\theta, \theta, \sigma)}{L(\theta) q(\theta, \theta^*, \sigma)} \right\}, \quad (5)$$

where $q(x, y, s)$ is a normal PDF value at x with mean y and covariance matrix s .

440 As a pre-emptive culling policy within a 1km radius of flocks reporting infection was introduced during the outbreak, we simulate this by reducing the number of susceptible flocks as a fraction given by the ratio of the area of a circle with 1km radius to the total area of the subdistrict in question. We assume that there was no re-population of any flocks after culling prior to the end of
445 the epidemic.

5.4. Model simulation

Posterior samples of parameters are used to generate predicted outbreaks. At every time t of the simulation (time step $\Delta t = 1$ day), every susceptible flock k has a probability of becoming infected:

$$p_k(t) = 1 - \exp(-\lambda_k(t)), \quad (6)$$

450 where the infection pressure depends on a particular model. We start the outbreak by seeding infection in the six subdistricts in Thailand that had reported infection during the first three days of the outbreak. The seed cases were the same for all simulations. The outbreak was assumed to end after seven days without new infected cases (or alternatively the simulation was terminated after
455 2000 days).

Modeling Intervention strategies

With vaccine capacity γ , the demographic factors on the infection pressure then are reduced so that:

$$\begin{aligned} nc_{k,i}^v &= \left(\left(1 - \frac{\gamma}{100} \right) \frac{NC_{k,i}}{NC} \right)^{p_C}, \\ nd_{k,i}^v &= \left(\left(1 - \frac{\gamma}{100} \right) \frac{ND_{k,i}}{ND} \right)^{p_C}. \end{aligned}$$

Financial disclosure

460 MJT and RR are supported by the Engineering and Physical Sciences Research Council and the Biotechnology and Biological Sciences Research Council [grant numbers EP/P511079/1 and BB/K010972/4]. XX, MG and MJT are supported by the National Institutes of Health (NIH grant 1R01AI101028-02A1). The funders had no role in study design, data collection and analysis,
465 decision to publish, or preparation of the manuscript.

Data availability

Data are available from the Department of Livestock Development in Thailand.

Acknowledgments

470 We thank Will Probert and Ed Hill for useful comments on this manuscript.

Author Contributions

The author(s) have made the following declarations about their contributions: Conceived and designed the experiments: MT, RR, CJ. Performed the experiments: RR, MT. Analyzed the data: RR, WT. Contributed reagents/materials/analysis
475 tools: TB, MG, XX, GZ, WT. Wrote the paper: RR, MT, MK, TB, MG, GZ, XX, CJ, WT.

References

- Artois, J., Lai, S., Feng, L., Jiang, H., Zhou, H., Li, X., Dhingra, M.S., Linard, C., Nicolas, G., Xiao, X., Robinson, T.P., Yu, H., Gilbert, M., 2017. H7n9
480 and h5n1 avian influenza suitability models for china: accounting for new poultry and live-poultry markets distribution data. *Stochastic Environmental Research Risk Assessment* 31, 393–402. doi:10.1007/s00477-016-1362-z.
- Auewarakul, P., 2008. The past and present threat of avian influenza in thailand, in: Lu, Y., Essex, M., Roberts, B. (Eds.), *Emerging Infections in Asia*.
485 Springer, pp. 31–44.
- de Bruin, E., Zhang, X., Ke, C., Sikkema, R., Koopmans, M., 2017. Serological evidence for exposure to avian influenza viruses within poultry workers in southern China. *ZOOZOSES AND PUBLIC HEALTH* 64, e51–e59. doi:10.1111/zph.12346}.
- 490 Buhnerkempe, M.G., Tildesley, M.J., Lindstrom, T., Grear, D.A., Portacci, K., Miller, R.S., Lombard, J.E., Werkman, M., Keeling, M.J., Wennergren, U., Webb, C.T., 2014. The impact of movements and animal density on continental scale cattle disease outbreaks in the united states. *PLOS ONE* 9. doi:10.1371/journal.pone.0091724.

