



HAL
open science

Monitoring the risk of Legionella infection using a general Bayesian network updated from temporal measurements in agricultural irrigation with reclaimed wastewater

Gaspar Massiot, Dominique Courault, Pauline Jacob, Isabelle Albert

► To cite this version:

Gaspar Massiot, Dominique Courault, Pauline Jacob, Isabelle Albert. Monitoring the risk of Legionella infection using a general Bayesian network updated from temporal measurements in agricultural irrigation with reclaimed wastewater. *Environmental Science: Water Research and Technology*, 2023, 9 (1), pp.176-192. 10.1039/D2EW00311B . hal-03930047v2

HAL Id: hal-03930047

<https://hal.science/hal-03930047v2>

Submitted on 3 Dec 2024

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.

Cite this: DOI: 00.0000/xxxxxxxxxx
Post-Print

Monitoring the risk of *Legionella* infection using general Bayesian network updated from temporal measurements in agricultural irrigation with reclaimed wastewater

Gaspar Massiot^{*a,b}, Dominique Courault^c, Pauline Jacob^d, Isabelle Albert^b

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Reuse of reclaimed wastewater for agricultural irrigation is an expanding practice worldwide. This practice needs to be monitored, partly because of pathogens that water may contain after treatments. More particularly, sprinkler irrigation is known to generate aerosols which may lead to severe health risks on the population close to irrigated areas in case of presence of *Legionella* bacteria in the water. A pilot experiment was conducted on two corn fields in South-Western France, irrigated with wastewater undergoing two different water treatments (ultra-filtration and UV). Water analyses have shown high levels of *Legionella* in the water even after standard WasteWater Treatment Plant (WWTP) cleaning process followed by the UV treatment (up to 10^6 GC/L in 2019). In this context, an updated General Bayesian Network (GBN), using discrete and continuous random variables, in Quantitative Microbial Risk Assessment (QMRA) is proposed to monitor the risk of *Legionella* infection in the vicinity of the irrigated plots. The model's originality is based on i) a graphical probabilistic model that describes the exposure pathway of *Legionella* from the WWTP to the population using observed and non-observed variables and ii) the model inference updating at each new available measurement. Different scenarios are simulated according to the exposition time of the persons, taking into account various distances from the emission source and a large dataset of climatic data. From the learning process included in the Bayesian principle, quantities of interest (contaminations before and after water treatments, inhaled dose, probabilities of infection) can be quantified with their uncertainty before and after the inclusion of each new data collected *in situ*. This approach gives a rigorous tool that allows to monitor the risks, facilitates discussions with reuse experts and progressively reduces uncertainty quantification through field data accumulation. For the two pilot treatments analyzed in this study, median annual risk of *Legionella* infection did not exceed the US EPA annual infection benchmark of 10^{-4} for any of the population at risk during the past months of the pilot experiment (DALYs are estimated up to 10^{-5}). The risk still bears watching with support from the method shown in this work.

1 Introduction

In the context of climate change, growing global water scarcity and more severe and frequent drought events have intensified the need to find alternative water resources^{1,2}. Reclaimed wastewater

is a renewable resource from which renewable energy can be produced, but also materials and water can be recovered^{3,4}. It is a possible alternative supply to address such water shortages⁵. The combination of climate change, population growth and urbanisation has led the European Parliament to foster increased water reuse to alleviate the stress on freshwater supply and prevent the fall in groundwater levels, due in particular to agricultural irrigation⁶ which represented more than half of the water used annually in Europe in 2017 according to the European Environment Agency.⁷

However, the wastewater reuse still remains uncommon in some countries. Several factors can explain this observation

*corresponding author.

^aUMR AgroParisTech/INRAE Silva, INRAE-AgroParisTech-Université de Lorraine, 54000, Nancy, France.

^bUniversité Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France.

^cUMR 1114 EMMAH, INRAE, Site Agroparc, CS 40509, 84914 Avignon, France.

