



HAL
open science

Inferring ASF transmission in domestic pigs and wild boars using a paired model iterative approach

G. Beaunée, François Deslandes, Elisabeta Vergu

► **To cite this version:**

G. Beaunée, François Deslandes, Elisabeta Vergu. Inferring ASF transmission in domestic pigs and wild boars using a paired model iterative approach. *Epidemics*, 2023, 42, pp.100665. 10.1016/j.epidem.2023.100665 . hal-03993608

HAL Id: hal-03993608

<https://hal.inrae.fr/hal-03993608v1>

Submitted on 6 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Inferring ASF transmission in domestic pigs and wild boars using a paired model iterative approach

G. Beaunée^{a,*}, F. Deslandes^b, E. Vergu^b

^a Oniris, INRAE, BIOEPAR, 44300, Nantes, France

^b Université Paris-Saclay, INRAE, MalAGE, 78350, Jouy-en-Josas, France

ARTICLE INFO

Keywords:

African swine fever
Inference
Mechanistic model
Stochastic model
Data-driven model

ABSTRACT

The rapid spread of African swine fever (ASF) in recent years has once again raised awareness of the need to improve our preparedness in preventing and managing outbreaks, for which modelling-based forecasts can play an important role. This is even more important in the case of a disease such as ASF, involving several types of hosts, characterised by a high case-fatality rate and for which there is currently no treatment or vaccine. Within the framework of the ASF challenge, we proposed a modelling approach based on a stochastic mechanistic model and an inference procedure to estimate key transmission parameters from provided data (incomplete and noisy) and generate forecasts for unobserved time horizons. The model is partly data driven and composed of two modules, corresponding to epidemic and demographic dynamics in domestic pig and wild boar (WB) populations, interconnected through the networks of animal trade and/or spatial proximity. The inference consists in an iterative procedure, alternating between the two models and based on a criterion optimisation. Estimates of transmission and detection parameters appeared to be of similar magnitude for each of the three periods of the challenge, except for the transmission rates in WB population through contact with infectious individuals and carcasses, higher during the first period. The predicted number of infected domestic pig farms was in overall agreement with the data. The proportion of positive tested WB was overestimated, but with a trend close to that observed in the data. Comparison of the spatial simulated and observed distributions of detected cases also showed an overestimation of the spread of the pathogen within WB metapopulation. Beyond the quantitative results and the inherent difficulties of real-time forecasting, we built a modelling framework that is flexible enough to accommodate changes in transmission processes and control measures that may occur during an epidemic emergency.

1. Introduction

Epidemic modelling and forecasting will play a key role in our preparedness to face the next emerging infectious disease outbreaks (Heesterbeek et al., 2015). Indeed, they have received particular attention during the recent public health emergencies. Whether it was the Ebola outbreak, influenza or the Covid-19 pandemic, these events highlighted the significant value of infectious disease forecasts in improving our understanding of spread dynamics and supporting decision making (Chretien et al., 2015; Moghadas et al., 2009; Brooks-Pollock et al., 2021). A reliable and timely assessment of an infectious disease outbreak, particularly in the early stages of the spread, enables decision-makers to make rapid and appropriate decisions needed to ensure epidemic control (Chowell et al., 2016; Kucharski et al., 2020; James et al., 2021). The same applies to animal epidemiology, where modelling has been widely used to develop herd management protocols

and assess surveillance and control strategies (Brooks-Pollock et al., 2021; Ezanno et al., 2020).

The rapid spread of the African swine fever (ASF) in recent years (Vergne et al., 2017) has once again raised awareness of the need to improve our preparedness to prevent and manage the spread in a given area of a pathogen already known but not yet present in the area, as well as the emergence of a new pathogen. It is therefore crucial to identify the critical points that could hamper our ability to respond quickly to an infectious disease outbreak in farm animals. This is all the more important in the case of a disease such as ASF, which is characterised by a high case-fatality rate and for which there is currently no treatment or vaccine (Dixon et al., 2019). Furthermore, the involvement of wild boar, themselves clinically affected by this disease, in the introduction and spread of the ASF virus between farms, further complicates the picture.

While epidemiological models are now recognised as a valuable tool to inform decision-making (Alahmadi et al., 2020), the use of

* Corresponding author.

E-mail address: gael.beaunee@inrae.fr (G. Beaunée).

model predictions is often challenging. This is partly due to the fact that the quality of the predictions is highly dependent on the model parameterisation, which in turn requires prior knowledge and sufficient available data of fairly good quality to be integrated. Indeed, data on the dynamics of infectious diseases are often scarce and incomplete due to the nature of the epidemic processes partially observed.

Epidemiological modelling has benefited from recent advances in computational power on the one hand and inference methods on the other (Sisson et al., 2018), allowing the development and calibration of large-scale disease spread models (Brooks-Pollock et al., 2014), and to perform a large number of simulations in a short time. However, progress remains to be made for model-based forecasting to become a fully operational tool. Partial lack of knowledge related in particular to limited data availability (Moran et al., 2016) is one of the potential factors that could hamper the effectiveness of model-based forecasting. Another would be related to the time inherent in research for the development of relevant models, which is sometimes insufficient in view of the responsiveness that decision-making authorities need to show to deal with health emergency or crisis.

The ASF modelling challenge (Picault et al., 2022) allowed teams of modellers to improve their readiness in terms of developing and calibrating models in a limited time frame when facing emerging infectious threats, and to assess their ability to advise policy makers in a timely manner. Specifically, for our team, the motivations for participating in this challenge were: (i) to test our ability to propose a relevant propagation model during the emergence of a new pathogen (as none of us had worked on ASF before), (ii) to improve our ability to handle new data on an outbreak and use it to calibrate models, and (iii) to evaluate how our experience in modelling infectious diseases of livestock helped us in the analysis of a different host-pathogen system.

Here, we present the modelling framework, model fits and forecasts that we generated as part of the blinded forecasting ASF challenge conducted in 2020. The challenge was based on three periods, each providing new synthetic data generated with a simulation model describing an ASF epidemic spreading among populations of domestic pigs and wild boar, on an island with 25 administrative units, and its own set of questions. We mainly focus on the methodological framework we developed and the results obtained. We also highlight some lessons learned from this experience.

2. Methods

The approach we considered is based on a stochastic mechanistic model consisting of two parts, hereafter referred to as modules (one for pigs and the other for wild boar related processes). The main reasons for splitting the model into two modules were: (i) to be able to use only one given the question and available information, (ii) to separate the computer implementation and therefore speed up the coding step and debugging, (iii) to provide the possibility of using separate inference methods for each module to implement procedures specific to the context and available data, and (iv) to provide a failure tolerant approach that allows predictions using only one of the two modules.

The developed approach is partly data-driven. This choice was mainly motivated by the lack of information to implement and calibrate some processes with a mechanistic approach.

In the following, we first describe the structures of the two modules, the domestic pig farms and wild boar metapopulations (Section 2.1), and how they are linked. Then, we present the modelling framework and its practical implementation (Section 2.2), followed by the inference procedure (Section 2.3) and prediction (Section 2.4). Finally, we give a brief description of the changes we made to our approach during the challenge.

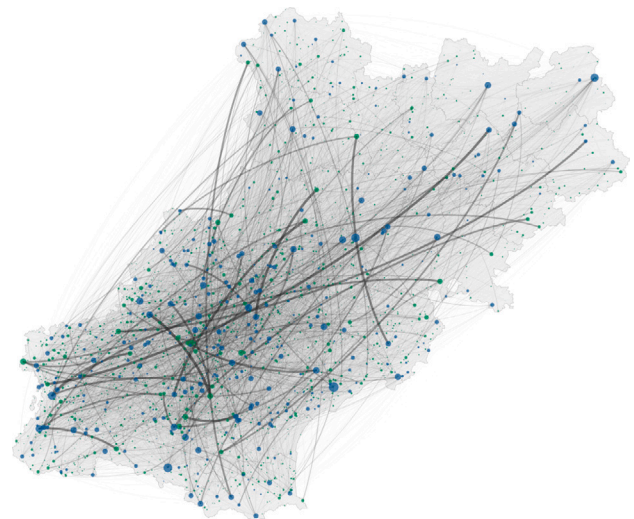


Fig. 1. Visualisation of the contact network through trade movements between domestic pig farms.

2.1. Metapopulation structure

Epidemic and demographic dynamics in the study area are described as a set of two metapopulations of epidemiological units, corresponding to domestic pig (DP) farms and wild-boar (WB) populations, respectively. Different levels of interactions are represented, some of which are specific to the units of the same metapopulation, and others that connect the units of both metapopulations.

The representation of each of the metapopulations are based on the data provided. These data are (i) the farm characteristics and the history of trade movements for the domestic pig farms, and (ii) the description of the land use (as urban, agricultural or forest) and the census of the numbers of animals hunted (hunting bag).

2.1.1. Domestic pig farms

The underlying connection structure of the metapopulation of domestic pig farms is naturally defined as the network of trade movements between farms (see Fig. 1). Within this metapopulation, each of the 4540 epidemiological units (node) corresponds to a farm with specific characteristics, which allow us to assign transmission routes through which it can get infected or infect other nodes. Thus, each farm: is either a breeder farm or a finisher farm or a breeder-finisher farm; has an activity listed as commercial or backyard farm; and it is known whether its animals have access to the outside or not.

2.1.2. Wild boar populations

The only information available regarding the wild boar population on the island is the hunting bag for the 25 administrative units, and the fact that during the hunting season about 50% of the wild boar are shot (*prop. WB shot*). The administrative units cover far too large areas to be suitable as epidemiological units of interest. It was therefore necessary to further subdivide the landscape in order to define the granularity of the wild boar metapopulation. We then used a hexagonal grid to discretise the landscape. The size of the tiles used varied according to the period of the challenge: the diameter of the circumscribed circle of the tile was 15,500 m for the first period, while the diameter of the inscribed circle was 15,000 m for the second and third periods, corresponding to 1045 and 854 tiles respectively.

