



**HAL**  
open science

# Modelling the effect of heterogeneity of shedding on the within herd spread and identification of key parameters by sensitivity analysis

Aurélie Courcoul, Hervé Monod, Mirjam Nielen, Don Klinkenberg, Lenny Hogerwerf, François Beaudeau, Elisabeta Vergu

## ► To cite this version:

Aurélie Courcoul, Hervé Monod, Mirjam Nielen, Don Klinkenberg, Lenny Hogerwerf, et al.. Modelling the effect of heterogeneity of shedding on the within herd spread and identification of key parameters by sensitivity analysis. *Journal of Theoretical Biology*, 2011, 284 (1), pp.130. 10.1016/j.jtbi.2011.06.017 . hal-00720862

**HAL Id: hal-00720862**

**<https://hal.science/hal-00720862>**

Submitted on 26 Jul 2012

**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.

## Author's Accepted Manuscript

Modelling the effect of heterogeneity of shedding on the within herd *Coxiella burnetii* spread and identification of key parameters by sensitivity analysis

Aurélie Courcoul, Hervé Monod, Mirjam Nielen, Don Klinkenberg, Lenny Hogerwerf, François Beaudeau, Elisabeta Vergu

PII: S0022-5193(11)00316-X  
DOI: doi:10.1016/j.jtbi.2011.06.017  
Reference: YJTBI6519

To appear in: *Journal of Theoretical Biology*

Received date: 21 September 2010  
Revised date: 16 April 2011  
Accepted date: 16 June 2011

Cite this article as: Aurélie Courcoul, Hervé Monod, Mirjam Nielen, Don Klinkenberg, Lenny Hogerwerf, François Beaudeau and Elisabeta Vergu, Modelling the effect of heterogeneity of shedding on the within herd *Coxiella burnetii* spread and identification of key parameters by sensitivity analysis, *Journal of Theoretical Biology*, doi:10.1016/j.jtbi.2011.06.017

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



[www.elsevier.com/locate/jtbi](http://www.elsevier.com/locate/jtbi)

**Modelling the effect of heterogeneity of shedding on the within herd *Coxiella burnetii* spread and identification of key parameters by sensitivity analysis**

Aurélie Courcoul<sup>1,2,\*</sup>, Hervé Monod<sup>3</sup>, Mirjam Nielen<sup>4</sup>, Don Klinkenberg<sup>4</sup>, Lenny Hogerwerf<sup>4</sup>, François  
Beaudeau<sup>1,2</sup>, Elisabeta Vergu<sup>3</sup>

<sup>1</sup>INRA, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Atlanpôle La Chantrerie, BP 40706,  
44307 Nantes, France

<sup>2</sup>LUNAM Université, Oniris, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Atlanpôle La  
Chantrerie, BP 40706, 44307 Nantes, France

<sup>3</sup>INRA, UR341 Mathématiques et Informatique Appliquées, Domaine de Vilvert, 78350 Jouy-en-Josas,  
France

<sup>4</sup>Utrecht University, Faculty of Veterinary Medicine, Yalelaan 7, 3584 CL Utrecht, The Netherlands

<sup>5</sup>Université Nantes, Angers, Le Mans, 19Bis rue Lanoué Bras de Fer, 44200 Nantes, France

E-mail addresses: [aurelie.courcoul@oniris-nantes.fr](mailto:aurelie.courcoul@oniris-nantes.fr), [herve.monod@jouy.inra.fr](mailto:herve.monod@jouy.inra.fr), [M.Nielen@uu.nl](mailto:M.Nielen@uu.nl),  
[D.Klinkenberg@uu.nl](mailto:D.Klinkenberg@uu.nl), [L.Hogerwerf@uu.nl](mailto:L.Hogerwerf@uu.nl), [francois.beaudeau@oniris-nantes.fr](mailto:francois.beaudeau@oniris-nantes.fr),  
[elisabeta.vergu@jouy.inra.fr](mailto:elisabeta.vergu@jouy.inra.fr)

\*Corresponding author: [aurelie.courcoul@oniris-nantes.fr](mailto:aurelie.courcoul@oniris-nantes.fr),

Tel: +33 (0) 240 68 76 74,

Fax: +33 (0) 240 68 77 68

**KEYWORDS:** Q fever; stochastic model; uncertainty and variability; PCA; ANOVA

### **ABSTRACT**

*Coxiella burnetii* is the bacterium responsible for Q fever, a worldwide zoonosis. Ruminants, especially cattle, are recognized as the most important source of human infections. Although a great heterogeneity between shedder cows has been described, no previous studies have determined which features such as shedding route and duration or the quantity of bacteria shed have the strongest impact on the environmental contamination and thus on the zoonotic risk. Our objective was to identify key parameters whose variation highly influences *C. burnetii* spread within a dairy cattle herd, especially those related to the heterogeneity of shedding.

To compare the impact of epidemiological parameters on different dynamical aspects of *C. burnetii* infection, we performed a sensitivity analysis on an original stochastic model describing the bacterium spread and representing the individual variability of the shedding duration, routes and intensity as well as herd demography. This sensitivity analysis consisted of a Principal Component Analysis followed by an ANOVA. Our findings show that the most influential parameters are the probability distribution governing the levels of shedding, especially in vaginal mucus and faeces, the characteristics of the bacterium in the environment (i.e. its survival and the fraction of bacteria shed reaching the environment), and some physiological parameters related to the intermittency of shedding (transition probability from a non shedding infected state to a shedding state) or to the transition from one type of

shedder to another one (transition probability from a seronegative shedding state to a seropositive shedding state).

Our study is crucial for the understanding of the dynamics of *C. burnetii* infection and optimization of control measures. Indeed, as control measures should impact the parameters influencing the bacterium spread most, our model can now be used to assess the effectiveness of different control strategies of Q fever within dairy cattle herds.

## 1. Introduction

Q fever is a worldwide zoonosis caused by *Coxiella burnetii*. This intracellular bacterium infects a wide range of animals and is associated with reproductive disorders in domestic ruminants [1-4]. Goats, sheep and cattle are recognized as the main source of human infection [5-8]. Infected animals shed bacteria through various routes (parturition products, faeces, urine, vaginal mucus, milk) [4,9,10]. As the bacterium survives very well in the environment, humans can get infected by inhaling contaminated dusts or aerosols [11-13]. This was recently experienced in the Netherlands where more than 3,500 human cases were reported since 2007 [14]. Although Q fever is asymptomatic in humans in more than 60% of cases, it can lead to acute or chronic infections and cause flu-like syndrome, hepatitis, pneumonia, endocarditis or abortions [15,16]. Hence, for public health and economic and animal health concerns, it is important to control *C. burnetii* infections in livestock herds.

In *C. burnetii* infections, a great heterogeneity between shedders has been described [17-19]: the shedding duration and routes, as well as the level of shedding (i.e. the quantities of bacteria shed) are variable between cows. According to Guatteo *et al.* [20], cows can shed sporadically or persistently, the shedding routes are rarely concomitant and the concentrations of bacteria shed in vaginal mucus or milk

can vary from less than 100 Bacteria/g to more than 1,000,000 B/g. Heterogeneity of shedding is known to affect infection dynamics in many diseases [21] but it is generally difficult to determine which of its aspects are the most influential. The length of shedding, its route, the quantity of bacteria shed, or other features may all have a strong impact on the environment contamination by *C. burnetii* and thus on the zoonotic risk in the case of Q fever infection.

The identification of key parameters characterizing the heterogeneity of shedding and the inclusion of this heterogeneity in representations of disease spread within a herd are thus critical for the understanding of the infection dynamics. In addition, for some diseases, the effectiveness of control measures was shown to be dramatically improved by targeting the individuals transmitting the pathogen most (e.g. in the case of *Escherichia coli* O157 infection [22], of measles epidemics [23] or of *Salmonella* transmission [24]). However, understanding and predicting the spread of *C. burnetii* in a herd or identifying such key parameters cannot be assessed by field experiments alone. In this context, mathematical models and statistical methods are useful tools for understanding how the infection spreads within the herd and how various inputs (such as epidemiological characteristics of infected animals) affect the dynamics [25]. Techniques such as sensitivity analysis allow assessing the impact of the uncertainty and variability in the parameters on models outputs and hence determining key factors [26]. It consists in studying how the variation in the outputs of the model can be apportioned to different sources of variation, and how the model depends upon the information fed into it.

The aim of our study is to determine, using a sensitivity analysis, the key parameters related to the heterogeneity of shedding whose variation highly influences *C. burnetii* infection dynamics within a dairy cattle herd. First, the characteristics of the original model we developed for representing the spread of this infection will be summarized. Then, the sensitivity analysis performed, accounting for factors

governed by probability distribution, will be detailed. This will be followed by the presentation and the discussion of the results.

## 2. Model

The original epidemic model that we developed is presented in details in the Supplementary material.

### 2.1. General description

In summary, the model describes the spread of *C. burnetii* within a dairy cattle herd and takes into account the heterogeneity of shedding: both the shedding routes and shedding levels (quantities of bacteria shed) are represented. Herd demography is also accounted for. This model is, to our knowledge, the first one proposed in the literature for Q fever spread coupling epidemiological aspects with herd population dynamics. Six different health statuses are defined function of the shedding status (i.e. excretion of bacteria or not), the shedding route if the animal is a shedder and the immunity status (presence or absence of antibodies). See Fig. 1 and Table 1 for parameter description. Each cow is then in one of the six mutually exclusive health states at a given time (Fig. 1):  $S$  (susceptible, non-shedder without antibodies),  $I^-$  (shedder without antibodies),  $I^+$  (shedder with antibodies),  $I^{+ \text{ milk pers}}$  (shedder with antibodies, shedding in milk at higher levels and for a longer period of time than  $I^+$ , as described by Guatteo *et al.* [20]),  $C^+$  (animal with antibodies, non-shedder but still infected),  $C^-$  (non-shedder without antibodies which was infected and had antibodies in the past). Besides, all shedding cows  $I^-$ ,  $I^+$  and  $I^{+ \text{ milk pers}}$  are subdivided according to their shedding routes: (1)  $I_1$ , milk only, (2)  $I_2$ , vaginal mucus and/or faeces, (3)  $I_3$ , both. Susceptible animals can become contaminated through bacteria shed in the environment by infected animals. Once infected, animals progress through the different health states according to specific probability distributions (Table 1). The environment has its own dynamics (as described in the Supplementary material). The necessity to consider herd demography is generated by

the interaction between lactation and gestation statuses and shedding routes and levels: cows which recently calved (during the 4 weeks post-calving) have different probability distributions governing shedding routes and levels than other cows. See Table 2 for description of herd demography related parameters.

The model is stochastic, individual-based and in discrete time with a time step of one week, which is appropriate for both epidemiological and herd management processes. The stochasticity has two main sources: for each individual, all the transitions between health states are supposed stochastic and the quantities of bacteria shed in the environment follow discrete distributions, different according to the shedding route.