^dVeolia Recherche & Innovation, Chemin de la Digue, 78600 Maisons-Laffitte, France.

among them the pathogens that water may contain, which resulted in strict reuse guidelines. Adegoke, *et al.* (2018)⁸ have reported the different pathogens found in wastewaters and the associated risks. Quality indicators are commonly used, but they are more and more debated on their representativeness (particularly for viruses and *Legionella e.g.*, they are not explicitly taken into account as indicators despite their wide occurrence in the environment). Pragmatic recommendations and guidelines were defined by the World Health Organization (WHO)^{9–12} on the basis of health risk considerations. The new European regulations on minimum requirements for water reuse for agricultural irrigation will apply from 26 June 2023.¹³

Sprinkler irrigation with reclaimed wastewater is widely practiced throughout the world but may enhance hazards relative to the dispersion of pathogens by bioaerosols.^{14–16} A collective expertise appraisal has been performed by ANSES in 2012¹⁷, which assessed various secondary risks (e.g. ingestion) and pointed out the lack of knowledge on health risks, particularly due to pathogen inhalation.

An experiment has been conducted in South-Western France to evaluate the impact of two different treatments on the water quality, reused for sprinkler irrigation of two corn fields (Smart-FertiReuse project, <https://www.smartfertireuse.fr/>, accessed on October 4th 2022) controlling the population risks. Among the different analyzed pathogens in the water, *Legionella* was found at high level both at the entrance and output of the WasteWater Treatment Plant (WWTP) (up to 10⁷ GC/L). In the new EU regulations, criteria have been defined for water quality. A minimum required values for *Legionella* has been set to <1000 CFU/L for a reclaimed water class A used for food culture irrigation. These rules are expected to stimulate and facilitate water reuse in the EU and shall also be accompanied by appropriate monitoring of the health risks linked to this practice.

Legionella is an opportunistic bacteria which may be inhaled by human through water aerosol.^{18–20} *Legionella* has around 50 species among them *Legionella pneumophila* which is responsible of the most severe infections such as Legionnaire's disease and pneumonia illness (Pontiac fevers)²¹. These infections due to *Legionella pneumophila* are particularly problematic for people who have a certain fragility or those who are exposed for a long period to these pathogens and may cause in some mortality cases. In France, an increasing trend of Legionellosis is observed since 1995*. A lot of cases were linked to cooling towers²² but some studies have shown contaminated aerosols by *Legionella pneumophila* in the vicinity of WWTP²³ and associated diseases have been observed for example in Netherlands.²⁴

Different approaches have been proposed to quantify health risk applied to *Legionella* (reviewed by Hamilton, *et al.* (2018)¹⁹ and Hamilton, *et al.* (2016)²⁰). Often there are few observations and numerous authors have shown a wide variability in both space and time of the punctual measurements.^{18,25} It is therefore

important to take into account this variability and the uncertainty in risk assessment.

Quantitative Microbial Risk Assessment (QMRA) is an established framework recommended by multiple international organizations (WHO, US EPA, EFSA) to monitor such risks of infection.²⁶ QMRA models have been widely used for food safety management²⁷, in the risk assessment of waterborne disease,^{28–30} for bioaerosol risk^{14,31–34} and for water reuse systems³⁵, and are often coupled with Monte Carlo (MC) simulation approaches which consist of drawing values for all the inputs in the model from distributions defined by external knowledge (e.g., experimental or historical data, expert knowledge) and then propagating these values to all the intermediate variables and outputs through the system in a way that takes the variability and uncertainty of the inputs into account. It is, by definition, a unidirectional approach that prevents the inversion of the dependency relationship between the variables. A Bayesian Network (BN) approach as an alternative to MC simulation approach in QMRA, overcomes this unidirectional path among other things.^{36,37} While QMRA may not be able to incorporate the many factors that influence the quantification of the risk (for example, epidemiological data), models based on BNs are an effective framework that addresses this limitation. BNs approaches in QMRA provide a powerful approach where the dependency relationship between variables can be inverted ('backward calculation', by which the states of the model's variables are updated using 'downstream' data).

