Supplement S1: Model details

S1.A. Detailed within-canton transmission

$\lambda_{int}(k,t)$ represents the force of vector-borne infection from female midges located in the canton that got infected locally while feeding on infectious ruminants in the previous time steps, that completed the extrinsic incubation period (EIP) required for BTV replication and dissemination up to the arthropod vector salivary glands, and survived up to time *t*. We made the assumption that, in a given canton, and during the vector activity period, the vector to host ratio was constant. Under this assumption, the vector-borne transmission can be represented by a non-Markovian force of infection, which accounts for the *Culicoides* cohorts that emerged in the preceding weeks.

$$\lambda_{int}(k,t) = \tau(k,t) * \sum_i (w_i * Prev(k,t-i))$$

with $\tau(k,t)$, the weekly effective contact rate at which vectors and hosts from canton *k* come into effective contact; π^{sp}, the relative preference of vectors for cattle or sheep (conditional on feeding on these species); $Prev(k,t\text{-}i)$, the proportion of infectious animals at time *t-i* weighted by vectors species-specific trophic preferences; w_i, the fraction of *Culicoides* vectors that have completed their EIP in *i* weeks and survived over that period.

$\tau(k,t)$, $Prev(k,t\text{-}i)$ and w_i are given by:

$$Prev(k,t\text{-}i) = \pi^c * Prev^c(k,t\text{-}i) + (1-\pi^c) * Prev^s(k,t\text{-}i)$$

with π^c, the preference for feeding on cattle *vs* sheep (if feeding of these species).

$$w_i = \begin{cases} 0 & \text{if } Tp(t\text{-}i,k) < T_{min} \text{ or } i=0 \\ (1\text{-}\mu^v)^{(i*7)} * x_{EIP}(i) & \text{if } Tp(t\text{-}i,k) \geq T_{min} \text{ and } i>0 \end{cases}$$

$$x_{EIP}(i) \quad (i*7)/EIP \quad \begin{cases} & \text{if } i*7 < EIP \end{cases}$$

23 *1* if $i*7 \geq EIP$

24 with *Tp(t-i,k)*, the temperature; $x_{EIP}(i)$, the fraction of *Culicoides* that have completed their EIP
25 in i weeks; $\mu^{v,}$ the daily mortality proportion of *Culicoides* vectors; and T_{min}, the threshold
26 temperature for virus replication.

$$\tau(k,t) = \beta_0 * Env(k) * b(k,t)$$

27 with $\beta_{0,}$ a coefficient that represents the baseline exposure of hosts to vectors, defined here as the
28 product of the baseline vector to host ratio, the host to vector and vector to host probabilities of
29 successful transmission, and the trophic preference of *Culicoides* for cattle and sheep *vs* other
30 warm-blooded species; *Env(k)*, the environmental variables used as proxy of host availability,
31 *Culicoides* presence and abundance; $b(k,t)$, the temperature dependent biting rate of *Culicoides*
32 at time *t* in canton *k* that represents the seasonal variation in *Culicoides* activity.

33 *Env(k)* was defined under the assumption that bluetongue transmission in a given area
34 depends on the proportion of surface covered in pastures, where hosts and vectors come into
35 contact. We modulated the effect of transmission that occurred on pastures by considering two
36 additional land cover metrics: *(i)* the edge density between pastures and arable lands where
37 manure is spread and provides suitable breeding sites for BTV vector species (Ninio, 2011); and
38 *(ii)* the edge density between pastures and forests/semi-natural areas which provide shelter to the
39 wild animals that may contribute to BTV sylvatic cycle (Rossi et al., 2014). Edge densities are
40 landscape diversity indicators, here defined as the cumulative length of borders (m) between two
41 types of land cover within a canton, divided by the surface area of the canton (hm^2).We attributed
42 coefficients β_a and β_f to the edge densities of pasture *vs* arable lands and *vs* forest respectively,
43 and expressed the *Env(k)* term as follows:

$$Env(k) = dPast(k) * \left(1 + \beta_A * e_{PA}(k) + \beta_F * e_{PF}(k)\right)$$

with $dPast(k)$, the proportion of the surface of canton *k* covered in pastures; $e_{PA}(k)$and $e_{PF}(k)$, the edge densities of pasture *vs* arable lands and *vs* forest respectively, in canton *k*.

S1.B. Detailed between-canton transmission

BTV transmission through the three contact networks depends on both the frequency of movements and prevalence of infection in the source canton. We represented BTV transmission due to midges dispersal by applying a fraction of the force of infection of a given canton to its neighbors on the pasture network. $\lambda_{vect}(k,t)$, the force of vector-borne infection applied to canton *k* at time *t* from all its neighbors was given by:

$$\lambda_{vect}(k,t) = \sum_{j \in \theta_{past}(k)} \Psi_P * \lambda_{int}(j,t)$$

with $\theta_{past}(k)$ the neighbors of canton *k* on the pasture network; Ψ_P, the proportion of canton surface that can be reached by vectors coming from a neighboring canton; $\lambda_{int}(j,t)$, the force of infection in canton *j*, neighbor of canton *k* on the pasture network.

We estimated $n^{sp}_{intro}(k,t)$, the number of infectious animals introduced in cell *k* at time *t*, as the sum of those introduced through the farm and trade networks ($n^{sp}_{farm}(k,t)$ and $n^{sp}_{trade}(k,t)$):

$$n^{sp}_{intro}(k,t) = n^{sp}_{farm}(k,t) + n^{sp}_{trade}(k,t)$$

$$n^{sp}_{intro}(k,t) = \sum_{j \in \theta_{farm}(k)} n^{sp}_{farm}(k,j,t) + \sum_{j \in \theta_{trade}(k,t)} n^{sp}_{trade}(k,j,t)$$

with $\theta_{farm}(k)$ and $\theta_{trade}(k)$, the neighbors of canton *k* on the farm and trade network respectively; $n^{sp}_{farm}(k,j,t)$ and $n^{sp}_{trade}(k,j,t)$, the number of infected animal introduced from canton *j* to canton *k* at time *t* on the trade and farm network respectively.

$$n^{sp}_{farm}(k,j,t) \sim Binom(N^{sp}_{farm}(k,j,t), Prev^{sp}(j,t))$$

where $N^{sp}_{farm}(k,j,t)$, the weekly number of animals moved on the farm network from canton *j* to canton *k* at time *t* is given by:

$$N^{sp}_{farm}(k,j,t) = \Psi_F * Q^{sp}_{farm}(k,j) * N^{sp}(j,t)$$

with Ψ_F, the weekly proportion of animals moved through the farm network; $Q^{sp}_{farm}(k,j)$, the proportion of cattle or sheep farms from canton *j* that have pastures in canton *k;* and $N^{sp}(j,t)$, the number of animals of each species in canton *j* at time *t.*

$$\begin{cases} n^{s}_{trade}(k,j,t) = 0 & \\ n^{c}_{trade}(k,j,t) \sim Binom(N^{c}_{trade}(k,j,t), Prev^{c}(j,t)) & \text{without movement restrictions} \\ n^{c}_{trade}(k,j,t) \sim Binom(N^{c}_{trade}(k,j,t), Prev^{c}(j,t)) & \text{with movement restrictions, } R\text{=0} \\ n^{c}_{trade}(k,j,t) = 0 & \text{with movement restrictions, } R\text{=1} \end{cases}$$

with $N^{c}_{trade}(k,j,t)$, the total number of cattle traded from canton *j* to *k* on week *t*, which could be fully informed by data. *R* was randomly drawn before each movement: *R ~ Bern(θ);* with *θ*, the probability that control measures are correctly implemented and complied with. Movement restrictions were applied for *R=1*.

