



# Human Q fever incidence is associated to spatiotemporal environmental conditions



J.P.G. Van Leuken<sup>a,b,\*</sup>, A.N. Swart<sup>a</sup>, J. Brandsma<sup>c</sup>, W. Terink<sup>c</sup>, J. Van de Kasstele<sup>a</sup>, P. Droogers<sup>c</sup>, F. Sauter<sup>d</sup>, A.H. Havelaar<sup>a,b,e</sup>, W. Van der Hoek<sup>a</sup>

<sup>a</sup> Centre for Infectious Disease Control (CIb), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

<sup>b</sup> Institute for Risk Assessment Sciences (IRAS), Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

<sup>c</sup> Future Water, Wageningen, The Netherlands

<sup>d</sup> Environmental Safety (M&V), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

<sup>e</sup> Emerging Pathogens Institute, University of Florida, Gainesville, Florida, United States

## ARTICLE INFO

### Article history:

Received 24 March 2015

Received in revised form 4 February 2016

Accepted 14 March 2016

Available online 4 April 2016

### Keywords:

Q fever

Environmental

Risk factors

Atmospheric dispersion model

Prediction

Multivariate

*Coxiella burnetii*

Concentration

Vegetation

Soil moisture

## ABSTRACT

Airborne pathogenic transmission from sources to humans is characterised by atmospheric dispersion and influence of environmental conditions on deposition and re-aerosolisation. We applied a One Health approach using human, veterinary and environmental data regarding the 2009 epidemic in The Netherlands, and investigated whether observed human Q fever incidence rates were correlated to environmental risk factors.

We identified 158 putative sources (dairy goat and sheep farms) and included 2339 human cases. We performed a high-resolution ( $1 \times 1$  km) zero-inflated regression analysis to predict incidence rates by *Coxiella burnetii* concentration (using an atmospheric dispersion model and meteorological data), and environmental factors – including vegetation density, soil moisture, soil erosion sensitivity, and land use data – at a yearly and monthly time-resolution.

With respect to the annual data, airborne concentration was the most important predictor variable (positively correlated to incidence rate), followed by vegetation density (negatively). The other variables were also important, but to a less extent. High erosion sensitive soils and the land-use fractions “city” and “forest” were positively correlated. Soil moisture and land-use “open nature” were negatively associated. The geographical prediction map identified the largest Q fever outbreak areas. The hazard map identified highest hazards in a livestock dense area.

We conclude that environmental conditions are correlated to human Q fever incidence rate. Similar research with data from other outbreaks would be needed to more firmly establish our findings. This could lead to better estimations of the public health risk of a *C. burnetii* outbreak, and to more detailed and accurate hazard maps that could be used for spatial planning of livestock operations.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

### 1.1. Background

Q fever is a worldwide livestock-associated disease caused by the gram-negative bacterium *Coxiella burnetii*. The largest Q fever outbreak ever described occurred in The Netherlands from 2007 to 2010 with over 4,000 notified human cases [1]. In this epidemic, human Q fever was mainly associated with dairy goats, and to a lesser degree with dairy sheep [2]. The main reservoirs of *C. burnetii* are the goats' and sheep's placentas and birth products; the main clinical symptom is late

abortion of lambs [3–5]. The birth products may contain up to about one billion *C. burnetii* bacteria per gram [4–6]. Other, but less important, excretion routes are milk [7–9] and feces and urine [9]. Excretion of the bacterium can last up to months after infection by *C. burnetii* [6].

Since goat stables are only partially enclosed, excreted bacteria can be transmitted from the indoor to the outdoor environment. *C. burnetii* has been detected in outdoor air samples in several particle size fractions in the surroundings of positive farms [10,11]. The bacterium and its spore-like forms may survive for weeks to months in the outdoor environment [12–14].

### 1.2. Human infection

Human infection generally occurs by inhaling the *C. burnetii* bacterium. The dose for 50% human illness (ID50) is about 1.18 bacteria

\* Corresponding author at: National Institute for Public Health and the Environment, Centre for Infectious Disease Control, P.O. Box 1, 3720 BA Bilthoven, The Netherlands.  
E-mail address: [jeroen.van.leuken@rivm.nl](mailto:jeroen.van.leuken@rivm.nl) (J.P.G. Van Leuken).

[15]. There is no evidence for human illness being transmitted by the foodborne pathway [14].

Human Q fever is a notifiable disease in The Netherlands since 1978. Prior to the large epidemic in 2007–2010, the number of notified cases varied between 5 and 20 per year, with a maximum estimated seroprevalence of antibodies against *C. burnetii* in the general population of 2.3% in 2006 [16]. The total number of (a)symptomatic human infections during the epidemic was estimated to be about 12.6 times larger than the total number of notified cases [17]. The seroprevalence in the centre of the epidemic was estimated at 12.2% in 2009 [18] and 10.7% in 2010–2011 [19]. The Q fever prevalence on large dairy goat farms (>50 animals) was estimated to be 43.1% in 2009 [20].

### 1.3. Problem description

The number of humans infected by an airborne pathogen depends on:

- (1) emission strength;
- (2) airborne transmission from source to receptor as a function of meteorological and environmental conditions [21–28]; and
- (3) human exposure: the exposure level is dependent on, among others, duration [29], location [30] and physical activity [31].

In the current study we mainly focused on the second point to investigate the possible effects of the transmission pathway (outdoor environment) between source (farms) and receptors (humans). The outdoor environment could play a threefold role: first, meteorological conditions (including wind speed/direction, precipitation and vertical motions in the atmosphere) influence the degree and pre-dominant direction of the airborne transmission of *C. burnetii* bacteria that were released from a farm's stable into the air [22,32]; secondly, landscape elements (including vegetation) are able to remove particles from the atmosphere by deposition (e.g., [33,34]). Thirdly, these landscape elements and specific environmental conditions possibly affect the re-aerosolisation of bacteria [35,36].

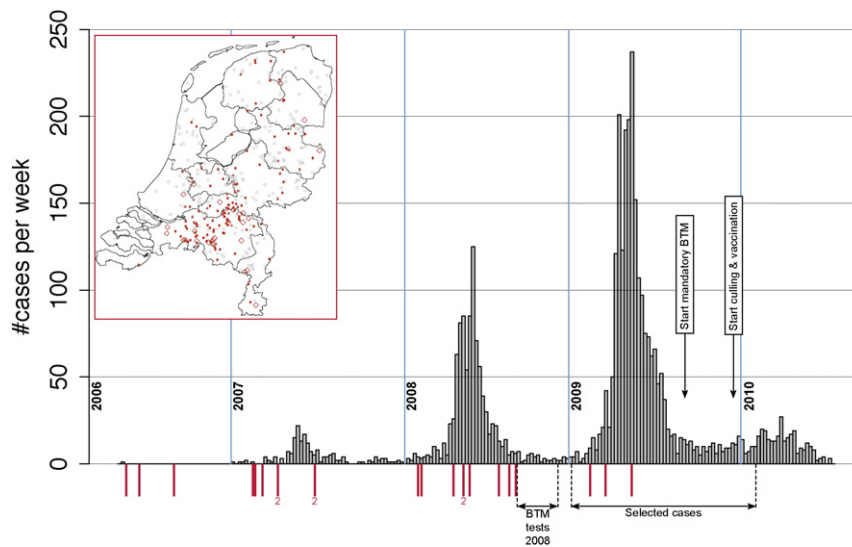
The Dutch Q fever cases were spatially clustered with large regional differences in incidence rates. Highest rates were reported in the south-east of The Netherlands [1]; in several other regions positive farms were identified without hardly any reported human cases nearby (Fig. 1). Therefore, an explorative statistical analysis was performed in 2011 to investigate the correlation between Q fever incidence rate and specific environmental conditions (vegetation greenness, land use, soil texture, soil moisture content, wind velocity, temperature, and global radiation) [27]. The authors concluded that relatively low vegetation greenness and low soil moisture content levels were correlated to higher transmission levels of *C. burnetii*.

Since that study was explorative, several important assumptions were made:

- Cases were assumed to have been infected only by the nearest positive farm.
- Only farms with reported abortion waves (i.e. >5% of the parturitions [37]) ( $n = 27$ ) were assumed to have emitted bacteria, while in practice many more had tested positive in one or multiple tests (see Methods section).
- Soil moisture content was based on static ground water level data with no temporal component.
- “Farm status” was dichotomised to months with “no transmission” (if the incidence rate was  $\leq 1/10,000$  within 5 km) or “transmission”.