- 495 Gilbert, M., Chaitaweesub, P., Parakarnawongsa, T., Premashtira, S., Tiensin, T., Kalpravidh, W., Wagner, H., Slingenbergh, J., 2006. Free-grazing ducks and highly pathogenic avian influenza, Thailand. *EMERGING INFECTIOUS DISEASES* 12, 227–234. doi:{10.3201/eid1202.050640}.
- Gilbert, M., Golding, N., Zhou, H., Wint, G.R.W., Robinson, T.P., Tatem, A.J., Lai, S., Zhou, S., Jiang, H., Guo, D., Huang, Z., Messina, J.P., Xiao, X., Linard, C., Van Boeckel, T.P., Martin, V., Bhatt, S., Gething, P.W., Farrar, J.J., Hay, S.I., Yu, H., 2014. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. *NATURE COMMUNICATIONS* 5. doi:{10.1038/ncomms5116}.
- 505 Gilbert, M., Newman, S.H., Takekawa, J.Y., Loth, L., Biradar, C., Prosser, D.J., Balachandran, S., Rao, M.V.S., Mundkur, T., Yan, B., Xing, Z., Hou, Y., Batbayar, N., Natsagdorj, T., Hogerwerf, L., Slingenbergh, J., Xiao, X., 2010. Flying Over an Infected Landscape: Distribution of Highly Pathogenic Avian Influenza H5N1 Risk in South Asia and Satellite Tracking of Wild Waterfowl. *ECOHEALTH* 7, 448–458. doi:{10.1007/s10393-010-0672-8}.
- Gilbert, M., Xiao, X., Chaitaweesub, P., Kalpravidh, W., Premashtira, S., Boles, S., Slingenbergh, J., 2007. Avian influenza, domestic ducks and rice agriculture in thailand. *Agriculture, Ecosystems and Environment* 119, 409–415. doi:10.1016/j.agee.2006.09.001.
- 515 Gilbert, M., Xiao, X., Pfeiffer, D.U., Epprecht, M., Boles, S., Czarnecki, C., Chaitaweesub, P., Kalpravidh, W., Minh, P.Q., Otte, M.J., Martin, V., Slingenbergh, J., 2008. Mapping h5n1 highly pathogenic avian influenza risk in southeast asia. *PNAS* 105, 4769–4774. doi:10.1073/pnas.0710581105.
- Hill, E.M., House, T., Dhingra, M.S., Kalpravidh, W., Morzaria, S., Osmani, M.G., Yamage, M., Xiao, X., Gilbert, M., Tildesley, M.J., 2017. Modelling H5N1 in Bangladesh across spatial scales: Model complexity and zoonotic transmission risk. *EPIDEMICS* 20, 37–55. doi:{10.1016/j.epidem.2017.02.007}.

- Jeong, O.M., Kim, M.C., Kim, M.J., Kang, H.M., Kim, H.R., Kim, Y.J., Joh,
525 S.J., Kwon, J.H., Lee, Y.J., 2009. Experimental infection of chickens, ducks
and quails with the highly pathogenic H5N1 avian influenza virus. *JOURN-
AL OF VETERINARY SCIENCE* 10, 53–60. doi:{10.4142/jvs.2009.10.
1.53}.
- Jewell, C.P., Keeling, M.J., Roberts, G.O., 2009. Predicting undetected infec-
530 tions during the 2007 foot-and-mouth disease outbreak. *Journal of The Royal
Society Interface* 6, 1145–1151. doi:10.1098/rsif.2008.0433.
- Jewell, C.P., Kypraios, T., Christley, R.M., Roberts, G., 2009. A novel approach
to real-time risk prediction for emerging infectious diseases: A case study in
Avian Influenza H5N1. *PREVENTIVE VETERINARY MEDICINE* 91, 19–
535 28. doi:{10.1016/j.prevetmed.2009.05.019}.
- Minh, P., Stevenson, M., Jewell, C., French, N., Schauer, B., 2011. Spatio-
temporal analyses of highly pathogenic avian influenza H5N1 outbreaks
in the Mekong River Delta, Vietnam, 2009. *Spatial and Spatio-temporal
Epidemiology* 2, 49–57. doi:{10.1016/j.sste.2010.11.001}.
- 540 Paul, M.C., Gilbert, M., Desvaux, S., Andriamanivo, H.R., Peyre, M., Khong,
N.V., Thanapongtharm, W., Chevalier, V., 2014. Agro-Environmental De-
terminants of Avian Influenza Circulation: A Multisite Study in Thailand,
Vietnam and Madagascar. *PLOS ONE* 9. doi:{10.1371/journal.pone.
0101958}.
- 545 Peiris, J.S.M., de Jong, M.D., Guan, Y., 2007. Avian influenza virus (H5N1):
a threat to human health. *Clin Microbiol Rev.* 20, 243+. doi:{10.1128/CMR.
.00037–06}.
- Probert, W., Shea, K., Fonnesebeck, C., Runge, C., Carpenter, T., Durf, S.,
Garner, M., Harvey, N., Stevenson, M., Webb, C., Werkman, M., Tildesley,
550 M., Ferrari, M., 2016. Decision-making for foot-and-mouth disease control:
Objectives matter. *Epidemics* 15, 10–19. doi:{10.1016/j.epidem.2015.11.
002}.

