Insights on paratuberculosis in cattle from a metapopulation model and inference via a composite-likelihood approximation

Gael Beaunée, Pauline Ezanno, Alain Joly, Pierre Nicolas, Elisabeta Vergu

Insights on paratuberculosis in cattle from a metapopulation model and inference via a composite-likelihood approximation

Gael Beaunée[1], Pauline Ezanno[1], Alain Joly[2], Pierre Nicolas[3] and Elisabeta Vergu[3]

Bovine paratuberculosis, caused by *Mycobacterium avium subsp. paratuberculosis* (*Map*), is a worldwide enzootic disease of considerable economic importance for dairy cattle producers, mainly introduced into herds by purchasing infected animals. Its screening in the field is difficult due to its long incubation period and the low sensitivity of diagnostic tests.

The main objective of this work was to estimate key parameters of a multilevel dynamic model of *Map* spread from long term serological survey data. This study focuses on dairy cattle herds located in Brittany (Northwestern France), region with a high density of cattle (85% of cows are dairy cows).

Fig. 1 Cattle trade network between dairy herds in Brittany region, 2005–2013.
Nodes = administrative communes, with size ~ number of animals exchanged (incoming mvts (purchase) in red and outgoing mvts (sale) in blue).
Links = trade relationships, thickness ~ number of animals exchanged.

1 BIOEPAR, INRA, Oniris, Nantes, France; 2 GDS Bretagne, France;
3 MaIAGE, INRA, Université Paris-Saclay, 78350, Jouy-en-Josas, France
Author for correspondence : gael.beaunee@inra.fr (www.gaelbn.com);
Pierre Nicolas and Elisabeta Vergu contributed equally to this work

Regional model of Map spread and data overview

Stochastic mechanistic regional model of *Map* spread between dairy cattle herds (Fig. 1)

Within-herd
— Demography (herd-specific size and management)
— Infection dynamics (through environment, vertical transmission, feeding).

Between-herd
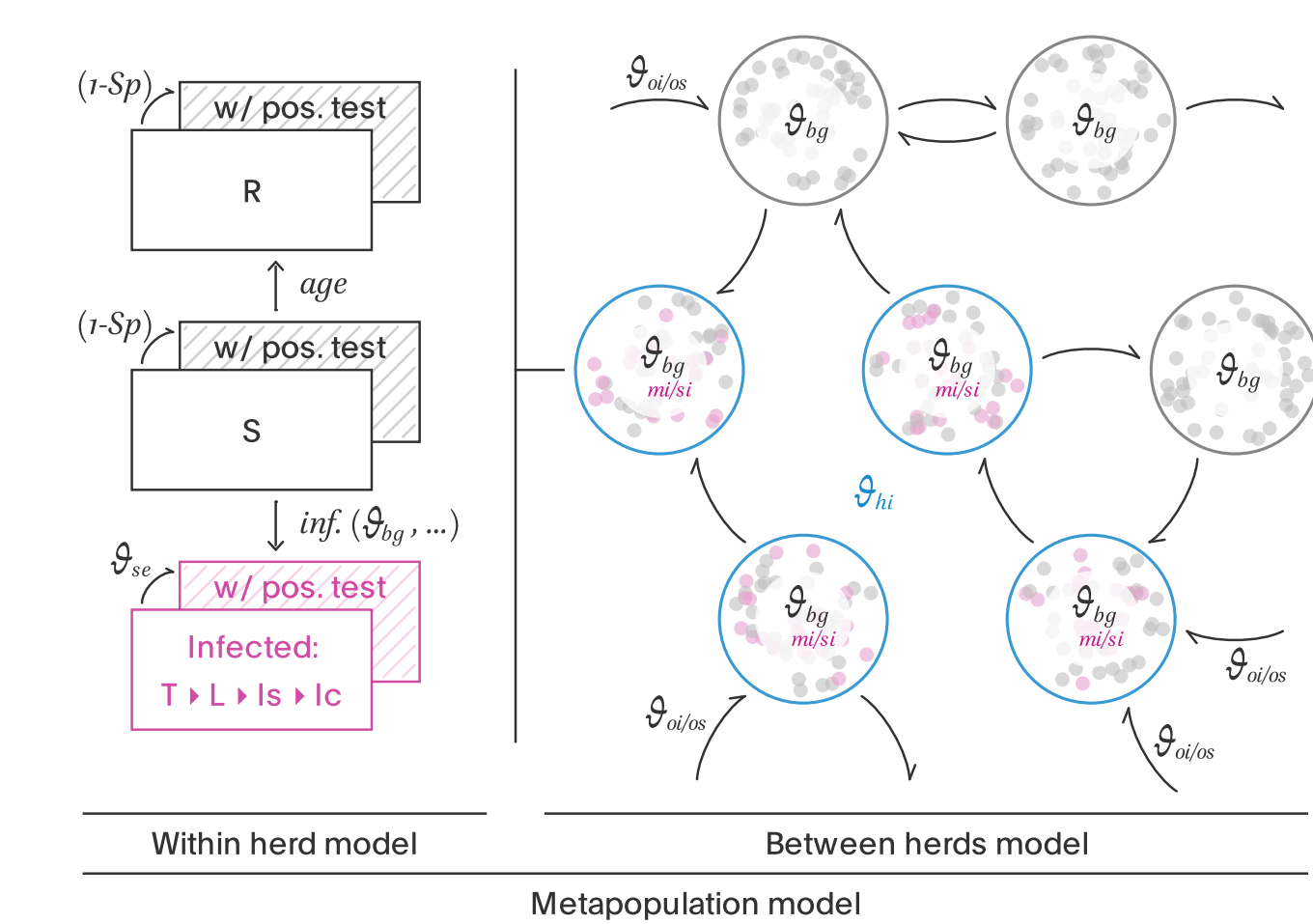— Trade animal movements as propagation route