Then, in order to allocate the wild boar to the epidemiological units (defined by the tiles) thus created, we used the available land-use data combined with the information provided by the challenge organisers. We then assumed that the spatial distribution of the wild

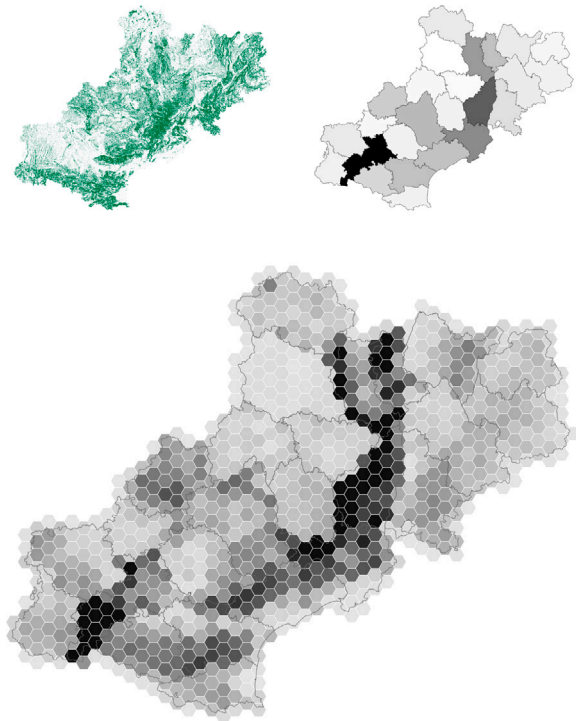


Fig. 2. Visualisation of the spatial density of wild boar using a hexagonal grid to delineate populations: distribution of forest landcover (top left), hunting bag (number of wild boar shot) by the administrative unit (top right), and the layout shown corresponds to the one used in periods 2 and 3, which consisted of 854 hexagonal tiles with an inscribed circle diameter of 15,000 m (bottom).

boar population was directly related to the proportion of forest within a tile, relative to the area of forest in the administrative unit to which the tile belongs (Supported by organiser data : WB spend 80% of their time in forest area). Thus, for a tile i located in the administrative unit u , the wild boar population size M_i is calibrated as: $M_i \sim \frac{forest_i}{forest_u} \times HuntingBag_u \times prop. WB shot$, where $forest_i$ and $forest_u$ are the forest areas in tile i and administrative unit u respectively, $HuntingBag_u$ the number of wild boar shot during the previous year in the administrative unit u , and $prop. WB shot$ the estimated proportion of wild boar shot during the hunting season (equal to 0.5). The resulting distribution of wild boar population in the tiles is shown in Fig. 2.

2.1.3. Contact network

The transmission routes between populations of domestic pigs and wild boar were described using a multi-layer network. This allows the description of several types of interactions, representing each transmission route as a network layer (the different interactions are illustrated in Fig. 3).

We considered two different types of potential transmission routes between the epidemiological units of the domestic pig metapopulation: through commercial movements and neighbourhood interactions. The data provided on trade movements can be naturally described as a temporal directed weighted network. The period covered starts 59 days before the detection of the first case until the end of each specific period of the challenge, which corresponds to a duration of 170 days for the last period. The neighbourhood network is defined by the spatial proximity: a link exists between two epidemiological units when the distance between them is less than or equal to 4 km, a value chosen to be of the same order of magnitude as the buffer defining the protection zone.

Concerning the wild boar metapopulation, only one potential transmission route was considered, defined by the spatial proximity between the epidemiological units: a link between two tiles exists as soon as they

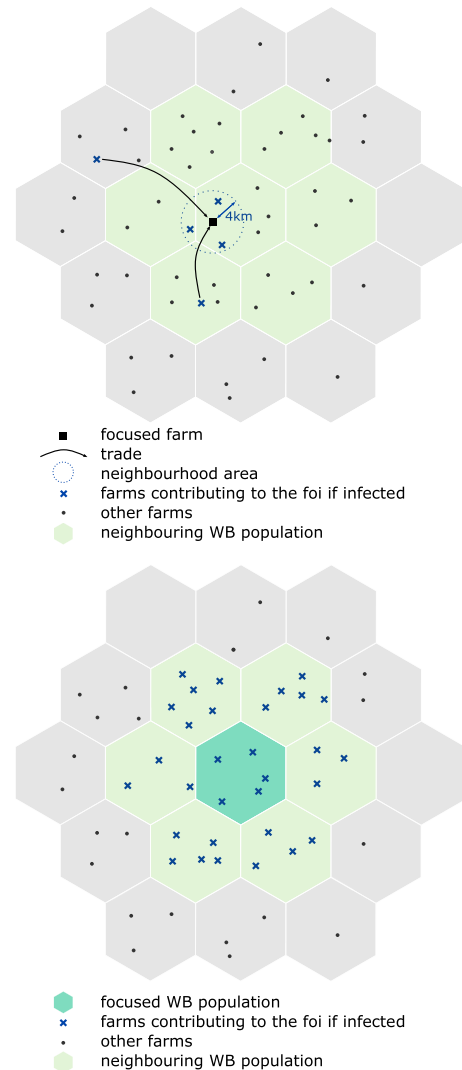


Fig. 3. Illustration of the interactions between populations of domestic pigs and wild boar, with (a) a focus on a domestic pig farm and (b) a focus on a wild boar population.

share a common border, therefore animals from one tile can come in contact with animals from directly adjacent tiles.

Regarding the contacts between the different epidemiological units of the metapopulations of domestic pigs on one hand and wild boar on the other, they are defined by two additional layers. For domestic pig farms, if the animals had access to the outside, contacts with wild boar were assumed possible in the tile where the farm is located and in the directly adjacent tiles, and thus corresponding links were generated. This hypothesis has been symmetrised for wild boar populations.

2.2. Model description

Our modelling approach is based on the combination of two stochastic mechanistic compartmental models, under a metapopulation framework. A brief description of the two models is given below.

2.2.1. Dynamics of domestic pig farms

Domestic pig farms are modelled by considering discretised health status, without explicitly describing intra-farm dynamics. The temporal evolution of farm health statuses includes 5 different states, S (Susceptible), E (Exposed), I_s (subclinically Infectious), I_c (clinically Infectious) and R (Removed), and transitions between these states (whose associated parameters are presented in Table 1).

Table 1
Transitions between the health states for the domestic pig model.

Processus		Rate
Infection	$(S \rightarrow S - 1, E \rightarrow E + 1)$	λ_D
Become infectious	$(E \rightarrow E - 1, I_s \rightarrow I_s + 1)$	δ_D
Onset of clinical signs	$(I_s \rightarrow I_s - 1, I_c \rightarrow I_c + 1)$	γ_D
Culling ^a	$(X \rightarrow X - 1, R \rightarrow R + 1)$	η_D
Repopulation ^a	$(R \rightarrow R - 1, S \rightarrow S + 1)$	τ_D

Note: X means any of the health statuses.

^aDeterministic transitions.

During the first and second periods of the challenge, apart from the *infection* which was implemented stochastically, all transitions were assumed deterministic. In the third (last) challenge period, only the *culling* and *repopulation* transitions were implemented deterministically. In the case of stochastic transitions, exponentially distributed durations were assumed. The average times spent in the *E*, *I_s* and *R* compartments, equal to $1/\delta_D$, $1/\gamma_D$ and $1/\tau_D$ respectively, are based on the information provided during the challenge. *Culling* events can occur under different conditions: upon detection of an infected farm and then occur 4 days after the farm has become *I_c*, or due to the application of control measures, and then occur the next day if the farm meets the regulations currently in place.

Regarding the *infection*, the force of infection for a domestic pig farm *i* is expressed as follows:

$$\lambda_D(i, t) = \underbrace{\lambda_D^t \sum_{k \in K_i(t)} \omega_k^i}_{\text{trade}} + \underbrace{\lambda_D^n N_i \sum_{l \in L_i^{dp}} \mathbb{1}_{\text{infect}} N_l}_{\text{neighbourhood}} + \underbrace{\alpha_i \lambda_D^{wb} N_i \sum_{j \in J_i^{dp}} \frac{I_{s_j}(t) + I_{c_j}(t)}{M_j(t)}}_{\text{contact with wild boar}}$$

where λ_D^t , λ_D^n and λ_D^{wb} are the transmission rates through trade, neighbourhood and contact with wild boar populations respectively, ω_k^i is the number of animals sold by farm *k* to farm *i* and N_i , N_l the sizes of farms *i* and *l*. $\mathbb{1}_{\text{infect}}$ is an indicator function equal to 1 when farm *l* is infected, 0 otherwise, α_i a boolean equal to 1 if animals of farm *i* have access to the outside, otherwise 0. I_{s_j} and I_{c_j} are the numbers of subclinical and clinical infectious animals in wild boar population *j* and M_j is the size of wild boar population *j*. $K_i(t)$ and L_i^{dp} represent the sets of infected domestic pig farms that exchange animals with domestic pig farm *i* or are located in its neighbourhood, respectively, and J_i^{dp} is the set of wild boar population in contact with herd *i*.

The status of domestic pig farms is data driven. A farm that is reported as infected will have its evolution set according to the reported infection date and following the status transitions described in Table 1. A farm that is not reported as infected, will have its status set to *S* until the final observation date minus the time spent in *E* and *I_s*. The status of farms not reported as infected will then be simulated according to the previously defined model.

2.2.2. Dynamics of wild boar populations

For wild boar populations, the within-population epidemic dynamics as well as the main components of the demographic dynamics are taken into account. The evolution of the health status of wild boar includes 6 different states: *S* (Susceptible), *E* (Exposed), *I_s* (subclinically Infectious), *I_c* (clinically Infectious), *D* (Dead) and *R* (Removed). The transitions between these states and the associated parameters are presented in Table 2.

All transitions are stochastic, assuming exponentially distributed duration, with a particularity for the *carcasses discovery* through passive and active search events, and *hunting*, which are partly data-driven when data available and simulated otherwise. Over the period for

Table 2
Transitions between the health states for the wild boar model.