## 2.2. Initial conditions and parameter values of the standard scenario

At  $t=0$ , the herd consists of 50 cows. To initiate the infection cycle, a primiparous  $I^+$  cow which has just calved is introduced into a wholly susceptible herd.

The epidemiological parameters are put at their standard values (Table 1): parameters  $m$  (transition probability  $I^- \Rightarrow S$ ),  $q$  (transition probability  $I^- \Rightarrow I^+$ ),  $r_1$  (transition probability  $I^+ \Rightarrow C^+$ ),  $s$  (transition probability  $C^+ \Rightarrow I^+$ ) and  $\mu$  (elimination rate of *C. burnetii*) come from a study where they were estimated through Bayesian inference using data from five French chronically infected dairy cattle herds [27];  $pp$  (proportion of cows going from  $I^-$  to  $I^+$  and becoming  $I^{+milk\ pers}$ ) and  $probav$  (probability of abortion after a transition  $S \Rightarrow I^-$ ,  $C^+ \Rightarrow I^+$  and  $C^- \Rightarrow I^+$ ) were qualitatively calibrated to match respectively the average number of  $I^{+milk\ pers}$  cows and the distribution of abortions observed in the field (R. Guatteo and A.F. Taurel, 2010, personal communication); for  $r_2$  (transition probability  $I^{+milk\ pers} \Rightarrow C^+$ ), it was assumed that the shedding duration of  $I^{+milk\ pers}$  cows (i.e.  $1/r_2$ ) was on average 10 times longer

than the shedding duration of  $I^+$  cows (i.e.  $1/r_1$ );  $\rho^{mf}$  (proportion of bacteria shed through mucus/faeces arriving into the environment compartment) and  $ratio \left( \frac{\rho^{milk}}{\rho^{mf}} \right)$  were calibrated from expert opinion to match the environmental bacterial load inferred in Courcoul et al. [27]; for  $\tau$  (transition probability  $C^+ \Rightarrow C^-$ ), it was assumed, based on Fournier et al. [28] and Plommet et al. [29], that the mean life duration of antibodies in cattle was 2 years; probability distributions of shedding related parameters,  $\alpha$ ,  $\beta$ ,  $\beta_{calv}$ ,  $\gamma$ ,  $\gamma_{calv}$  (all five representing probability distributions of the shedding routes for the different types of shedders) and  $Q1$ ,  $Q2$ ,  $Q3$ ,  $Q4$  and  $Q5$  (all five representing probability distributions of the shedding levels for the different types of shedders) were qualitatively calibrated from field data (R. Guatteo 2009, personal communication).; The parameters governing the demography and herd management (Table 2) were chosen to represent a standard French dairy cattle herd.

To account for the variability in Q fever infections, 200 repetitions of the standard scenario were run over a 5-year simulation period.

### 3. Sensitivity analysis

We conducted a sensitivity analysis to identify the parameters that mostly contributed to the output variability. This analysis, similar to that performed in Lurette *et al.* [30] following the methodology developed in Lamboni *et al.* [31] involves an original aspect (not present in [30]): some factors are governed by probability distributions (such as heterogeneity related parameters) and are not point values, as usual in this kind of methods. Various scenarios were run, each of them being characterized by a specific combination of parameter values, in order to relate the variability obtained for the outputs to that induced by the input parameters.

#### 3.1. Outputs and factors

Eight outputs were considered (Table 3): (i) the environmental bacterial load of the main buildings and close pastures for lactating cows  $E_{building}$ , (ii) the environmental bacterial load of the specific pastures for dry cows  $E_{dry}$ , (iii) the prevalence of milk shedders, (iv) the prevalence of mucus/faeces shedders, (v) the prevalence of shedders in milk in a persistent way, (vi) the seroprevalence, (vii) the number of abortions per herd per year, and (viii) the extinction rate. All these outputs except the number of abortions and the extinction rate were computed weekly over a 5-year period.

Parameters related to the herd demography were fixed at their nominal values of Table 2 since demography and herd management processes are considered as well known. The sensitivity of the model outputs was evaluated with respect to the epidemiological parameters, which are those given in Table 1, except for  $\tau$ . These 19 parameters presented in Table 1 are thus the inputs of the sensitivity analysis and they will be called factors in the rest of the paper. They belong to two categories: parameters concerning the transitions between health states and the environment ( $m, q, p/p, r_1, r_2, s, \mu$ ) and parameters directly related to the heterogeneity in shedding (i.e.  $\alpha, \beta, \beta_{calv}, \gamma, \gamma_{calv}, Q1, Q2, Q3, Q4, Q5, \rho^{mf}$  and  $ratio \left( \frac{\rho^{milk}}{\rho^{mf}} \right)$ ). The parameters in the first category were estimated from field data previously [27], but some uncertainty still remains (due for instance to the limitation of the data). They were included in the sensitivity analysis but with relatively limited ranges of variation. In this study we focused on the latter category of parameters because they directly describe heterogeneity related aspects, which represented our main objective.

### 3.2. Design of experiments

All the designs were generated using R 2.10.1 [32] and PLANOR R package [33].

#### 3.2.1. First experiment

We used a fractional factorial experiment design, with four values (called levels) per factor related to the shedding and two levels for the other factors (values in Table 1). As our model is stochastic, we ran the model for each combination of factor levels 30 times. Since the complete factorial design would lead to too many combinations (exactly  $30 \times 4^{12} \times 2^7$  simulations), a fractional factorial design of resolution V was chosen. Such a design allows to estimate the main effects and two-factor interactions, provided higher order interactions are assumed to be negligible [34,35]. In the present case, a design was obtained with 4,096 scenarios. Thus, we ran 122,880 realizations of the model (i.e. 30 repetitions for each of the 4,096 scenarios).

### 3.2.2. Second experiment

A complete factorial design for the eight most influential factors according to the first experiment was performed. This enabled us to more accurately quantify the impact of the interactions between these eight factors and also to disentangle potential confounded main effects and interactions. Besides, we determined in this experiment the factors that mostly contributed to the variability of the extinction rate between repetitions. The remaining 11 parameters were put to their standard value (Table 1). For this second study, we ran 2,048 scenarios with 30 repetitions each.

### 3.2.3. Third experiment

In a third analysis, the influence of the probability distributions of the shedding levels (factors  $Q$ ) was specifically explored. Since these probability distributions depend on the type of shedder ( $I^-$ ,  $I^+$  or  $I^{+ \text{ milk pers}}$ ) and route of shedding (milk versus mucus/faeces), this analysis enabled us to explore which type of shedders and which type of shedding route played a major role in the variability of the outputs. Ten new factors (called  $Q^*$ ) were defined: each factor  $Q^*$  corresponds to a given shedding level probability distribution for a type of shedder and a shedding route. These 10 factors  $Q^*$  replaced the former 5

factors  $Q$  used in first and second experiments (where a same shedding level probability distribution  $Q$  could be used for different types of shedders or different shedding routes). Factors  $Q^*$  were numbered as follows: for milk and mucus/faeces respectively,  $Q1^*$  and  $Q2^*$  refer to the probability distributions of the shedding levels for the  $I^-$ ,  $Q3^*$  and  $Q4^*$  refer to those for the  $I^+$  after 4 weeks post-calving,  $Q5^*$  and  $Q6^*$  refer to those for the  $I^+$  during the 4 first weeks post-calving,  $Q7^*$  and  $Q8^*$  refer to those for the  $I^{+ \text{ milk pers}}$  after 4 weeks post-calving, and finally  $Q9^*$  and  $Q10^*$  refer to those for the  $I^{+ \text{ milk pers}}$  during the 4 first weeks post-calving. Thus, former factor  $Q1$  of first and second experiments corresponds to new factors  $Q1^*$ ,  $Q2^*$  and  $Q4^*$ , former factor  $Q2$  to  $Q3^*$ , former factor  $Q3$  to  $Q5^*$  and  $Q6^*$ , former factor  $Q4$  to  $Q8^*$  and former factor  $Q5$  to  $Q7^*$ ,  $Q9^*$  and  $Q10^*$ .

A fractional factorial design for the 10 new factors  $Q^*$  with four levels each was generated. These four levels were (0.85, 0.15, 0), (0.6, 0.4, 0), (0.25, 0.25, 0.5) and (0.15, 0.6, 0.25) for the probability to be in (low, mid, high) shedding level respectively. The other parameters were put to their standard values given in Table 1. For this third experiment, we ran 1,024 scenarios with 30 repetitions each.

### 3.3. Analysis of the temporal outputs (of the first, second and third experiments)

In order to compare the influence of factors on the seven outputs which exhibit temporal dynamics (all outputs except the extinction rate), we applied a method developed by Lamboni *et al.* [31]. The results are recorded as tables with one row for each scenario and one column for each output time point (260 weekly time points for outputs (i) to (vi) and five annual time points for the abortion number, output (vii)).

This method allows to simultaneously analyze potentially correlated variables (here the successive time points of a given output). It consists in two main steps. First, a Principal Component Analysis (PCA) is operated in order to provide linear combinations (or components) of the initial variables (here the

columns of our tables) explaining the maximum of inertia (i.e. variability) between scenarios. Only the first three principal components (PC) were kept since they are sufficient to cover most variability amongst simulations. The PCA provides to each line of the tables a score on each component. The second step involves an ANOVA, including the main effects and the two-factor interactions for all factors and carried out on the scores of each of the components considered. Sensitivity indices (SI), corresponding to the main effect or to interactions, and total sensitivities (TS), corresponding to the sum of the main effect and the interactions, were calculated for each factor and for each component. This analysis was performed with R 2.10.1 [32] and multisensi R package [36].

The analyses were performed on both the mean and standard deviation of the 30 repetitions of each scenario, in order to assess the two sources of variability influencing the outputs: the model intrinsic stochasticity and the parameter variability generated by the factorial designs.

#### **3.4. Analysis of the extinction rate (of the second experiment)**

An ANOVA was performed to assess the influence of the eight most influential factors on the extinction rate. It was calculated for each scenario defined by the complete factorial design of the second experiment.

#### **3.5. Analysis of the outputs at the time point 260 (of the first, second and third experiments)**

In order to determine the factors with the highest influence on the output variability as a whole, we performed a joint analysis on the values of the six dynamic outputs (first six lines of Table 3) at the last simulation time step (week 260). This time point was chosen to illustrate the long-term steady-state of the system. Thus, the two-step analysis (PCA followed by ANOVA) was performed twice on six output variables. The first analysis was done on the mean and the second on the standard deviation of the 6 dynamic outputs at time 260.