S1.C. Discussion of modelling assumptions

We attributed weights to the prevalence of infectious animals in the previous time steps (w_i) to account for the proportion of *Culicoides* that had achieved their EIP and survived between the time when they got infected and the current time step. These weights were calculated using survival probabilities specific to *Culicoides Obsoletus* that had been measured in a unique setting (17-25°C, Goffredo et al., 2004). The vast majority of authors who modeled BTV transmission in the 2006/09 European outbreak (Gubbins et al., 2008; Szmaragd et al., 2009; Guis et al., 2012; Graesbøll et al., 2012; Sumner et al., 2017) represented *Culicoides* survival using the same formulae (Gerry and Mullens, 2000) that accounts for the temperature-dependency of survival.

However, that formulae originates from a study conducted under a different climate (Southern California) and on *C. sonorensis* (Gerry and Mullens, 2000), a species considered as the main BTV vector in North America (Ninio C., 2011), but not in Europe where the *Obsoletus group* is thought to play a major role in BTV transmission (Ninio C., 2011). These vector species have different life history traits: *C. sonorensis* have a short life expectancy that highly varies with temperatures: from a few days in summer, up to >10 days in winter (Gerry and Mullens, 2000)*;* by contrast, the variation in the number of *Culicoïdes* from the *Obsoletus group* between seasons may rather be attributed to the impact of temperatures on larval development than on survival probabilities (Birley and Boorman, 1982), and higher life expectancies have been found in this group: 10% survival (N=150/1,500) over 40 days at 17-25°C and a maximal survival period of 92 days (Goffredo et al., 2004).

We thus calculated the w_i weights based on survival probabilities that had been estimated in laboratory conditions at 17-25°C. These temperatures matched the average ones recorded in France in the months when most BT cases were reported in 2007 (July- September). We added a 12°C threshold below which vectors could not complete their EIP (Carpenter et al., 2011), and accounted for the influence of temperature variations on the dynamics of transmission by adding a temperature-dependent biting rate.

Supplement S2: Details on fixed parameters

S2.A. Demography

The size of cattle and sheep population by canton was matched to real data. For cattle, we updated the number of animals and births per canton every week. For sheep, we only had information on population sizes per canton in 2010, so that we assumed a constant size in each canton and applied a weekly renewal proportion μ^s (4‰, Institut de l'élevage, 2016, 2017). We randomly attributed the newborns to the C^{sp} and S^{sp} compartments proportionally to the species-specific seroprevalence in the canton. Animals from the C^{sp} compartment then transitioned towards the S^{sp} one after the disappearance of colostral antibodies at a rate α_1^{sp}. Animals moving out of the canton were considered as randomly distributed between all compartments, proportionally to the number of animals in each of them.

S2.B. Detection

Cantons with infected animals could be detected by passive clinical surveillance. We called Δ the probability that infectious animals could show clinical signs and be detected, so that:

$$n_{detect}^{sp}(k,t) \sim Binom(n_{nI1.I2}^{sp}(k,t), \Delta)$$

with $n_{detect}^{sp}(k,t)$, the number of animals detected in canton *k* at time *t*; $n_{nI1.I2}^{sp}(k,t)$, the number of ruminants moving from the first to the second infectious stages in canton *k* at time *t*. We used that transition because the incubation period lasts about a week in cattle (Guyot et al., 2008) and to ensure that each animal would be counted only once.

We used a retrospective serological study conducted in French cattle soon after BTV introduction to a new area in 2007 to identify the likely order of magnitude of Δ (Durand et al., 2010; Courtejoie et al., 2018). In that area, it was estimated that 40% of all cattle was likely to

have been infected (*i.e.* ~120,000 cattle heads), whereas 155 bovine herds (estimated 3,000 animals with clinical signs, Mounaix, B. et al., 2010) had been notified since BTV introduction. We thus considered as plausible a 2% probability of detection of infected cases upon clinical suspicion in newly infected areas (Δ), and varied this value between 1 and 5% in a sensitivity analysis.

S2.C. Proportion of canton surface reachable by Culicoides from neighboring cantons

We combined data on *Culicoides* flight distances, French cantons surface areas and pasture network topology to infer plausible values for Ψ_P. *Culicoides* dispersal at the farm level is still poorly understood and often assumed to remain at short distance from the breeding site (EFSA, 2017). However, recent studies showed that the dispersion of *Culicoides* may be higher than originally thought: engorged females of *C. chiopterus*, positive for cattle, were captured in sheep farms with no cattle less than 2 km away (Garros et al., 2011); a mark-release-recapture experiment estimated a possible daily dispersal distance of 1.75 km for *C. punctatus* and *C. pulicaris* (Kirkeby et al., 2009). In a similar study conducted in the *Obsoletus group*, thought to play a major role in BTV transmission in Europe (Ninio C., 2011), Kluiters et al. (2015) estimated a mean distance travelled of 2.15 km for female midges and 2.5 km for all midges, in two and three nights respectively. Here, we thus considered a weekly flight distance of 5 km and estimated a value of 0.4 for Ψ_P. as follows:

We approximated cantons by disks of radius R_{canton} ($=\sqrt{\frac{A_{canton}}{\pi}}$), with A_{canton}, the average surface of a canton. We considered the area covered by *Culicoides* coming from canton *k* as a circle of radius R_{culi} ($=R_{canton} + d_{flight}$), and area A_{culi} ($= \pi*(R_{culi}^2 - R_{canton}^2)$), with d_{flight}, the average distance that can be covered by *Culicoides* flight in a week. Hence, the proportion of area of any

146 neighbor of canton *k* on the pasture network, that can be reached every week by *Culicoides*
147 coming from canton *k* is Ψ_P ($= \frac{A_{culi}}{n_{neigh} * A_{canton}}$) , with n_{neigh}, the average number of neighbors (*i.e.*
148 the average degree) in the pasture network.

149 Knowing the average canton surface (A_{canton}=150 km^2), the average degree of the pasture
150 network (N=5), and considering a weekly flight distance of 5 km for female *Culicoides* (Kluiters
151 et al., 2015), we assumed that Ψ_P would take the value of 0.4, and varied this value (0.2, 0.3, 0.5
152 *i.e.* dispersal of 3, 4 and 6 km) in a sensitivity analysis.

Supplement S3: ABC-RF for model comparison

To avoid unnecessary complexity, we checked whether model fit would benefit from a detailed description of BTV transmission processes: we investigated the need for land-cover variables in the representation of within-canton transmission, and the need for the different contact networks in the representation of between-canton transmission.

S3A. Model selection by random forest

To assess the required level of complexity, we built separate models including various combinations of the variables and networks of interest and compared them using random forest classification methods (Pudlo et al., 2016). We sampled 10,000 sets of parameter values and used them to simulate surveillance and serological data with each model. We grew a forest of 1,000 trees to train the RF classifier on the simulated data, and applied the results of the RF computation to the observed surveillance and serological data. We obtained classification votes for each model, representing the number of times they were selected in the forest. The model collecting most votes was the one providing the best fit to the observed data. These results allowed us to choose a model and a set of variables used for all subsequent analyses.

S3.B. Environmental variables included in BTV within-canton transmission

To select the land-cover metrics that should be included in the model, we built five separate models. One of them did not include any environmental variable ($model_0$). The proportion of pastures per canton was added to the four remaining ones. Its effect on the force of infection, as given by the $\beta_0 * Env(k)$ term, could be modulated by both edge densities ($model_{AF}$), by only one of them ($\beta_F=0$ or $\beta_A=0$, $model_A$ and $model_F$ respectively), or by none of them ($\beta_A = \beta_F = 0$, $model_1$).

The models differed in the expression of the term *Env(k)*:

- $model_0$: *Env(k) = 0*
- $model_1$: *Env(k) = dPast(k)*
- $model_A$: $Env(k) = dPast(k) * (1 + \beta_A * e_{PA}(k))$
- $model_F$: $Env(k) = dPast(k) * (1 + \beta_F * e_{PF}(k))$
- $model_{AF}$: $Env(k) = dPast(k) * (1 + \beta_A * e_{PA}(k)) + \beta_F * e_{PF}(k)))$

with $dPast(k)$, the proportion of the surface of canton *k* covered in pastures; $e_{PA}(k)$and $e_{PF}(k)$, the edge densities of pasture *vs* arable lands and *vs* forest respectively, in canton *k*.