Processus		Rate
Infection	$(S \rightarrow S - 1, E \rightarrow E + 1)$	λ_W
Become infectious	$(E \rightarrow E - 1, I_s \rightarrow I_s + 1)$	δ_W
Onset of clinical signs	$(I_s \rightarrow I_s - 1, I_c \rightarrow I_c + 1)$	γ_W
Death	$(I_c \rightarrow I_c - 1, D \rightarrow D + 1)$	μ_W
Carcasses no more infectious	$(D \rightarrow D - 1, R \rightarrow R + 1)$	ν_W
Carcasses discovery ^a	$(D \rightarrow D - 1, R \rightarrow R + 1)$	ρ_W
Hunting ^a	$(X \rightarrow X - 1, R \rightarrow R + 1)$	η_W

^aData-driven if data available and simulated otherwise.

which data are available, the number of carcasses found and animals hunted are calibrated on the data. In the latter case, the health status of the animals are randomly drawn according to the population composition. The values for the average time spent in *E*, *I_s*, *I_c* and *D* (if not discovered) compartments, $1/\delta_W$, $1/\gamma_W$, $1/\mu_W$ and $1/\nu_W$ respectively, are based on the information provided during the challenge. The force of *infection* for wild boar in population *i* is expressed as follows:

$$\lambda_W(i, t) = \underbrace{\lambda_W^{Int} \frac{I_{s_i}(t) + I_{c_i}(t)}{M_i(t)}}_{\text{within pop. } I_s+I_c} + \underbrace{\lambda_W^{Ext} \sum_{l \in L_i} \frac{I_{s_l}(t) + I_{c_l}(t)}{M_l(t)}}_{\text{neighbourhood } I_s+I_c} + \underbrace{\lambda_W^{Dint} D_i(t)}_{\text{within pop. } D} + \underbrace{\lambda_W^{Dext} \sum_{l \in L_i^{wb}} D_l(t)}_{\text{neighbourhood } D} + \underbrace{\lambda_W^{dp} \sum_{j \in J_i^{wb}} \mathbb{1}_{\text{infect}}}_{\text{contact with domestic pigs}}$$

where λ_W^{Int} , λ_W^{Ext} , λ_W^{Dint} , λ_W^{Dext} and λ_W^{dp} are the transmission rates through contacts with infectious wild boar and carcasses within and outside the population and through contacts with infected domestic pig farms in the neighbourhood. M_x is the total size and I_{s_x} , I_{c_x} and D_x are the numbers of subclinical and clinical infectious animals and carcasses in wild boar population *x* ($M_x = S_x + E_x + I_{s_x} + I_{c_x}$). $\mathbb{1}_{\text{infect}}$ is an indicator function equal to 1 when domestic pig farm *j* is infected, 0 otherwise. L_i^{wb} and J_i^{wb} are the sets of wild boar populations and domestic pig farms in the neighbourhood of wild boar population *i*.

By the second period, the control measures to be assessed involved the location of wild boar carcasses found by active and/or passive searches, as well as those of hunted animals that tested positive. We then implemented a random generation of geographical coordinates to simulate the spatial positions of the relevant animals within the perimeter of the tiles.

2.2.3. Control measures

Throughout the challenge, different modalities of control measures had to be tested. Around infected domestic pig farms, regulated zones were defined: “protection zone” as 3 km around an infected farm for 40 days, and “surveillance” as 10 km around an infected farm for 30 days.

During the first period, we had to assess the efficacy of implementing fences around the forest nearby the primary case, combined or not with overhunting.

For the second period, we had to assess the efficacy of new alternative measures: the culling of pig farms located less than 3 km from a detected pig farm, the culling of pig farms traced as having had contact with a detected pig farm less than 3 weeks before detection, the extension of the surveillance zone from 10 to 15 km, and the increase of the area of active search around detected wild boar carcasses from 1 to 2 km.

The culling of pig farms located less than 3 km from a wild boar carcass detected as infected was not tested during the second period but implemented during the third period.

The different layers of the contact network (defined in Section 2.1) can be altered by the implementation of the control measures. For example, fencing off the initial outbreak in wild boar resulted in the

removal of links between the epidemiological units inside and outside the area. The establishment of regulated zones around infected farms (protection and surveillance zones) led to movement bans.

2.2.4. Implementations

Both modules have been implemented independently, in C++ for computational efficiency reasons. In the case of the domestic pig module, in discrete time, the stochastic events follow a random sampling using binomial or exponential distributions. For the wild boar module, the tau-leap algorithm was used and continuous simulated trajectories were further discretised in time, with a daily time step, as for the domestic pig module, for synchronisation reasons. Due to time constraints, scheduling and task sharing were important considerations in order to provide timely forecasts. That is one of the reasons why this modelling framework has been designed this way, allowing simultaneous implementation of both modules, combined in a second step. Thus, the two modules can be used jointly through an interface linking them, or alone. The source code of the model can be found on the following git repository: <https://forgemia.inra.fr/gaelbeaunee/asf-challenge-inrae-team>.

2.3. Parameters inference

The inference approach implemented aimed to estimate the values of the different transmission rates to be further used for providing forecasts over four weeks following the end of the current period.

It is an iterative procedure alternating between the two models (Fig. 4). The first run of this inference procedure starts with an estimation of the transmission parameters of the wild boar epidemic dynamics, based on a predefined scenario of the epidemic dynamics of domestic pig farms, and which corresponds to the observed trajectories in the provided data. Based on the estimated values, it is then possible to define a scenario of the epidemic dynamics of wild boar populations using the model, which is then frozen for the estimation of the transmission parameters of the domestic pig epidemic dynamics. It is then possible to simulate, using the newly estimated parameters, a new predefined scenario of the epidemic dynamics of domestic pig farms, which can be used to re-estimate the transmission parameters of the epidemic dynamics of wild boar populations.

A predefined scenario is a complete description of epidemiological units status over time. This corresponds to the hidden state of the system. It is therefore assumed in the first run that we have a comprehensive observation of the health status of domestic pig farms. Thus, the initial state of the model is defined by the data provided and corresponds to the first detection of an infected domestic pig farm observed in the data.

Depending on the period, different estimation methods were used, all based on a common principle, i.e. optimising a simulation-based criterion. In this section, we briefly describe the different methodologies that were part of our calibration procedure.

2.3.1. Estimation of transmission rates for domestic pig farms

The data on the number of farms detected over time was used to estimate the specific transmission rates of infection of domestic pig farms: λ_D^i , λ_D^n and λ_D^{wb} .

Due to the continuous evolution of the information available on domestic pig farms during the challenge, an adaptive inference approach was preferred, in order to adapt the method used to the data provided for each period.

In the first period, due to the design of the model, the data available were deemed insufficient for estimating a non-zero between-farm transmission and did not allow for an estimate of the overall transmission rate. Only the transmission rate due to contact with wild boar populations (λ_D^{wb}) could be estimated, the other two specific transmission rates (through neighbourhood and trade movements) assumed to be null. Indeed, the data provided included only three infected farms between

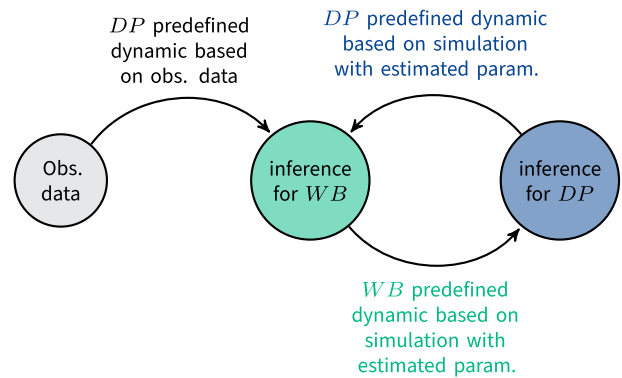


Fig. 4. Illustration of the whole inference procedure, as an iterative process alternating between estimation for each model, domestic pigs (DM) or wild boar (WB), using a predefined dynamic based on data or simulation for the other module (WB or DM, respectively).

which there were no links on the trade and neighbourhood layers of the contact network. The selected method was the minimisation (golden-search + interpolation) of a likelihood on the probability to be infected.

For the second period of the challenge, the provided data prevented the estimation of the transmission rates using the method designed for the first period. The time constraints did not allow us to undertake the implementation of a new estimation method, the transmission rates were then manually calibrated, based on estimation results obtained using data from the first period.

During the last period of the challenge, in order to overcome the problems encountered during the second period, we used a new estimation method, based on the minimisation, using the simplex algorithm (Nelder and Mead, 1965), of a least squares criterion (based on the total number of infected pig farms).

2.3.2. Estimation of transmission rates for wild boar populations

The wild boar model was fit to the data provided including incidence of carcasses found and the number of hunted wild boar testing positive. The five transmission rates, λ_W^{Int} , λ_W^{Ext} , λ_W^{Dint} , λ_W^{Dext} and λ_W^{dp} , were estimated for each period of the challenge.

Irrespective of the period, we used the same estimation method: the minimisation of a least squares criterion using an adapted simplex algorithm, but adjusting the design of the criterion used.

For the first period, the criterion was defined by combining normalised observations of the number of carcasses found and of the number of hunted wild boar testing positive, per five-day time window, differentiated according to three distinct zones. The zoning was defined on the basis of the observations as (i) all tiles where infected boars or domestic pigs were observed, (ii) all tiles adjacent to the previous ones, and (iii) all remaining tiles. Thus, the criterion (C_W^1) corresponds to the sum of squared differences between the observed and simulated number of infected carcasses found (y^c and \bar{y}^c , respectively) and the differences between the observed and simulated proportion of hunted wild boar testing positive (y^h and \bar{y}^h , respectively), per time window (k) and zones (z), computed as:

$$C_W^1 = \sum_{k,z} \frac{(y_{k,z}^c - \bar{y}_{k,z}^c)^2}{\hat{\sigma}_{(c),z}^2} + \frac{(y_{k,z}^h - \bar{y}_{k,z}^h)^2}{\hat{\sigma}_{(h),z}^2}$$

where $\hat{\sigma}_{(c),z}^2$ is the variance of the observed infected carcasses found and $\hat{\sigma}_{(h),z}^2$ the variance of the observed proportion of hunted wild boar testing positive, over all time windows for the zone z .

For the second and third periods, the criterion formula has been updated by splitting the data into two time periods: before and after the fence installation. Thus, the criterion (C_W^{23}) corresponds to the sum

of squared of differences per time window (k), zones (z) and fencing periods (p).

$$C_W^{23} = \sum_{k,z,p} \frac{(y_{k,z,p}^c - \bar{y}_{k,z,p}^c)^2}{\hat{\sigma}_{(c),z,p}^2} + \frac{(y_{k,z,p}^h - \bar{y}_{k,z,p}^h)^2}{\hat{\sigma}_{(h),z,p}^2}$$

where $\hat{\sigma}_{(c),z,p}^2$ is the variance of the observed infected carcasses found and $\hat{\sigma}_{(h),z,p}^2$ the variance of the observed proportion of hunted wild boar testing positive, over all time windows for the zone z and fencing period p .