A BN is a flexible graphical model based on directed acyclic graph (DAG) where conditional dependencies between the variables are qualitatively represented by the directed arcs on the graph^{38,39} and quantitatively defined with local transition probabilities or deterministic equations. BNs offer transparency and the ability to support causal reasoning because based on causal connections. The visual representation in a BN of large quantities of complex information provides an information platform for improved communications with experts through their graphical representation.^{36,37,40,41} BN as a system approach encompasses both holistic and modular views, enables the synthesis of knowledge of the parts to help understand the whole and makes the complex system more manageable⁴². BNs can be applied to a wide range of application domains such as environmental modelling,⁴³ artificial intelligence⁴⁴ and as mentioned above to quantitative microbial risk assessment (for a review, see Beaudeau, *et al.* (2015)⁴⁵), and water reuse and QMRA.^{41,42,46–48} While QMRA are often constrained by the availability of required data (for example, in the dose-response assessment), system models based on BNs are an effective framework that tackle the lack of data. As all data in the network contributes to the whole, accurate predictions can be made with incomplete data or quite small sample size⁴⁸. BNs were also used for scenario assessment⁴⁶, risk minimisation³⁶ and work with uncertainty and variability as a probabilistic tool^{41,49}

One of the objectives of this paper is the combination of General Bayesian Network (GBN) and QMRA. GBNs, known also as Graphical Independence Networks in the context of continuous and discrete random variables, are the more general case of

* <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/legionellose> (accessed on October 4th 2022)

the BN in which each variable in the data is modelled with the random variable that best suits it rather limiting ourselves to multinomial and normal distributions as in other more classical BNs^{39,50,51}. GBNs can be used as an attractive way of modelling complex systems because it most closely models the expert knowledge and data. Due to their ability to incorporate diverse data types, a GBN enables a complex, multivariate statistical problem (such as QMRA) to be efficiently addressed where classical statistical methods are inept. The method allows the integration of a broad range of quantitative or semi-quantitative information, which is particularly useful in domains where traditional experimental and observational data are missing, inaccurate, sparse, or costly. In this paper, compared to other applications in water reuse and QMRA context previously cited, there is no discretization of the continuous variables in the BN for the learning algorithm. The joint posterior distribution of all the random variables in the BN are obtained by a MCMC (Markov Chain Monte Carlo) algorithm⁵² as used in computational Bayesian model inference. The powerful capacity of conceptualization available by the BN approach via the graphical network is used to construct the core model³⁶ and the knowledge is updated via computational Bayesian inference via data in the spirit of Bayesian model approach (see Verbyla, *et al.* (2016)⁵³ in the water reuse QMRA context).

Another objective was the development and application of an updated GBN to monitor health risks associated with exposure to *Legionella pneumophila* during irrigation events with reclaimed wastewater. The impact of two types of water treatments were analyzed (described in the following section). Keeping the steady monitoring aspect in mind, the constructed GBN is reevaluated with each arrival of new data. At each step posterior distributions for all the variables in the model were computed using priors defined as the posterior distributions at the previous step. The final purposes were to assess different risk scenarios according to the population activities, accounting their distance from the irrigated areas and the environmental conditions in order to get infection probabilities for each configuration.

In the following, material and methods section describes the study site, the data and the models implemented for the developed risk assessment monitoring tool. The update process of the risk quantification as well as three scenarios on simulated data are described in the results section. The update can occur whenever desired and discussed with the experts. The discussion section discusses the benefits of the developed approach for the continuous monitoring of health risks in the context of QMRA, emphasizing over the necessity of informed use of such tools when few data are available.