## 4. Results

### 4.1. Infection dynamics of the standard scenario

Over the 200 repetitions of the standard scenario, 34 led to the extinction of infection (defined as the absence of shedders (animals in  $I$  states) or chronically infected animals (in  $C^+$  state) in the herd) occurring in week 60 (median; see Figure S2 in the Supplementary material) after the introduction of the initial infected cow (min: week 2, max: week 255). The mean seroprevalence and the mean prevalences of shedders increased with time (Fig 2) to reach respectively 34.4% [0 – 57% for the percentiles 2.5% and 97.5% respectively] and 36.3% [0 – 60% for the percentiles 2.5% and 97.5% respectively] five years after the initial infection. The ratio between the mean prevalence of shedders and the mean prevalence of milk shedders was around 2.6 in the first weeks of simulation and then it decreased to reach 1.9 at the end of the simulation time. The mean environmental bacterial load  $E_{building}$  increased with time corresponding to a mean transition probability from  $S$  to  $I^-$ ,  $p_{building}$ , equal to 0.45 at the end of the simulation time. On the contrary, the mean environmental bacterial load  $E_{dry}$  was close to 0 for the 5 years of simulation (results not shown). The median number of abortions varies from 2 per herd in the first year to 4 per herd in the 5<sup>th</sup> year, and a large variability surrounded these values [0-10 for the percentiles 2.5% and 97.5% respectively]. In addition, as shown in Figure 3, the route and the level of shedding of a shedder cow had a great impact on the contamination of the environment. This result is an unsurprising consequence of the model parameterization. As expected, the ‘low level’ shedding category, although the most common, did not contribute much to the increase of the environmental bacterial load. On the contrary, shedders in mucus/faeces of the ‘mid level’ category and shedders in milk of the ‘high level’ category filled the environment in a non negligible way. Above all, shedders in mucus/faeces of the ‘high level’ category (both non aborting and aborting cows) had the greatest impact.

## 4.2. Influence of the epidemiological factors on the model outputs

The results obtained with 30 runs for each parameter set were robust: the mean and the percentiles 2.5, 50 and 97.5 of our outputs were similar to those obtained with 200 runs (results not shown). The outputs variability due to the model stochasticity (within scenario variability) is illustrated Figure S5 of the Supplementary material.

The variability between scenarios was important. As an example, for the first experiment, the percentiles 2.5%, 50% and 97.5% of the seroprevalence (mean over the 30 repetitions of a given scenario) at the end of the 5-year simulation period were of 0.0%, 45.4% and 94.5% respectively. Those of the environmental bacterial load (mean over the 30 repetitions of a given scenario) were of 0.0, 2.1 and 51.0 respectively (Figure S4 of the Supplementary material).

### 4.2.1. First experiment

As shown in Table 4, since the inertia obtained for the first principal component (PC) was very high for each of the model outputs (except the joined analysis at time 260), only the results on the first PC are presented.

For the two mean environmental bacterial loads, the factors  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for the  $I^+$  shedding in mucus/faeces after 4 weeks post-calving),  $\mu$  (elimination rate of *C. burnetii*) and  $\rho^{mf}$  (proportion of bacteria shed through mucus/faeces arriving into the environment) were the most influential ones.

For the mean prevalences of mucus/faeces and milk shedders, the most sensitive factors were  $q$  (transition probability from  $I^-$  to  $I^+$ ),  $s$  (transition probability from  $C^+$  to  $I^+$  representing the intermittency of shedding) and  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for

some of the  $I^+$  in mucus/faeces), whereas the mean prevalence of milk shedders in a persistent way was mostly impacted by  $p/p$  (the proportion of cows going from  $I^-$  to  $I^+$  and becoming  $I^{+milk\ pers}$ ),  $q$  (transition probability from  $I^-$  to  $I^+$ ) and  $r_2$  (transition probability from  $I^{+milk\ pers}$  to  $C^+$ ). Concerning the mean seroprevalence, the factor  $q$  (transition probability from  $I^-$  to  $I^+$ ) had a TS higher than 60%.

Lastly, the most influential factors of the mean abortion number were  $q$  (transition probability from  $I^-$  to  $I^+$ ),  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for some of the  $I^+$  in mucus/faeces),  $s$  (transition probability from  $C^+$  to  $I^+$ ),  $\mu$  (elimination rate of *C. burnetii*) and  $\rho^{mf}$  (proportion of bacteria shed through mucus/faeces arriving into the environment).

Globally, the most sensitive two-factor interactions (with a SI higher than 5%) were  $Q1:q$  on the variability of the number of abortions,  $Q1:\mu$ ,  $\rho^{mf}:Q1$ ,  $\rho^{mf}:\mu$  on the variability of the environmental bacterial loads and  $q:p/p$ ,  $q:r_2$ ,  $p/p:r_2$  on the variability of the prevalence of milk shedders in a persistent way.

Concerning the variability of the standard deviations of the outputs, the same factors as above were identified as the most influential ones for the environmental bacterial loads, the prevalence of milk shedders in a persistent way and the abortion number. The main effect of the factors was always very low (no SI higher than 5%) on the prevalences of mucus/faeces shedders, whereas the part of two-factor interactions was much more important. The most sensitive factors were  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for some of the  $I^+$  in mucus/faeces),  $\rho^{mf}$  (proportion of bacteria shed through mucus/faeces arriving into the environment),  $q$  (transition probability from  $I^-$  to  $I^+$ ) and  $Q_3$  (probability distribution of the shedding levels for all the  $I^+$  during the 4 first weeks post-calving).

Regarding the standard deviation of the prevalence of milk shedders, the most sensitive factors were  $q$  (transition probability from  $I^-$  to  $I^+$ ),  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for some of the  $I^+$  in mucus/faeces) and  $\alpha$  (probability distribution of the shedding routes for the  $I^-$  cows).

For the joined analysis on six of the dynamic outputs at time 260, the inertia obtained for the first PC was much lower and the second PC had to be taken into account. For the means, the most influential factors were  $Q1$  (probability distribution of the shedding levels for all the  $I^-$  and for some of the  $I^+$  in mucus/faeces),  $q$  (transition probability from  $I^-$  to  $I^+$ ),  $s$  (transition probability from  $C^+$  to  $I^+$ ),  $\mu$  (elimination rate of *C. burnetii*) and  $\rho^{mf}$  (proportion of bacteria shed through mucus/faeces arriving into the environment) on the first PC and  $Q1$ ,  $\mu$ ,  $\rho^{mf}$  and  $q$  on the second one, by order of importance. For the standard deviation, the most sensitive factors were  $q$ ,  $Q1$ ,  $Q3$  (probability distribution of the shedding levels for all the  $I^+$  during the 4 first weeks post-calving),  $r_2$  (transition probability from  $I^{+ \text{ milk pers}}$  to  $C^+$ ) and  $p/p$  (the proportion of cows going from  $I^-$  to  $I^+$  and becoming  $I^{+ \text{ milk pers}}$ ) on the first PC and  $Q1$ ,  $\rho^{mf}$  and  $\mu$  on the second.

#### 4.2.2. Second experiment

The eight factors chosen in the second experiment (appearing as the most influential according to the findings of the first experiment) were two of the five shedding level probability distributions,  $Q1$  and  $Q3$ , three transition probabilities  $q$ ,  $s$ , and  $r_2$ , the elimination rate of the bacterium in the environment  $\mu$ , the proportion of bacteria shed through mucus/faeces arriving into the environment,  $\rho^{mf}$  and the proportion of cows going from  $I^-$  to  $I^+$  and becoming  $I^{+ \text{ milk pers}}$ ,  $p/p$ . The results obtained were globally similar to the results of the first experiment described above (same most influential factors, sometimes

in a slightly different order), suggesting that no important interactions were confounded with the main effects in the first analysis.

The most sensitive factors on the extinction rate (frequency distribution amongst scenarios in Figure S3 of Supplementary material) were firstly  $Q1$ ,  $\mu$  and  $\rho^{mf}$  (with a SI higher than 10%) then  $q$ ,  $s$ , and  $Q3$  (with a lesser SI, but higher than 5%). The most sensitive two-factor interactions were  $Q1:q$  and  $Q1:\mu$  (SI higher than 5%).

#### 4.2.3. Third experiment

Amongst the ten probability distributions of the shedding levels  $Q^*$ , the most influential factor on the variability of the outputs was  $Q2^*$ , the probability distribution of the shedding levels for the  $I^-$  shedding in mucus/faeces (Table 5 and Figure 5).  $Q4^*$  and  $Q6^*$ , the probability distributions of the shedding levels for the  $I^+$  shedding in mucus/faeces before and after the 4<sup>th</sup> week post-calving, respectively, also had a significant impact and, to a lesser extent,  $Q1^*$ , the probability distribution of the shedding levels for the  $I^-$  shedding in milk, and  $Q8^*$ , the one for the  $I^{+ \text{ milk pers}}$  shedding in mucus/faeces. Moreover, the only interaction among the five most sensitive terms was  $Q2^*:Q4^*$ . Overall, the factors with the greatest impact were probability distributions of the shedding levels in mucus/faeces.

## 5. Discussion

The main objective of this work was the identification, using a sensitivity analysis, of the key parameters related to the heterogeneity of shedding highly impacting *C. burnetii* infection dynamics within a dairy cattle herd. The underlying step of the achievement of this goal was the elaboration of a model describing the bacterium spread.

We then proposed, in the first part of our study, the first model of *C. burnetii* spread within a dairy cattle herd taking into account the individual variability of shedding, defined through duration, routes and intensity. Simulated infection dynamics are consistent with field data: at the last time point of the simulated time series (five years after the introduction of the initial infectious case), the median seroprevalence is around 31% [11.0% – 41.2% for the 25<sup>th</sup> and 75<sup>th</sup> percentiles respectively], which is globally consistent with the mean observed seroprevalence in cows (mean: 40%, 25<sup>th</sup> and 75<sup>th</sup> percentiles: 25% and 51% respectively) of 56 naturally infected French dairy herds [37]. At the same simulated time point, the median prevalence of shedders is around 32.3% [0 – 56.9% for the percentiles 2.5<sup>th</sup> and 97.5<sup>th</sup> respectively] whereas in the field, the apparent proportion of shedder cows is 45.5% in Guatteo *et al.* [9] and 38.9% in Rodolakis *et al.* [19].

The second and main part of our study consisted in the sensitivity analysis. This approach aims at improving the understanding of complex systems such as stochastic epidemiologic models and at suggesting possible targeted control strategies in livestock infections [38-40] or assessing the effect of varying the input parameters on the economic impact associated with the disease [41,42].