Fig. 2 Conceptual model of *Map* spread within and between herds. Each of the populations (circles, right) connected by animal trade (arrows) is represented by its own demographic and infectious dynamics (left). S, susceptible; R, resistant; T, transiently infectious; L, latently infected; Is, moderate shedders; Ic, heavy shedders and clinically affected animals; Compartments with hatched line correspond to animals with a positive test result.

Animal trade data, Brittany, 2005–2013
Animal movements and demography, based on observed data for 12 857 herds.

Serological data set, Brittany, 2005–2013
Long-term longitudinal follow-up in herds: serological tests on animals over 2 years of age.
Selection of the sampling points used for estimation: [≥ 30] or [≥ 50% & ≥ 10] of animals tested in the herd at the same date (5434 screening events for 2278 herds).

Inference method

Parameters to be estimated
— The proportion of initially infected herds (θ_{hi})
— The within-herd prevalence distribution, mean (θ_{mi}) and standard deviation (θ_{si}), in initially infected herds
— The probability to purchase an infected animal from outside the metapopulation (baseline value θ_{oi} and slope θ_{os})
— The within-herd transmission rate through the general environment (θ_{bg})
— The sensitivity of the diagnostic test (θ_{se})

Inference procedure
— Comparison of observations and simulations using a composite log-likelihood criterion
— Adequate exploration of the parameter space: optimization using the Nelder-Mead Simplex algorithm