2 Material and methods

2.1 Study site and data

A pilot experiment was conducted on an area near Tarbes (43°15'45.78"N, 0° 5'16.95"E) where irrigated corn fields from two farms having different agricultural managements are monitored for various quality criteria since 2018 in the framework of

the SmartFertireuse project (FUI SmartFertiReuse project[†]). Irrigation generally occurs from end of June up to the end of August, with 3, 4 or 5 water supplies of 30-40 mm/ha. Two different water treatment pilots (ultra-filtration (A) and UV (B)) were installed at the wastewater treatment plant outlet in order to reach the required log reduction after standard treatment of the initial WWTP (stopping only at the secondary level). The different water compartments were sampled at each irrigation dates in order to monitor the classical indicators for water quality¹³ and to quantify *Legionella* and additional pathogens such as enteric viruses. For each sample the following were collected: 100 ml at the entrance of WWTP (6 samples with 2 replicates each), 250 ml after standard WWTP treatments (6 samples with 2 replicates each), 250 ml after A treatment (4 samples with 2 replicates each), 250 ml after B treatment (6 samples with 2 replicates each), and 400 ml at the groundwater (4 samples at each of 2 different sites on the field and with 2 replicates each) for PCR analysis. Additionally, 2 biosamplers (AGI4-Ace Glass incorporated USA, connected to a 12 L/min pump) containing 40 ml of ultrapure Milli-Q water, were used to sample bioaerosols in the air above corn during irrigation (see Figure 1).

All samples were analyzed to quantify *Legionella* from molecular biology (qPCR from the standard NFT90-471[‡]). All *Legionella* species are considered potentially pathogenic for humans, but *Legionella pneumophila* is the etiological agent responsible for most reported cases of community-acquired and nosocomial legionellosis.^{54,55} Additionally, some cultivations were made for samples showing high values. Figure 2 shows the data collected for *Legionella spp.* presence using qPCR method and differentiated by sampling zone. The graph presents a large variability both between the different analyzed compartments and for the three studied years. As expected the raw water at the entrance of the WWTP presented the highest values of *Legionella* and the groundwater the lowest values from 2019 to 2021. The impact of the two tertiary treatments (A and B) tend to decrease the quantities compared to values found in raw water. A large variability was also observed according to the dates in 2021 which is difficult to explain. Many factors among them climatic and water storage conditions, can lead variability. These points will be discussed in the results.

In parallel to these analyses, the meteorological conditions are monitored thanks to a climatic station set up on the area, which records continuously wind speed and direction, global radiation, air temperature and moisture. To complete this dataset, daily data acquired from the Meteo-France weather station (Tarbes (43°11'12"N, 0°00'00"E) have been used between May and September over the period from 04/29/2010 to 06/18/2020 in order to better take into account the temporal variability of environmental conditions. The crop height was also measured during the irrigation events to assess the aerodynamic roughness (z_0) occurring on the atmospheric transport. Values of the crop height varied from 80 cm to 3.2 m for the studied period.

[†] <https://www.smartfertireuse.fr/les-moyens> (accessed on October 4th 2022)

[‡] <https://tinyurl.com/5n6v4y38> (accessed on October 4th 2022)