To perform sensitivity analyses, we used complete and fractional factorial designs. Alternative approaches are available (see for example [43] or [44]), but factorial designs are very convenient to control which main effects and interactions can be estimated and to manage a mixture of qualitative and quantitative factors. In the present study, some factors (the probability distributions of the shedding routes and levels for the different types of shedders  $\alpha$ ,  $\beta$ ,  $\beta_{calv}$ ,  $\gamma$ ,  $\gamma_{calv}$ ,  $Q1$ ,  $Q2$ ,  $Q3$ ,  $Q4$  and  $Q5$ ) were not scalars but probability distributions. Performing a sensitivity analysis with such types of factors is to our knowledge fairly new. We used multinomial distributions with three classes and defined the modalities of each such factor as four alternative sets of multinomial probabilities. This choice allowed flexibility as well as a fine control in the probability distributions that were simulated.

To cope with the dynamic and multivariate outputs of the model, the PCA+ANOVA approach [31] offered additional insight compared to single sensitivity analyses. Multivariate decomposition methods other than PCA could be used, but the key idea is that sensitivity analyses are now performed on synthetic and meaningful output variables. As the infection dynamics is composed of two phases (see on Figure 2 an initial phase of rapid progression followed by a kind of steady state), the parameters influencing the dynamics may not be the same between the phases. We additionally performed a preliminary sensitivity analysis on the very initial phase of infection (first 26 weeks; results not shown). The most sensitive factors during the first 15 weeks were highly influenced by the initial conditions but very fast afterwards, the same factors as in experiments one or two were identified as the most influential ones. We then chose to conduct our study on the whole simulation period.

Another aspect concerns the stochasticity of our model which generates complex dynamics. Hence, attention has to be paid to this point in order to appropriately apply sensitivity analysis approaches that are mostly developed for deterministic models. Here, to be able to study with confidence how the variation in the outputs of our model can be apportioned to the uncertainty of epidemiological factors, we checked that the variability due to the model stochasticity (i.e. the within scenario variability) was negligible compared to the variability due to the input parameters variability (i.e. the between scenarios variability). More specifically, to provide reliable analysis on trends, means were calculated on 30 repetitions. Standard deviations were also considered as they can provide complementary information on the most influential factors.

In summary, the method used has the major advantage of allowing to consider temporal varying outputs and thus to identify parameters influencing the dynamics over the 5 year simulation period. Moreover, it allows properly dealing with the model stochasticity.

According to our findings (first and second experiments), we can classify the eight parameters that have most influence on the *C. burnetii* infection dynamics in three categories, depending on the dynamics aspects they are involved in. These parameters should be precisely assessed in further studies to improve the model accuracy and the understanding of the infection spread. Control strategies impacting them would also strongly influence the infection dynamics.

The first group of key parameters, comprising the parameters related to the transition between health states, slightly influences the different prevalences and the abortion number. The transition probability from  $I^-$  to  $I^+$  ( $q$ ) is a physiological parameter associated to seroconversion. It was estimated based on data from a follow-up of five chronically infected herds [27]. However, we can assume that those five herds do not cover the whole potential range of variability of this factor, especially at the beginning of the infection when this parameter probably takes a different value depending on how recently the infection occurred in the herd. Further experimental or survey studies focusing on the start of the infection dynamics are needed to improve the knowledge of this parameter, and therefore the accuracy of the model. However, it is unlikely that a control measure could impact this seroconversion parameter. The transition probability from  $C^+$  to  $I^+$  ( $s$ ) represents the intermittency of shedding. It was inferred from data in the same study [27] and the estimated values were herd-dependent. It is biologically plausible to assume that a control measure such as vaccination could decrease this parameter and then have an impact on the prevalence of shedders. However, since, to our knowledge, no data on the impact of vaccination on the intermittency of shedding is currently available, further studies are needed. The transition probability from  $I^{+ \text{ milk pers}}$  to  $C^+$  ( $r_2$ ), and the proportion of cows going from  $I^-$  to  $I^+$  and becoming  $I^{+ \text{ milk pers}}$  ( $p/p$ ) have an impact on the variability in the prevalence of persistent milk shedders ( $I^{+ \text{ milk pers}}$ ). These parameters were not estimated from data, but calibrated so that simulated trajectories of prevalence of persistent milk shedders are consistent with field observations. Indeed, following Guatteo

*et al.* [20], almost one milk shedder out of three was detected as persistent shedder over three months. In our case, the mean prevalence of milk shedders at time point 260 weeks is 3.3 times the mean prevalence of persistent milk shedders. While the proportion of cows becoming  $I^{+ \text{milk pers}}$  seems difficult to be decreased, the transition probability from  $I^{+ \text{milk pers}}$  to  $C^+$  ( $r_2$ ), which rules the shedding duration of persistent milk shedders, could probably be modulated by control strategies. Although according to Astobiza *et al.* [45], an oxytetracycline treatment would not limit the duration of bacterial excretion in a dairy sheep flock, further studies are needed in order to determine if vaccination could decrease the duration of the shedding period.

The second group of parameters that influenced the infection dynamics most is related to the characteristics of *C. burnetii* in the environment. In fact, the proportion of bacteria shed through mucus/faeces and arriving into the environment ( $\rho^{mf}$ ) and the elimination rate of *C. burnetii* in the environment of the farm ( $\mu$ ) have a strong impact on the environmental bacterial load and to a lesser extent on the abortion number. It is very difficult to quantify in practice which proportion of bacteria contained in milk, vaginal mucus or faeces contaminates the environment. Thus, we calibrated the proportion of bacteria shed reaching the environment to match the environmental bacterial load estimated by Courcoul *et al.* [27]: the posterior medians for the environmental bacterial loads of each of the five chronically infected herds were comprised between 0.044 [0.005–0.143 for the 95% credible interval (CI)] and 0.558 units of environment [0.201–1.278 for the 95% CI]. Since those herds did not exhibit any clinical signs, we assumed that the simulated environmental bacterial load in herds with abortions should be slightly higher. The median environmental bacterial load at time step 260 weeks is then 0.57 (with the percentiles 2.5<sup>th</sup> and 97.5<sup>th</sup> equal to 0.00 and 1.49 respectively). Concerning the elimination rate of the bacterium in the environment, given that *C. burnetii* withstands harsh environmental conditions [13], its life expectancy ( $1/\mu$ ) within the farm in an infectious form was

assumed to be five weeks in the standard scenario and two or 13 weeks (two extreme values) in the sensitivity analysis. However, this assumption is difficult to verify. Based on Dutch studies, it seems that within a month, more than 75% of the manure does not contain viable *C. burnetii* anymore but that the bacterium survives only a few days in the outer layer of the manure [46]. It is then difficult to calibrate the elimination rate of *C. burnetii* in the environment which represents both the natural mortality of the bacterium and its removal due to the periodic cleaning of the cattle housing carried out by the farmer. However, it seems possible to influence it by implementing environmental control measures such as increased cleaning of the farm.

Lastly, as suggested by the results of the first experiment (Table 4) and detailed in the third experiment (Table 5), the parameters which have the greatest impact on the infection dynamics are those governing the shedding levels (through their associated probability distributions  $Q^*$ ), especially in mucus/faeces. The specific effects of mean, variance or form of these distributions have not been separately evaluated when adjusting for constant overall transmissibility; the influence of these distributions can therefore be considered as an effect of the mean quantities of bacteria shed. As shedding in mucus/faeces contaminates the environment much more than shedding in milk, it could seem surprising that the parameters governing the probability distribution of the shedding routes ( $\alpha$ ,  $\beta$  and  $\gamma$ ) do not appear to influence the model outputs. This could partly be explained by the parameter values used, especially for the quantity of bacteria released by shedders in low, mid and high levels respectively ( $Qty$ ). The quantity of bacteria released by a high level shedding cow is so high (compared to mid or low levels), that the probability distributions governing the levels (such as  $Q$  in experiments one and two) are by construction more important than those related to the shedding routes. However, although exact values of the quantities of bacteria shed by shedders in low, mid and high levels are unknown, the standard values used in this study were calibrated according to field data, which tend to support our findings.

This third experiment was of high interest as it allows highlighting the importance of a sub-category of animals. Indeed, the distribution of the levels of bacteria shed in mucus/faeces by seronegative shedders (factor  $Q2^*$ ) has a strong impact on all the outputs, including the shedder prevalences and the environmental bacterial loads. The distributions of the levels of bacteria shed in mucus/faeces by seropositive shedders (factors  $Q4^*$  and  $Q6^*$ ) also have a significant impact on the variability of the shedder prevalences. The predominance of  $Q2^*$  over  $Q4^*$  and  $Q6^*$  can be explained by the larger simulated number of seronegative shedders than seropositive ones. In fact, due to the low standard value of the transition probability from  $I^-$  to  $I^+$  ( $q$ ) and the high standard value of the transition probability from  $I^-$  to  $S$  ( $m$ ), the number of seropositive cows is very low during the three first years of simulation compared to the number of seronegative cows. The standard values of these two transition probabilities, which were estimated in chronically infected herds, are perhaps not perfectly appropriate to describe the initial phase of an infection and then could lead to overestimation of the influence of seronegative shedders.

As suggested by Matthews *et al.* for *Escherichia coli* O157 [22], identifying factors such as age, genetics, reproductive status or other management factors that might predispose an animal to high levels of shedding would be of undisputable interest. Moreover, control measures should aim at reducing the proportion of high shedders in mucus/faeces, such as phase I vaccines seem to do. According to Arricau-Bouvery *et al.* [47], vaccination dramatically reduced both abortions and excretion of bacteria in the milk, vaginal mucus and faeces. In Rousset *et al.* [48], the vaccine was effective in reducing massive bacterial shedding from a heavily infected goat herd. In the same way, Hogerwerf *et al.* [49] found that in uterine fluid, vaginal mucus and milk of vaccinated dairy goats, both the prevalence of shedders as well as the concentration of bacteria were reduced.

## 6. Conclusion

This work led to the identification of key parameters in *C. burnetii* infection dynamics based on a methodology combining novel aspects of sensitivity analysis and an original model describing the bacterium spread within a dairy cattle herd composed of animals with heterogeneous shedding characteristics. The most influential parameters are associated to the probabilities governing the levels of shedding, especially for mucus/faeces shedders and to the characteristics of the bacterium in the environment. Some physiological parameters related to the intermittency of shedding or to the transition from one type of shedder to another one also play a non negligible role. Our study also highlights parameters that should be precisely assessed and then further investigated to improve the model accuracy and the understanding of the infection spread. Besides, interventions impacting those key parameters would be of great interest. The model developed here can be further used to assess over a longer time scale the effectiveness of different control strategies for *C. burnetii* infection, such as vaccination.

## 7. Acknowledgements

The authors would like to thank Alain Joly and Raphaël Guatteo from the Oniris-INRA group 'Bioaggression, Epidemiology and Risk Analysis' for fruitful discussions on the epidemiological model and on the way to take into account the heterogeneity of shedding. The collaboration between the Faculty of Veterinary Medicine of Utrecht and the French National Institute for Agricultural Research (INRA) was financially supported by the Netherlands Organisation for Scientific Research (NWO) and the French Ministry of Foreign and European Affairs through the Van Gogh Programme.