The RF classifier was trained on simulated data generated from parameter values sampled in uniform distributions. For β_0 , Ψ_F, θ, we used the prior distribution found in Table S3 and described in more details in Supplement S4.C. The bounds for the prior distributions of β_A, β_F, were chosen from the distributions of edge densities $e_{PA}(k)$ and $e_{PF}(k)$in all cantons, so that the effect of edge densities would take the minimal value of -1, corresponding a force of infection of zero. We thus allowed β_A and β_F to vary uniformly between -3.5 and 3.5 when only one edge density was accounted for ($model_A$, $model_F$), and between -1.75 and 1.75 when they were accounted for simultaneously ($model_{AF}$).

Table S1: Comparison of the models including various sets of land-cover variables: first of the five models (left), then of two of them, given the results of the first step (right).

Proportion of votes (/1,000 trees)					**Proportion of votes (/1,000 trees)**	
$model_0$	$model_1$	$model_A$	$model_F$	$model_{AF}$	$model_1$	$model_A$
0.004	0.213	0.291	0.232	0.260	0.444	0.556
Post probability					**Post probability**	
0.416					0.519	

The results (Table S1) showed that the proportion of pastures was crucial to representing BTV within-canton transmission. We also compared the simplest model that included the proportion of pastures, with the one that collected the most scores, to assess whether

complexifying the model provided a true improvement to model fit. As the two models performed equally (Table S1), we concluded that there was no additional benefit to model fit with the inclusion of any of the edge densities. We selected the most parsimonious model (*model$_1$*), with the proportion of pastures as only environmental variable, and used it for all subsequent analyses.

S3.C. Contact networks included in BTV between-canton transmission.

To assess which of the three contact networks was needed to represent BTV spread between French cantons, we built seven separate models, for all combinations of the three networks:

- *model$_{past}$*, *model$_{farm}$*, *model$_{trade}$*, for each network independently
- *model$_{past_farm}$*, *model$_{past_trade}$*, *model$_{farm_trade}$*, for all pairs of networks
- *model$_{pft,}$* for all networks together

The distributions used for sampling 10,000 sets of parameter values ($\beta_{0,}$ $\Psi_{F,}$ θ) are the same as those used in the previous section (Supplement S3.B). The results clearly showed that no network on its own was enough to represent BTV spread to new areas and that a combination of the networks was needed (Table S2). The best fit was obtained when all networks were combined. As the model including only the farm and pasture networks came second, we compared the two models that collected the most votes to assess whether complexifying the model provided a true improvement to model fit (Table S2). In the second step of model comparison, the model including all networks collected >75% of the votes, so that we kept it and used it for all subsequent analyses.

Table S2: Comparison of the models including various combinations of contact networks: first of the seven models (top), then of the two ones providing the best model fit (bottom).

	Proportion of votes (/1,000 trees)	**Post**

$model_{past}$	$model_{farm}$	$model_{trade}$	$model_{past\_farm}$	$model_{past\_trade}$	$model_{farm\_trade}$	$model_{pft,}$	**probability**
0.003	0.003	0.008	0.433	0.035	0.005	0.513	0.52

219

Proportion of votes (/1,000 trees)		**Post probability**
$model_{farm\_trade}$	$model_{pft}$	
0.236	0.764	0.791

220 **Supplement S4: Details on parameter estimation**

221 *S4.A. Adaptive population Monte-Carlo approximate Bayesian computation method*

222 ABC likelihood-free methods only require being able to sample from the likelihood. They
223 are useful for complex, stochastic models where estimating the full likelihood is difficult or
224 impossible (Pritchard et al., 1999). They are based on the generation of sets of model parameters
225 values (particles) initially sampled from the joint prior distribution of each parameter, followed
226 by the selection of the particles for which the model outputs (summary statistics) satisfy a
227 proximity criterion with the target data. Posterior distributions for each parameter are then
228 obtained from the selected particles.

229 Sampling the whole parameter space is poorly efficient and computationally demanding;
230 numerous methods have been developed to improve the basic rejection algorithm (Marjoram et
231 al., 2003; Del Moral et al., 2006; Beaumont MA et al., 2009; Wegmann et al., 2009; Drovandi
232 and Pettitt, 2011). The required number of simulations can be reduced by spending more time in
233 the areas of the parameter space for which model outputs are frequently close to the target ones.
234 The algorithm proposed by Lenormand et al. (2013) starts by the generation of a set of particles,
235 to which a weight is attributed given their importance for the inference combined with the prior
236 probabilities, as in Beaumont et al. (2009). In the next estimation steps, the particles are
237 resampled based on the weights from the previous one and perturbed according to a pre-defined
238 perturbation kernel and a new set of weights is calculated. The tolerance controlling the matching

is decreased at each step, and the algorithm proposed by Lenormand et al. (2013) has the additional advantage of providing automatically computed thresholds, defined as the α-quantile of the distance between simulated and observed summary statistics in the previous sample set. This algorithm provides an explicit stopping criterion as computation stops when the proportion of particles satisfying the tolerance level among the newly generated particles is below a chosen minimal acceptance value. The final result is the last set of particles generated, with their associated weights.

S4.B. Summary statistics

The three summary statistics used for inference were built from surveillance and seroprevalence data from the 2007 epizootic wave. For surveillance data, we attributed a score of one for departments with reporting cantons by winter 2007/08, zero otherwise. In each simulation, we extracted the number of departments with a score of zero (*FF*) and one (*TT*) in both simulated and observed data, and we calculated an L2 distance (*Surv*). For seroprevalence, we calculated species-specific L2 distances of the number of seropositive animals detected in each department sampled in the serosurvey conducted in winter 2007/08 ($Sero^{sp}$). The summary statistics were the following:

$$Surv = \sqrt{(TT_{sim} - TT_{obs})^2 + (FF_{sim} - FF_{obs})^2}$$

$$Sero^{sp} = \sqrt{\sum_{dpt^{sp}} (NPos_{sim}^{sp} - NPos_{obs}^{sp})^2}$$

With *sim* and *obs* for simulated and observed measures respectively; *TT*, the number of departments with reporting cantons in both simulated and observed data, *FF*, in neither simulated

nor observed data; dpt^{sp}, the seven and four departments where cattle and sheep had been sampled respectively; $NPos^{sp}$, the number of seropositive animals detected in these departments.

S4.C Prior distributions

We used uniform distributions for all parameters: an uninformative prior for the proportion of animals moved weekly through the functional network for which we had no previous knowledge, and informative priors for the other two. The ranges were extrapolated from official outbreak records or external data, as described here below.

Table S3: Ranges for the uniform prior distribution of the estimated parameters:

Notation	Description	Possible range	Prior range distribution	Source
β_0	Baseline exposure of hosts to vectors	[0 - +∞[	[1,000 – 10,000]	External data
Ψ_F	Proportion of animals moved weekly through the farm network	[0 - 1]	[0 - 1]	Non informative prior
θ	Probability of control measures being implemented on movements of cattle through the trade network	[0 - 1]	[0.9 – 1]	Outbreak record (Drouet M., 2010)

The bounds of the prior distribution for β_0 were chosen arbitrarily after having verified that they covered all plausible values of the parameter space. Given the general modelling assumption that within-canton transmission was the main driver of infection in infected areas, we focused on a subset of 22 cantons for which we knew the dates of detection and the levels of seroprevalence in winter 2007/08. We explored the parameter space by simulating infection at different dates and with various β_0 values. We plotted the evolution in time of the proportion of *S*, *I* and *R* animals in each canton and scenario and concluded that there would hardly be any within-canton transmission for β_0 below 1,000, while there would be saturation for β_0 above 10,000.