### 1.4. Aim

The aim of this research was to refine the study of [27] to test the hypothesis that transmission of *C. burnetii* from farms to humans was associated to environmental conditions and their spatial and temporal heterogeneity using a zero inflated regression analysis. In addition, we included additional variables: concentration and deposition of *C. burnetii* modeled by means of an hourly-based atmospheric dispersion model (which includes among others precipitation amount/duration, atmospheric stability, and wind speed/direction, and the relative geographical position of cases with respect to the position of putative sources), leaf area index, soil wind erosion sensitivity, and time-dependent soil



**Fig. 1.** Time course of the Q fever epidemics in The Netherlands with the number of notified cases per week (gray bars). Red bars indicate the week in which an abortion wave was registered at one (or two: '2') farms. The non-systematic bulk tank milk (BTM) tests were performed in the autumn of 2008; the mandatory and systematic BTM tests were performed from September 2009. Mandatory and systematic vaccination and culling started at the end of December 2009. The geographic map of The Netherlands shows the location of all large dairy goat farms with abortion waves (♦), all other positive (•) and negative (○) farms.

moisture levels at a high spatial and temporal resolution. Results of this study might lead to environmental transmission risk mapping.

## 2. Method

### 2.1. Data description

#### 2.1.1. Research period

The Q fever epidemic lasted four years (2007–2010) [1] (Fig. 1), but in the current study we focused primarily on the epidemic in 2009, since:

- 1) A systematic farm monitoring program was not organised until September 2009, hence the infection status was not known for all goat farms during the previous years;
- 2) Dairy goats and dairy sheep were systematically and mandatory vaccinated or culled from the beginning of 2010 onward, resulting in sharply decreasing numbers of human cases in that year; and
- 3) The number of human notifications was largest in 2009, both due to the larger number of positive farms (hence a larger infection pressure) and a possibly increased awareness.

We performed our analyses at a monthly resolution and for the full research year (1 January 2009–31 December 2009).

#### 2.1.2. Human data

The regional Municipal Health Services made the dates of disease onset of notified human Q fever cases available, including their six-figure zip codes (referred to as PC6, i.e. at street-level resolution). Combining these dates of disease onset with the Q fever incubation period of 8–33 days (95% CI,  $\mu = 20.7$  days) [38], we selected all notified human Q fever cases with a date of onset of disease between 9-Jan.-2009 and 2-Feb.-2010 ( $n = 2339$ ). Monthly incidence rate were based on all cases that could have been infected symptomatically during that particular month (Fig. 2), and we assumed that notified cases were infected within their street of residence [30].

In addition, we used population density data at the PC6-level (reference date 1 January 2010), made available by Statistics Netherlands (CBS). We aggregated the population density data and case data to a 1-km<sup>2</sup> grid and selected all cells with one or more inhabitants (“inhabited grid cells”) ( $n = 25,339$ ).

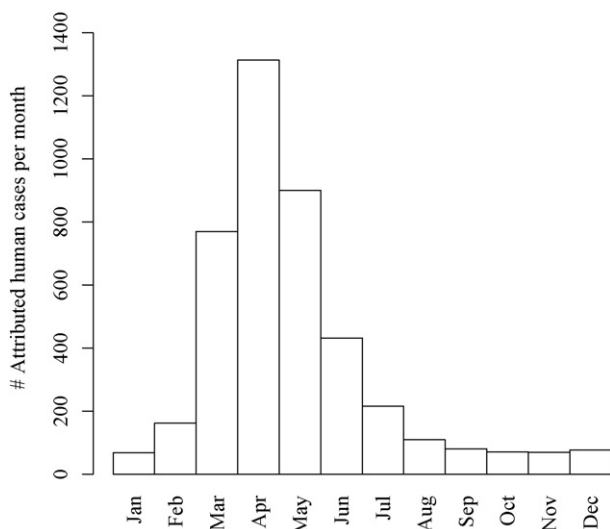


Fig. 2. Timeline of the number of attributed human Q fever cases per month based on a case's day of onset of disease and the Q fever incubation period of 8–33 days (95% CI) [38].

#### 2.1.3. Veterinary data

The Ministry of Economic Affairs made the coordinates available of all goat and sheep farms in The Netherlands ( $n = 42,509$ ) and the number of goats and sheep per farm in November 2008 and November 2009. The majority of these farms were small hobby farms. We selected all large dairy goat farms (i.e. >50 animals;  $n = 373$  in 2008 and  $n = 405$  in 2009), since those were associated to the Q fever outbreaks [37].

We assumed a farm was positive for *C. burnetii* in 2009 if at least one of the following conditions was met (most farms met multiple conditions):

- The farm's tank milk tested positive in the non-systematic and non-mandatory Bulk Tank Milk (BTM) sampling program implemented by the Animal Health Service (GD) in the autumn of 2008. This program included approximately 66% of the large dairy goat farms [8].
- The farm's tank milk tested positive in the mandatory BTM sampling program, systematically conducted at all large dairy goat and sheep farms by the Animal Health Service on authority of The Netherlands Food and Consumer Product Safety Authority (nVWA) from September 2009 [2].
- An abortion rate of >5% was reported by the Animal Health Service either in the years 2007, 2008 or 2009 [37].
- At least one vaginal or environmental swab was confirmed positive in a non-systematic sampling investigation [39,40] in 2008 and 2009.

This resulted in a list of 158 positive dairy goat farms (Fig. 1), of which most kept >50 goats in both 2008 and 2009 (Supplementary Table A1). Some also kept sheep; thirteen farms started after November 2008; two farms terminated their activities after November 2008.

#### 2.1.4. Concentration data (airborne transmission)

We correlated observed Q fever incidence rate to the concentration modeled by an atmospheric dispersion model (airborne transmission), modeled deposition levels, and other environmental conditions (related to deposition and re-aerosolisation) (Fig. 3).

We calculated airborne concentration levels of *C. burnetii* by means of the OPS-ST model (Operational Priority Substances – Short Term, version 10.3.2), which is an atmospheric dispersion model (ADM) developed by The Netherlands National Institute for Public Health and the Environment (e.g., [41–44]).

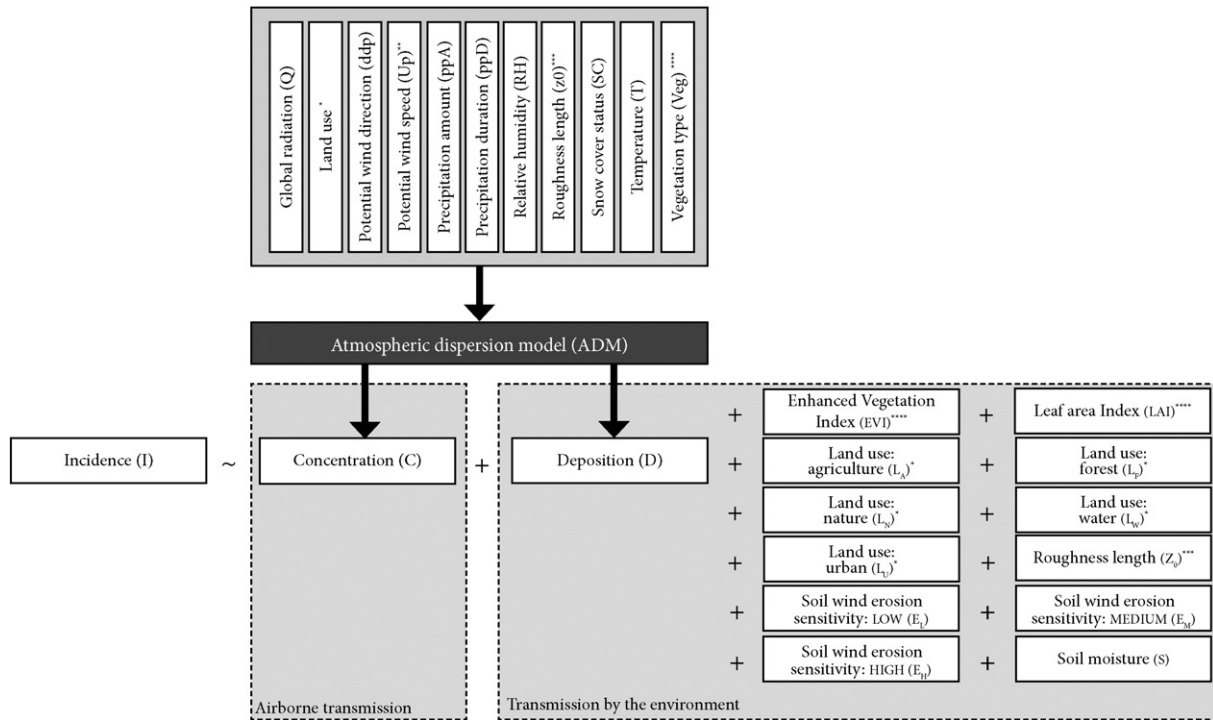
In a previous study it was concluded that modeled log-transformed ADM-concentrations correlated better to observed Q fever incidence rate than either a distance-based model or a model with no predictors [21]. That is, meteorological conditions (including wind and precipitation) are correlated with observed Q fever incidence rates.

We applied their ADM-configurations and considered each positively classified farm as a putative source of *C. burnetii*. Although the bulk tank milk status of the majority of the farms was known, it was unclear how to convert these values to emission strengths [45]. Therefore, we used each farm's herd size as a proxy for the steady-state emission strength:

$$E_j = \max(A_{j,2008}, A_{j,2009}). \tag{1}$$

With  $A_{j,2008}$  as the number of animals at farm  $j$  in November 2008, and  $A_{j,2009}$  in November 2009.