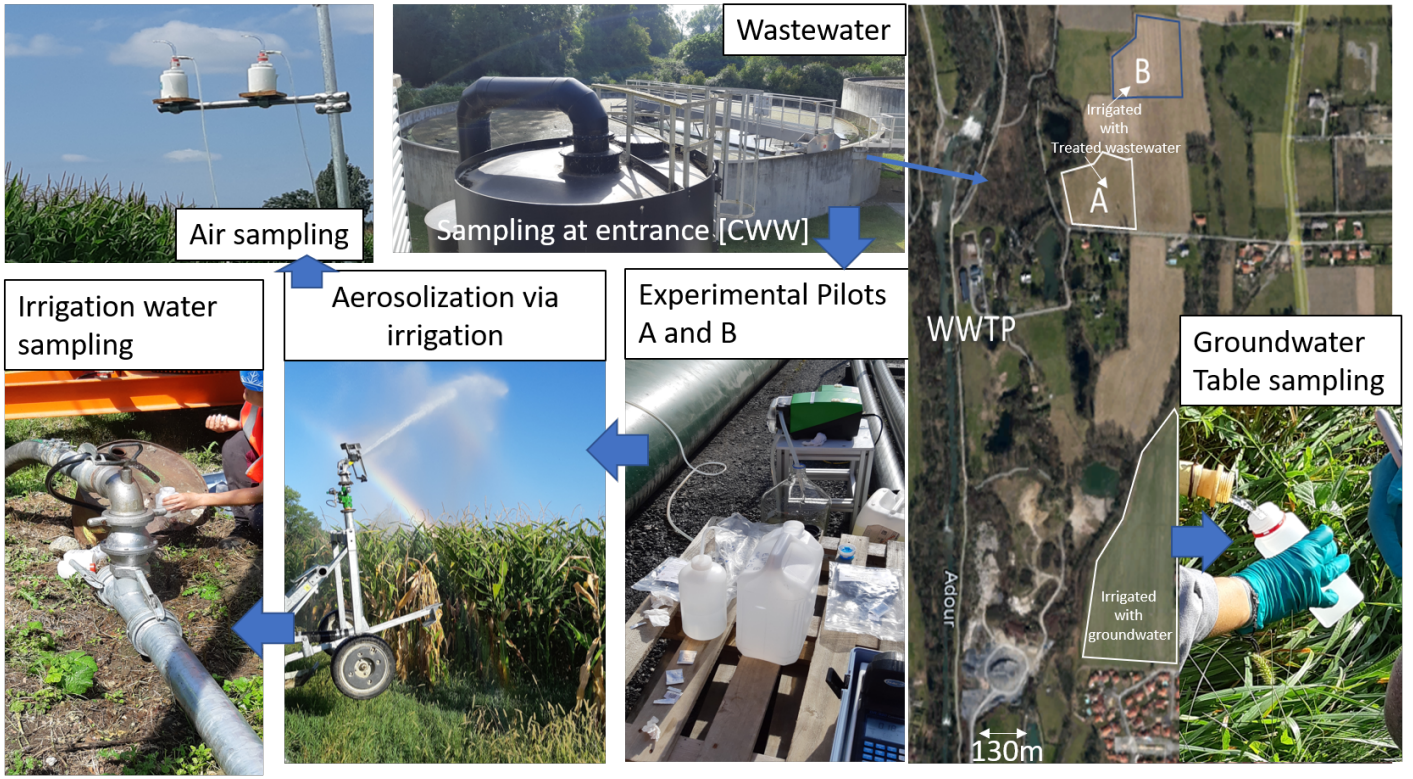


Fig. 1 Location of the study site with the monitored fields and the different water sampling. Two fields were irrigated with treated wastewater (the blue arrows show the water transfert from the WWTP then the pilots A or B, to the irrigated plots).

2.2 The QMRA model

The first step of the modelling is to construct a QMRA model that describes the transmission of *Legionella* along the water exposure pathway described in Figure 3.

2.2.1 Water quality model

Three different exposure pathways were modeled for the water used for the irrigation of the fields:

1. Irrigation directly from the groundwater table with quality C_{GT} [genome copies per liter (GC/L)], currently in use;
2. Irrigation from wastewater sequentially treated by standard WWTP treatments and experimental pilot A with quality C_A [GC/L];
- 2b. Irrigation from wastewater sequentially treated by standard WWTP treatments and experimental pilot B with quality C_B [GC/L].