Numerical exploration
— Empirical identifiability analysis on simulated data
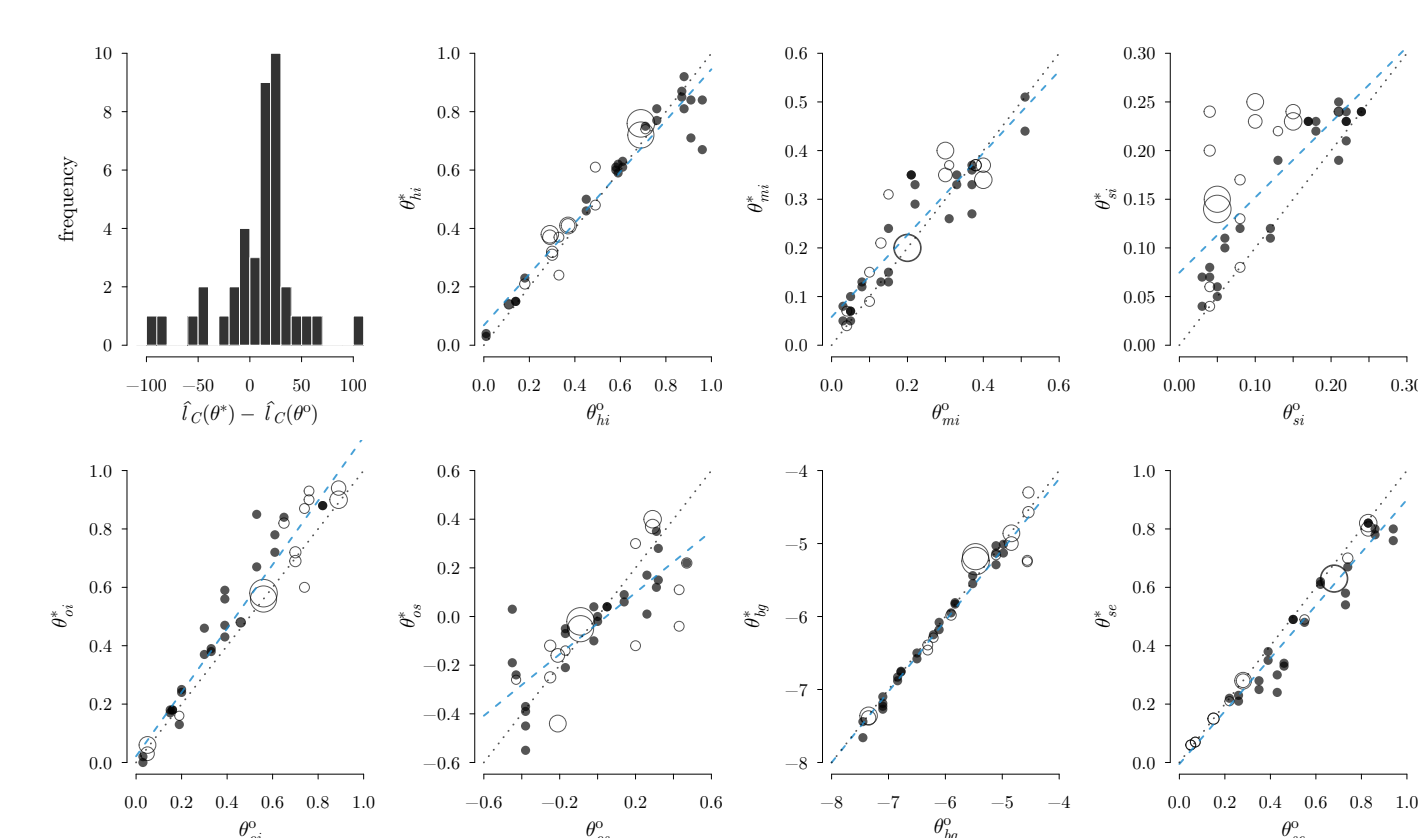— Definition of Confidence intervals using a profile likelihood approach

Fig. 3 Numerical identifiability analysis. Comparison between true and estimated parameter values. (dotted line: bisecting line, dashed line: simple regression).