## 8. References

- [1] M.C. M. Plommet, J. Gestin, G. Renoux Fièvre Q expérimentale des bovins, Ann. Rech. vétér. 4 (1973) 325-346.

- [2] I.D. Aitken, Clinical aspects and prevention of Q fever in animals, *European J. Epidemiology*. 5 (1989) 420-424.
- [3] R.J. Bildfell, G.W. Thomson, D.M. Haines, B.J. McEwen, N. Smart, *Coxiella burnetii* infection is associated with placentitis in cases of bovine abortion, *J Vet Diagn Invest*. 12 (2000) 419-25.
- [4] M. Berri, A. Souriau, M. Crosby, D. Crochet, P. Lechopier, A. Rodolakis, Relationships between the shedding of *Coxiella burnetii*, clinical signs and serological responses of 34 sheep, *Vet Rec*. 148 (2001) 502-5.
- [5] O. Lyytikainen, T. Ziese, B. Schwartlander, P. Matzdorff, C. Kuhnhen, C. Jager, L. Petersen, An outbreak of sheep-associated Q fever in a rural community in Germany, *Eur J Epidemiol*. 14 (1998) 193-9.
- [6] J.H. McQuiston, J.E. Childs, Q fever in humans and animals in the United States, *Vector Borne Zoonotic Dis*. 2 (2002) 179-91.
- [7] B. Schimmer, R. Ter Schegget, M. Wegdam, L. Zuchner, A. de Bruin, P.M. Schneeberger, T. Veenstra, P. Vellema, W. van der Hoek, The use of a geographic information system to identify a dairy goat farm as the most likely source of an urban Q-fever outbreak, *BMC Infect Dis*. 10 (2010) 69.
- [8] A. Wallensten, P. Moore, H. Webster, C. Johnson, G. van der Burgt, G. Pritchard, J. Ellis-Iversen, I. Oliver, Q fever outbreak in Cheltenham, United Kingdom, in 2007 and the use of dispersion modelling to investigate the possibility of airborne spread, *Euro Surveill*. 15 (2010)
- [9] R. Guatteo, F. Beaudeau, M. Berri, A. Rodolakis, A. Joly, H. Seegers, Shedding routes of *Coxiella burnetii* in dairy cows: implications for detection and control, *Vet Res*. 37 (2006) 827-33.

- [10] N. Arricau-Bouvery, A. Souriau, P. Lechopier, A. Rodolakis, Excretion of *Coxiella burnetii* during an experimental infection of pregnant goats with an abortive goat strain CbC1, *Ann N Y Acad Sci.* 990 (2003) 524-6.
- [11] W.D. Tigertt, A.S. Benenson, W.S. Gochenour, Airborne Q fever, *Bacteriol Rev.* 25 (1961) 285-93.
- [12] N. Arricau-Bouvery, A. Rodolakis, Is Q fever an emerging or re-emerging zoonosis?, *Vet Res.* 36 (2005) 327-49.
- [13] T.J. Marrie, Q fever - a review, *Can. Vet. J.* 31 (1990) 555-563.
- [14] W. van der Hoek, F. Dijkstra, B. Schimmer, P.M. Schneeberger, P. Vellema, C. Wijkmans, R. ter Schegget, V. Hackert, Y. van Duynhoven, Q fever in the Netherlands: an update on the epidemiology and control measures, *Euro Surveill.* 15 (2010)
- [15] D. Raoult, H. Tissot-Dupont, C. Foucault, J. Gouvernet, P.E. Fournier, E. Bernit, A. Stein, M. Nesri, J.R. Harle, P.J. Weiller, Q fever 1985-1998. Clinical and epidemiologic features of 1,383 infections, *Medicine (Baltimore).* 79 (2000) 109-23.
- [16] A. Gikas, S. Kokkini, C. Tsioutis, Q fever: clinical manifestations and treatment, *Expert Rev Anti Infect Ther.* 8 (2010) 529-39.
- [17] M.P. Durand, Lacteal and placental excretion of *Coxiella burnetii*, agent of Q fever, in the cow. Importance and prevention, *Bull Acad Natl Med.* 177 (1993) 935-45; discussion 945-6.
- [18] N. Arricau Bouvery, A. Souriau, P. Lechopier, A. Rodolakis, Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes, *Vet Res.* 34 (2003) 423-33.

- [19] A. Rodolakis, M. Berri, C. Hechard, C. Caudron, A. Souriau, C.C. Bodier, B. Blanchard, P. Camuset, P. Devillechaise, J.C. Natorp, J.P. Vadet, N. Arricau-Bouvery, Comparison of *Coxiella burnetii* shedding in milk of dairy bovine, caprine, and ovine herds, *J Dairy Sci.* 90 (2007) 5352-60.
- [20] R. Guatteo, F. Beaudeau, A. Joly, H. Seegers, *Coxiella burnetii* shedding by dairy cows, *Vet Res.* 38 (2007) 849-60.
- [21] L. Matthews, M. Woolhouse, New approaches to quantifying the spread of infection, *Nat Rev Microbiol.* 3 (2005) 529-36.
- [22] L. Matthews, J.C. Low, D.L. Gally, M.C. Pearce, D.J. Mellor, J.A. Heesterbeek, M. Chase-Topping, S.W. Naylor, D.J. Shaw, S.W. Reid, G.J. Gunn, M.E. Woolhouse, Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control, *Proc Natl Acad Sci USA.* 103 (2006) 547-52.
- [23] T. Garske, C.J. Rhodes, The effect of superspreading on epidemic outbreak size distributions, *J Theor Biol.* 253 (2008) 228-37.
- [24] C. Lanzas, S. Brien, R. Ivanek, Y. Lo, P.P. Chapagain, K.A. Ray, P. Ayscue, L.D. Warnick, Y.T. Grohn, The effect of heterogeneous infectious period and contagiousness on the dynamics of *Salmonella* transmission in dairy cattle, *Epidemiol Infect.* 136 (2008) 1496-510.
- [25] M. J. Keeling, P. Rohani, *Modeling infectious diseases in Humans and animals*, Princeton University Press, Princeton, 2008.
- [26] A. Saltelli, K. Chan, E.M. Scott, *Sensitivity Analysis*, Wiley, Chichester, 2000.
- [27] A. Courcoul, E. Vergu, J.B. Denis, F. Beaudeau, Spread of Q fever within dairy cattle herds: key parameters inferred using a Bayesian approach, *Proc Biol Sci.* 277 (2010) 2857-65.
- [28] P.E. Fournier, T.J. Marrie, D. Raoult, Diagnosis of Q fever, *J Clin Microbiol.* 36 (1998) 1823-34.

- [29] M. Plommet, M. Capponi, J. Gestin, G. Renoux, Fièvre Q expérimentale des bovins, *Ann Rech Vet.* 4 (1973) 325-346.
- [30] A. Lurette, S. Touzeau, M. Lamboni, H. Monod, Sensitivity analysis to identify key parameters influencing *Salmonella* infection dynamics in a pig batch, *J Theor Biol.* 258 (2009) 43-52.
- [31] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate global sensitivity analysis for dynamic crop models, *Field Crop. Res.* 113 (2009) 312-320.
- [32] R Development Core Team (2009). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.R-project.org>.
- [33] A. Bouvier, A. Kobilinsky, H. Monod, PLANOR : an R library for the automatic generation of regular fractional factorial designs. Technical Report, INRA Jouy en Josas, 2010.
- [34] R.A. Bailey, *Design of comparative experiments*, Cambridge University Press, Cambridge, 2008.
- [35] A. Kobilinsky, Les plans factoriels, In: Dreesbeke, J.-J., Fine, J., Saporta, G. *Plans d'expériences : applications à l'entreprise*. Editions Technip, Paris, 1997, pp. 69–209.
- [36] M. Lamboni, H. Monod (2010). Multisensi: Multivariate Sensitivity Analysis. R package version 1.0-3, <http://CRAN.R-project.org/package=multisensi>.
- [37] A.F. Taurel, R. Guatteo, A. Joly, H. Seegers, F. Beaudeau, Q fever within-herd seroprevalence in infected dairy herds: assessment using an ELISA applied to bulk tank milk. [http://www.svepm.org.uk/posters/2010/Taurel\\_Q%20fever%20within-herd%20seroprevalence%20in%20infected%20dairy%20herd\\_assessment%20using%20an%20ELISA%20applied%20to%20bulk%20tan.pdf](http://www.svepm.org.uk/posters/2010/Taurel_Q%20fever%20within-herd%20seroprevalence%20in%20infected%20dairy%20herd_assessment%20using%20an%20ELISA%20applied%20to%20bulk%20tan.pdf). Accessed on 17/08/2010.

- [38] P. Ezanno, C. Fourichon, A.F. Viet, H. Seegers, Sensitivity analysis to identify key-parameters in modelling the spread of bovine viral diarrhoea virus in a dairy herd, *Prev Vet Med.* 80 (2007) 49-64.
- [39] J. Turner, R.G. Bowers, M. Begon, S.E. Robinson, N.P. French, A semi-stochastic model of the transmission of *Escherichia coli* O157 in a typical UK dairy herd: dynamics, sensitivity analysis and intervention/prevention strategies, *J Theor Biol.* 241 (2006) 806-22.
- [40] Z. Lu, Y.H. Schukken, R.L. Smith, Y.T. Grohn, Stochastic simulations of a multi-group compartmental model for Johne's disease on US dairy herds with test-based culling intervention, *J Theor Biol.* 264 (2010) 1190-201.
- [41] M.F. Weber, M. Nielen, A.G. Velthuis, H.J. van Roermund, Milk quality assurance for paratuberculosis: simulation of within-herd infection dynamics and economics, *Vet Res.* 39 (2008) 12.
- [42] M.R. Bruijnjs, H. Hogeveen, E.N. Stassen, Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model, *J Dairy Sci.* 93 (2010) 2419-32.
- [43] A. Saltelli, K. Chan, E.M. Scott, *Sensitivity Analysis*, Wiley, Chichester, 2000.
- [44] H. Monod, D. Makowski, C. Naud, Uncertainty and sensitivity analysis for crop models, In: D. Wallach, D. Makowski, J.W. Jones (Eds), *Working with Dynamic Crop Models, Evaluation, Analysis, Parameterization, and Applications*. Elsevier, Amsterdam, 2006, pp. 55-100.
- [45] I. Astobiza, J.F. Barandika, A. Hurtado, R.A. Juste, A.L. Garcia-Perez, Kinetics of *Coxiella burnetii* excretion in a commercial dairy sheep flock after treatment with oxytetracycline, *Vet J.* 184 (2009) 172-5.