Movement restrictions applied to the trade network had most likely been correctly implemented (Drouet, 2010). In a preliminary analyses, we simulated 100 outbreaks with θ

277 values comprised between 50 and 100%. We obtained outbreaks that were compatible with those
278 observed for θ values above 90%. Otherwise, there was saturation in the summary statistics as the
279 whole French territory was infected by winter 2007 in all simulations. We thus limited the prior
280 distribution for the θ parameter to the [0.9-1] interval.

281 We validated these choice of prior distributions in the assessment of the ability of our
282 framework to estimate known parameter values (Supplement S6).

Supplement S5: Network analysis

Introduction

Network-based approaches have been developed to face the many information on livestock movements now available through the systematic building of exhaustive databases. They rely on the study of the relationships among farms or livestock operations, and allow the study of sequences of movements producing paths on which infectious diseases can spread (Dubé et al., 2009). Networks analyses have proven useful in understanding the implications of long-range host movements by exploring the vulnerability of the French network of cattle movements to the spread of pathogens (Rautureau et al., 2012), or assessing the risk of a large epidemic of foot-and-mouth disease in the UK (Kao et al., 2006). More recently, network analyses have been used to explore the implications of short-range host movements and biosecurity measures on disease spread in French cattle (Palisson et al., 2017).

Methods

We used three distinct contact networks, in which the nodes were cantons, to represent the likely movements of hosts and vectors: *(i)* the pasture network, that represents midges flight from pastures, located in different cantons but less than one km apart, a distance used by Palisson et al. (2017) to represent the most likely routes of vector-borne disease transmission across the densely connected French pastures; *(ii)* the farm network, that represents the movements of cattle or sheep between pastures belonging to the same farm, but located in different cantons; *(iii)* the trade network, that represents movements of cattle traded between farms located in different cantons. The links were aggregated at the canton-level, so that a link existed between two cantons: *(i)* in the pasture network, if at least two pastures from each canton were close enough (N=8,009 links); *(ii)* in the farm network, if at least one farm had pastures in both cantons (N=22,905 links); *(iii)* in the trade network, if cattle had been traded between at least two farms

307 located in each canton (N=174,702 in the 2nd semester of 2007). The pasture and farm networks
308 were static with links existing at all times and movements through these links as likely to go
309 either way. The trade network was temporal and oriented, linking different donors and recipients
310 every week. The topological analysis of this network was performed on the time period of the
311 first epizootic wave: by aggregating all links from the second semester of 2007, or by aggregating
312 all links on each week of this semester.

313 We computed classical network indicators for all of them to better grasp their specific
314 topological properties, and understand their likely impact on transmission patterns: the average
315 degree (*i.e.* mean number of connections from a node to all others); the average path length (*i.e.*
316 the most typical separation of one pair of nodes, an indicator of the effective size of the network);
317 the clustering coefficient (*i.e.* proportion of one's neighbors who are also neighbors of one
318 another, a measure of how the nodes tend to cluster together). We also studied the fragmentation
319 of the networks, an indicator of the network vulnerability, by looking for the presence of giant
320 components, *i.e.* subnetworks in which all nodes are linked, meaning that if any node is infected,
321 all other nodes may be subsequently reached (Kao et al., 2006).

322 **Results**

323 The pasture network had an average degree of 5.0, an average path length of 27, and a
324 clustering coefficient of 37.1%. The farm network had an average degree of 13.9, an average path
325 length of 9.9 and a clustering coefficient of 43.2%. In addition, with an average path length of the
326 same magnitude than that of a random network of the same size (3.4), but a clustering coefficient
327 >100 times bigger, the farm network showed small world properties.

328 The trade network had the smallest average path length of all three networks: 2.9 if
329 aggregating all links of the second semester of 2007 (and 6.1 in a week), and an average degree
330 of 54.8 (8.6 in a week). In addition, the number of connections from each nodes was highly

variable with a power-law distribution, which gave that network scale-free properties. Its clustering coefficient was 23.5%.

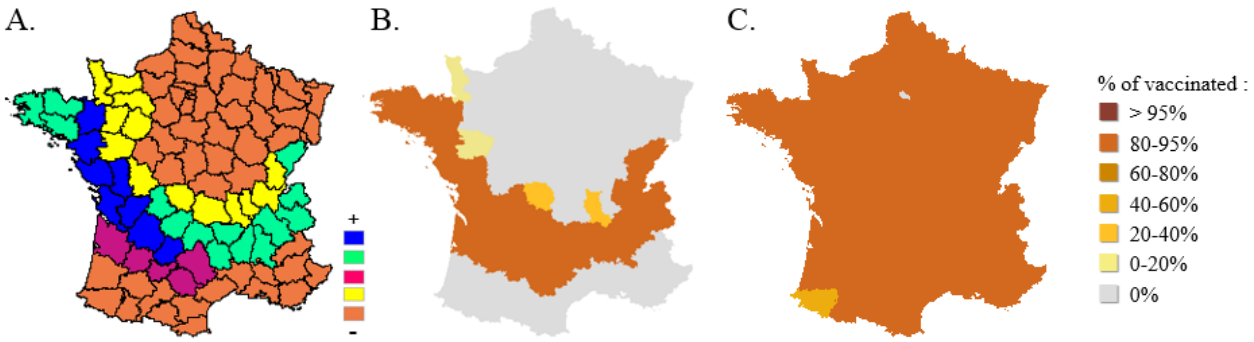
None of these networks was fragmented: they had big giant components including 96.3% and 98.3% of the nodes in the pasture and farm networks respectively. All nodes of the trade network (99.9%) were part of its weakest giant component, and 91.2% of its strongest one.

Discussion

BTV could reach the whole French territory through either one of the three contact networks, but their different topological properties had different implications in terms of impact on disease spread (Dubé et al., 2009; Rautureau et al., 2012; Palisson et al., 2017): the pasture network was basically a grid, with all nodes located at close distance and having few neighbors; the farm network was a grid with additional small-world properties: most pastures from the same farm were located at close distance, but some of them were further apart, which introduced shortcuts in transmission; and the trade network had scale-free properties: there were hubs that, if reached, spread the infection fast and far. By comparing models in which transmission could occur through different combinations of networks, we showed that the representation of BTV-8 transmission was significantly improved by combining these networks rather than considering them individually.

Figure S1: Alternative strategy for the 2008 emergency vaccination campaign.

A. Vaccination schedule in which the order of priority for the distribution of the limited number of vaccine doses in cattle and sheep was defined to create a buffer zone beyond the previously affected areas ("AFFSA scenario, AFSSA, 2008). The order of priority is indicated by the color code, vaccination was spread out between May and September 2008. We vaccinated all departments by order of priority until reaching the target coverage of 95% as we considered that about 5% of the population was not eligible for vaccination (aged<10 weeks, vaccination exemptions). B, C. vaccination coverage achieved by the end of the campaign (October 2008) in cattle (B) and sheep (C).



Supplement S6: Validation of the framework (POC)

To assess the ability of our framework to estimate parameter values using the chosen summary statistics, we randomly sampled a hundred particles from a subset of prior distributions. The subsets had been chosen to explore preferentially realistic areas in the parameter space ($2{,}000<\beta_0<5{,}000$, $0.2<\Psi_F<0.8$, $\theta>0.92$). Using these particles, we generated summary statistics, considered as pseudo-observations. We then used the ABC-APMC procedure and quantified its reliability. We used uniform prior distributions on the following intervals: [1,000 – 10,000], [0 - 1] and [0,9 - 1] for β_0, Ψ_F and θ respectively.

The mean estimated values were close to the input ones for most β_0 and Ψ_F values tested (Figure S2). However we identified some saturation in the model for extreme values of β_0 and Ψ_F (β_0<3,000, Ψ_F<0.3, Ψ_F>0.6), with little variation in the pseudo-observed summary statistics above –and below- these thresholds and estimated values regressing towards the mean of the prior distribution.

Figure S2: Validation of the framework in an *in-silico* analysis.