To account for the stochasticity of the model in the evaluation of our criterion, the Nelder–Mead algorithm (Nelder and Mead, 1965) has been adapted according to the guidelines proposed by Barton and Ivey (1991). In short, it consists in re-evaluating the criterion at all points of the simplex after a shrink step, to select a small amplitude for the shrink step (coefficient of 0.9 instead of 0.5), and restarting several times the algorithm by constructing a fresh simplex centred on the current best point.

Carcasses discovery rates through passive search inside or outside the fenced area and through active search event ($\rho_W = \{\rho_W^{in}, \rho_W^{out}, \rho_W^a\}$) were estimated during the first period of the challenge and determined separately via a calibration step according to the available data on passive search inside/outside the fenced area and active search events.

2.4. Model forecasts

The two models are coupled in order to simulate the epidemic dynamics of the whole system (domestic pig farms and wild boar populations) at the same time. For this purpose, each time step is divided into two parts. The first consists of the computation of the epidemic dynamics of the different epidemiological units, each with their dedicated module. The second part is the synchronisation of the relevant information between the two modules for use during the next time step computation.

For the period where data are available, some processes (hunting, carcass discovery rates) are data-driven (as for the inference part), then after this period the processes are fully simulated, using calibrated and estimated parameter values. Except transmission rates and passive/active search event discovery rates, parameters were set according to the information provided by the organisers.

Six outputs were chosen to describe the infection dynamics across both types of epidemiological units. Four of them match the observed data: (i) the observed health status of domestic pig farms, (ii) the number of positive tested hunted wild boar, (iii) the number of positive tested carcasses, through active and passive search, and (iv) the location of infected wild boar, detected through hunting or carcasses discovery through active/passive search, all of them for each time step and epidemiological unit. The last two outputs represent the hidden (unobserved) state of the system and are: (v) the health status of domestic pig farms over time, and (vi) the comprehensive dynamics of wild boar populations (headcounts + events) over time. The information on trade movements of domestic pigs for the forecasted time horizon were obtained as sampling from the observed movements, by taking advantage of their periodicity.

3. Results

During the challenge, the varying conditions between the periods and the relatively short deadlines led us to provide different results in terms of quality and quantity for the different periods. Indeed, the envisaged inference methods could not be carried out for each period, leading sometimes to partial results, and code errors could be corrected during or after the challenge. Here we briefly present the results of the parameter estimates and predictions obtained during the challenge. We also include a second set of results, obtained using the modelling and inference framework we arrived at by the end of the challenge, in order to evaluate the relevance of the approach finally developed.

Table 3

Estimated and calibrated parameter values for the three different periods (P1, P2, P3) of the challenge using the different version of the modelling framework produced during the challenge. These three periods cover, respectively, 50, 80 and 110 days since the first detection of an infected domestic pig farm (at $t = 0$).

θ	P1	P2	P3
Estimated parameters			
λ_W^{int}	0.22	0.16	0.19
λ_W^{Dint}	2.1e−04	9.8e−05	2.4e−06
λ_W^{ip}	0.06	0.06	0.05
λ_W^{ext}	3.2e−04	2.0e−07	2.8e−05
λ_W^{Dest}	1e−04	6.3e−17	5.5e−17
λ_D^i	NA	–	4.6e−06
λ_D^n	NA	–	3.4e−05
λ_D^{wb}	NA	–	3.0e−05
ρ_W^{in}	–	–	–
ρ_W^{out}	1.4e−03	–	–
ρ_W^a	0.08	–	–
Calibrated parameters			
ρ_W^{in}	–	8.9e−04	1.5e−03
ρ_W^{out}	–	4.9e−06	1.5e−06
ρ_W^a	–	0.49	0.56
Fixed parameters			
λ_D^i	–	1e−3	–
λ_D^n	–	1e−4	–
λ_D^{wb}	–	2e−5	–

Note: ρ_W^{in} is equal to ρ_W^{out} before the fence is in place.

3.1. Challenge results

During the challenge, it was not possible to estimate the transmission parameters for domestic farms for periods 1 and 2 within the time frame available. Therefore, we were not able to produce forecasts on the domestic pig farms for period 1. For period 2, the provided forecasts were made on the basis of parameter values fixed through manual calibration, based on estimates using data from the first period, obtained at the beginning of the second period.

The estimated parameter values for the different periods show similar orders of magnitude (Table 3). The transmission rate through contact with infected individuals is, for all periods, higher than the transmission rates through contact with carcasses. The same is true for the transmission rate through contact with a domestic pig farm, estimated to be between 0.049 and 0.060 depending on the period considered. The estimates suggest a likely more important role of live infected individuals on the epidemic dynamics.

Due to the unavailability of transmission parameter estimates for domestic pig farms during the first period and the use of a data-driven approach for periods 2 and 3, only the epidemic dynamics of wild boar over time are presented here (Fig. 5). In order to compare the dynamics for scenarios similar to the data, the simulations shown in Figs. 5 and 6 cover the time horizons for which the data were available. Indeed, the predictions we made during the challenge (Fig. 7) did not necessarily incorporate the right control measures or in a different way than expected because of a misunderstanding on how to implement them.

The visual comparison between the observed data and the simulations for wild boar populations (Fig. 5) shows a likely overestimation of the pathogen transmission, leading to a larger proportion of positive test results. This can also be seen spatially, with a different distribution of detected cases and a much or slightly wider spread, respectively for period 1 and periods 2 and 3 (Fig. 6). The predictions made during the challenge for periods 2 and 3 (Fig. 7) suggested a significant extension of the pathogen to the west of the fenced area. The predictions at $t = 110$ made with the parameters estimated for period 2 (Fig. 7A) show a wider spread than the observed data (Fig. 6G), but whose uncertainty was deemed high given the estimation issues encountered and the parameter values used.

A more detailed comparison including the forecasts carried out during the challenge is presented in Ezanno et al. (2022).

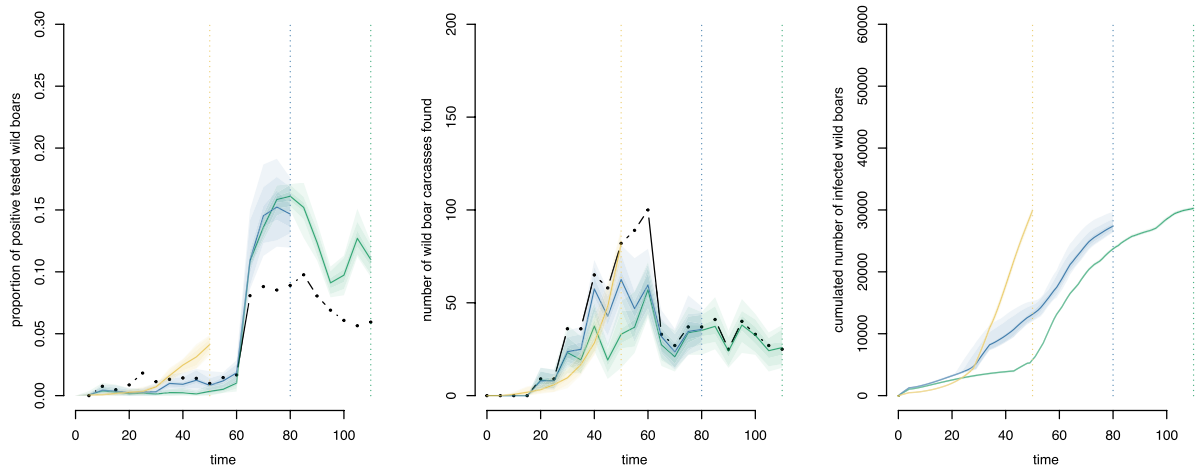


Fig. 5. Epidemic dynamic in wild boar populations over time. Data in black and simulations based on parameter values estimated during the challenge for period 1 (yellow), 2 (blue) and 3 (green) with time horizon of respectively 50, 80 and 110 days after first detection. Mean dynamic (dark line) and credibility intervals (50%, 90%, lower and upper bounds, represented by the shaded areas), aggregated over a 5 days time window (as for the computation of the criterion used in the inference part). Results based on 100 runs for the first period and 50 runs for the second and third periods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Post-challenge model fitting and forecasts

The second set of results, outlined below, corresponds to the adjustments and predictions using the final framework that we produced at the end of the challenge.

An important difference with the results obtained during the challenge lies in the inference part which could not be provided in the first and second periods.

The estimates and calibrations of parameters carried out after the challenge led to different parameter values for the wild boar part for periods 1 and 2, and quite similar for period 3, compared to the results obtained during the challenge. Thus, for periods 1 and 2, we observe values more than twice as high for the rate of transmission by contact with infected individuals from the same population. The rate of transmission through contact with domestic pig farms is almost twice as small. The estimated value of transmission parameters for domestic pig farms are similar across the periods (Table 4). By comparing with the values estimated for period 3 during the challenge, we observe that the transmission rate by contact with wild boar is of the same order of magnitude, while the other transmission rates are slightly lower. The calibration of the wild boar carcass discovery rates leads to similar values for the different periods. The values estimated during and after the challenge are also close to each other, except for the active search, where the values are lower. This can be explained by a correction made to avoid overestimating this parameter. Over the course of the challenge, improvements were made to the modelling framework, including some code fixes. This may partly explain the differences between the estimates produced during and after the challenge.

The comparison between observed and simulated data, over the specific time horizons for which data were available for each period, highlights the same pattern as during the challenge. An overestimation of the pathogen spread among wild boar, through a larger proportion of positive tested wild boar, is noted (Fig. 9). Spatial spread seems to be more consistent in period 1, compared to the results obtained during the challenge, with a higher concentration of cases in the same tiles. For periods 2 and 3 the results obtained during and after the challenge were similar. As an exception, cases north of the fenced area around an infected farm in the data did not appear in the predictions made during the challenge.

One of the motivations for carrying out these post-challenge estimates and forecasts was the opportunity to assess the reliability of the predictions. The results presented in Figs. 8 and 9, as well as the maps in Figs. 10 and 11, allow for a comparison of observed and

Table 4

Estimated and calibrated parameter values for the three different periods (P1, P2, P3) of the challenge using the modelling framework produced at the end of the challenge. These three periods cover, respectively, 50, 80 and 110 days since the first detection of an infected domestic pig farm (at $t = 0$).