Required meteorological data as input data for the ADM included hourly averaged global radiation, precipitation amount and duration, relative humidity, snow cover status, temperature, and wind speed and direction (Fig. 3), retrieved from the Royal Netherlands Meteorological Institute (KNMI) and interpolated to farm levels [21]. Precipitation data were available at a 1 km<sup>2</sup> resolution; we spatially interpolated the other hourly data to get values at farm levels. We aggregated the hourly output concentration and deposition maps to monthly and yearly averaged values at a 1 km<sup>2</sup> resolution.



**Fig. 3.** Overview of all variables. Incidence rate is explained by concentration (airborne transmission) and variables related to deposition and re-aerosolisation. Concentration and deposition were calculated by the OPS-ST model using meteorological data and data of land use, roughness length, and vegetation type. Variables indicated by one or more asterisks (\*) are related to each other, but are not necessarily equal.

### 2.1.5. Environmental data related to deposition and re-aerosolisation

Important environmental variables potentially influencing the re-aerosolisation and deposition of particles/bacteria/spores are: soil texture, roughness elements, soil moisture content, and wind speed [35]. Therefore, we added soil moisture content [%] [46], vegetation greenness (Enhanced Vegetation Index) [–] [47], vegetation density (Leaf Area Index) [m<sup>2</sup> leaf per m<sup>2</sup> ground surface] [48], soil erosion sensitivity [categorical: low/moderate/high] [49], land use (categorical: agriculture/forest/open nature/water/urban) [50], and roughness length [41] to the model (Fig. 3). Data were averaged to yearly and monthly values (Supplementary Text A). Note that vegetation, land use and roughness length are related to both deposition and re-aerosolisation.

Assuming that the notified cases were infected at their residential addresses [30], not only the environmental conditions within an inhabited grid cell of 1 km<sup>2</sup>, but also the conditions in the surrounding of a grid cell will have positive or negative effects on deposition and re-aerosolisation. Since the risk of a Q fever infection has been estimated to be largest within 5 km of a positive farm [51], we spatially averaged the monthly and yearly averaged numerical environmental parameters within 5 km to a single numerical value per inhabited grid cell per month and per year (Supplementary Text A). In addition, we calculated fractions of the categorical parameters (land use and soil erosion sensitivity classes) within the 5 km radius (Supplementary Text A).

## 2.2. Statistical analyses

### 2.2.1. Preparation

In order to prevent too high skewness of the predictor variables, we transformed the concentration and environmental data to normal distributions (e.g., log-transformation or (square) root) by visual inspection and by application of a Kolmogorov–Smirnov test on normality ( $p < 0.05$ ).

In addition, we determined the correlation coefficients of all combinations of (transformed) predictor variables. If two variables highly correlated ( $r \geq 0.7$ ) and if there was a plausible physical explanation for their high correlation, we excluded one of them from the analysis to prevent collinearity.

### 2.2.2. Zero inflated regression

We correlated the number of observed cases, corrected for population density, to all remaining environmental variables. To deal with the large number of zero incidence rates in the inhabited grid cells, we used a zero inflated regression model (using R version 3.0.3 and package ‘pscl’) [52,53], which accounts for large amounts of true zeros (not exposed, so not infected) and accidental zeros (exposed but not infected).

The model consists of a mixture of two components: firstly, a ‘zero’ (binomial) model to distinguish between exposure and no exposure, i.e. the probability of observing  $\geq 1$  cases:

$$Y = \{Z > 0\} \sim \text{Bin}(1, p) \quad (2)$$

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n \quad (3)$$

where  $Z$  is the number of cases,  $Y$  is the binomial event that  $>0$  cases were notified (or the event that ‘exposure’ occurs), and  $X_i$  is the value of predictor variable  $i$ ;  $p$  is the binomial probability;

Secondly, a ‘count’ model is included to determine the correlation of predictors to the actual number of observed cases, given exposure. For the count model we applied a negative binomial model to allow for over-dispersion (extreme incidence rates):

$$Z|Y \sim \text{NegBin}(\lambda, \theta) \quad (4)$$

$$\lambda = n \cdot \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n). \quad (5)$$



With  $\lambda$  being the expected number of cases,  $n$  being the number of inhabitants per inhabited grid cell, and  $\theta$  being the dispersion parameter. If no fit could be found, a Poisson model was applied, which is basically a negative binomial distribution with  $\theta = 0$ :

$$Z|Y \sim \text{Pois}(\lambda) \quad (6)$$

We first performed univariate analyses at the month and year level to detect all correlating variables at the 80% confidence level ( $p < 0.20$ ). Secondly, we performed a multivariate analysis at the 95% confidence level ( $p < 0.05$ ) using the *step* function in R (package 'stats') for forward and backward model selection.

Subsequently, we retrieved the fitted values of the coefficients to create a geographical prediction map per month and for the full research period using the *predict.zeroinfl()* function in R. We extracted the model's Akaike's Information Criterion (AIC) that is a measure for the model's goodness of fit. We also extracted the differences in AIC ( $\Delta\text{AIC}$ ) of the model with respect to the same model in case each of the variables would be removed. The AIC-differences are a measure for the contribution of a variable to the prediction of Q fever incidence rate.

### 2.2.3. Validation

We applied validation of the results by

(1) visual inspection of the geographical observed, predicted, and residual incidence rate maps;

(2) performing a linear regression analysis of the log-transformed predicted and observed incidence rates for all grid cells with  $>0$  observed cases. We applied orthogonal linear regression and assumed  $0.3 < |r| < 0.6$  (Pearson's correlation coefficient) to correspond to a moderate, and  $|r| \geq 0.6$  to correspond to a strong linear relationship. In addition, we determined the 95% confidence intervals of the slope and intercept of the regression line by bootstrapping 10,000 times; and

(3) applying spatial cross-validation to the yearly averaged data. That is, we trained the model based on the annually averaged data with all variables that were significant ( $p < 0.05$ ) in the negative binomial model by randomly selecting 2/3 of the inhabited grid cells and applied that model to the remaining 1/3 of the inhabited grid cells. We repeated this process 1,000 times and calculated the mean predicted number of cases in each inhabited grid cell, both for the binomial and the negative binomial model. By applying cut off values (from 0.001 through 1.00 with steps of 0.001) to the binomial probability  $p$ , we created a plot of the sensitivity versus the specificity and extracted the area under the curve (AUC). We then retrieved the 'optimal cut-off' (i.e. the value corresponding to the maximum sum of the sensitivity and specificity). We then calculated the mean squared error (MSE) as a measure for the difference between the observed and predicted number of cases of these true positive grid cells.

### 2.2.4. Environmental hazard maps

Finally, we calculated an environmental hazard map using the fitted multivariate model, under absence of positive (*C. burnetii* emitting) farms and human cases. These hazard maps reflect the potential risk for an infection if *C. burnetii* would be emitted. They could give insight in the spatial variation in potential relative risks for transmission.

## 3. Results

### 3.1. Data preparation

#### 3.1.1. Transformation of variables

Supplementary Table 2 shows the transformation functions that we applied to the environmental data.

Supplementary Figs. 1–8 show geographical plots of the yearly and monthly averaged concentration, LAI, and soil moisture, and the aggregated land use and soil wind erosion sensitivity map.

#### 3.1.2. One to one regression analysis

From the one to one regression analysis we excluded the variables deposition, Enhanced Vegetation Index, and roughness length from any further analysis due to collinearity. The land-use fraction "agriculture" and the soil erosion sensitivity fraction "moderate" were considered as reference variables (in the intercept). See Supplementary Fig. 9 for the Pearson's regression coefficients ( $r$ ) of all combinations of predictor variables (yearly averaged).

#### 3.1.3. Univariate analyses

Supplementary Figs. 10 and 11 show the effects of the variables on the incidence rates in case of a one-unit increase of the (transformed) variable. Values larger respectively smaller than 1 correspond to positive respectively negative effects on the incidence rate.

- (1) *Model part 1: zero (binomial) model* (Supplementary Fig. 10): no clear, consistent correlations at the 80% confidence level are found. Concentration is never correlated to incidence rate. In general, values are very high or very low, indicating high uncertainties.
- (2) *Model part 2: count (negative binomial/Poisson) model* (Supplementary Fig. 11): concentration is positively correlated to incidence rate (mean effect of  $\pm 2.0$ ). The leaf area index is negatively correlated to incidence rate. ( $\pm 0.75$ .) Soil moisture shows inconsistent results with mostly very negative effects. High soil erosion sensitivity is positively correlated ( $\pm 1.9$ ). The land use fractions are alternately correlated, with 'water' and 'city' generally consistent.

We excluded variables not significant at the 80% confidence level (in case of white cells in both Supplementary Figs. 10 and 11) for the multivariate analysis.