Posterior distributions and input value obtained for the three estimated parameters: β_0 (A), Ψ_F (B) and θ (C).

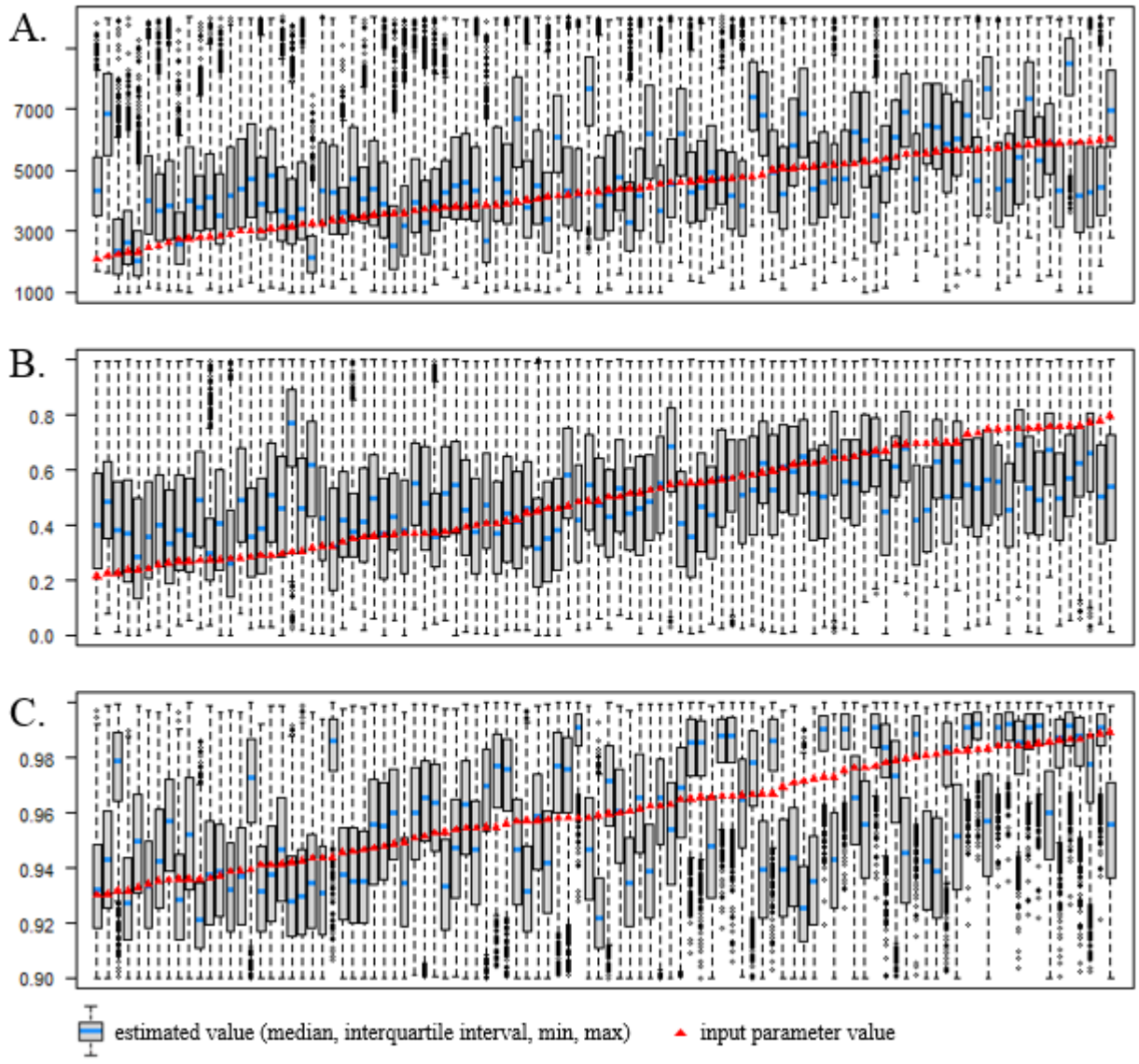


Figure S3: Model fit: comparison of observed *vs* simulated summary statistics.

Summary statistics generated after sampling 1,000 particles in the joint posterior distributions. A. Simulated and observed number of samples with detectable antibodies against BTV among those collected in the serosurvey conducted in winter 2007/08 in cattle and sheep, in seven and four departments respectively; B. Number of French departments with the same status by winter 2007/08 (*i.e.* with or without BTV detection based on clinical suspicion) in both observed and simulated data. CI95%, 95% confidence intervals; dpts, departments; Nb, number of.

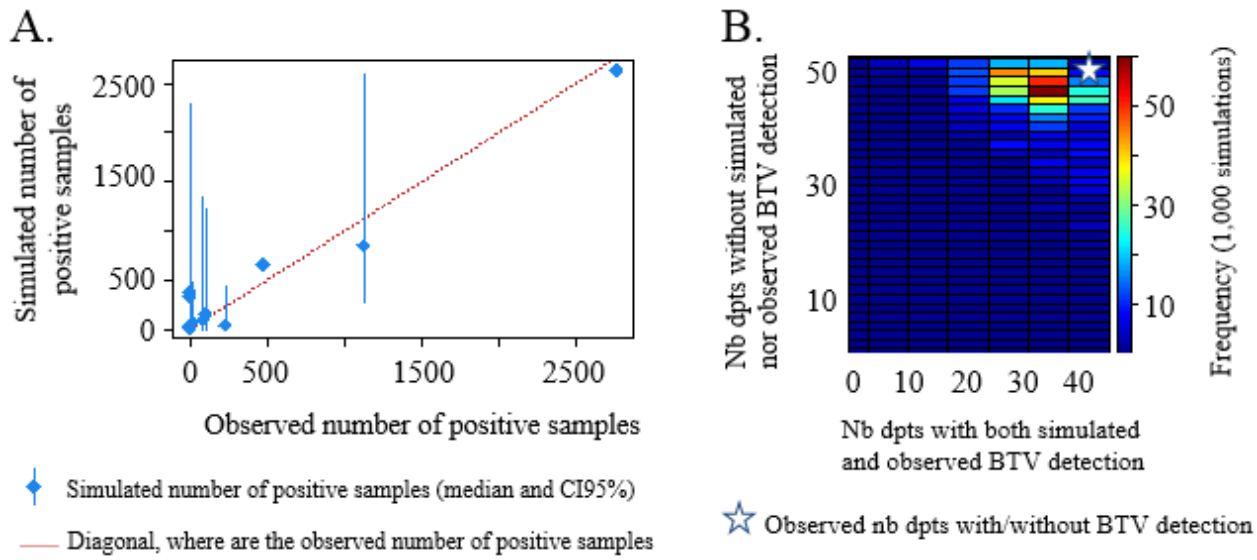


Figure S4: External validation: spatio-temporal pattern of the apparent infection.

Illustration of the ability of the parametrized model to reconstruct the epizootic wave that crossed France in 2007 and 2008, by reconstructing one possible spatio-temporal pattern of detection, using one of the particle with the highest weights. Reporting cantons were mapped every six weeks from mid-July 2007 until late October 2008 (A, C) and confronted with observed data (B., D.). A, B. 2007 epizootic wave; C, D. 2008 epizootic wave.