In the process currently used (part 1 above), the groundwater is not treated in WWTP before being used for irrigation. In the experimental processes (parts 2 and 2b above) reclaimed wastewater is used. The initial wastewater coming to the WWTP is sequentially treated by standard WWTP treatments and experimental pilot (A or B). These treatments effects on the water quality are modelled by the following logarithmic decays formulas:

$$C_A = C_{TP}/10^{k_A}, \quad (1)$$

$$C_B = C_{TP}/10^{k_B}, \quad (2)$$

$$\text{with } C_{TP} = C_{WW}/10^{k_{TP}}, \quad (3)$$

where C_{TP} denotes the concentration of *Legionella* in wastewater treated by standard WWTP treatments [GC/L], C_{WW} denotes the concentration of *Legionella* in the raw water at the entrance of the WWTP (wastewater) [GC/L], k_{TP} denotes the decay rate of the filtration process of the WWTP [unitless], and k_A (resp. k_B) denotes the decay rate of the filtration process of experimental pilot A (resp. B) [unitless].

Concentration in *Legionella pneumophila* $C_{Lp,P}$ [GC/L] in each treatment train $P \in \{GT, A, B\}$ was deduced from the concentrations C_P above by the following formula¹⁹:

$$C_{Lp,P} = C_P f_{Lp}, \quad (4)$$

where f_{Lp} denotes the fraction of *Legionella pneumophila* among the measured *Legionella* spp.

2.2.2 Spray irrigation

Two different sprinklers are used on the fields. They are characterized by their respective flow rate F_1 and F_2 [m^3/h] from which a portion p_{150} of particles is smaller than $150\mu m$ (portion of aerosols in respirable range, $< 150\mu m$). Experimental pi-