### 3.2. Multivariate analysis

Table 1 shows the AIC-differences ( $\Delta\text{AIC}$ ) if each of the individual variables would be removed from the final multivariate model. These results show that concentration is by far the most important predictor for Q fever incidence rate. The second most important predictor is the leaf area index (both yearly and monthly data). For the yearly averaged data all variables had  $p < 0.05$ . For the monthly averaged data most correlations with  $p < 0.05$  are found during the months April and May (reflected by the highest number of human cases (Fig. 2)).

Fig. 4 shows the effects of the predictor variables on the incidence rate in case of a one-unit increase of the back-transformed predictor variables (count model, i.e. negative binomial/Poisson). It shows that concentration is positively correlated to observed incidence rate (mean effect of  $\pm 1.8$ ), and that leaf area index is negatively correlated ( $\pm 0.7$ ). The other variables did not always have  $p < 0.05$ : soil moisture, and the fraction land use 'open nature' are generally negatively correlated; high soil erosion sensitivity and the fraction land use 'forest' are positively correlated. High numbers refer to large standard errors. Supplementary Fig. 12 shows the effects of the predictor variables for the binomial model. Similar to the univariate analyses, these results are inconsistent and the standard errors are large.

### 3.3. Validation

At the 1 km<sup>2</sup> level, the observed incidence rates are highly heterogeneous, whereas the predicted incidence rates are much more smoothed (Supplementary Fig. 13). To reduce the occurrence of random spatial attribution, we aggregated the observed and predicted incidence rates to different geographical levels: 5 km<sup>2</sup> (Supplementary Fig. 14), 10 km<sup>2</sup> (Supplementary Fig. 15), and to the municipality level (Fig. 5). These figures are based on yearly averaged environmental data.

**Table 1**  
Difference in Akaike's Information Criterion ( $\Delta AIC$ ) if each of the individual variables would be removed from the multivariate model per month. A difference of approximately 3 points in AIC is considered to be significant. White cells correspond to non-significance in the zero-inflated model.

	YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Concentration	947	190	235	384	657	552	530	188	145	100	63	64	58
Leaf area index	186	61	49	135	60	81	118	98	78	66	69	28	66
Land use: open nature	57				65	26		7.2	5.2				
Soil moisture	49		19	3.1	55			16		23		6.0	21
Land use: forest	45	9.1			61	14		12	3.4		14		
Erosion sensitivity: low	40		22		22	12			6.1	6.7	3.0	27	19
Land use: city	36		27		46	7.4			22	28	5.5		15
Land use: water	36	5.2		38		9.2	19	24					
Erosion Sensitivity: high	34			13	49	12	14	6.8			18	3.9	

By visual inspection of Fig. 5, regions with the highest observed incidence rate (in the southeast of The Netherlands) are identified by the multivariate model, although the observed incidence rate numbers are smaller (negative residuals). Overestimation of the incidence rate is found around this high incidence rate area.

Fig. 6 shows Pearson's correlation coefficient ( $r$ ) and the slopes and intercepts of the orthogonal regression curves as function of time at four geographical aggregation levels. The correlation coefficient is highly influenced by the geographical aggregation level. Lowest values are found at the 1 km<sup>2</sup> level (0.3–0.4); at the other geo-aggregation levels the differences are smaller with values around  $r = 0.6$ . Supplementary Table 3 shows the values of Fig. 6 in a table, including the 95% confidence intervals of the slopes and intercepts.

The slopes of the regression curves are close to 1 in the yearly analysis and during the months with the highest number of cases (March – June), indicating that there is a positive, one-to-one, correlation between predicted and observed incidence rates. During the rest of the year and at the 1 km<sup>2</sup> aggregation level the values are generally much smaller. The intercepts are generally negative, but best values (i.e. close to zero) are found at the 10 km<sup>2</sup> and municipality level.

Finally, Supplementary Fig. 17 shows the curves of the sensitivity versus the specificity, resulting from the spatial cross-validation analysis (as a function of a cut-off value of 0.001...1.00 with steps of 0.001). The resulting area under the curve is equal to 0.453 (1 km resolution),

0.60 (5 km), 0.63 (10 km), and 0.73 (municipality level). The mean squared error (MSE) is equal to 18.9, 320.1, 1295, and 1227 respectively.

3.4. Environmental hazard maps

Fig. 7 shows a dimensionless environmental hazard map based on yearly averaged data. These hazards are related to the effect of the environment (deposition and reaerosolisation), under absence of any emissions. Regions with the highest hazards are located in the southeast, and are related to low vegetation density (Supplementary Fig. 3), low soil moisture levels (Supplementary Fig. 5), and high soil erosion sensitivity (Supplementary Fig. 8).

4. Discussion

4.1. Conclusions

We performed a spatiotemporal regression analysis to investigate the correlation between observed human incidence rates and modeled environmental risk factors during the large Q fever epidemic in The Netherlands in 2009. We conclude that *C. burnetii* concentration is positively correlated to incidence and that it is the best predictor variable for human Q fever, followed by the mean spatial vegetation density (Leaf Area Index) (negative). High soil erosion sensitivity fractions (positive), soil moisture (generally negative), and

(Intercept)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	# 3.8e+10 #	# 2.3e+09 #		0.0	
Concentration	1.8	2.5	2.1	1.5	1.9	1.8	1.8	1.8	# 1.3 #	# 1.7 #	1.6	1.5	# 1.6 #
Leaf Area Index	0.8	0.4	0.8	0.7	0.9	0.8	0.8	0.8	# 0.6 #	# 0.6 #	0.6	0.6	# 0.5 #
Soil moisture	0.1		0.0		118.3			0.0				0.0	# 6.1e+03 #
Soil Erosion Sensitivity: High	1.3			1.4	1.4	1.2	1.3						
Soil Erosion Sensitivity: Low			6.9						# 0.0 #	# 0.0 #		143.7	# 0.0 #
Land use fraction: forest	1.6				1.8			2.0			6.5		
Land use fraction: nature	0.7			0.7	0.6	0.6			# 0.4 #				
Land use fraction: water				1.5									
Land use fraction: city	1.8				3.7	1.6			# 0.0 #	# 0.0 #	0.1		# 0.1 #
	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

**Fig. 4.** Mean effects of the back-transformed predictor variables in the multivariate zero-inflated negative binomial or Poisson regression analysis. All values were the result of a negative binomial regression analysis, except for those indicated by hash tags (#) (Poisson). White cells refer to variables with  $p \geq 0.05$ ; gray cells refer to variables excluded from the analysis (since  $p \geq 0.20$  in the univariate analysis). Values represent the effect of the predictor variable in case of a one-unit increase. The results of the binomial part of the regression analysis are shown in Supplementary Fig. 12.

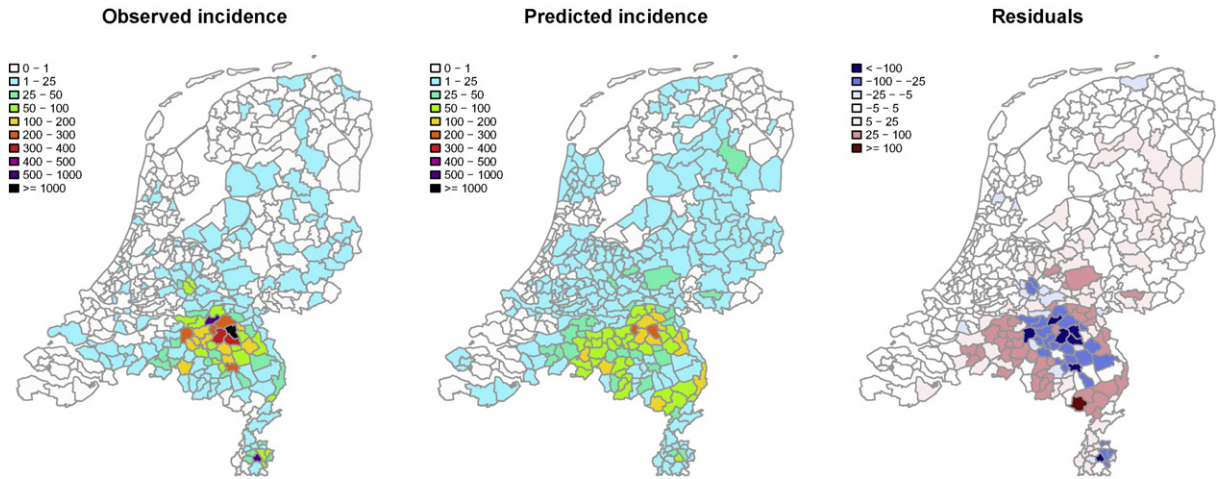


Fig. 5. (Left, Middle) Observed and predicted incidence rate per 100,000 inhabitants, aggregated to the municipality level; (Right) Predicted incidence rate minus the observed incidence rate per 100,000 inhabitants per municipality.