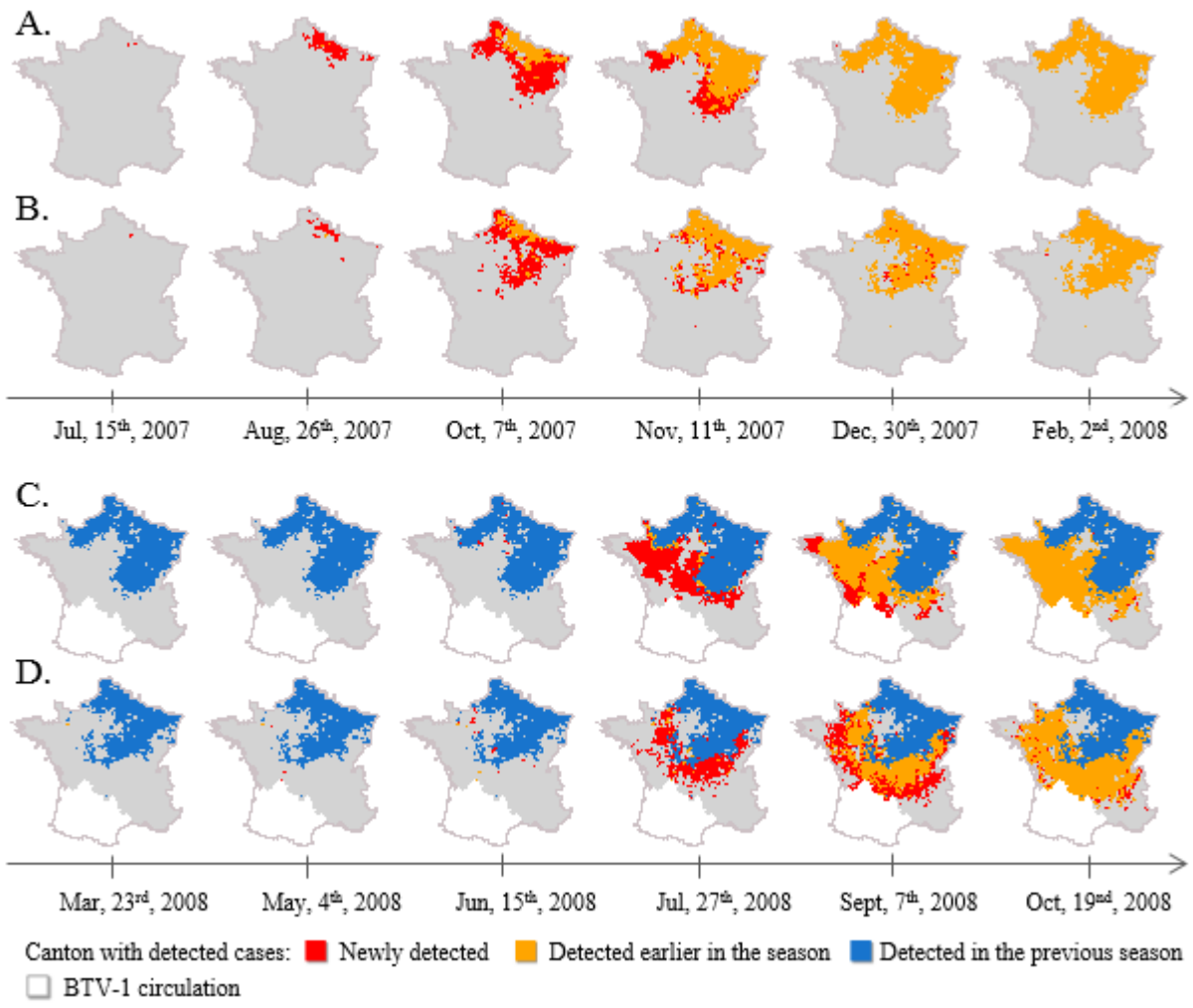
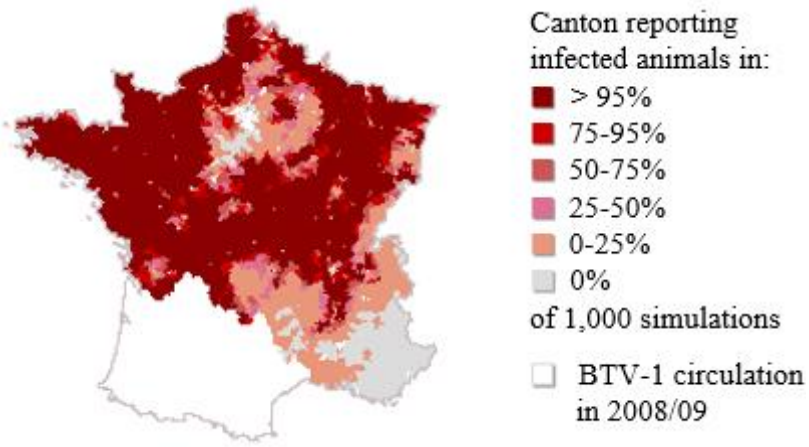


Figure S5: Reporting cantons after the 2007/09 outbreak.

Frequency of detection of at least one infected animals from 2010 in each canton in 1,000 simulations. Almost all cantons had reported infected cases in the 2007/09 outbreak and the last infected case was detected in 2009. To see whether the simulated levels of infection in 2010 would have been detected in our setting, we computed a date of first detection after the official end of the outbreak, *i.e.* from January 2010. We applied the same probability of detection (Δ) and showed that there would have been reporting cantons in most of the French territory in most simulations.



Supplement S7: Sensitivity analyses

S7.A. Sensitivity analysis on parameter estimates

Methods

A sensitivity analysis was performed to evaluate the effect of two key parameters (Ψ_P, Δ, Table 1) on the estimated parameter values. We analyzed the first-order effects using 5 different values (fixed deviation of 25%) for each fixed parameter, *i.e.* 25 combinations and 25 posterior distributions per parameter. For Ψ_P, the baseline value of 0.4 corresponded to weekly flight distance of 5 km for *Culicoides*, and we investigated alternative values of 0.2, 0.3, 0.5 and 0.6, corresponding respectively to flight distances of 3, 4, 6 and 7 km. For Δ, the baseline value was 0.02 and we investigated alternative values of 0.01, 0.03, 0.04 and 0.05. We used three generalized linear model (GLMs), one per estimated parameter with a full factorial design, to predict the effect of an increase of 25% of each fixed parameter on the average values of each estimated parameter. We compared the relative error (RE) induced by a 25% change of fixed parameters with the coefficient of variation (CV) of the posterior distributions obtained with the default values.

Results

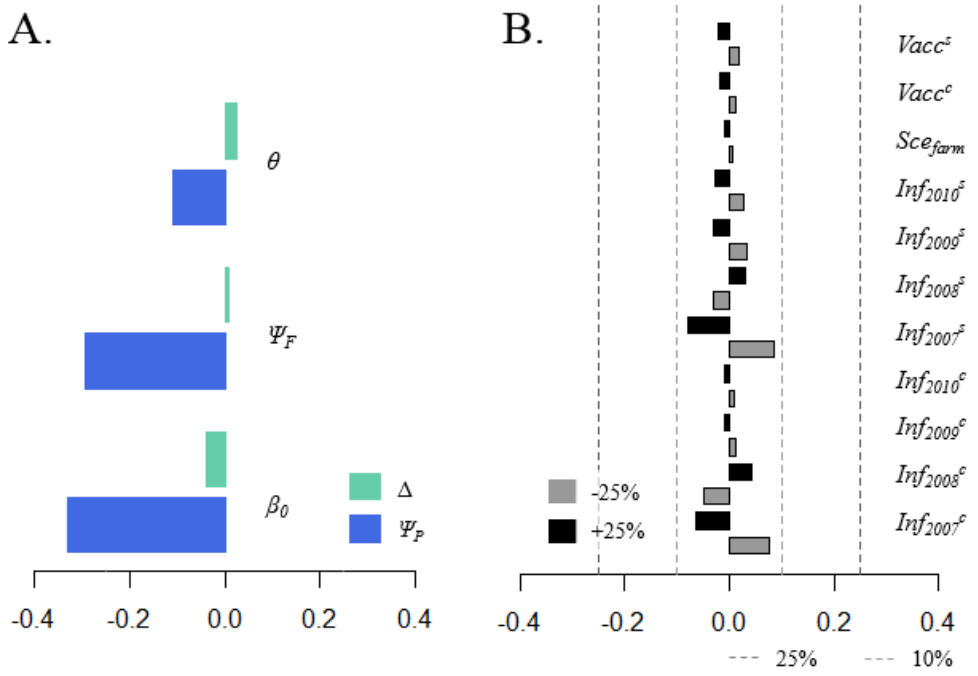
The ratios RE/CV were always below one, with little effect of a 25% variation of Δ but a stronger effect of a 25% of variation of Ψ_P on both Ψ_F and β_0 estimates (Figure S6.A). However, for each couple of Ψ_P and associated parameter estimates, we showed little variation of model predictions (Figure S6.B).