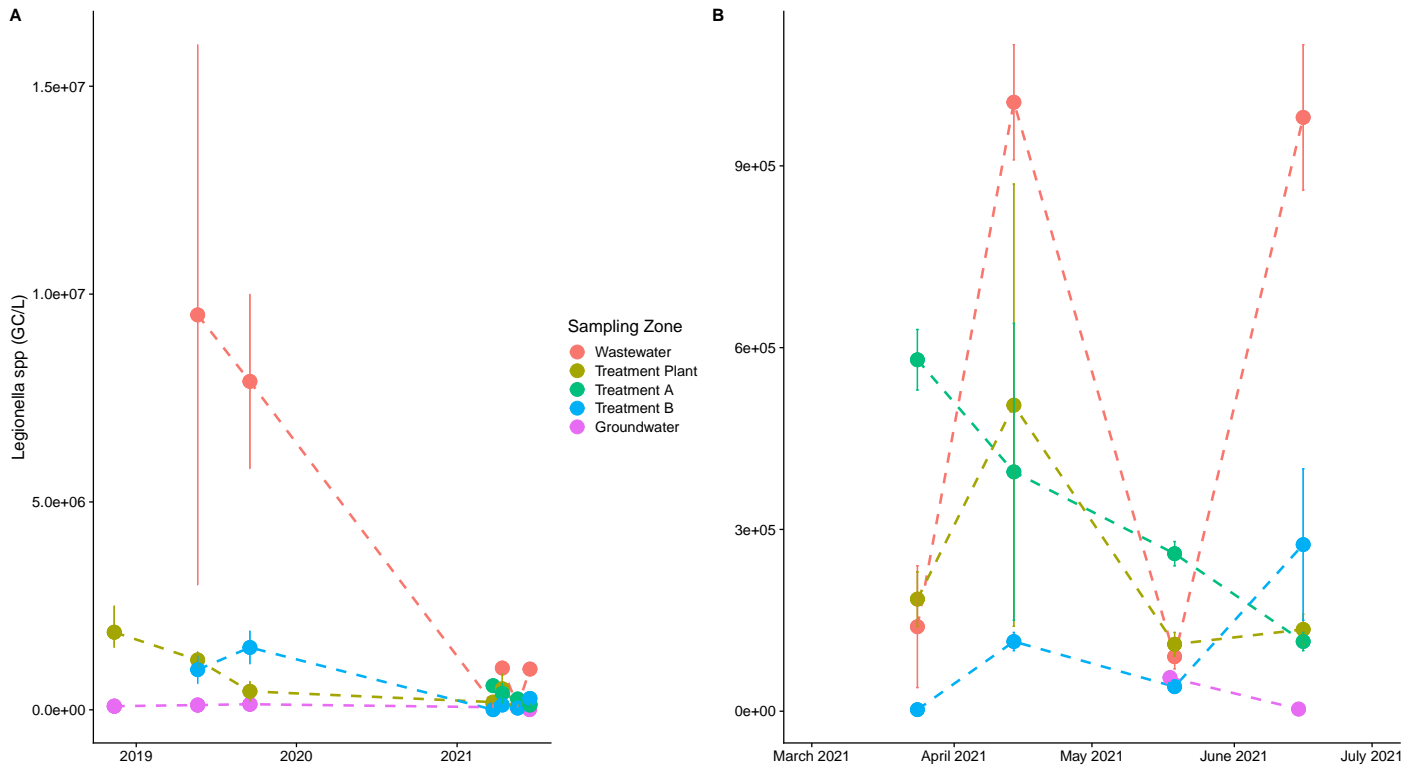


Fig. 2 Data collected on site differentiated by sampling zone. Points represent mean concentration of *Legionella spp.* for replicates and vertical bars represent their variability. A Complete data from November 2018 to June 2021. B Zoom on year 2021.

lot A is associated with sprinkler 1 and experimental pilot B is associated with sprinkler 2. The following formula models the quality $Q_{Lp,P}$ [GC/m^3] of the water used for the irrigation after aerosolization:

$$Q_{Lp,P} = F_{S_P} P_{150} C_{Lp,P} 3600 / 10^3, \quad (5)$$

where $C_{Lp,P}$ is the concentration of *Legionella pneumophila* in the water used for irrigation defined in equation (4) above [GC/L], $Q_{Lp,P}$ denotes the concentration of *Legionella pneumophila* in the aerosolized water used for irrigation [GC/m^3], S_P denotes the sprinkler associated to the treatment train $P \in \{GT, A, B\}$, with $F_{S_{GT}} = (F_1 + F_2)/2$, $F_{S_A} = F_1$ and $F_{S_B} = F_2$ and the 10^3 and 3600 multipliers ensure the unit for $Q_{Lp,P}$ (in GC/m^3).

2.2.3 Atmospheric dispersion model for spray irrigation

A modified Gaussian plume atmospheric transport model was used to describe the concentration of *Legionella pneumophila* $C_{x,P}$ [GC/m^3] at $x \in \{100; 300; 500; 1000\}$ meters downwind from the irrigation sprinkler for each treatment train P:

$$C_{x,P} = \frac{Q_{Lp,P}}{2\pi u \sigma_y \sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2}\right) \left\{ \exp\left(-\frac{(z-z_0)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z+z_0)^2}{2\sigma_z^2}\right) \right\} \exp\left(-\frac{\lambda x}{u}\right), \quad (6)$$

where $Q_{Lp,P}$ is defined in equation (5) [GC/m^3], y is the horizontal distance perpendicular to wind [m], z is the downwind

receptor breathing zone height [m]; z_0 is the estimated aerodynamic rugosity [m], u is the wind speed [m/s]; λ is the microbial decay coefficient [s^{-1}], and σ_y , σ_z denote respectively the horizontal and vertical dispersion coefficients [m] and are calculated by the following equations: $\sigma_y = R_y x^{r_y}$ and $\sigma_z = R_z x^{r_z}$, where R_y , r_y , R_z and r_z are constants depending on the wind speed u [m/s] and the insolation value E [J/cm^2] (see Seinfeld, *et al.* (2016)⁵⁶ and Tables 1 and 2 in Supplementary Materials).

2.2.4 Exposure model

Three exposure models were constructed depending on the activity of the population at risk:

- passersby with a supposed average $t_{passerby}$ of 1 minute of exposure per irrigation day¹⁷,
- residents with a supposed average $t_{resident}$ of 2 seconds of exposure over the 2.27 hours spent outside per irrigation day^{17,57}, and
- farmers with a supposed average t_{farmer} of 30 minutes of exposure per irrigation day (chosen according to the local farmer behavior knowledge).