θ	P1	P2	P3
Estimated parameters			
λ_W^{Int}	0.52	0.50	0.24
λ_W^{Dint}	$7.4e-04$	$3.2e-06$	$3.8e-07$
λ_W^{Ip}	0.02	0.03	0.04
λ_W^{Iext}	$1.1e-03$	$2.3e-08$	$1.6e-06$
λ_W^{Dext}	$3.7e-07$	$2.4e-22$	$1.9e-20$
λ_D^i	–	$6.4e-07$	$9.4e-07$
λ_D^n	–	$9.7e-08$	$2.9e-08$
λ_D^{nb}	$2.6e-05$	$6.5e-05$	$5.3e-05$
Calibrated parameters			
ρ_W^{in}	–	$1.4e-03$	$1.2e-03$
ρ_W^{out}	$1.1e-05$	$1.9e-06$	$3.6e-06$
ρ_W^a	0.143	0.11	0.13

Note: ρ_W^{in} is equal to ρ_W^{out} before the fence is in place.

simulated data beyond the time sequences that were used for estimation for periods 1 and 2. The predicted number of infected domestic pig farms appears to be in agreement with the data, with a slightly different dynamic over time for the second period. The proportion of positive tested wild boar remains overestimated, but with a trend close to that observed in the data. While the dynamics of the number of infected carcasses discovered is quite close to the observed data for the first period, it is largely overestimated for the second period.

Comparison of the spatial simulated and observed distributions of detected cases spanning over similar time periods (until day 80 - Fig. 11 A with Fig. 10 D and until day 110 - Fig. 11 C with Fig. 10 G) highlights an overestimation of the spread of the pathogen between wild boar populations. In particular, there is an apparent spread of the pathogen among domestic pig farms and wild boar in the northern area that is not apparent in the data. Contrary to the predictions made for period 2 during the challenge, no cases were found in the southeast. This difference can be explained by the change in the method used to augment the farm-to-farm movement data, and a lower value for the transmission rate through trade movements.

4. Discussion

The main objective of this work was to develop, in a relatively short period of time and in a predefined context in terms of mechanisms to

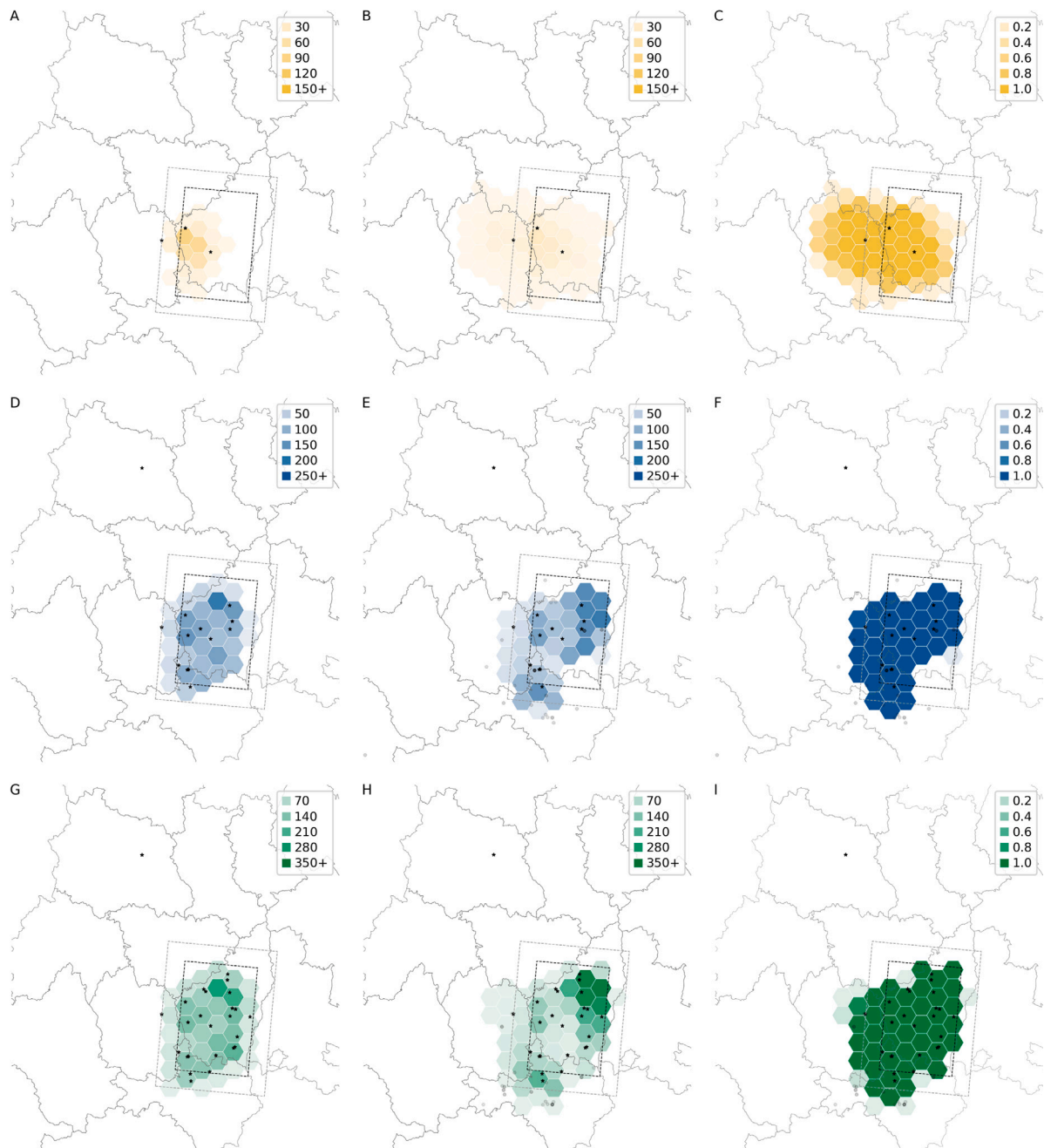


Fig. 6. Spatial distribution of infection in domestic pig farms (circles and stars) and wild boar populations (hexagonal tiles): observed real situation (left column), mean observed simulated situation (centre column) and probability to observe infected animals (right column) based on 10 runs for the first period and 50 runs for the second and third periods. Simulations are based on parameter values estimated during the challenge (Table 3) for period 1 (yellow A, B, C), 2 (blue D, E, F) and 3 (green G, H, I) with time horizon of respectively 50, 80 and 110 days after first detection. For the first period, due to the unavailability of transmission parameter estimates for domestic pigs farms, only the epidemic dynamics of wild boar was simulated and is therefore shown, domestic pig farms epidemic dynamic being only based on data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be considered and constrained concerning the data, a ‘good’ model, which can be defined by an appropriate balance between accuracy, transparency and flexibility (Keeling and Rohani, 2011). Accuracy is essential and can be defined as the ability to produce reliable predictions. Transparency represents the ability to understand how the different elements of a model influence the infection dynamics. As the number of elements in a model increases, it becomes more difficult to assess the role of each element and their possible interactions. Transparency can therefore be in direct opposition to accuracy. Flexibility represents the ability of the model to be adapted to new situations, which is essential if it is intended for the evaluation of control strategies to support decision making.

We have chosen to use a relatively simple modelling framework, neglecting the demographic dynamics within the epidemiological units. This choice avoids the potential hurdles arising from the calibration of a complex model in the context of limited data, which can lead to parameter identifiability issues. A suitable approach to address this issue would be to perform a numerical identifiability analysis or a thorough sensitivity analysis of the model and calibrate only the influential parameters. Given the evolving nature of the challenge and the effort that this type of analysis may involve, we have instead preferred to have a limited number of parameters to estimate, while keeping open the possibility of increasing the complexity of the model should the need arise.