Results

— Empirical identifiability was verified on simulated data (Fig. 3).
— Point estimates and profile likelihoods (Fig. 4 et 5) indicate a very large proportion (0.97) of infected herds with a low within-herd prevalence (right skewed distribution shape) at the initial time (2005), a moderate and stable risk of introducing an infected animal from outside the metapopulation and a low sensitivity of the diagnostic test (0.21).
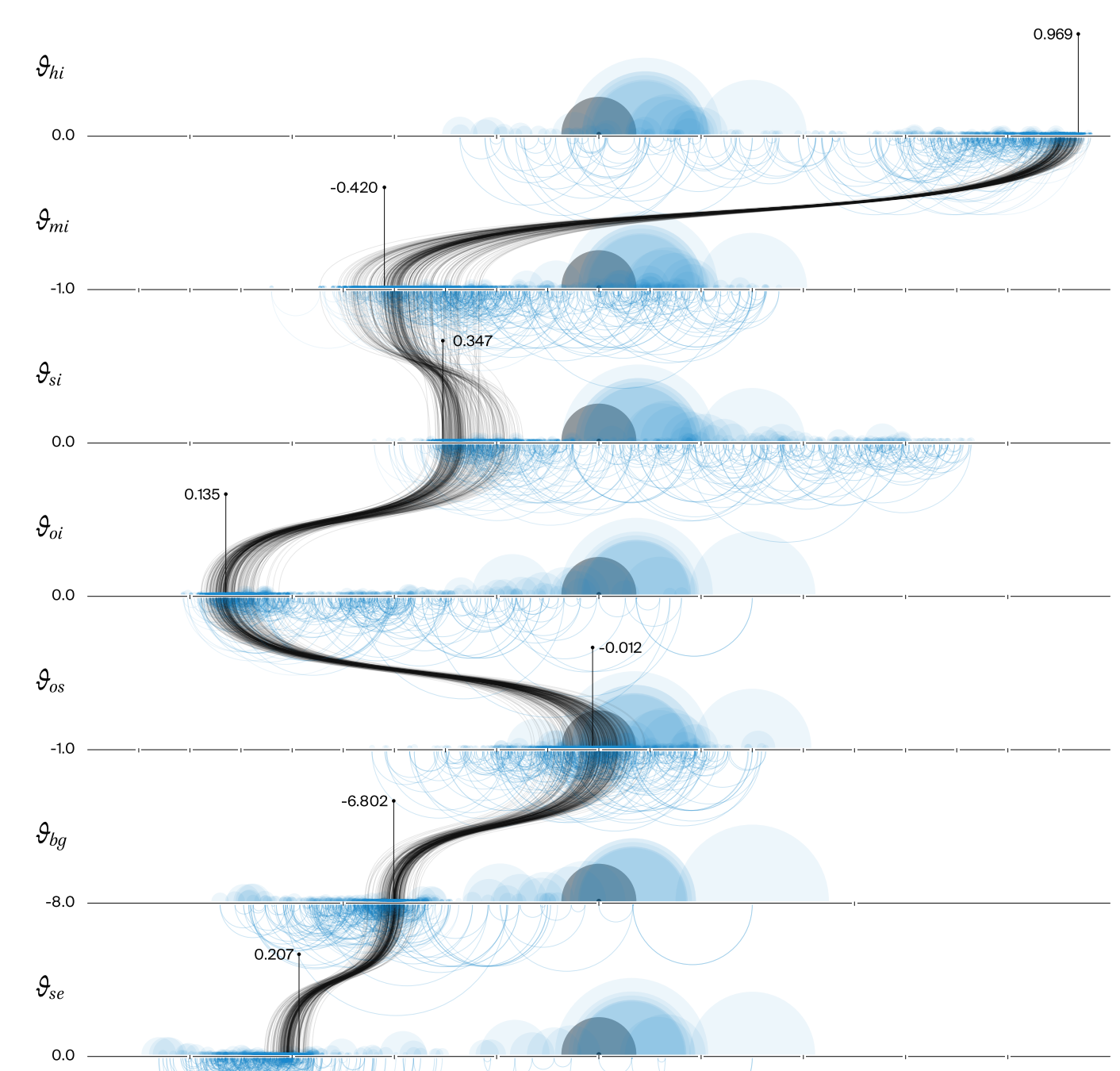

Fig. 4 Visualization of the successive iterations performed during the inference procedure. The half colored discs (above the x-axis) correspond to the quality of the estimation with a size proportional to the distance from the best trajectory (the smaller the radius, the better the value of the criterion). The half discs colored in gray correspond to the initial values used for the estimation. The arcs (below the x-axis) correspond to the jumps in the parameter space, realized during the successive iterations of the algorithm. The set of parameters leading to the best criterion value is indicated by the vertical lines. The gray lines correspond to the sets of parameter within the credibility intervals.