- Sharkey, K.J., Bowers, R.G., Morgan, K.L., Robinson, S.E., Christley, R.M.,
2008. Epidemiological consequences of an incursion of highly pathogenic H5N1
555 avian influenza into the British poultry flock. *PROCEEDINGS OF THE
ROYAL SOCIETY B-BIOLOGICAL SCIENCES* 275, 19–28. doi:{10.1098/
rsjb.2007.1100}.
- Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A., 2002. Bayesian mea-
sures of model complexity and fit. *JOURNAL OF THE ROYAL STATISTI-
560 CAL SOCIETY SERIES B-STATISTICAL METHODOLOGY* 64, 583–616.
doi:{10.1111/1467-9868.00353}.
- Ssematimba, A., Hagenaars, T.J., de Jong, M.C.M., 2012. Modelling the Wind-
Borne Spread of Highly Pathogenic Avian Influenza Virus between Farms.
PLOS ONE 7. doi:{10.1371/journal.pone.0031114}.
- 565 Stegeman, J.A., Bouma, A., de Jong, M.C.M., 2011. Epidemiological mod-
els to assist the management of highly pathogenic avian influenza. *Revue
scientifique et technique (International Office of Epizootics)* 30, 571–579.
- Tiensin, T., Chaitaweesub, P., Songserm, T., Chaisingh, A., Hoonsuwan,
W., Buranathai, C., Parakamawongsa, T., Premashthira, S., Amonsin, A.,
570 Gilbert, M., Nielen, M., Stegeman, A., 2005. Highly pathogenic avian in-
fluenza h5n1, thailand, 2004. *Emerging Infectious Diseases* 1, 1664–1672.
- Truscott, J., Garske, T., Chis-Ster, I., Guitian, J., Pfeiffer, D., Snow, L., Wile-
smith, J., Ferguson, N.M., Ghani, A.C., 2007. Control of a highly pathogenic
H5N1 avian influenza outbreak in the GB poultry flock. *PROCEEDINGS
575 OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES* 274, 2287–2295.
doi:{10.1098/rsjb.2007.0542}.
- Walker, P., Cauchemez, S., Hartemink, N., Tiensin, T., Ghani, A.C., 2012.
Outbreaks of h5n1 in poultry in thailand: the relative role of poultry pro-
duction types in sustaining transmission and the impact of active surveil-
580 lance in control. *Journal of The Royal Society Interface* 9, 1836–1845.
doi:10.1098/rsif.2012.0022.

Xiao, X., Boles, S., Frohling, S., Li, C., Babu, J., Salas, W., Moore, B.,
2006. Mapping paddy rice agriculture in south and southeast asia using
multi-temporal modis images. *Remote Sensing of Environment* 100, 95–113.
585 doi:10.1016/j.rse.2005.10.004.

Yu, Z., Song, Y., Zhou, H., Xu, X., Hu, Q., Wu, H., Zhang, A., Zhou, Y., Chen,
J., Dan, H., Luo, Q., Li, X., Chen, H., Jin, M., 2007. Avian influenza (H5N1)
virus in waterfowl and chickens, central China. *EMERGING INFECTIOUS
DISEASES* 13, 772–775. doi:{10.3201/eid1305.061209}.

Table 1: Estimated parameters with posterior mean and 95% CI for random process model (A), spatial model (B), spatial rice model (C), and spatial duck model (D).