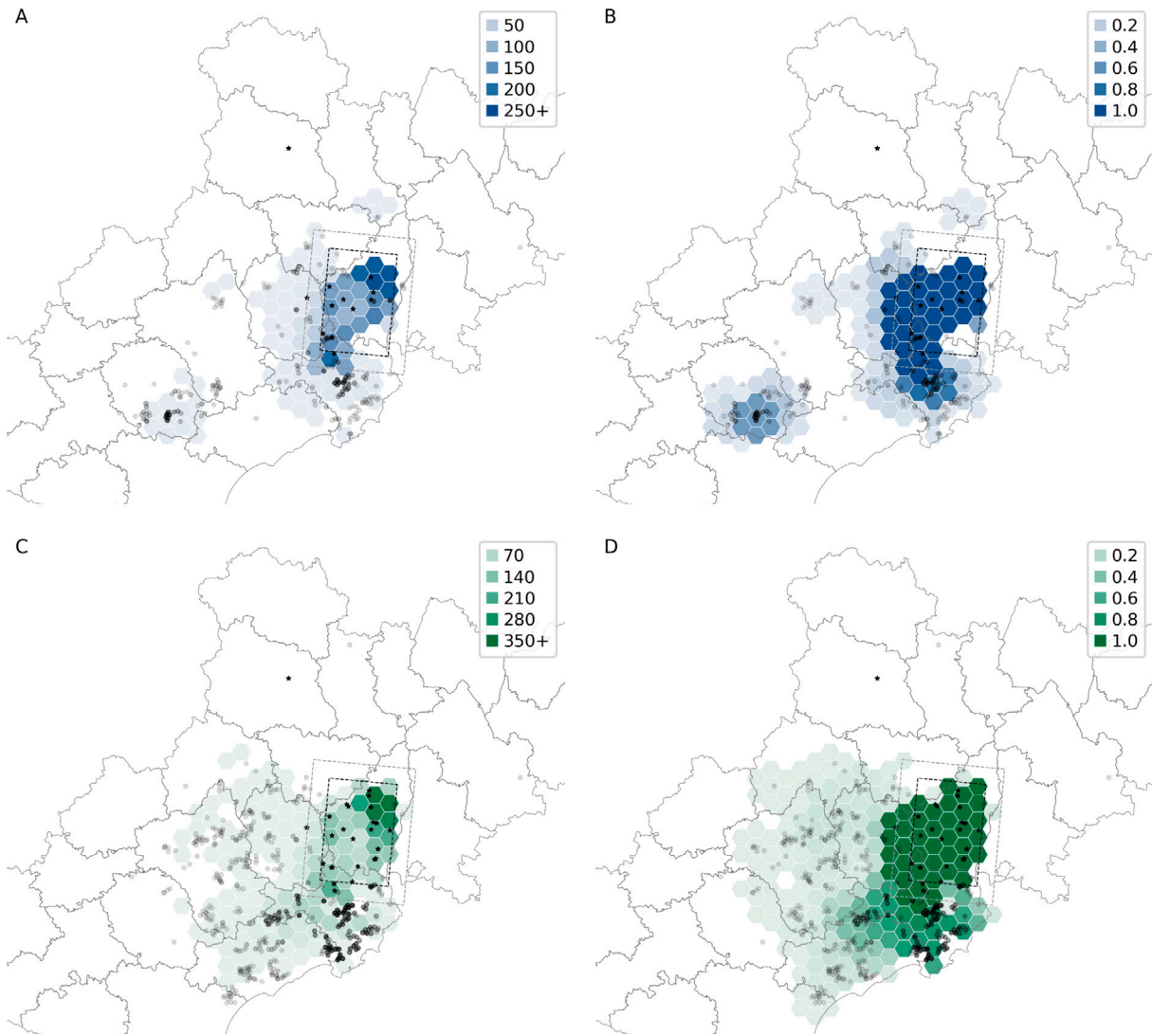


Fig. 7. Spatial distribution of infection in domestic pig farms (circles and stars) and wild boar populations (hexagonal tiles): mean observed simulated situation (left column) and probability to have infected animals (hidden situation) (right column) based on 50 repetitions. Simulations are based on parameter values estimated during the challenge for period 2 (blue A, B) and 3 (green C, D) with time horizon of respectively 110 and 200 days after first detection. The results presented are the same as those provided during the challenge, therefore period 1 is not included here as the results produced did not allow for the quantification of these outputs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

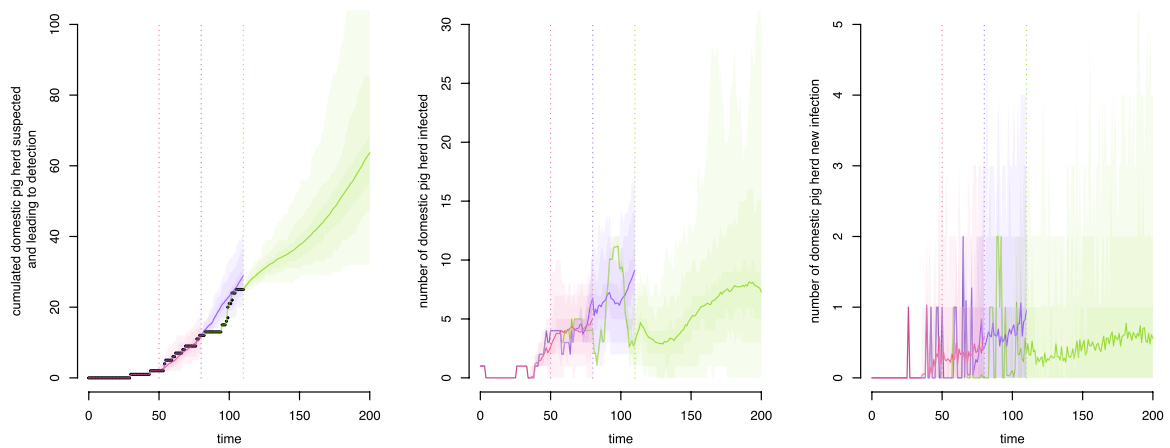


Fig. 8. Epidemic dynamics in domestic pig farms over time. Data in black and simulations based on parameter values estimated using the final modelling framework, for period 1 (magenta), 2 (purple) and 3 (green), with time horizon of respectively 80, 110 and 200 days after first detection. Mean dynamic (dark line) and credibility intervals (50%, 90%, lower and upper bounds, represented by the shaded areas) calculated on 100 runs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

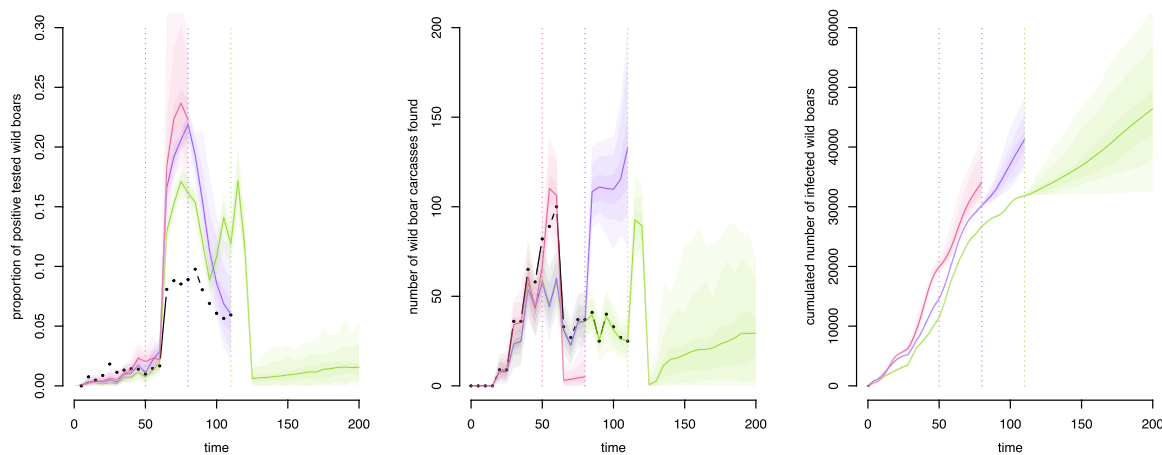


Fig. 9. Epidemic dynamics in wild boar populations over time. Data in black and simulations based on parameter values estimated using the final modelling framework, for period 1 (magenta), 2 (purple) and 3 (green), with time horizon of respectively 80, 110 and 200 days after first detection and calculated on 100 runs. Mean dynamic (dark line) and credibility intervals (50%, 90%, lower and upper bounds, represented by the shaded areas), aggregated over a 5 days time window (as for the computation of the criterion used in the inference part). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Determining the initial state is key to understanding the dynamics of an epidemic. Here, we make the assumption that the outbreak started on a domestic pig farm at the time of first detection, assuming that the pathogen was introduced to the island via the importation of an infected domestic pig, and that there were no infected wild boars at that time. This simplifying assumption implies a known initial state, and thus contributes to reduce the complexity of the inference procedure. Estimating the date, host type for the first infection and location (possibly several locations) was too complex to be achieved in the time available for the challenge. However, if the epidemic started in the wild boar population and before the first detection, this assumption introduced a bias in the epidemic dynamics that could have had an impact on the parameter estimates, and thus on the predictions.

As for the other modelling choices made during the challenge, the discretisation of the wild boar population spatial distribution into epidemiological units using a hexagonal grid implies a trade-off between realism and parsimony. The selected tile size is relatively large and may not be suitable to reflect the living/dispersal patterns of wild boar and/or to capture landscape features. However, the size chosen allows for a limited number of tiles and minimises the computation time, thus leaving more time for inference, predictions and analysis of the results. The same applies to the modelling of the neighbourhood transmission route for domestic pig farms, defined by a set of farms that are potential sources of infection using a buffer around each farm. This choice has the advantage of being simple, easy to implement and fast in terms of computing time. However, this implementation might be too simple to represent the neighbourhood interactions accurately enough. The conditions of the challenge did not allow us to reconsider this choice during the challenge. We preferred a modelling approach that adds complexity as we go along, which seemed the most relevant in the context of the challenge, wanting to keep the model as parsimonious as possible given the time constraints and forces involved.

Our modelling framework is flexible enough to accommodate changes in transmission processes and control measures that may have occurred during the challenge. This is particularly useful when dealing with epidemic control in real-time, where a combination of changing available information and evolving control strategies requires the modelling framework to be adjusted continuously during the outbreak. The design of our modelling framework makes it possible, for example, to replace the epidemic dynamics for one of the two host population with a predefined scenario or a completely different model at any time, as long as the information to be exchanged between the two modules can be generated.

A better understanding of the demographic dynamics of potential host populations is key to improve our preparedness. This modelling

challenge has highlighted that the quality of the available data is a significant concern. The lack of accurate data on the demography of the wild boar population was an important source of uncertainty, which has to be filled by assumptions and approximations. In order to use robust modelling approaches as a decision support tool when facing the emergence of new pathogens, it appears essential to have access to both qualitative and quantitative information on the demography of species of interest. Such databases already exist for livestock at different scales of detail (Gilbert et al., 2018; Nöremark et al., 2011; Vernon, 2011; Dutta et al., 2014), and need to be established for relevant wildlife species as well (Bengis et al., 2004).

The ability to react quickly to the emergence of a pathogen partly relies on our capacity to re-use and/or implement previous work. Over the last few decades, the development of modelling in epidemiology has led to the production of a large number of models, of varying complexity, for a wide variety of pathosystems and at different scales (Thompson and Brooks-Pollock, 2019). Recent efforts by the scientific community to make their work available on an open access basis aim to facilitate the reproduction and reuse of these models (open-access publishing, use of public code repositories). Thus, when a similar study is to be carried out or a model is to be used to answer the same question for which it was designed, it is usually relatively easy to reuse existing work. However, when investigating a new question, or studying a new or (slightly) different pathosystem, it is clear that it remains relatively difficult to quickly reuse existing models, which were moreover developed by another research team. Indeed, it can quickly become more time-consuming to reuse a model (effort required to understand someone else's code and to update it, to solve compatibility problems and computer dependencies), especially when it is necessary to continuously adapt the modelling framework to the epidemiological situation (scenario proposed during the challenge). There are initiatives to make model reuse and sharing easier (Sandve et al., 2013; List et al., 2017; Lee, 2018), such as tools that allow to focus on the conceptual part rather than on the programming part (Broeck et al., 2011; Bui et al., 2019; Picault et al., 2019). These types of approaches obviously provide solutions that can help us to improve our responsiveness to the emergence of new pathogens. However, these frameworks still have some weaknesses (such as the use of a dedicated language, or their computational inefficiency) that can make their use potentially difficult in a health crisis response situation.