Fig. 5 Credibility intervals. Profile likelihood were drawn using local optimizations by fixing, one at a time, the parameters at different values.

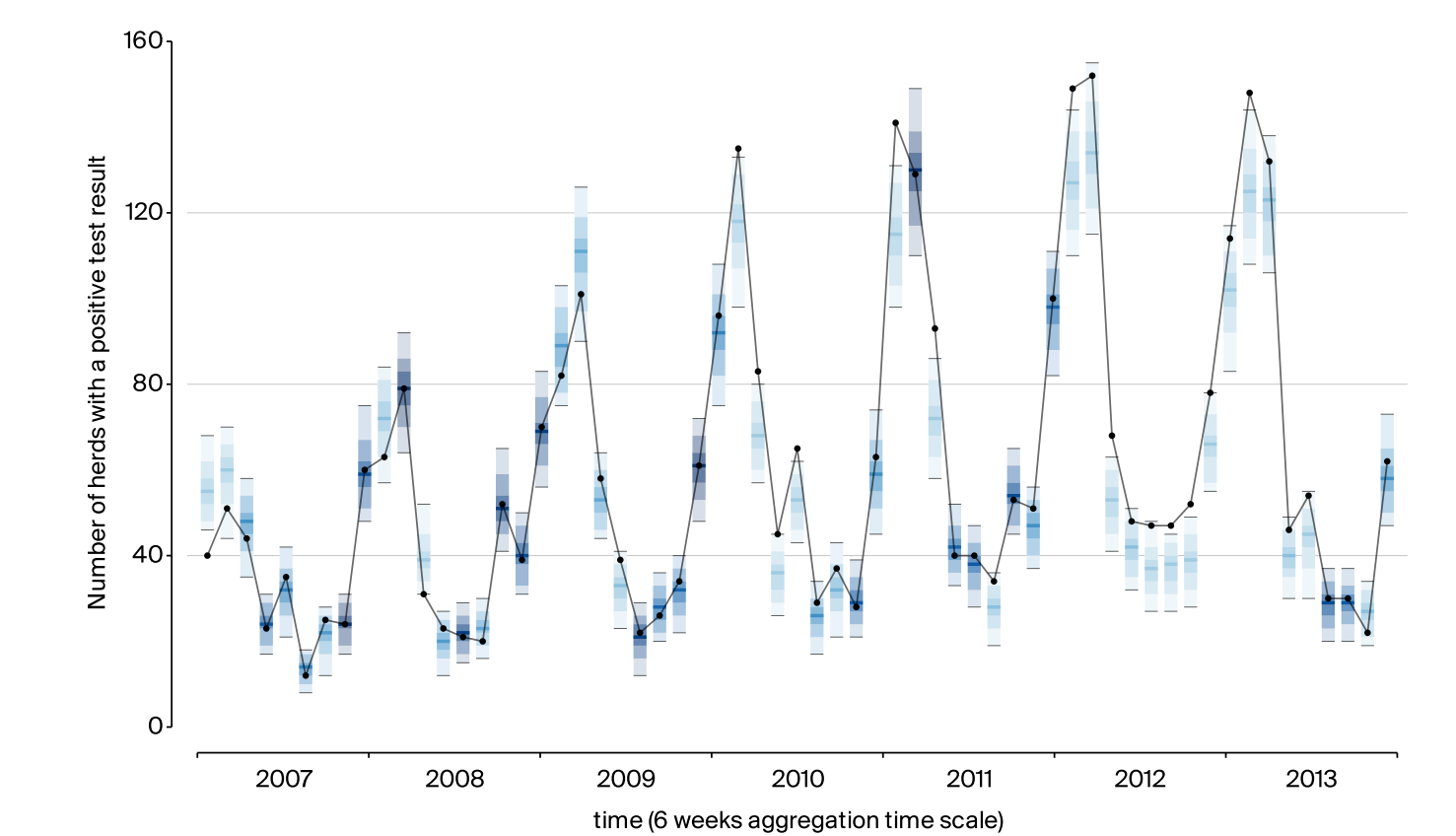## Observation vs. simulation

Fig. 6 Proportion of herds with a positive test result. Each bar is divided into four levels of color intensity, corresponding, from darkest to lightest, to the median, and to the quantiles 25-75, 5-95 and 0-100, computed from a simulation of 200 runs with the estimated parameter values. Black line corresponds to the observed data and blue bars correspond to simulations.