- [46] P. Vellema, R. Van den Brom, D. Dercksen, L. Moll, H.J. Roest, Q fever in the Netherlands: One Health in relation to Q fever, in humans and animals. [www.minInv.nl/txmpub/files/?p\\_file\\_id=2000277](http://www.minInv.nl/txmpub/files/?p_file_id=2000277). Accessed on 17/08/2010.
- [47] N. Arricau-Bouvery, A. Souriau, C. Bodier, P. Dufour, E. Rousset, A. Rodolakis, Effect of vaccination with phase I and phase II *Coxiella burnetii* vaccines in pregnant goats, *Vaccine*. 23 (2005) 4392-402.
- [48] E. Rousset, B. Durand, J.L. Champion, M. Prigent, P. Dufour, C. Forfait, M. Marois, T. Gasnier, V. Duquesne, R. Thiéry, M.F. Aubert, Efficiency of a phase 1 vaccine for the reduction of vaginal *Coxiella burnetii* shedding in a clinically affected goat herd, *Clinical Microbiology and Infection*. 9999 (2009)
- [49] L. Hogerwerf, R. van den Brom, H.I.J. Roest, A. Bouma, P. Vellema, M. Pieterse, D. Dercksen, M. Nielen, Reduction of *Coxiella burnetii* prevalence by vaccination of goats and sheep, the Netherlands, *Emerging Infectious Diseases*. 17 (2011) 379-86.

## 9. Table and figure captions

Figure 1. Flow diagram describing the modelled spread of *C. burnetii* within a cattle herd. The health states are:  $S$ , susceptible, non-shedder cow without antibodies,  $I^-$ , shedder cow without any antibodies,  $I^+$ , shedder cow with antibodies,  $I^{+ \text{ milk pers}}$  shedder cow with antibodies shedding in milk in a persistent way,  $C^+$ , non-shedder cow with antibodies and  $C^-$ , non-shedder cow without antibodies which was infected and had antibodies in the past.  $I^-$  and  $I^+$  cows are in the shedding route category 1 if they shed in milk only, 2 if they shed in vaginal mucus/faeces only and 3 if they shed in milk and vaginal mucus/faeces.  $I^{+ \text{ milk pers}}$  cows are in the shedding category 1 if they shed in milk only and 3 if they shed in milk and vaginal mucus/faeces.  $E$  represents the environmental bacterial load. The model

parameters are presented in Table 1.  $\varepsilon_1$ ,  $\varepsilon_2$  and  $\varepsilon_3$  are the quantities of bacteria shed during a time step by an individual  $I^-$ ,  $I^+$  and  $I^{+ \text{ milk pers}}$  respectively and contaminating the environment. These quantities are the sum of all quantities  $Qty$  of bacteria shed by all the shedders through all the shedding routes times  $\rho$ , the fraction of bacteria shed reaching the environment of the herd.

Figure 2. Temporal dynamics of the seroprevalence, prevalence of shedders, prevalence of milk shedders and environmental bacterial load  $E_{building}$  (the environmental bacterial load of the main buildings and close pastures for lactating cows) for the standard scenario: mean (grey plain line), median (black plain line) and percentiles 2.5 and 97.5% (black dotted lines). 200 repetitions were run.

Figure 3. Number of shedders for each shedding route and each shedding level and their contributions in terms of contamination of the environment. Illustration based on the results at a snapshot in time for a single run.

Figure 4. Sensitivity analysis on the mean prevalence in mucus/faeces shedders over time: results of the ANOVA performed for the first component (inertia: 94.1%). (A) Loadings defining the principal component for each time variable (in abscissa) and total sensitivities for the 10 most influential factors ranked in descending order. Sensitivities are split in main effect (black) and two-factor interactions (grey). (B) Sensitivity indices of the 15 main factorial terms (main effects or interactions) in descending order.

Figure 5. Sensitivity analysis of the means of six of the outputs (all except the abortion number) for the last simulation time point (week 260): results of the ANOVA performed for the first principal component (inertia: 80.5%). (A) Loadings defining the principal component for each time variable (in abscissa) and total sensitivities for the 10 probability distributions  $Q$  ranked in descending order. Sensitivities are split in main effect (black) and two-factor interactions (grey). (B) Sensitivity indices of the 15 main factorial terms (main effects or interactions) in descending order.

Table 1. Parameters of the epidemiological model: description, standard values and values tested in the first experiment of the sensitivity analysis.

Table 2. Description of the parameters of the herd demography model and their standard values.

Table 3. Description of the outputs of the sensitivity analysis.

Table 4. Results of the first experiment of the sensitivity analysis: for the mean and standard deviation of each output, the proportion of total inertia represented by the first principal component and the corresponding most influential factors are given. Factors with total sensitivities higher than 10%: one cross, factors with total sensitivities  $\geq 25\%$ : two crosses, factors with total sensitivities  $\geq 50\%$ : three crosses.

Table 5. Results of the third experiment of the sensitivity analysis: for the mean and standard deviation of each output, the proportion of total inertia represented by the first principal component and the corresponding most influential factors are given. Factors with total sensitivities higher than 10%: one cross, factors with total sensitivities  $\geq 25\%$ : two crosses, factors with total sensitivities  $\geq 50\%$ : three crosses.