In addition to these methodological considerations, the relevance of the context is paramount. Many previous studies have focused on developing models of ASF spread in order to answer a wide range of questions (Hayes et al., 2021). However, at the time of the challenge,

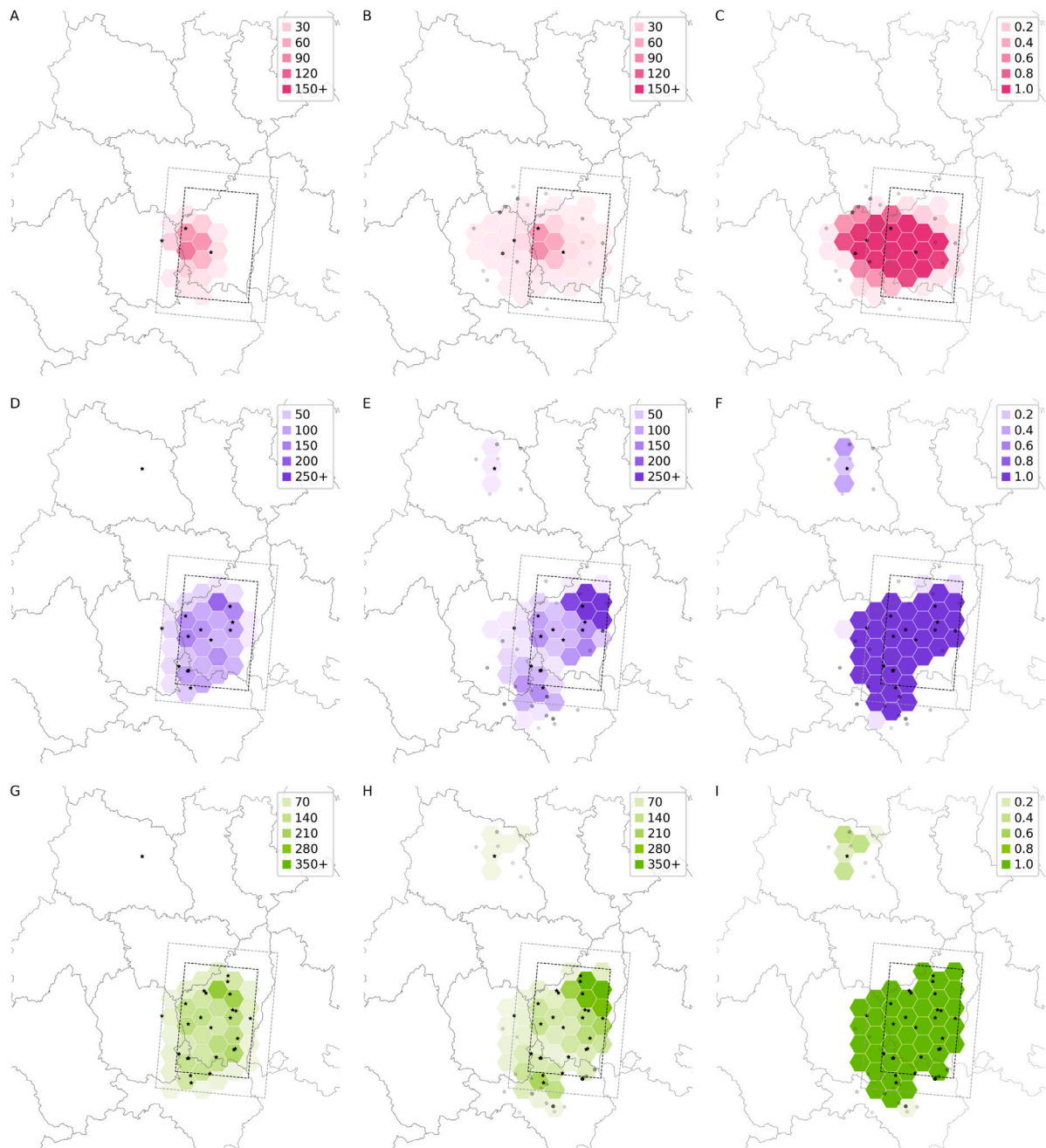


Fig. 10. Spatial distribution of infection in domestic pig farms (circles and stars) and wild boar populations (hexagonal tiles): observed real situation (left column), mean observed simulated situation (centre column) and probability to observe infected animals (right column), calculated on 100 runs. Simulations based on parameter values estimated using the final modelling framework for period 1 (magenta A, B, C), 2 (purple D, E, F) and 3 (green G, H, I) with time horizon of respectively 50, 80 and 110 days after first detection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

no model considered the complete transmission between domestic pigs and wild boars, only one study (Taylor et al., 2021) came close, neglecting the transmission from domestic pigs to wild boars. Thus the specific context of the challenge required the development of an original model. The context of the challenge also makes it difficult to compare our results with the literature. Indeed, the data provided by the organisers correspond to a particular synthetic trajectory, chosen according to different criteria, corresponding in part to constraints linked to the challenge narrative, giving little meaning to a comparison with previous studies on real data.

Besides the design of the model, the major difficulty of this challenge laid in the parameter estimation. Changes in the proposed scenarios and available data during the different periods of the challenge

required the adaptation of the inference methods used to meet the increasing complexity of the model. During the first period of the challenge, the choices of model design and inference procedure did not allow us to estimate all the parameters, highlighting the importance of the possible constraints associated with the modelling choices and/or the inference methods. The characteristics of the models and the data have thus oriented the choice of methods towards likelihood-free procedures. Recent advances in likelihood-free inference methods and increasing computing power allow for the fitting of complex large-scale models. However, difficulties arise when it comes to selecting the information that best represents the processes to be estimated, to be used for comparing observations and simulations. Proposing appropriate summary statistics (i.e. containing enough information to

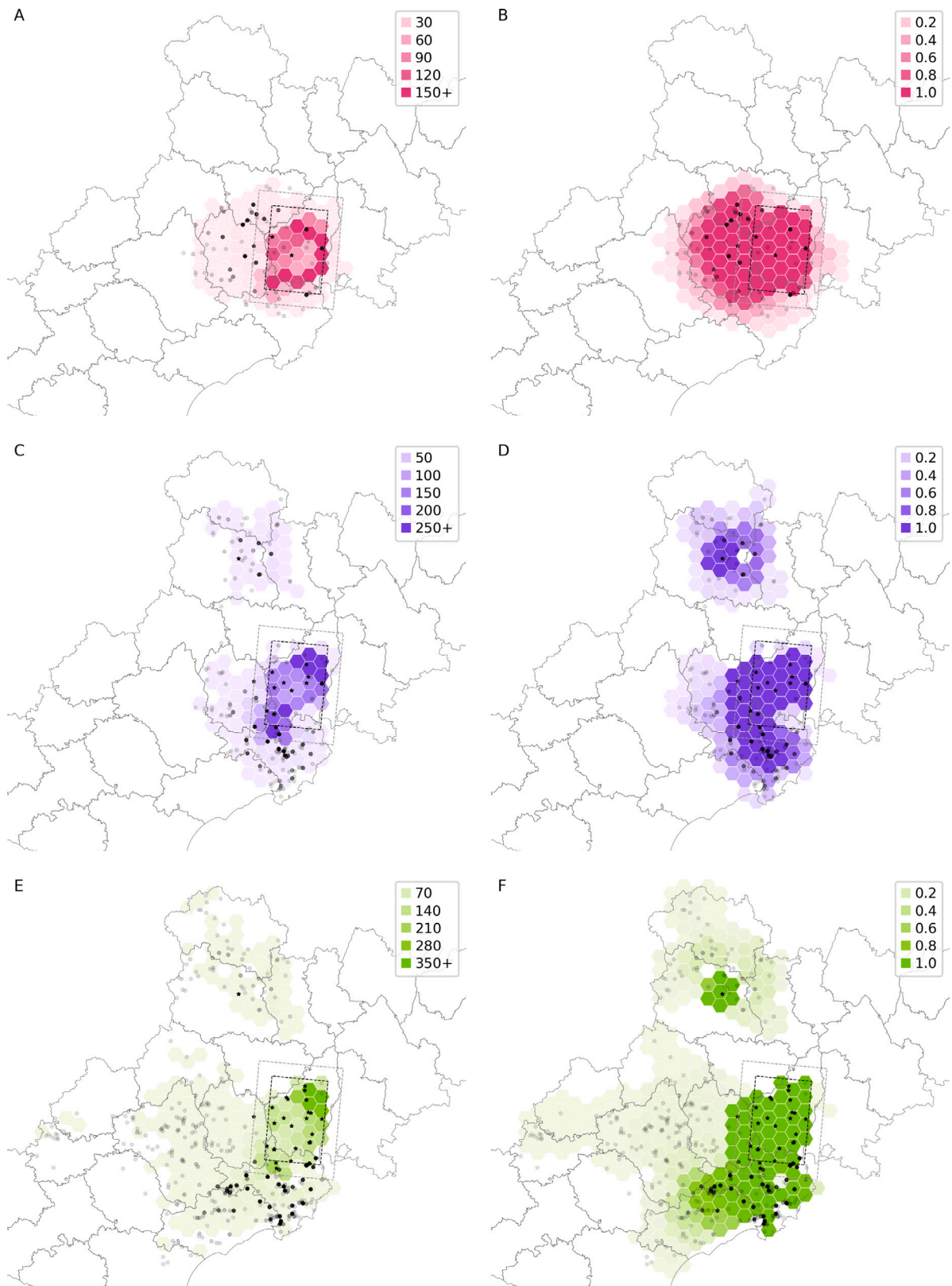


Fig. 11. Spatial distribution of infection in domestic pig farms (circles and stars) and wild boar populations (hexagonal tiles): mean observed simulated situation (left column) and probability to have infected animals (hidden situation) (right column) calculated on 100 repetitions. Simulations based on parameter values estimated using the final modelling framework for period 1 (magenta A, B), 2 (purple C, D) and 3 (green E, F) with time horizon of respectively 80, 110 and 200 days after first detection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be a discriminating criterion) is a key step, which can impact on the inference results. There is limited work on how to select the most useful information in the data when working with these methods (Sisson et al., 2018; Joyce and Marjoram, 2008; Nunes and Balding, 2010; Barnes et al., 2012). Furthermore, they require a large number of simulations and therefore potentially prohibitive computation times when it comes to providing a rapid response on the spread of a pathogen on a large scale. Therefore, there is a need for prospective work on existing data and synthetic scenarios in order to be able to narrow down more quickly the choice of criteria to be used, depending on the characteristics of the system (types of hosts and pathogens) and the data available.