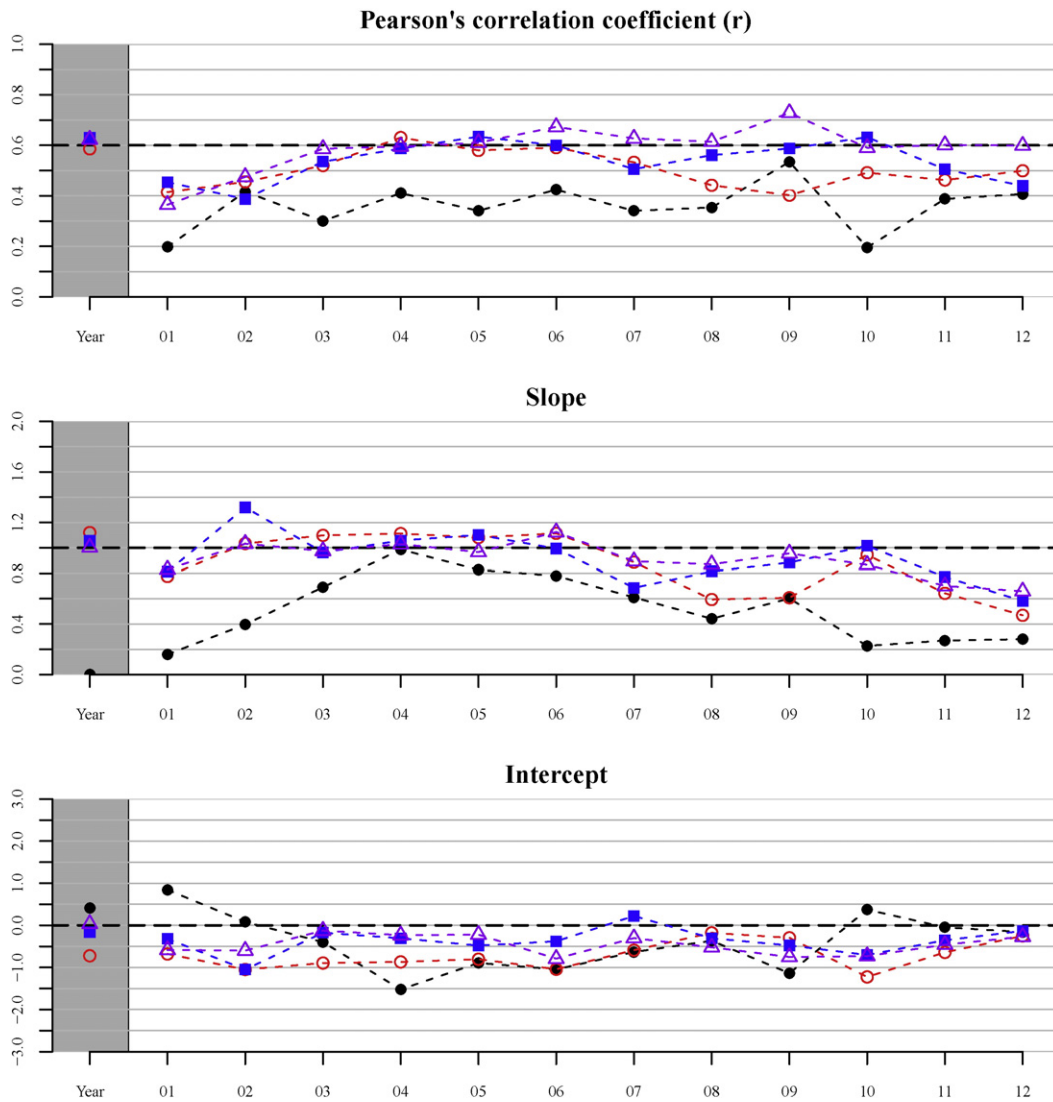


Fig. 6. Statistics for four geo-aggregation levels (1 km<sup>2</sup>, 5 km<sup>2</sup>, 10 km<sup>2</sup>, and municipality level) for yearly and monthly averaged data: Pearson's correlation coefficient (r), and slopes and intercepts of the linear regression line. Most ideal values for slope and intercept, and the minimum level for a 'strong' correlation (r = 0.60) are indicated by a dashed horizontal line. Symbols: (•) 1 km<sup>2</sup> level, (○) 5 km<sup>2</sup> level, (■) 10 km<sup>2</sup> level, and (△) municipality level.

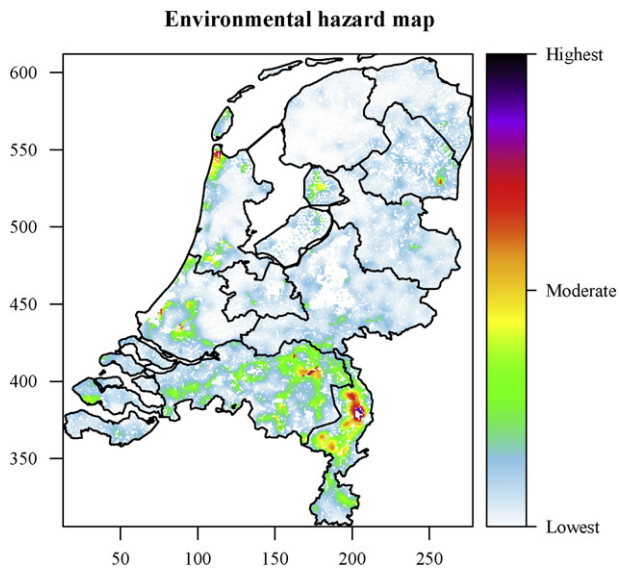


Fig. 7. Environmental hazard map for yearly averaged data at the 1 km<sup>2</sup> level.

the land use classes 'forest' (positive) and open nature (negative) had  $p < 0.05$  as well, but less frequent in time. Low soil erosion sensitivity and the land use fractions 'water' and 'city' did generally not have  $p < 0.05$  or at least not consistently. The effects of the land use classes 'forest' and 'open nature' on incidence are counter-intuitive, which is caused by the coincidental presence of large forest areas and absence of large open nature areas in the major Q fever outbreak area. Best results are found for the 2009-averaged data and the monthly averaged data of March, April and May – which corresponds to the large number of notified human cases.

In other words, airborne transmission is the most powerful predictor for observed human Q fever incidence rates. The effect of the environment on the transmission (deposition and reaerosolisation, mainly covered by LAI, but also by soil moisture, soil erosion sensitivity and the land use class fractions) is also of importance, but to a lesser degree. In general, our results are in line with those of the previous pilot study [27], where it was concluded that vegetation, soil moisture and arable land were risk factors for *C. burnetii* exposure.

#### 4.2. Research context

Our results are consistent with previously published literature on the physical effects of meteorological conditions (e.g., atmospheric stability, wind speed/direction, temperature, and precipitation) on survival and transmission of airborne pathogens [22,44,54–56], and the effects of vegetation, objects, and surface roughness on deposition and reaerosolisation [35]. For instance, precipitation increases the amount of wet deposition, while a smoother surface results in further spread. For the Dutch situation, this means that most exposure could be expected in regions with little vegetation, dry conditions, and arable land. Despite the fact that presence of *C. burnetii* in outdoor environments has been demonstrated [10,11,57,58], studies concerning the possible effects of vegetation and soil conditions on airborne pathogens do exist [59–62], but are scarce. The study of [63] is comparable to our study, although it does not include an atmospheric dispersion model.

#### 4.3. Uncertainties

Uncertainties in the predictor data will be relatively small, for they are environmental data that are systematically measured at high spatial and temporal resolutions by official institutes (e.g., Royal Netherlands Meteorological Institute, NASA, Wageningen University & Research Centre). Uncertainties in the human and farm data, however, will be

much larger: human Q fever data do not include non-notified infected persons. With respect to the farm data, a systematic and mandatory test was not implemented before September 2009.

Fig. 5 showed that the residual incidence rates (predicted minus observed) were spatially clustered. In the area with the highest Q fever incidence rates the observations are underestimated; in other areas overestimations occurred. The total number of predicted cases is however almost identical to the total number of observed cases (data not shown). However, the curves of the sensitivity versus the specificity (based on the binomial model) might be further improved (Suppl. Fig. 17). Possible explanations might be that:

- 1) The incidence rates are extremely peaked: for the yearly averaged data and during springtime the variance was much larger than the mean indicating that a zero inflated Poisson model was not applicable. Instead, we applied a zero inflated negative binomial model to account for extreme values. Nevertheless, this could not prevent underestimations still to occur.
- 2) Large differences in emissions between farms might have occurred (as also suggested by [64] in the 2006-Bluetongue outbreak in northwest Europe). This includes that large shedding animals might have been present at the positive farms, possibly represented by 'abortion farms': the 27 farms at which abortion waves have occurred during the years 2007–2009 [37] (Fig. 1). Since, abortion waves are a symptomatic indication for the presence of *C. burnetii* [3–5], 'abortion farms' were associated to high emission strengths of *C. burnetii* [2]. However, their geographical location does not always coincide to high-observed incidence rates (Fig. 4A).
- 3) An over-simplified emission profile: judging from the epidemic curve in Fig. 1, a steady-state emission profile is hard to defend. Although it was previously shown that such a steady-state profile fitted better to the human data than a lognormal temporal emission profile [21], the current study included multiple sources. That is, several lognormal emission profiles from multiple sources could have led to a more temporarily smoothed exposure curve.