Accepted manuscript

Figure 1

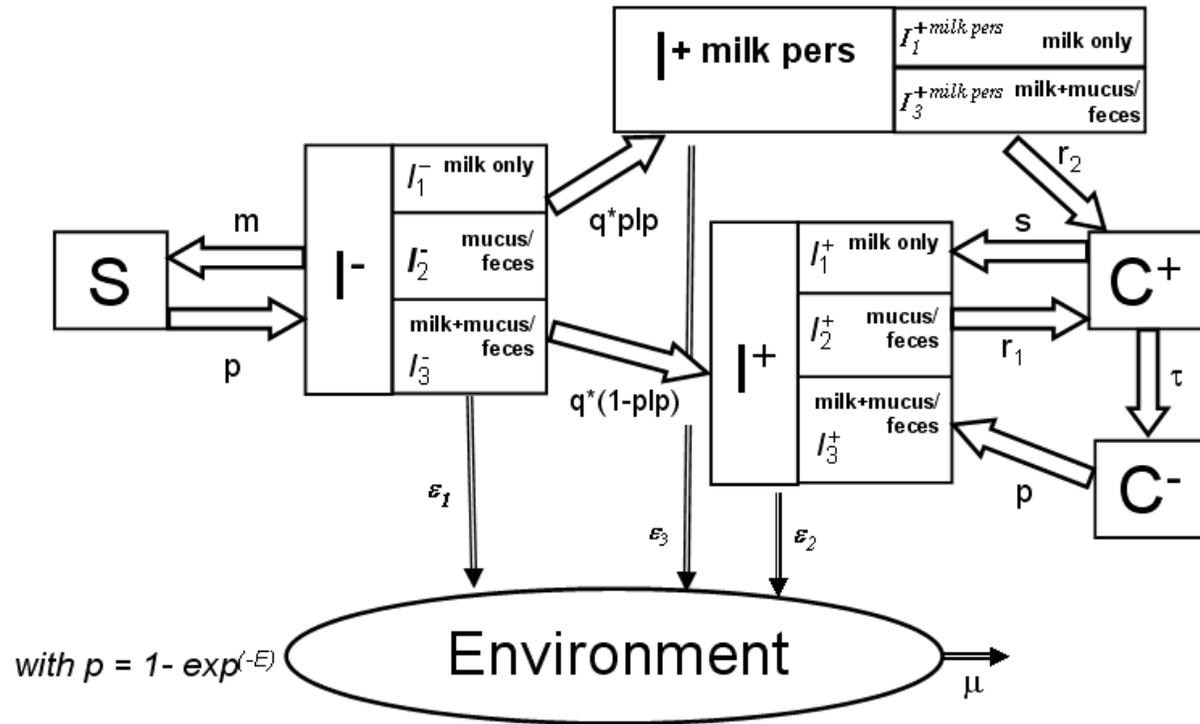


Figure 2

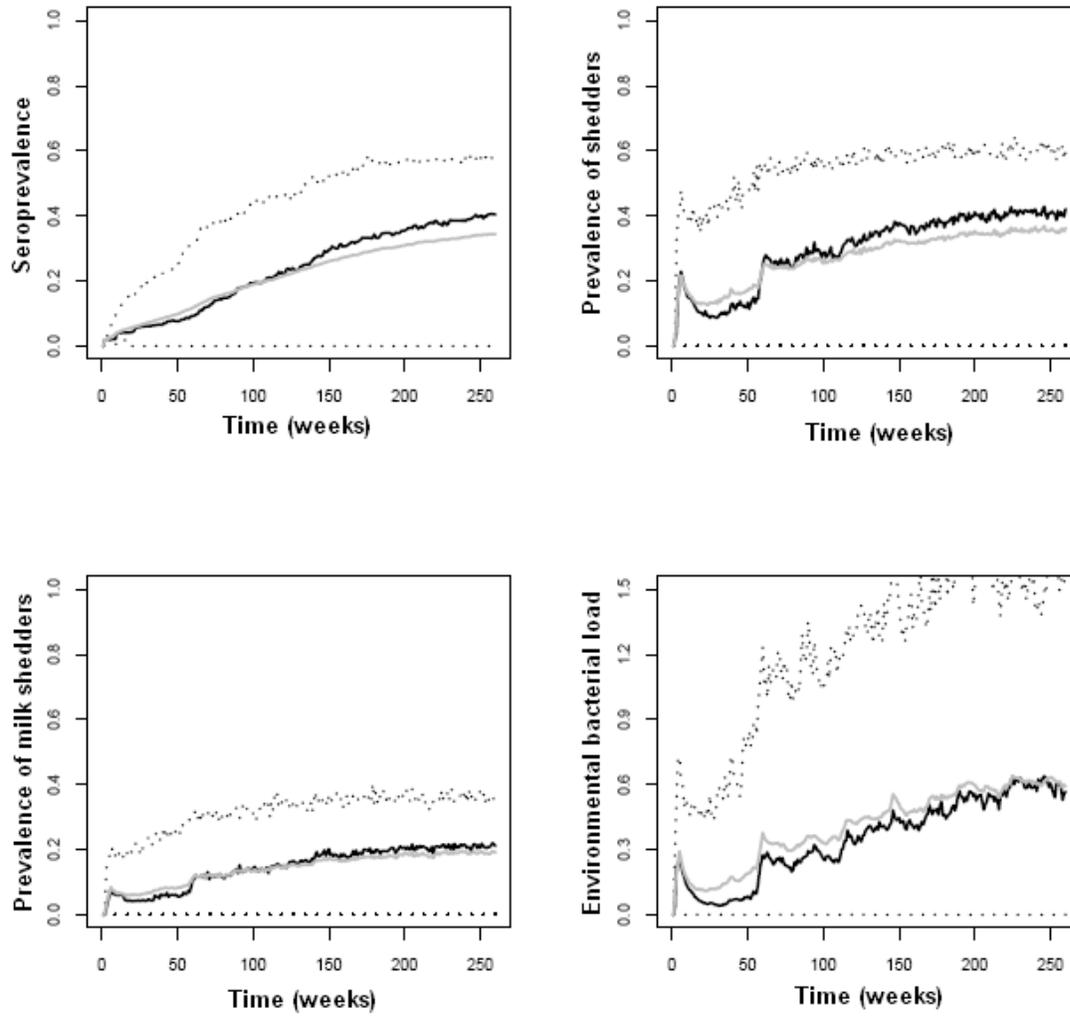


Figure 3

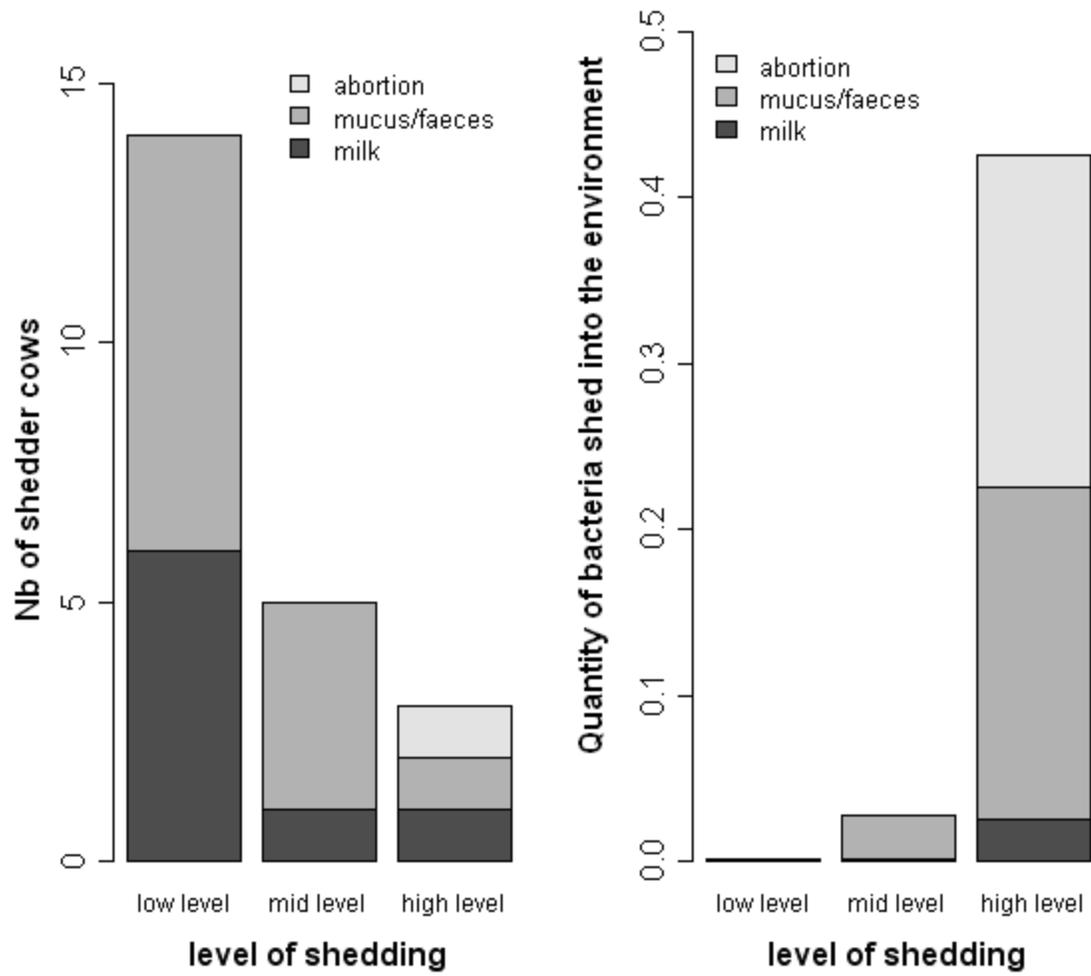


Figure 4

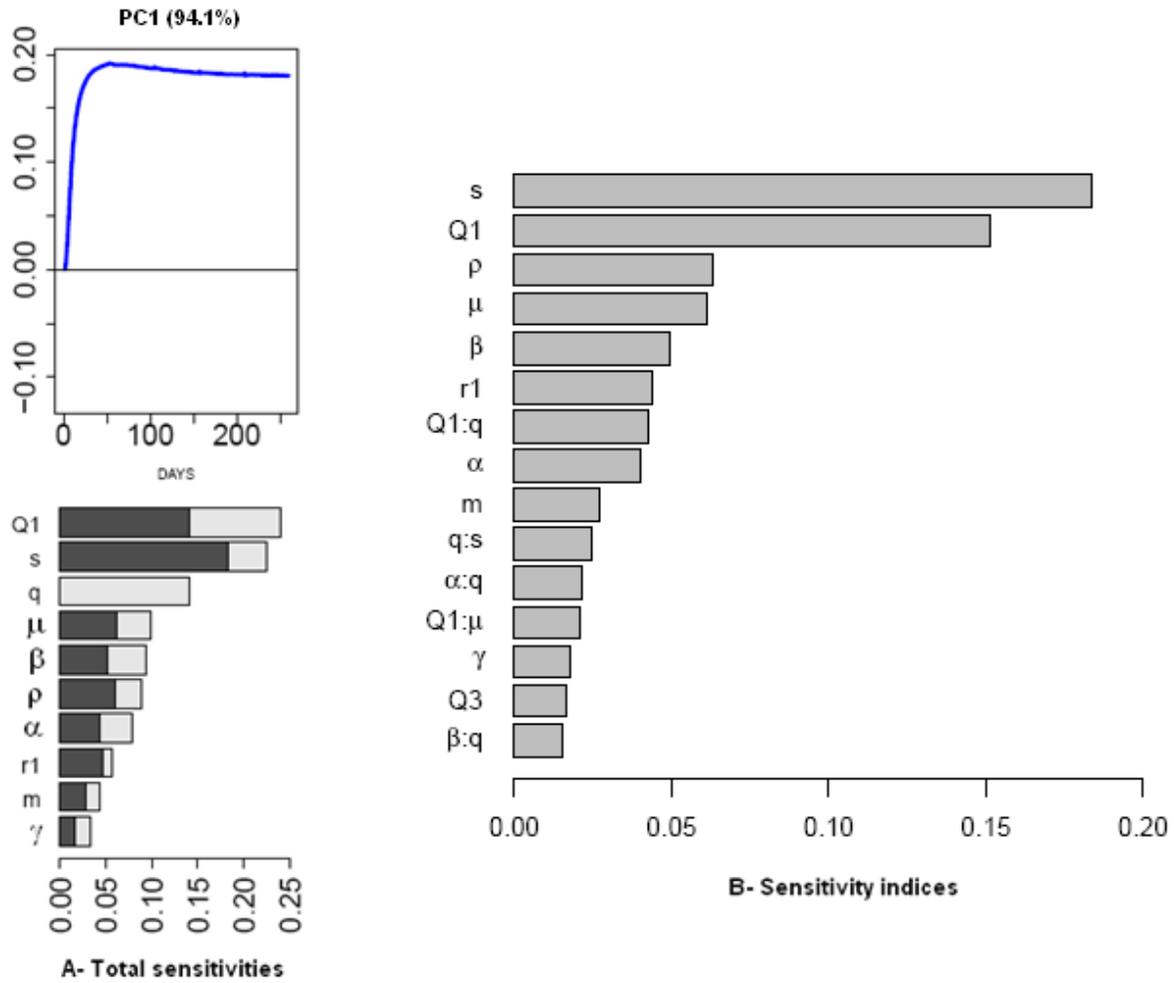
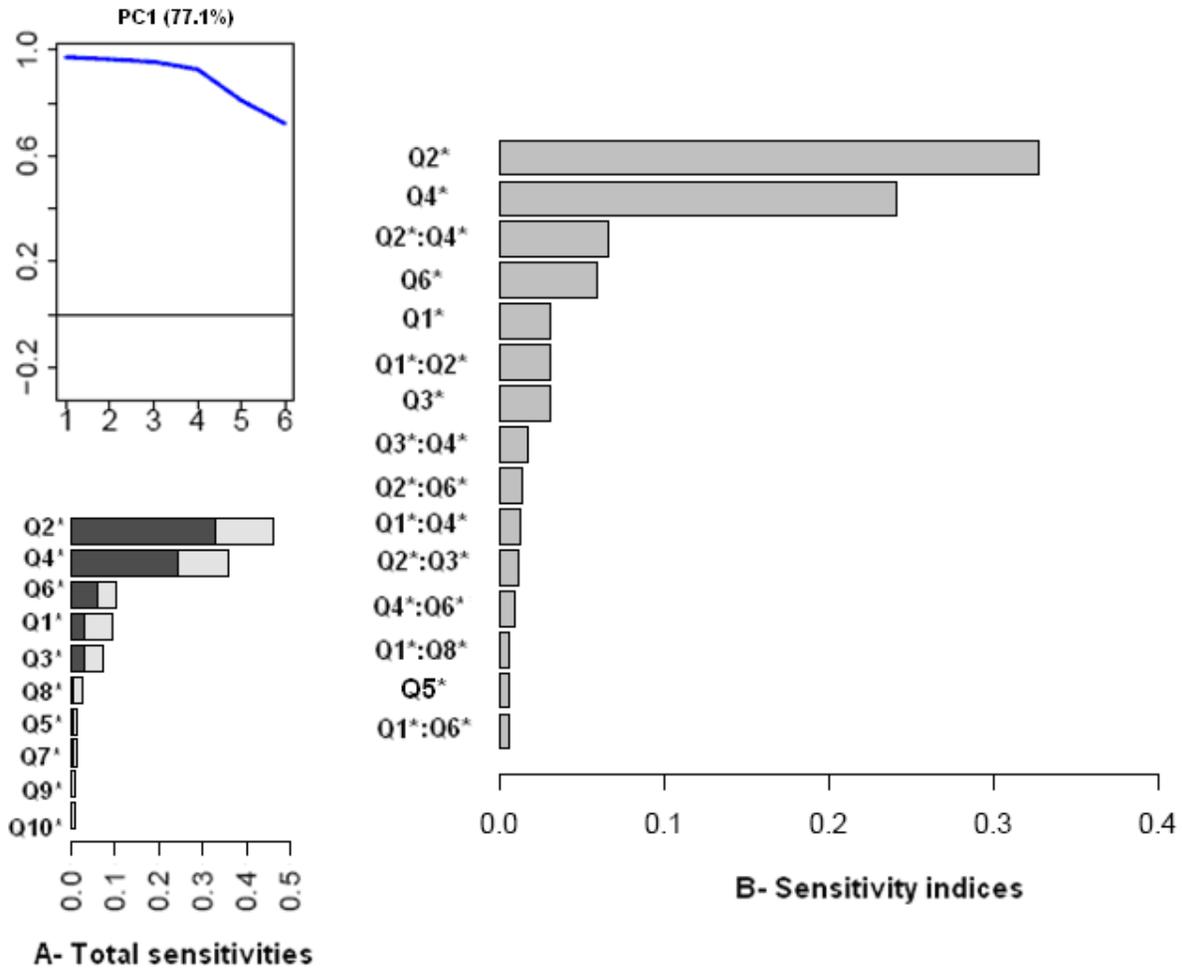