Discussion

We showed in this sensitivity analysis that parameter estimation was impacted by the fixed proportion of canton surface reachable by midges from the neighboring cantons, a value

extrapolated from flight distances measured in a capture/recapture assay of *Culicoides obsoletus* around farms (Kluiters, 2015). This proportion is applied to the force of infection in neighboring cantons as a multiplication factor of the baseline exposure of hosts to vectors (β_0), hence the identifiability issue. Yet, we also showed that even if we had considered slightly smaller or larger weekly *Culicoides* flight distances, we would have ended with similar model predictions.

Figure S6: Results of the sensitivity analysis of parameter estimates and model predictions to variations of two key parameters which values were fixed: the probability of detection of infectious animals with clinical signs (Δ) and the proportion of canton surface that can be reached by vectors coming from neighboring cantons *(Ψ_P)*. A. Ratios of the coefficient of variation of each parameter estimated with a 25% increase of Δ and Ψ_P, and of the coefficient of variation of each parameter estimate in the model with the default Δ and Ψ_P values. B. Variation (%) of the average model predictions compared to that obtained with a +/-25% variation of Ψ_P and associated parameter estimates. We investigated the variation from the baseline scenario of: $Vacc^{sp}$, the proportion of vaccine doses administered to already immune animal in the 2008 emergency vaccination campaign; Sce_{farm}, the proportion of first infections (in canton) attributed to the farm network; $Inf_{2007\text{-}10}^{sp}$, the number of infected animals in each year; *c* for cattle, *s* for sheep.



S7.B. Sensitivity analysis on model predictions

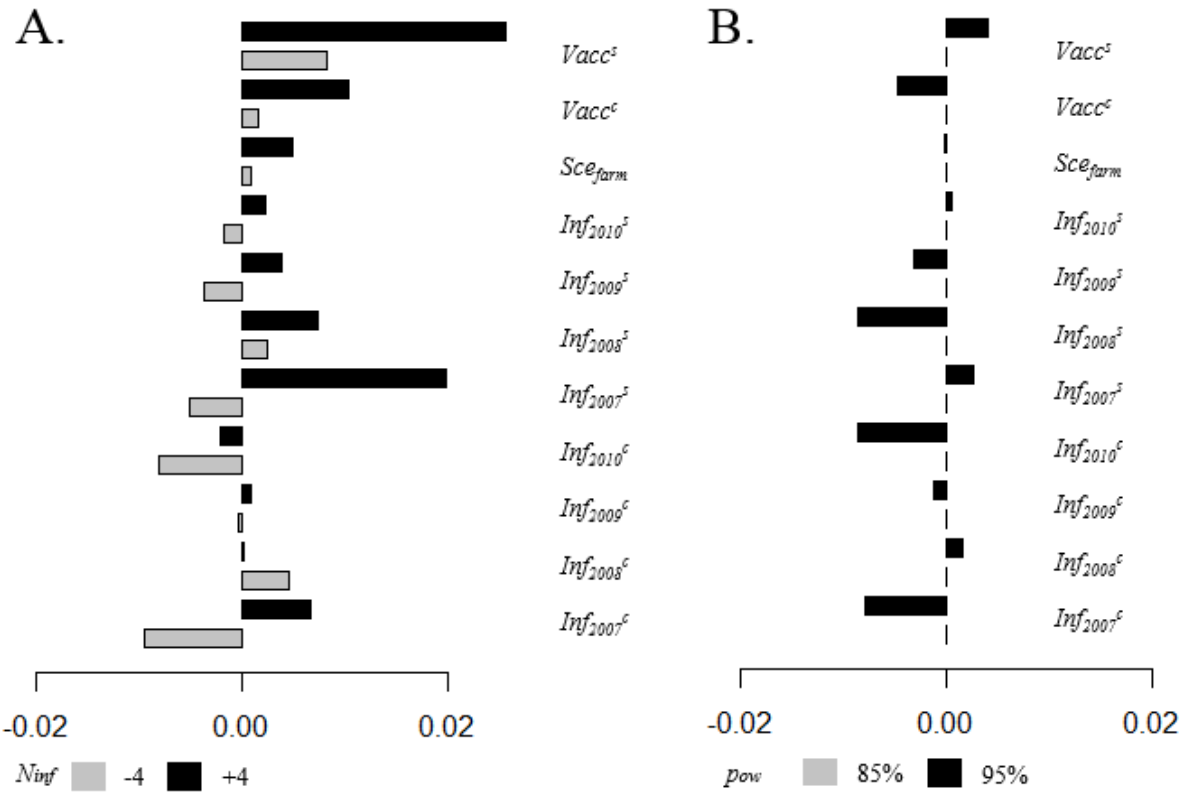
Methods

We did not use the same approach for the sensitivity analysis of model predictions to initial conditions as we could not assume a linear effect of these conditions on the outputs. We investigated the amplitude of deviation of 11 model predictions from their default value with deviations of initial conditions around their fixed values (N_{inf}, p_{ow}). These predictions were: the species-specific proportion of all vaccine doses administered to already immune animals ($Vacc^{sp}$), the species-specific total number of infected animals per season of circulation (Inf_{year}^{sp}), and the proportion of BTV introduction to new areas attributed to the farm network (Sce_{farm}). N_{inf}, the number of infected cattle introduced in the cantons where infection was seeded took the baseline value of 5. We investigated alternative values +/- 4 cattle. The cantons where BTV was reintroduced in season *n+1* were those where BTV was still circulating on the date when temperatures dropped below the T_{min} threshold in p_{ow} of the cantons in season *n*. p_{ow} took the average value of 90% and we investigated alternative values +/- 5%.

Results

We showed only little effect of variations of initial conditions on the average model predictions (<2%) (Figure S7), which may be even less if accounting for the standard deviation around the fixed value in the 1,000 simulations. This confirms that our predictions are robust to our assumptions on initial conditions, within a reasonable range of variation.

Figure S7: Sensitivity analysis of model predictions to variations of the initial conditions (N_{inf} and p_{ow}). Variation (%) of the default average model predictions compared to that obtained with a variation of N_{inf} (A) and a p_{ow} (B) in 1,000 simulations. We investigated the variation from the baseline scenario of: $Vacc^{sp}$, the proportion of vaccine doses administered to already immune animal in the 2008 emergency vaccination campaign; Sce_{farm}, the proportion of first infections (in canton) attributed to the farm network; $Inf_{2007\text{-}10}^{sp}$, the number of infected animals in each year; *c* for cattle, *s* for sheep.



ADDITIONAL REFERENCES

AFSSA, 2008. Avis de l'Agence française de sécurité sanitaire des aliments sur le risque de diffusion de la fièvre catarrhale ovine à sérotypes 1 et 8 en France et les mesures associées pour en diminuer le niveau, Saisine 2008-SA-0033.

Beaumont MA, Cornuet JM, Marin JM, Robert CP, 2009. Adaptive approximate Bayesian computation. Biometrika 96, 983–990.