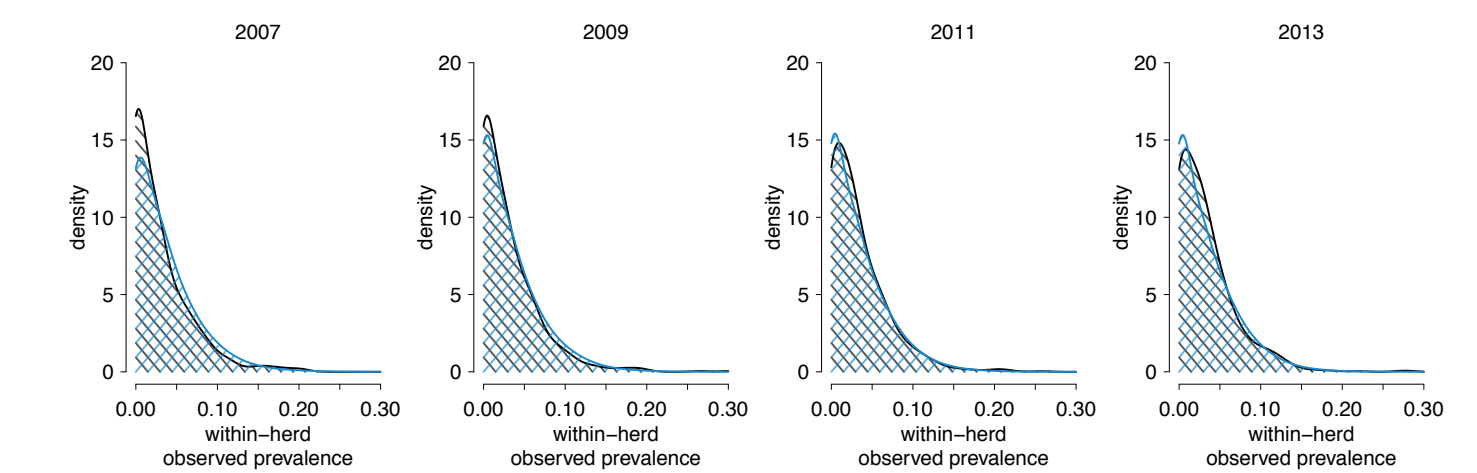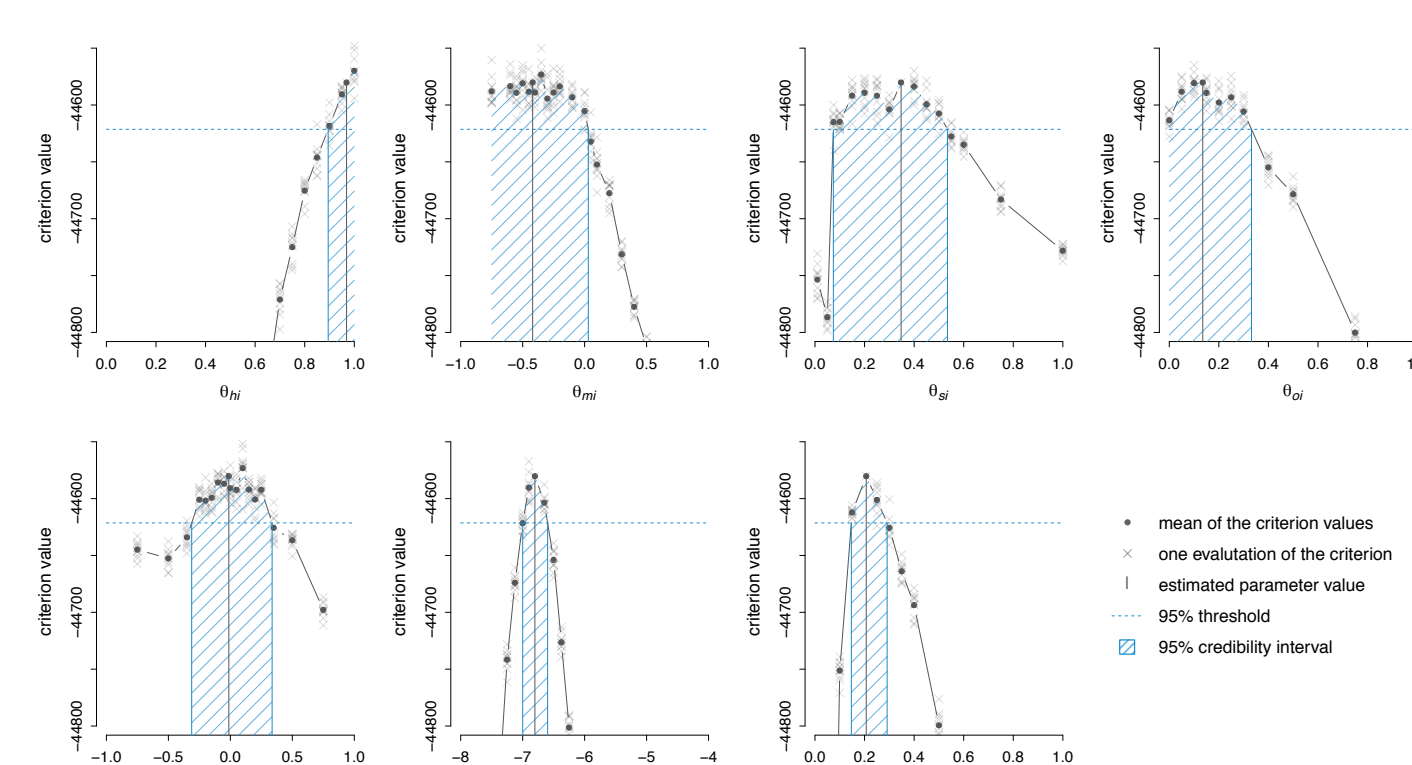

Fig. 7 Within-herd prevalence per year. Distribution in black and blue, correspond, respectively, to the observed data and simulated predictions of the apparent within-herd prevalence among the sampled herds, for a given year.

Conclusion and perspectives

Estimations of previously unknown key parameters provide, for the first time, quantitative insights on *Map* spread at the regional scale, mainly showing a high proportion of infected herds and a low sensitivity of the diagnostic test, in agreement with qualitative opinions of experts.

This inference method can be used to estimate key features of other spatio-temporal infection dynamics, even if imperfectly observed, especially for long-lasting endemic diseases. It is of particular interest when ABC-like inference methods fail due to difficulties in defining relevant summary statistics.

References
Beaunée, G., et al. (2015). Modelling of paratuberculosis spread between dairy cattle farms at a regional scale. Veterinary Research, 46 (111).

We are interested in collaborations to find other applications for this model (adapt it to another region) and/or this inference method.

INRA SCIENCE & IMPACT