Although environmental conditions could either have a positive or negative effect on the transmission of *C. burnetii*, the source term remains a very important variable. Thus, further investigations on the source strength and its temporal behavior is recommended.

Also, this study does not necessarily reflect direct causal relationship between the outcome (incidence rates) and its predictors. However, detected relationships are supported by literature [35], augmenting the probability of a true, direct, physical correlation as well as the confidence in the hazard maps.

#### 4.4. Binomial and count model

Strictly speaking, the population is divided into two groups in case of a large-scale outbreak: those who were exposed and those who were not. Non-exposed humans cannot have been infected. Those exposed could either be infected or not be infected; those infected could either developed clinical symptoms or not. Since the majority of the people were not infected [18,19] and a large amount of them were not exposed as well (Supplementary Fig. 2), we applied a zero inflated regression analysis. After all, zeros could either be explained by the absence of exposure, or by the probability of non-infection or asymptomatic infection despite exposure [53].

This might be an explanation for the variable 'concentration' to have  $p \geq 0.05$  in the binomial (zero) part of the univariate and multivariate analysis (Supplementary Figs. 10 and 12), whereas it had  $p \ll 0.05$  in the negative binomial/Poisson (count) part (Supplementary Figs. 11 and 4). After all, non-zero concentrations were modeled in many inhabited grid cells (Supplementary Figs. 1 and 2), whereas in many of these grid cells the observed incidence rates were equal to zero: that is, the specificity in a dichotomous situation is relatively low. There is, however, a clear relationship between the modeled



concentration and the number of observed cases (count model). Note that the effects of each variable represent mean effects, as the values were back-transformed to original values that were not normally distributed.

#### 4.5. Time-dependency

The correlation analysis in the current study was performed quasi-temporally. That means that we performed a regression analysis for each month separately and for the year 2009. The reason for that is that it is complicated to include the variable 'time'. For example, 'temperature' is a variable that is highest values during the summer; if now the emission strengths were also highest in summer, there still would not be a causal relationship yet. High temperatures and high emission strengths happen to occur at the same time; nevertheless, conditions could be different compared to climatological averages, resulting in a possible positive or negative effect.

'Soil moisture' is a typical variable for which its effect should be tested in a time-dependent analysis. In the current study design, only its spatial differences are included. That is, we included the spatial heterogeneity of the variable into the model, but not the difference with respect to other months or other years, thereby excluding the effect of a dry season. In a multivariate statistical model the absolute values of a variable are not as important as the deviation of individual values with respect to their common mean.

If a time-dependent analysis would be performed (comparable to [55]), it would be advisable to focus on specific outbreak areas (e.g., those of [30]) to exclude spatial variations and to be able to compare the effects in different areas. A Fourier-transformation applied to the temporal variables would be one of the possibilities [65].

#### 4.6. Hazard map

The hazard map should be interpreted prudently. It shows the spatial potential risk for a *C. burnetii* infection modulated by environmental factors. That is, neither meteorological conditions nor emission strengths were included, but the possible effects of spatial variations in vegetation density, land use, soil moisture and soil erosion sensitivity were.

Based on this study, the highest hazards were found in the south-eastern Netherlands, with additional hotspots in the northeast and the coastal areas. These maxima were in general correlated to relatively low vegetation densities, dry soil conditions and high soil erosion sensitivities.

The Dutch livestock industry is, however, highly concentrated in the southeast of The Netherlands, partly due to the relatively poor arable farming conditions there (oligotrophic soils and dry soil conditions). The combination of high livestock densities and 'high' hazards for spatial spread of airborne pathogens as a function of environmental conditions could result in enhanced human incidence rates in case of an outbreak among animals.

#### 4.7. Recommendations

In particular, we would recommend focusing on the uncertainties in the human and farm data for future research. As noticed, the human data set included only those who passed through four stages: (1) symptomatic infection; (2) occurrence of disease; (3) seeking medical care; and (4) being diagnosed as Q fever case. All other infected persons were not included in the data set (this group is estimated to be 92% of all infected humans [17]). Serological data could be used as well, but only if they cover a large geographical area.

Next to the human data, the source data is neither complete, despite all efforts. We would recommend formulating a protocol in case of a future veterinarian outbreak concerning zoonotic pathogens with an airborne transmission route. Such a protocol could include criteria on

(intensive) air sampling within and outside infected stables. In case of *Coxiella burnetii*, it could include criteria on vaginal sampling and tank milk monitoring and their temporal dynamics as well. That way, more effort could be invested in determining time-dependent emission rates for *C. burnetii*. Such, more detailed modeled concentrations could be calculated (with more realistic between-farm variations) and a subsequent exposure assessment could be elaborated.

In addition, similar research with data from other outbreaks would be needed to more firmly establish our findings. This could lead to better estimations of the public health risk of a *C. burnetii* outbreak, and to more detailed and accurate hazard maps that could be used for spatial planning of livestock operations.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.onehlt.2016.03.004>.

#### Acknowledgements

We thank Dr. P. Vellema and R. Van den Brom DVM (Animal Health Service) for providing data on abortion waves and bulk tank milk data of 2008 and 2009. We thank W. Boasson, M.Sc. (Agricultural University of Applied Sciences (HAS), 's-Hertogenbosch, The Netherlands) for his PostgreSQL and PostGIS advises. Also, we thank F. De Vries, B.Sc. (Wageningen University, Wageningen, The Netherlands) for providing data on soil erosion sensitivity, Dr. A. De Bruin (National Institute for Public Health and the Environment) for providing the vaginal swab data, and B. Bom (National Institute for Public Health and the Environment) for aggregation of the land use map.

This work was supported by the National Institute for Public Health and the Environment [project S/210106/01/RQ] ([www.rivm.nl](http://www.rivm.nl)) and ZonMW [project 205520010] ([www.zonmw.nl](http://www.zonmw.nl)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### References

- [1] F. Dijkstra, W. Van der Hoek, N. Wijers, B. Schimmer, A. Rietveld, C.J. Wijkman, et al., The 2007–2010 Q fever epidemic in The Netherlands: characteristics of notified acute Q fever patients and the association with dairy goat farming, *FEMS Immunol. Med. Microbiol.* 64 (2012) 3–12, <http://dx.doi.org/10.1111/j.1574-695X.2011.00876.x>.
- [2] H.J.J. Roest, J.J.H.C. Tilburg, W. Van der Hoek, P. Vellema, F.G. Van Zijderveld, C.H.W. Klaassen, et al., The Q fever epidemic in The Netherlands: history, onset, response and reflection, *Epidemiol. Infect.* 139 (2011) 1–12, <http://dx.doi.org/10.1017/S0950268810002268>.
- [3] E. Rousset, M. Berri, B. Durand, P. Dufour, M. Prigent, T. Delcroix, et al., *Coxiella burnetii* shedding routes and antibody response after outbreaks of Q fever-induced abortion in dairy goat herds, *Appl. Environ. Microbiol.* 75 (2009) 428–433, <http://dx.doi.org/10.1128/AEM.00690-08>.
- [4] M. Berri, A. Souriau, M. Crosby, A. Rodolakis, Shedding of *Coxiella burnetii* in ewes in two pregnancies following an episode of *Coxiella* abortion in a sheep flock, *Vet. Microbiol.* 85 (2002) 55–60, [http://dx.doi.org/10.1016/S0378-1135\(01\)00480-1](http://dx.doi.org/10.1016/S0378-1135(01)00480-1).
- [5] H.-J. Roest, B. van Gelderen, A. Dinkla, D. Frangoulidis, F. van Zijderveld, J. Rebel, et al., Q fever in pregnant goats: pathogenesis and excretion of *Coxiella burnetii*, *PLoS One* 7 (2012), <http://dx.doi.org/10.1371/journal.pone.0048949>.
- [6] N. Arricau Bouvery, A. Souriau, P. Lechopier, A. Rodolakis, Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes, *Vet. Res.* 34 (2003) 423–433, <http://dx.doi.org/10.1051/vetres:2003017>.
- [7] S.G. Kim, E.H. Kim, C.J. Lafferty, E. Dubovi, *Coxiella burnetii* in bulk tank milk samples, United States, *Emerg. Infect. Dis.* 11 (2005) 619–621.
- [8] R. Van den Brom, E. Van Engelen, S. Luttikholt, L. Moll, K. Van Maanen, P. Vellema, *Coxiella burnetii* in bulk tank milk samples from dairy goat and dairy sheep farms in The Netherlands in 2008, *Vet. Rec.* 170 (2012) 310, <http://dx.doi.org/10.1136/vr.100304>.
- [9] A. Rodolakis, M. Berri, C. Héchar, C. Caudron, A. Souriau, C.C. Bodier, et al., Comparison of *Coxiella burnetii* shedding in milk of dairy bovine, caprine, and ovine herds, *J. Dairy Sci.* 90 (2007) 5352–5360, <http://dx.doi.org/10.3168/jds.2006-815>.
- [10] L. Hogerwerf, F. Borlée, K. Still, D. Heederik, B. Van Rotterdam, A. de Bruin, et al., Detection of *Coxiella burnetii* DNA in inhalable airborne dust samples from goat farms after mandatory culling, *Appl. Environ. Microbiol.* 78 (2012) 5410–5412, <http://dx.doi.org/10.1128/AEM.00677-12>.
- [11] A. De Bruin, I. Janse, M. Koning, L. De Heer, R.Q.J. Van der Plaats, J.P.G. Van Leuken, et al., Detection of *Coxiella burnetii* DNA in the environment during and after a large Q fever epidemic in The Netherlands, *J. Appl. Microbiol.* 114 (2013) 1395–1404, <http://dx.doi.org/10.1111/jam.12163>.