Birley MH, Boorman JPT, 1982. Journal Article Estimating the Survival and Biting Rates of Haematophagous Insects, with Particular Reference to the Culicoides obsoletus Group (Diptera, Ceratopogonidae) in Southern England Published by: British Ecological Society DOI: https://www.jstor.org/stable/4315 Page Count: 14. J. Anim. Ecol. 51, 135–148. https://doi.org/10.2307/4315

Carpenter, S., Wilson, A., Barber, J., Veronesi, E., Mellor, P., Venter, G., Gubbins, S., 2011. Temperature dependence of the extrinsic incubation period of orbiviruses in Culicoides biting midges. PloS One 6, e27987. https://doi.org/10.1371/journal.pone.0027987

Courtejoie, N., Salje, H., Durand, B., Zanella, G., Cauchemez, S., 2018. Using serological studies to reconstruct the history of bluetongue epidemic in French cattle under successive vaccination campaigns. Epidemics. https://doi.org/10.1016/j.epidem.2018.05.005

Del Moral, P., Doucet, A., Jasra, A., 2006. Sequential Monte Carlo Samplers. J. R. Stat. Soc. Ser. B Stat. Methodol. 68, 411–436.

Drouet M., 2010. Synthèse sur l'évolution des mesures de «policesanitaire » mises en place vis-à-vis de la FCO en France. Bull. Épidémiologique 35, 13–14.

Drovandi, C.C., Pettitt, A.N., 2011. Estimation of parameters for macroparasite population evolution using approximate bayesian computation. Biometrics 67, 225–233. https://doi.org/10.1111/j.1541-0420.2010.01410.x

Dubé, C., Ribble, C., Kelton, D., McNab, B., 2009. A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development. Transbound. Emerg. Dis. 56, 73–85. https://doi.org/10.1111/j.1865-1682.2008.01064.x

Durand, B., Zanella, G., Biteau-Coroller, F., Locatelli, C., Baurier, F., Simon, C., Le Dréan, E., Delaval, J., Prengère, E., Beauté, V., Guis, H., 2010. Anatomy of bluetongue virus serotype 8 epizootic wave, France, 2007-2008. Emerg. Infect. Dis. 16, 1861–1868. https://doi.org/10.3201/eid1612.100412

EFSA AHAW Panel (EFSA Panel on Animal Health and Welfare), 2017. Scientific opinion on bluetongue: control, surveillance and safe movement of animals. EFSA J. 15(3):4698. https://doi.org/10.2903/j.efsa.2017.4698

Garros, C., Gardès, L., Allène, X., Rakotoarivony, I., Viennet, E., Rossi, S., Balenghien, T., 2011. Adaptation of a species-specific multiplex PCR assay for the identification of blood meal source in Culicoides (Ceratopogonidae: Diptera): applications on Palaearctic biting midge species, vectors of Orbiviruses. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 11, 1103–1110. https://doi.org/10.1016/j.meegid.2011.04.002

Gerry, A.C., Mullens, B.A., 2000. Seasonal abundance and survivorship of Culicoides sonorensis (Diptera: Ceratopogonidae) at a southern California dairy, with reference to potential bluetongue virus transmission and persistence. J. Med. Entomol. 37, 675–688.

Goffredo, M., Romeo, G., Monaco, F., Di Gennaro, A., Savini, G., 2004. Laboratory survival and blood feeding response of wild-caught Culicoides obsoletus Complex (Diptera: Ceratopogonidae) through natural and artificial membranes. Vet. Ital. 40, 282–285.

Graesbøll, K., Bødker, R., Enøe, C., Christiansen, L.E., 2012. Simulating spread of Bluetongue Virus by flying vectors between hosts on pasture. Sci. Rep. 2, 863. https://doi.org/10.1038/srep00863

Gubbins, S., Carpenter, S., Baylis, M., Wood, J.L.N., Mellor, P.S., 2008. Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. J. R. Soc. Interface 5, 363–371. https://doi.org/10.1098/rsif.2007.1110

Guis, H., Caminade, C., Calvete, C., Morse, A.P., Tran, A., Baylis, M., 2012. Modelling the effects of past and future climate on the risk of bluetongue emergence in Europe. J. R. Soc. Interface 9, 339–350. https://doi.org/10.1098/rsif.2011.0255

Guyot, H., Mauroy, A., Kirschvink, N., Rollin, F., Saegarman, C., 2008. Clinical aspects of bluetongue in ruminants, Bluetongue in Northern Europe. OIE Publ. 34–52.

Institut de l'élevage, 2017. Dossier annuel ovins (No. 488), Economie de l'élevage.

Institut de l'élevage, 2016. Cas type ob11, Références.

Kao, R.R., Danon, L., Green, D.M., Kiss, I.Z., 2006. Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. Proc. Biol. Sci. 273, 1999–2007. https://doi.org/10.1098/rspb.2006.3505

Kirkeby C, Bødker R, Stockmarr A, Enøe C, 2009. Association between land cover and Culicoides (Diptera: Ceratopogonidae) breeding sites on four Danish cattle farms. Entomol. Fenn. 20, 228–232.

Kluiters, G., Swales, H., Baylis, M., 2015. Local dispersal of palaearctic Culicoides biting midges estimated by mark-release-recapture. Parasit. Vectors 8, 86. https://doi.org/10.1186/s13071-015-0658-z

Lenormand, M., Jabot, F., Deffuant, G., 2013. Adaptive approximate Bayesian computation for complex models. Comput. Stat. 28, 2777–2796. https://doi.org/10.1007/s00180-013-0428-3

Marjoram, P., Molitor, J., Plagnol, V., Tavare, S., 2003. Markov chain Monte Carlo without likelihoods. Proc. Natl. Acad. Sci. U. S. A. 100, 15324–15328. https://doi.org/10.1073/pnas.0306899100

Mounaix, B., Caillaud, D., Echevarria, L., Reynaud, D., Fraboulet, M., Gorceix, M., Dupont, L., David, V., Lucbert, J., 2010. Estimation des impacts technico-économiques de la FCO-8 en 2007 au niveau de l'élevage. Bull. Épidémiologique 35, 17–19.

Ninio C., 2011. Fièvre catarrhale ovine dans les Ardennes : étude de la biologie des Culicoides et de leur rôle épidémiologique. Université Reims Champagne-Ardennes.

Palisson, A., Courcoul, A., Durand, B., 2017. Analysis of the Spatial Organization of Pastures as a Contact Network, Implications for Potential Disease Spread and Biosecurity in Livestock, France, 2010. PloS One 12, e0169881. https://doi.org/10.1371/journal.pone.0169881

Pritchard, J.K., Seielstad, M.T., Perez-Lezaun, A., Feldman, M.W., 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol. Biol. Evol. 16, 1791–1798. https://doi.org/10.1093/oxfordjournals.molbev.a026091

Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., Robert, C.P., 2016. Reliable ABC model choice via random forests. Bioinforma. Oxf. Engl. 32, 859–866. https://doi.org/10.1093/bioinformatics/btv684

Rautureau, S., Dufour, B., Durand, B., 2012. Structural vulnerability of the French swine industry trade network to the spread of infectious diseases. Anim. Int. J. Anim. Biosci. 6, 1152–1162. https://doi.org/10.1017/S1751731111002631

Rossi, S., Pioz, M., Beard, E., Durand, B., Gibert, P., Gauthier, D., Klein, F., Maillard, D., Saint-Andrieux, C., Saubusse, T., Hars, J., 2014. Bluetongue dynamics in French wildlife: exploring the driving forces. Transbound. Emerg. Dis. 61, e12–24. https://doi.org/10.1111/tbed.12061

Sumner, T., Orton, R.J., Green, D.M., Kao, R.R., Gubbins, S., 2017. Quantifying the roles of host movement and vector dispersal in the transmission of vector-borne diseases of livestock. PLoS Comput. Biol. 13, e1005470. https://doi.org/10.1371/journal.pcbi.1005470

Szmaragd, C., Wilson, A.J., Carpenter, S., Wood, J.L.N., Mellor, P.S., Gubbins, S., 2009. A modeling framework to describe the transmission of bluetongue virus within and between farms in Great Britain. PloS One 4, e7741. https://doi.org/10.1371/journal.pone.0007741

Wegmann, D., Leuenberger, C., Excoffier, L., 2009. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. Genetics 182, 1207–1218. https://doi.org/10.1534/genetics.109.102509