Parameter	Model (A)	Model (B)	Model (C)	Model (D)
$\delta \times 10^{-8}$	2.55 (2.43, 2.68)	1.23 (1.14, 1.30)	1.23 (1.16, 1.30)	0.77 (0.69, 0.84)
$\beta_W \times 10^{-6}$		3.6 (2.85, 4.52)	3.45 (3.22, 3.65)	0.99 (0.76, 1.12)
$\beta_B \times 10^{-5}$		1.71 (1.38, 2.22)	1.08 (1.03, 1.13)	0.41 (0.34, 0.52)
p_C		0.38 (0.29, 0.47)	0.34 (0.30, 0.38)	0.41 (0.32, 0.48)
p_D		0.26 (0.22, 0.31)	0.21 (0.18, 0.24)	0.03 (0.01, 0.08)
ξ		3.17 (2.28, 3.92)	3.63 (3.43, 3.92)	2.78 (2.31, 3.19)
α_D			0.71 (0.65, 0.74)	
ϵ_D			0.61 (0.58, 0.63)	
n_D			3.25 (3.14, 3.34)	
α_R				13.5 (11.2, 16.7)
ϵ_R				3.45 (3.23, 3.71)
n_R				5.92 (5.03, 6.86)

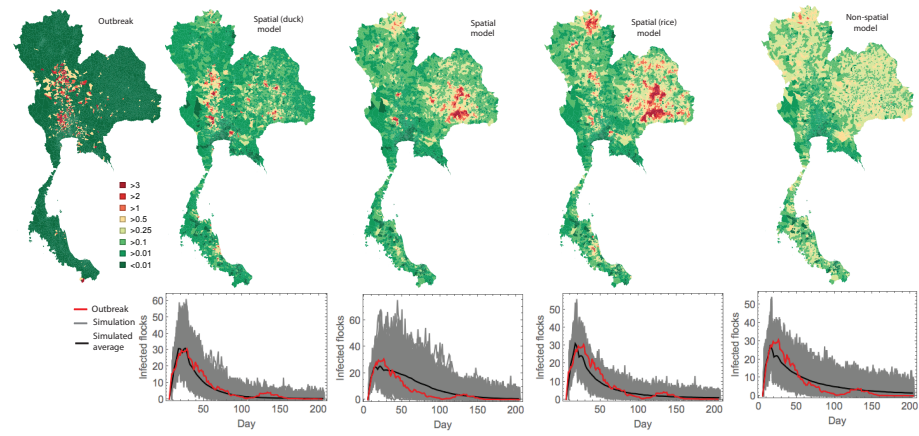


Figure 1: Predicted and observed outbreak distributions. Maps show the mean number of infected flocks in each subdistrict.

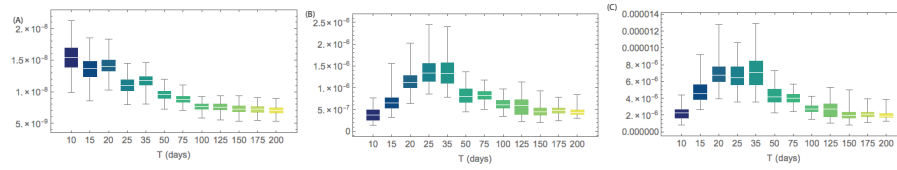


Figure 2: Distribution of fitted parameters for the spatial (ducks) model D: δ (A), β_W (B), and β_B (C). Sequential parameter estimation performed for outbreak data censored at time T days shown on x -axis.