- [12] T.F. McCaul, J.C. Williams, Developmental cycle of *Coxiella burnetii*: structure and morphogenesis of vegetative and sporogenic differentiations, *J. Bacteriol.* 147 (1981) 1063–1076.
- [13] M. Maurin, D. Raoult, Q fever, *Clin. Microbiol. Rev.* 12 (1999) 518–553.
- [14] EFSA, Scientific Opinion on Q fever, EFSA J. (2010) 8 (Available: <http://www.efsa.europa.eu/de/scdocs/doc/1595.pdf>).
- [15] R.J. Brooke, M.E.E. Kretzschmar, N.T. Mutters, P.F. Teunis, Human dose response relation for airborne exposure to *Coxiella burnetii*, *BMC Infect. Dis.* 13 (2013) 488, <http://dx.doi.org/10.1186/1471-2334-13-488>.
- [16] B. Schimmer, D.W. Notermans, M.G. Harms, J.H.J. Reimerink, J. Bakker, P. Schneeberger, et al., Low seroprevalence of Q fever in The Netherlands prior to a series of large outbreaks, *Epidemiol. Infect.* 140 (2012) 27–35, <http://dx.doi.org/10.1017/S0950268811000136>.
- [17] W. Van der Hoek, B.M. Hogema, F. Dijkstra, A. Rietveld, C.J. Wijkmans, P.M. Schneeberger, et al., Relation between Q fever notifications and *Coxiella burnetii* infections during the 2009 outbreak in The Netherlands, *Eurosurveillance* 17 (2012).
- [18] B.M. Hogema, E. Slot, M. Molier, P.M. Schneeberger, M.H. Hermans, E.J. Van Hanne, et al., *Coxiella burnetii* infection among blood donors during the 2009 Q-fever outbreak in The Netherlands, *Transfusion* 52 (2012) 144–150, <http://dx.doi.org/10.1111/j.1537-2995.2011.03250.x>.
- [19] L.M. Kampschreur, J.C.J.P. Hagenaars, C.C.H. Wielders, P. Elsman, P.J. Lestrade, O.H.J. Koning, et al., Screening for *Coxiella burnetii* seroprevalence in chronic Q fever high-risk groups reveals the magnitude of the Dutch Q fever outbreak, *Epidemiol. Infect.* 141 (2013) 847–851, <http://dx.doi.org/10.1017/S0950268812001203>.
- [20] B. Schimmer, S. Luttikholt, J.L.A. Hautvast, E.A.M. Graat, P. Vellema, Van Duynhoven YTHP. Seroprevalence and risk factors of Q fever in goats on commercial dairy goat farms in The Netherlands, *BMC Vet. Res.* 7 (2011) 7, <http://dx.doi.org/10.1186/1746-6148-7-81>.
- [21] J. Van Leuken, J. Van de Kastele, F.J. Sauter, W. Van der Hoek, D. Heederik, A. Havelaar, et al., Improved correlation of human Q fever incidence to modelled *C. burnetii* concentrations by means of an atmospheric dispersion model, *Int. J. Health Geogr.* 14 (2015), <http://dx.doi.org/10.1186/s12942-015-0003-y>.
- [22] A.M. Jones, R.M. Harrison, The effects of meteorological factors on atmospheric bioaerosol concentrations — a review, *Sci. Total Environ.* 326 (2004) 151–180, <http://dx.doi.org/10.1016/j.scitotenv.2003.11.021>.
- [23] H. Tissot-Dupont, M. Amadei, M. Nezzi, D. Raoult, Wind in November, Q fever in December, *Emerg. Infect. Dis.* 10 (2004) 1264–1269.
- [24] R.M. Harrison, A.M. Jones, P.D.E. Biggins, N. Pomeroy, C.S. Cox, S.P. Kidd, et al., Climate factors influencing bacterial count in background air samples, *Int. J. Biometeorol.* 49 (2005) 167–178, <http://dx.doi.org/10.1007/s00484-004-0225-3>.
- [25] G.D. Hess, M.G. Garner, X. Yang, A sensitivity analysis of an Integrated modelling approach to assess the risk of wind-borne spread of Foot-and-mouth disease Virus from infected premises, *Environ. Model. Assess.* 13 (2008) 209–220, <http://dx.doi.org/10.1007/s10666-007-9097-3>.
- [26] K. Nygård, Ø. Werner-Johansen, S. Rønnes, D.a. Caugant, Ø. Simonsen, A. Kanestrøm, et al., An outbreak of legionnaires disease caused by long-distance spread from an industrial air scrubber in Sarpsborg, Norway, *Clin. Infect. Dis.* 46 (2008) 61–69, <http://dx.doi.org/10.1086/524016>.
- [27] W. Van der Hoek, J. Hunink, P. Vellema, W. Droogers, Q fever in The Netherlands: the role of local environmental conditions, *Int. J. Environ. Health Res.* (2011), <http://dx.doi.org/10.1080/09603123.2011.574270>.
- [28] A. Ssematimba, T.J. Hagenaars, M.C.M. De Jong, Modelling the wind-borne spread of highly pathogenic Avian influenza virus between farms, *PLoS One* 7 (2012), <http://dx.doi.org/10.1371/journal.pone.0031114>.
- [29] J.M. Puijol, J.E. Eisenberg, C.N. Haas, J.S. Koopman, The effect of ongoing exposure dynamics in dose response relationships, *PLoS Comput. Biol.* 5 (2009), <http://dx.doi.org/10.1371/journal.pcbi.1000399>.
- [30] J.P.G. Van Leuken, A.H. Havelaar, W. Van der Hoek, G.A.F. Ladbury, V.H. Hackert, A.N. Swart, A model for the early identification of sources of airborne pathogens in an outdoor environment, *PLoS One* 8 (2013) 1–9, <http://dx.doi.org/10.1371/journal.pone.0080412>.
- [31] M. Nicas, A.E. Hubbard, R.M. Jones, A.L. Reingold, The infectious dose of Variola (smallpox) virus, *Appl. Biosaf.* 9 (2004) 118–127.
- [32] N. Pica, N.M. Bouvier, Environmental factors affecting the transmission of respiratory viruses, *Curr. Opin. Virol.* Elsevier B.V. 2 (2012) 90–95, <http://dx.doi.org/10.1016/j.coviro.2011.12.003>.
- [33] M.L. Wesely, B.B. Hicks, A review of the current status of knowledge on dry deposition, *Atmos. Environ.* 34 (2000) 2261–2282.
- [34] L. Zhang, S. Gong, J. Padro, L. Barrie, A size-segregated particle dry deposition scheme for an atmospheric aerosol module, *Atmos. Environ.* 35 (2001) 549–560.
- [35] Y. Shao, Physics and modelling of wind erosion [Internet], Springer, Köln, Germany, 2008 (Available: <http://www.springerlink.com/index/10.1007/978-1-4020-8895-7>).
- [36] J.A. Layschock, B. Pearson, K. Crockett, M.J. Brown, S. Van Cuyk, W.B. Daniel, et al., Reaerosolization of *Bacillus* spp. in outdoor environments: a review of the experimental literature, *Biosecur. Biodefense Strateg. Pract. Sci.* 10 (2012) 299–303, <http://dx.doi.org/10.1089/bsp.2012.0026>.
- [37] R. Van den Brom, P. Vellema, Q fever outbreaks in small ruminants and people in The Netherlands, *Small Rumin. Res.* 86 (2009) 74–79, <http://dx.doi.org/10.1016/j.smallrumres.2009.09.022>.
- [38] K. Porten, J. Rissland, A. Tigges, S. Broll, W. Hopp, M. Lunemann, et al., A super-spreading ewe infects hundreds with Q fever at a farmers' market in Germany, *BMC Infect. Dis.* 6 (2006), <http://dx.doi.org/10.1186/1471-2334-6-147>.
- [39] A. De Bruin, R.Q.J. Van der Plaats, L. De Heer, R. Paauwe, B. Schimmer, P. Vellema, et al., Detection of *Coxiella burnetii* DNA on small-ruminant farms during a Q fever outbreak in The Netherlands, *Appl. Environ. Microbiol.* 78 (2012) 1652–1657, <http://dx.doi.org/10.1128/AEM.07323-11>.
- [40] A. De Bruin, A. De Groot, L. De Heer, J. Bok, P.R. Wielinga, M. Hamans, et al., Detection of *Coxiella burnetii* in complex matrices by using multiplex quantitative PCR during a major Q fever outbreak in The Netherlands, *Appl. Environ. Microbiol.* 77 (2011) 6516–6523, <http://dx.doi.org/10.1128/AEM.05097-11>.