As highlighted by this challenge, and others before it (Ajelli et al., 2018; Valle et al., 2018), real-time forecasting can be particularly difficult, as it involves producing accurate trends while properly accounting for the uncertainties surrounding both the data and the epidemic dynamics of the pathogen, in a relatively short period of time. Under the conditions imposed by the challenge, mirroring an epidemic emergency situation, the amount of time available becomes a very important factor and brings an additional constraint. In addition, data availability and robust inference remains a key issue in our ability to provide robust forecasts and guidance. There is no doubt that lessons learned during this ASF modelling challenge will help to strengthen our preparedness to face emerging threats.

CRedit authorship contribution statement

G. Beaunée: Conceptualization, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **F. Deslandes:** Conceptualization, Methodology, Software, Validation, Writing – review & editing. **E. Vergu:** Conceptualization, Methodology, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the organisers of the ASF modelling Challenge for their thoughtful design and efficient organisation. We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing computing and storage resources.

References

Ajelli, M., Zhang, Q., Sun, K., Merler, S., Fumanelli, L., Chowell, G., Simonsen, L., Viboud, C., Vespignani, A., 2018. The RAPIDD Ebola forecasting challenge: Model description and synthetic data generation. *Epidemics* 22, 3–12.

Alahmadi, A., Belet, S., Black, A., Cromer, D., Flegg, J.A., House, T., Jayasundara, P., Keith, J.M., McCaw, J.M., Moss, R., Ross, J.V., Shearer, F.M., Tun, S.T.T., Walker, J., White, L., Whyte, J.M., Yan, A.W., Zarebski, A.E., 2020. Influencing public health policy with data-informed mathematical models of infectious diseases: Recent developments and new challenges. *Epidemics* 32, 100393.

Barnes, C.P., Filippi, S., Stumpf, M.P., Thorne, T., 2012. Considerate approaches to constructing summary statistics for ABC model selection. *Stat. Comput.* 22 (6), 1181–1197.

Barton, R.R., Ivey, Jr., J.S., 1991. Modifications of the Nelder-Mead simplex method for stochastic simulation response optimization. In: *Proceedings of the 23rd Conference on Winter Simulation. WSC '91, IEEE Computer Society, Washington, DC, USA*, pp. 945–953.

Bengis, R.G., Leighton, F.A., Fischer, J.R., Artois, M., Mörner, T., Tate, C.M., 2004. The role of wildlife in emerging and re-emerging zoonoses. *Rev. Sci. Et Tech. (International Office of Epizootics)* 23 (2), 497–511.

Brock, W.V.d., Gioannini, C., Gonçalves, B., Quaggiotto, M., Colizza, V., Vespignani, A., 2011. The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infect. Dis.* 11 (1), 37.

Brooks-Pollock, E., Danon, L., Jombart, T., Pellis, L., 2021. Modelling that shaped the early COVID-19 pandemic response in the UK. *Philos. Trans. R. Soc. B* 376 (1829), 20210001.

Brooks-Pollock, E., Roberts, G.O., Keeling, M.J., 2014. A dynamic model of bovine tuberculosis spread and control in Great Britain. *Nature* 511 (7508), 228.

Bui, T.M.A., Papoulias, N., Stinckwich, S., Ziane, M., Roche, B., 2019. The Kendrick modelling platform: Language abstractions and tools for epidemiology. *BMC Bioinformatics* 20 (1), 312.

Chowell, G., Sattenspiel, L., Bansal, S., Viboud, C., 2016. Mathematical models to characterize early epidemic growth: A review. *Phys. Life Rev.* 18, 66–97.

Chretien, J.P., Riley, S., George, D.B., 2015. Mathematical modeling of the West Africa Ebola epidemic. *ELife* 4, e09186.

Dixon, L.K., Stahl, K., Jori, F., Vial, L., Pfeiffer, D.U., 2019. African swine fever epidemiology and control. *Annu. Rev. Animal Biosci.* 8 (1), 221–246.

Dutta, B.L., Ezanno, P., Vergu, E., 2014. Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period. *Prevent. Vet. Med.* 117 (1), 79–94.

Ezanno, P., Andraud, M., Beaunée, G., Hoch, T., Krebs, S., Rault, A., Touzeau, S., Vergu, E., Widgren, S., 2020. How mechanistic modelling supports decision making for the control of enzootic infectious diseases. *Epidemics* 32, 100398.

Ezanno, P., Picault, S., Bareille, S., Beaunée, G., Boender, G.J., Dankwa, E.A., Deslandes, F., Donnelly, C.A., Hagenaars, T.J., Hayes, S., Jori, F., Lambert, S., Mancini, M., Munoz, F., Pleydell, D.R., Thompson, R.N., Vergu, E., Vignes, M., Vergne, T., 2022. The African swine fever modelling challenge: Model comparison and lessons learnt. *Epidemics* 40, 100615.

Gilbert, M., Nicolas, G., Cinardi, G., Boeckel, T.P.V., Vanwambeke, S.O., Wint, G.R.W., Robinson, T.P., 2018. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Sci. Data* 5 (1), 180227.

Hayes, B.H., Andraud, M., Salazar, L.G., Rose, N., Vergne, T., 2021. Mechanistic modelling of African swine fever: A systematic review. *Prevent. Vet. Med.* 191, 105358.

Heesterbeek, H., Anderson, R.M., Andreasen, V., Bansal, S., Angelis, D.D., Dye, C., Eames, K.T.D., Edmunds, W.J., Frost, S.D.W., Funk, S., Hollingsworth, T.D., House, T., Isham, V., Klepac, P., Lessler, J., Lloyd-Smith, J.O., Metcalf, C.J.E., Mollison, D., Pellis, L., Pulliam, J.R.C., Roberts, M.G., Viboud, C., Collaboration, I.N.I.I., 2015. Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347 (6227), aaa4339.

James, L.P., Salomon, J.A., Buckee, C.O., Menzies, N.A., 2021. The use and misuse of mathematical modeling for infectious disease policymaking: Lessons for the COVID-19 pandemic. *Med. Decis. Mak.* 41 (4), 379–385.

Joyce, P., Marjoram, P., 2008. Approximately sufficient statistics and Bayesian computation. *Stat. Appl. Genet. Mol. Biol.* 7 (1).

Keeling, M.J., Rohani, P., 2011. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press.

Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., group, Centre for Mathematical Modelling of Infectious Diseases COVID-19 working, Sun, F., Jit, M., Munday, J.D., Davies, N., Gimma, A., Zandvoort, K.v., Gibbs, H., Hellewell, J., Jarvis, C.I., Clifford, S., Quilty, B.J., Bosse, N.I., Abbott, S., Klepac, P., Flasche, S., 2020. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet. Infect. Dis.* 20 (5), 553–558.

Lee, B.D., 2018. Ten simple rules for documenting scientific software. *PLoS Comput. Biol.* 14 (12), e1006561.

List, M., Ebert, P., Albrecht, F., 2017. Ten simple rules for developing usable software in computational biology. *PLoS Comput. Biol.* 13 (1), e1005265.

Moghadas, S.M., Pizzi, N.J., Wu, J., Yan, P., 2009. Managing public health crises: The role of models in pandemic preparedness. *Influenza Other Respir. Viruses* 3 (2), 75–79.

Moran, K.R., Fairchild, G., Generous, N., Hickmann, K., Osthus, D., Priedhorsky, R., Hyman, J., Valle, S.Y.D., 2016. Epidemic forecasting is messier than weather forecasting: The role of human behavior and internet data streams in epidemic forecast. *J. Infect. Dis.* 214 (suppl_4), S404–S408.

Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *Comput. J.* 7 (4), 308–313.

Nöremark, M., Håkansson, N., Lewerin, S.S., Lindberg, A., Jonsson, A., 2011. Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk based surveillance. *Prevent. Vet. Med.* 99 (2–4), 78–90.

Nunes, M.A., Balding, D.J., 2010. On optimal selection of summary statistics for approximate Bayesian computation. *Stat. Appl. Genet. Mol. Biol.* 9 (1).

Picault, S., Huang, Y.L., Sicard, V., Arnoux, S., Beaunée, G., Ezanno, P., 2019. EMULSION: Transparent and flexible multiscale stochastic models in human, animal and plant epidemiology. *PLoS Comput. Biol.* 15 (9), e1007342.

Picault, S., Vergne, T., Mancini, M., Bareille, S., Ezanno, P., 2022. The African swine fever modelling challenge: Objectives, model description and synthetic data generation. *Epidemics* 40, 100616.

Sandve, G.K., Nekrutenko, A., Taylor, J., Hovig, E., 2013. Ten simple rules for reproducible computational research. *PLoS Comput. Biol.* 9 (10), e1003285.

Sisson, S.A., Fan, Y., Beaumont, M., 2018. *Handbook of Approximate Bayesian Computation*. CRC Press.

Taylor, R.A., Podgórski, T., Simons, R.R.L., Ip, S., Gale, P., Kelly, L.A., Snary, E.L., 2021. Predicting spread and effective control measures for african swine fever—Should we blame the boars? *Transbound. Emerg. Dis.* 68 (2), 397–416.

- Thompson, R.N., Brooks-Pollock, E., 2019. Detection, forecasting and control of infectious disease epidemics: Modelling outbreaks in humans, animals and plants. *Philos. Trans. R. Soc. B* 374 (1775), 20190038.
- Valle, S.Y.D., McMahon, B.H., Asher, J., Hatchett, R., Lega, J.C., Brown, H.E., Leany, M.E., Pantazis, Y., Roberts, D.J., Moore, S., Peterson, A.T., Escobar, L.E., Qiao, H., Hengartner, N.W., Mukundan, H., 2018. Summary results of the 2014–2015 DARPA Chikungunya challenge. *BMC Infect. Dis.* 18 (1), 245.
- Vergne, T., Chen-Fu, C., Li, S., Cappelle, J., Edwards, J., Martin, V., Pfeiffer, D.U., Fusheng, G., Roger, F.L., 2017. Pig empire under infectious threat: Risk of African swine fever introduction into the People's Republic of China. *Vet. Rec.* 181 (5), 117.
- Vernon, M.C., 2011. Demographics of cattle movements in the United Kingdom. *BMC Vet. Res.* 7 (1), 31.