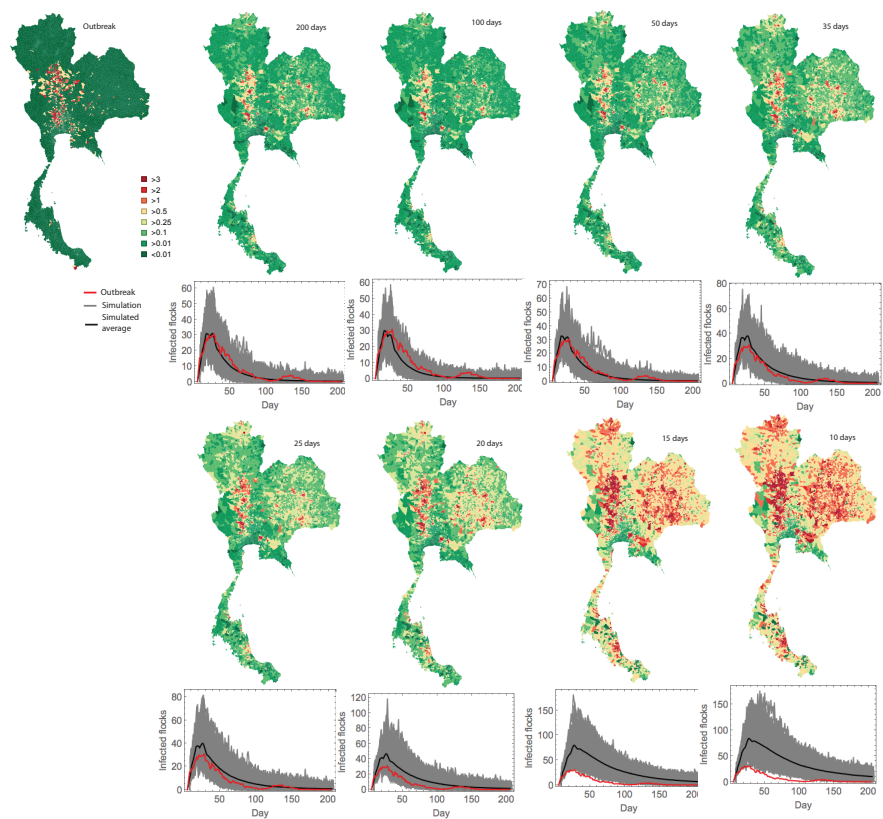


Figure 3: Predicted and observed spatial and temporal outbreak distributions. Each map show the mean number of infected flocks in each subdistrict as predicted by the model fitted at the given time in the outbreak. For each temporal profile, the observed 2004 outbreak is given by the red line, whilst the solid black line gives the mean model prediction.

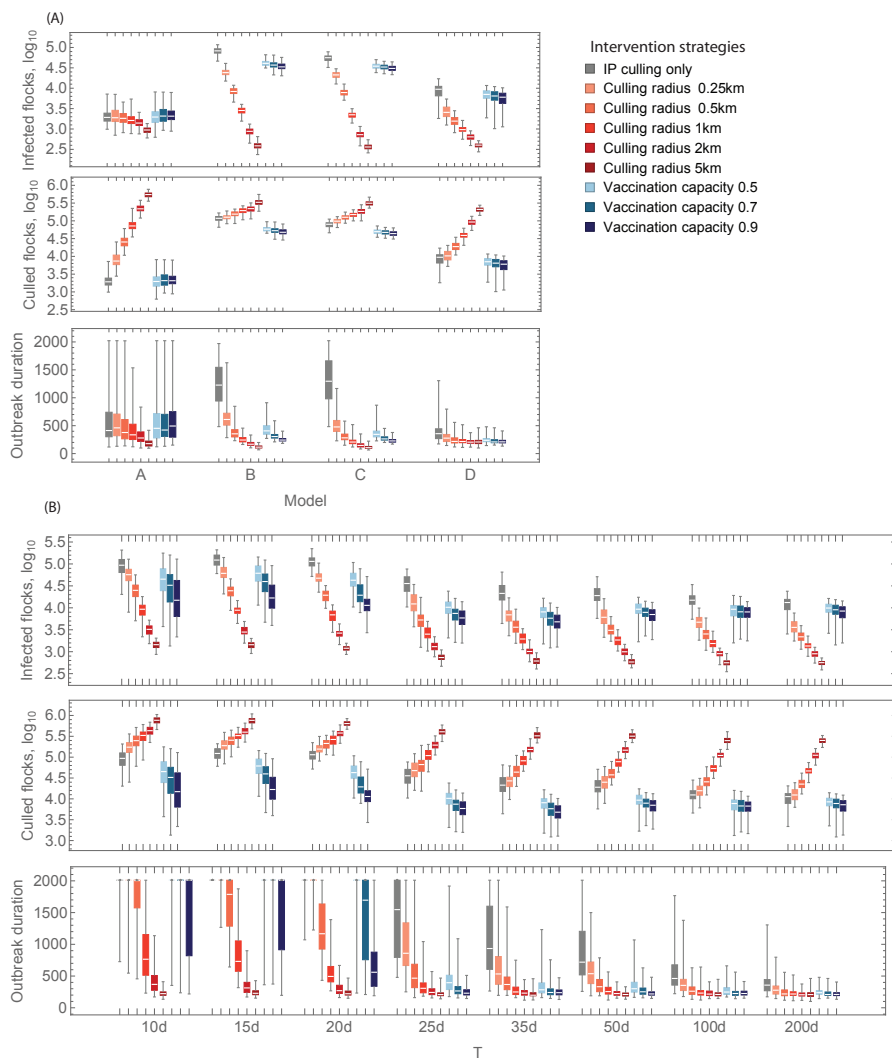


Figure 4: **Effect of local controls.** Projected number of infected flocks, number of culled flocks and duration of outbreak for (A) the four models, (B) for sequential parameter estimation using Model D with outbreak data censored at time T days.