Figure 5



1 Table 1

Factor name	Description	Standard value	Values tested in the sensitivity analysis			
$m$	Transition probability $I^- \Rightarrow S$	0.7 <sup>a</sup>	0.33	0.8		
$q$	Transition probability $I^- \Rightarrow I^+$	0.02 <sup>a</sup>	0.01	0.2		
$plp$	Proportion of cows going from $I^-$ to $I^+$ and becoming $I^{+ \text{ milk pers}}$	0.5	0.2	0.7		
$r_1$	Transition probability $I^+ \Rightarrow C^+$	0.2 <sup>a</sup>	0.11	0.33		
$r_2$	Transition probability $I^{+ \text{ milk pers}} \Rightarrow C^+$	0.02	0.01	0.06		
$s$	Transition probability $C^+ \Rightarrow I^+$	0.15 <sup>a</sup>	0.04	0.4		
$\tau$	Transition probability $C^+ \Rightarrow C^-$	0.0096				
$\mu$ (week <sup>-1</sup> )	Elimination rate of <i>C. burnetii</i> *	0.2 <sup>a</sup>	0.08	0.5		
$probav$	Probability of abortion after a transition $S \Rightarrow I^-$ , $C^+ \Rightarrow I^+$ or $C^- \Rightarrow I^+$	0.02				
$\rho^{mf}$	Proportion of bacteria shed through mucus/faeces filling the environment compartment	0.28	0.05	0.15	0.35	0.5
ratio $\rho^{milk} / \rho^{mf}$	Where $\rho^{milk}$ = proportion of bacteria shed through milk filling the environment	0.125	0.0625	0.125	0.25	0.5

		compartment					
$\alpha$	$\alpha_1$ , milk		0.31 <sup>b</sup>	0.31	0.6	0.06	0
	$\alpha_2$ , mucus/faeces	Probability distribution of the shedding routes for the $I^-$ cows	0.62 <sup>b</sup>	0.62	0.3	0.75	1
	$\alpha_3$ , milk+mucus/faeces		0.07 <sup>b</sup>	0.07	0.1	0.19	0
$\beta$	$\beta_1$ , milk		0.61 <sup>b</sup>	0.61	0.14	0.2	0.77
	$\beta_2$ , mucus/faeces	Probability distribution of the shedding routes for the $I^+$ cows after 4 weeks post-calving	0.33 <sup>b</sup>	0.33	0.5	0.7	0.09
	$\beta_3$ , milk+mucus/faeces		0.06 <sup>b</sup>	0.06	0.36	0.1	0.14
$\beta_{calv}$	$\beta_{calv1}$ , milk		0.14 <sup>b</sup>	0.14	0.61	0.2	0
	$\beta_{calv3}$ , mucus/faeces	Probability distribution of the shedding routes for the $I^+$ cows in the 4 first weeks post-calving	0.5 <sup>b</sup>	0.5	0.33	0.7	0.33
	$\beta_{calv3}$ , milk+mucus/faeces		0.36 <sup>b</sup>	0.36	0.06	0.1	0.67
$\gamma$	$\gamma_1$ , milk	Probability distribution of the shedding routes for the $I^{+milk\ pers}$ cows after 4 weeks post-calving	0.83 <sup>b</sup>	0.83	0.25	0.5	1
	$\gamma_3$ , milk+mucus/faeces		0.17 <sup>b</sup>	0.17	0.75	0.5	0
$\gamma_{calv}$	$\gamma_{calv1}$ , milk	Probability distribution of the shedding routes for the $I^{+milk\ pers}$ cows in the 4 first weeks post-calving	0.25 <sup>b</sup>	0.25	0.83	0.5	0
	$\gamma_{calv3}$ , milk+mucus/faeces		0.75 <sup>b</sup>	0.75	0.17	0.5	1
$Q1$	low level		0.85 <sup>b</sup>	0.85	0.6	0.25	0.15
	mid level	Probability distribution of the shedding levels for all the $I^-$ and for the $I^+$ shedding in	0.15 <sup>b</sup>	0.15	0.4	0.25	0.6

	high level	mucus/faeces after 4 weeks post-calving	0 <sup>b</sup>	0	0	0.5	0.25
Q2	low level	Probability distribution of the shedding levels for the $I^+$ shedding in milk after 4 weeks post-calving	0.4 <sup>b</sup>	0.4	0.85	0.25	0.15
	mid level		0.5 <sup>b</sup>	0.5	0.15	0.25	0.6
	high level		0.1 <sup>b</sup>	0.1	0	0.5	0.25
Q3	low level	Probability distribution of the shedding levels for all the $I^+$ in the 4 first weeks post-calving	0.25 <sup>b</sup>	0.25	0.85	0.6	0.15
	mid level		0.25 <sup>b</sup>	0.25	0.15	0.4	0.6
	high level		0.5 <sup>b</sup>	0.5	0	0	0.25
Q4	low level	Probability distribution of the shedding levels for the $I^{+milk\ pers}$ shedding in mucus/faeces after 4 weeks post-calving	0.6 <sup>b</sup>	0.6	0.85	0.25	0.15
	mid level		0.4 <sup>b</sup>	0.4	0.15	0.25	0.6
	high level		0 <sup>b</sup>	0	0	0.5	0.25
Q5	low level	Probability distribution of the shedding levels for all the $I^{+milk\ pers}$ shedding in milk and for the $I^{+milk\ pers}$ in the 4 first weeks post-calving	0.15 <sup>b</sup>	0.15	0.85	0.6	0.25
	mid level		0.6 <sup>b</sup>	0.6	0.15	0.4	0.25
	high level		0.25 <sup>b</sup>	0.25	0	0	0.5

<sup>a</sup>from Courcoul *et al.* [27]

<sup>b</sup>calibrated to match field data (R. Guatteo 2009, personal communication)

\*includes both the natural mortality of the bacterium and its removal in relation to the periodic cleaning of the cattle housing

Accepted manuscript

4 Table 2

Description		Standard value
Replacement rate (year <sup>-1</sup> )		0.355
Culling rate (week <sup>-1</sup> )	lactation 1	0.0057
	lactation 2	0.0052
	lactation 3	0.0065
	lactation 4	0.0067
	lactations 5&6	0.0161
Probability distribution at time 0 for the lactation numbers of the cows	lactation 1	0.337
	lactation 2	0.252
	lactation 3	0.173
	lactation 4	0.11
	lactation 5	0.088
	lactation 6	0.04

5

6

Accepted manuscript

7 Table 3

Output name	Description
$E_{building}$	Environmental bacterial load of the main buildings and close pastures
$E_{dry}$	Environmental bacterial load of the specific pastures for the dry cows
Prevalence of milk shedders	Proportion of animals of the herd shedding in milk, i.e. $\frac{I_1^- + I_3^- + I_1^+ + I_3^+ + I_1^{+ milk pers} + I_3^{+ milk pers}}{N} (*)$
Prevalence of mucus/faeces shedders	Proportion of animals of the herd shedding in vaginal mucus and/or faeces, i.e. $\frac{I_2^- + I_3^- + I_2^+ + I_3^+ + I_3^{+ milk pers}}{N} (*)$
Prevalence of shedders in milk in a persistent way	Proportion of animals $I^{+ milk pers}$ , i.e. $\frac{I_1^{+ milk pers} + I_3^{+ milk pers}}{N} (*)$
Seroprevalence	Proportion of animals with antibodies, i.e. $\frac{I_1^+ + I_2^+ + I_3^+ + I_1^{+ milk pers} + I_3^{+ milk pers} + C^+}{N} (*)$
Number of abortions per herd per year	
Extinction rate	Proportion of runs of a particular scenario leading to an extinction of

---

the infection (\*\*)

---

8 \*  $N$  denotes the herd size

9 \*\* the infection is assumed extinct when there is no  $I$  and  $C^f$  left until the end of the simulation time

Accepted manuscript

10 Table 4

	Environment $E_{building}$		Environment $E_{dry}$		Prevalence of milk shedders		Prevalence of m/f shedders		Prevalence of milk shedders in a persistent way		Seroprevalence		Abortion number		Joined analysis time point 260	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
inertia	99.1%	96.3%	98.4%	96.8%	94.4%	87.5%	94.1%	80.7%	93.1%	87.7%	97.0%	86.8%	94.0%	88.0%	49.0%	43.8%
PC1																
$m$																
$q$					++	+	+	+	++	++	+++	++	++	++	+	+
$plp$									++	++						+
$r_1$																
$r_2$									++	++						+
$s$						+		+				+	+	+	+	
$\mu$	++	++	++	++									+	+	+	
$\rho^{mf}$	++	++	++	++				+					+	+	+	
<i>ratio</i>																
$\rho^{milk}/\rho^{mf}$																
$\alpha$							+									

Accepted manuscript

$\beta$

$\beta_{calv}$

$\gamma$

$\gamma_{calv}$

Q1

++ ++ ++ +++ + + + + + ++ ++ + +

Q2

Q3

+ + +

Q4

Q5

12 Table 5

	Environment $E_{building}$		Environment $E_{dry}$		Prevalence of milk shedders		Prevalence of m/f shedders		Prevalence of milk shedders in a persistent way		Seroprevalence		Abortion number		Joined analysis time point 260	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
inertia PC1	96.6%	91.0%	86.3%	84.1%	89.7%	73.7%	86.2%	76.3%	83.8%	45.3%	95.8%	76.4%	80.8%	47.9%	77.1%	54.5%
Q1*					+	+	+	+	+	+	+		+	+		+
Q2*	+++	+++	+++	+++	+++	+	+++	+	+++	++	+++	+	+++	+	++	+
Q3*															+	
Q4*		+	++	++	++		++	++	++	+	++		+	+	++	
Q5*																
Q6*					+	++	+	++	+	+	+	++	+	+	+	+
Q7*																
Q8*										+					+	
Q9*																
Q10*																

13

- 14 We model *Coxiella burnetii* spread within a dairy cattle herd.
- 15 We identify key parameters whose variation highly influences the infection dynamics.
- 16 Shedding levels, mainly in mucus/faeces, are the most influential parameters.
- 17 Characteristics of the bacterium in the environment have also a great impact.
- 18

Accepted manuscript