- [41] J.A. Van Jaarsveld, The operational priority substances Model, Description and Validation of OPS-Pro 4.1. [Internet], National Institute for Public Health and the Environment, Bilthoven, The Netherlands, 2004 (Report No.: 500045001/2004. Available: <http://www.rivm.nl/dsresource?type=pdf&disposition=inline&objectid=rivmp:249101&versionid=&subobjectname=>).
- [42] J.A. Van Jaarsveld, D. Klimov, Modelling the impact of sea-salt particles on the exceedances of daily PM10 air quality standards in The Netherlands, *Int. J. Environ. Pollut.* 44 (2011) 217–225.
- [43] E. Van der Swaluw, W.A.H. Asman, H. Van Jaarsveld, R. Hoogerbrugge, Wet deposition of ammonium, nitrate and sulfate in The Netherlands over the period 1992–2008, *Atmos. Environ.* 45 (2011) 3819–3826, <http://dx.doi.org/10.1016/j.atmosenv.2011.04.017>.
- [44] J.P.G. Van Leuken, A.N. Swart, A.H. Havelaar, A. Van Pul, W. Van der Hoek, D. Heederik, Atmospheric dispersion modelling of bioaerosols that are pathogenic to humans and livestock – a review to inform risk assessment studies, *Microbiol. Risk Anal. Elsevier Ltd.* 000 (2015) 1–21, <http://dx.doi.org/10.1016/j.mran.2015.07.002>.
- [45] A.F. Tarel, R. Guatteo, A. Joly, F. Beaudou, Relationship between the level of antibodies in bulk tank milk and the within-herd seroprevalence of *Coxiella burnetii* in cows, *Epidemiol. Infect.* 140 (2012) 1710–1713, <http://dx.doi.org/10.1017/S0950268811000275>.
- [46] W. Terink, P. Droogers, W. Immerzeel, G. Van den Eertwegh, SPHY – EEN hydrologisch model gericht op de berekening van bodemvocht in de wortelzone en de actuele verdamping [Internet], Future Water, Wageningen, The Netherlands, 2012 (Report No.: 115. Available: [http://www.futurewater.nl/wp-content/uploads/2012/08/Rapport\\_SPHY\\_v04.pdf](http://www.futurewater.nl/wp-content/uploads/2012/08/Rapport_SPHY_v04.pdf)).
- [47] USGS, US Geological Survey, Land Processes Distributed Active Archive Center [Internet], 2013 (Available: [https://lpdaac.usgs.gov/products/modis\\_products\\_table/mod13q1](https://lpdaac.usgs.gov/products/modis_products_table/mod13q1)).
- [48] USGS, US Geology Survey, Land Processes Distributed Active Archive Center [Internet], 2013 (Available: [https://lpdaac.usgs.gov/products/modis\\_products\\_table/mod15a2](https://lpdaac.usgs.gov/products/modis_products_table/mod15a2)).
- [49] M.J.D. Hack-ten Broeke, C.L. Van Beek, T. Hoogland, M. Knotters, J.P. Mol-Dijkstra, R.L.M. Schils, et al., Kaderrichtlijn Bodem [Internet], Wageningen University and Research Centre, Wageningen: Alterra, 2009 (Report No.: 2007. Available: <http://edepot.wur.nl/137524>).
- [50] G.W. Hazeu, C. Schuiling, G.J. Dorland, J. Oldengarm, H.A. Gijsbertse, Landelijk Grondgebruiksbestand Nederland versie 6 (LGN6) [Internet], Alterra, Wageningen University and Research Centre, Wageningen, The Netherlands, 2012 (Report No.: 2012. Available: <http://edepot.wur.nl/137531>).
- [51] B. Schimmer, R. Ter Schegget, M. Wegdam, L. Züchner, A. De Bruin, P.M. Schneeberger, et al., 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) 1–7, <http://dx.doi.org/10.1186/1471-2334-10-69>.
- [52] A. Zeileis, C. Kleiber, S. Jackman, Regression models for count data in R, CRAN [Internet], 2014 (Available: <http://cran.r-project.org/web/packages/pscl/vignettes/countreg.pdf>).
- [53] M.C. Hu, M. Pavlicova, E.V. Nunes, Zero-inflated and Hurdle models of count data with Extra zeros: examples from an HIV-Risk Reduction Intervention Trial, *Am. J. Drug Alcohol Abuse* 37 (2011) 367–375, <http://dx.doi.org/10.3109/00952990.2011.597280>.
- [54] R.S. Dungan, Board-invited review: fate and transport of bioaerosols associated with livestock operations and manures, *J. Anim. Sci.* 88 (2010) 3693–3706, <http://dx.doi.org/10.2527/jas.2010-3094>.
- [55] S.M. Firestone, N. Cogger, M.P. Ward, J.-A.L.M.L. Toribio, B.J. Moloney, N.K. Dhand, The influence of meteorology on the spread of influenza: survival analysis of an equine influenza (A/H3N8) outbreak, *PLoS One* 7 (2012), <http://dx.doi.org/10.1371/journal.pone.0035284>.
- [56] C.E. Dunn, B. Rowlingson, R.S. Bhopal, P. Diggle, Meteorological conditions and incidence of Legionnaires' disease in Glasgow, Scotland: application of statistical modelling, *Epidemiol. Infect.* 141 (2013) 687–696, <http://dx.doi.org/10.1017/S095026881200101X>.
- [57] G.J. Kersh, T.M. Wolfe, K.A. Fitzpatrick, A.J. Candee, L.D. Oliver, N.E. Patterson, et al., Presence of *Coxiella burnetii* DNA in the environment of the United States, 2006 to 2008, *Appl. Environ. Microbiol.* 76 (2010) 4469–4475, <http://dx.doi.org/10.1128/AEM.00042-10>.
- [58] W.M. Bamberg, W.J. Pape, J.L. Beebe, C. Nevin-Woods, W. Ray, H. Maguire, et al., Outbreak of Q fever associated with a horse-boarding ranch, Colorado, 2005, *Vector Borne Zoonotic Dis.* 7 (2007) 394–402, <http://dx.doi.org/10.1089/vbz.2007.0104>.
- [59] J.P. Brooks, B.D. Tanner, C.P. Gerba, C.N. Haas, I.L. Pepper, Estimation of bioaerosol risk of infection to residents adjacent to a land applied biosolids site using an empirically derived transport model, *J. Appl. Microbiol.* 98 (2005) 397–405, <http://dx.doi.org/10.1111/j.1365-2672.2004.02484.x>.
- [60] R.C. Baptista-Rosas, J. Catalán-Dibene, A.L. Romero-Olivares, A. Hinojosa, T. Cavazos, M. Riquelme, Molecular detection of *Coccidioides* spp. from environmental samples in Baja California: linking Valley fever to soil and climate conditions, *Fungal Ecol.* 5 (2012) 177–190, <http://dx.doi.org/10.1016/j.funeco.2011.08.004>.
- [61] J. Lindemann, C.D. Upper, Aerial dispersal of epiphytic bacteria over bean plants, *Appl. Environ. Microbiol.* 50 (1985) 1229–1232.
- [62] R.M. Bowers, S. McLetchie, R. Knight, N. Fierer, Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments, *ISME J.* 5 (2011) 601–612, <http://dx.doi.org/10.1038/ismej.2010.167>.

- [63] C. Faes, Y. Van der Stede, H. Guis, C. Staubach, E. Ducheyne, G. Hendrickx, et al., Factors affecting bluetongue serotype 8 spread in Northern Europe in 2006: the geographical epidemiology, *Prev. Vet. Med.* Elsevier B.V. 110 (2013) 149–158, <http://dx.doi.org/10.1016/j.prevetmed.2012.11.026>.
- [64] L. Sedda, H.E. Brown, B.V. Purse, L. Burgin, J. Gloster, D.J. Rogers, A new algorithm quantifies the roles of wind and midge flight activity in the bluetongue epizootic in Northwest Europe, *Proc. R. Soc. B* 279 (2012) 2354–2362, <http://dx.doi.org/10.1098/rspb.2011.2555>.
- [65] A. Swart, A. Ibáñez-Justicia, J. Buijs, S.E. Van Wieren, T.R. Hofmeester, H. Sprong, et al., Predicting tick presence by environmental risk mapping, *Front Public Health* 2 (2014) 1–8, <http://dx.doi.org/10.3389/fpubh.2014.00238>.