For each population, the fraction of day this population was exposed was computed and an uncertainty of a half minute around it was considered. For example $t_{passerby}$ was taken as a Gaussian distribution centered around $1/(60 \times 24) = 6.94e-4$ with standard deviation $1/(\sqrt{2} \times 60 \times 24) = 4.91e-4$ truncated on $[0; 1]$.

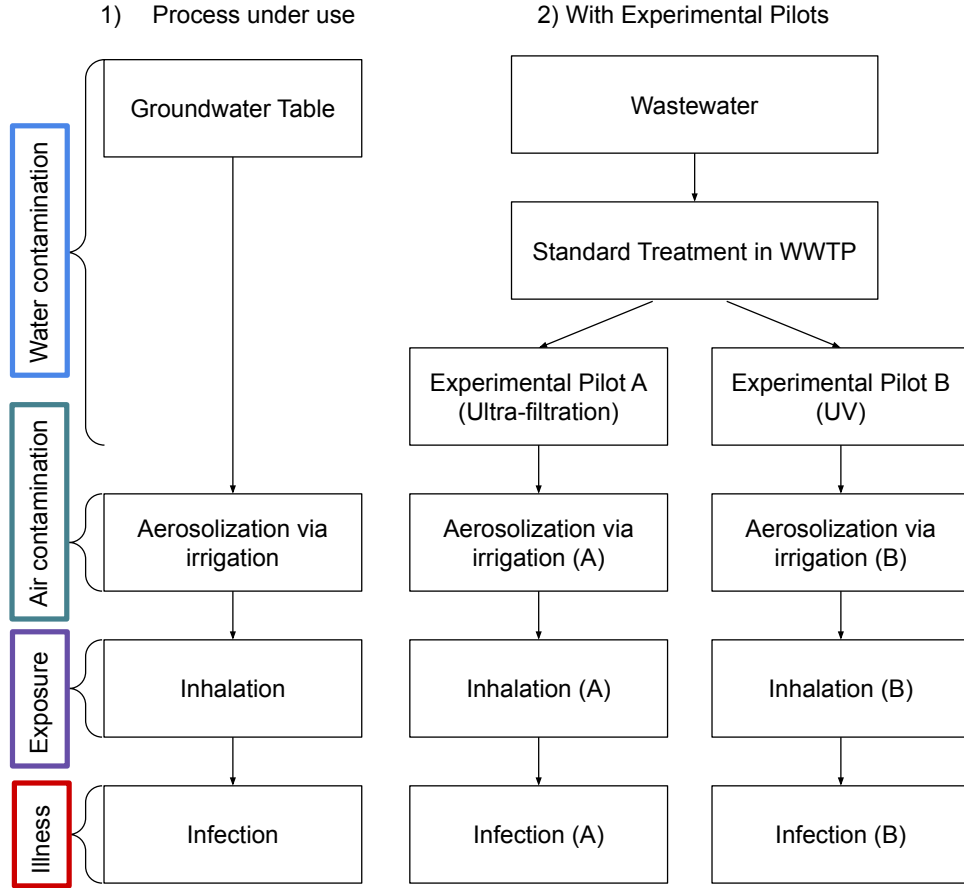


Fig. 3 Overview of the exposure pathway of *Legionella pneumophila* in the context of agricultural irrigation of the experimental plots in Tarbes.

The inhaled dose $D_{A,x,P}$ [genome copies per day] was calculated for each treatment train P , distance x and activity A taking values in $\{passerby; resident; farmer\}$ as follows:

$$D_{A,x,P} = C_{x,P} I t_A, \quad (7)$$

where $C_{x,P}$ is defined in equation (6) [GC/m^3], I denotes the mean inhalation rate [m^3 air/day], and t_A the fraction of day each population at risk A is exposed [unitless].

2.2.5 Risk characterization

In the chosen dose-response model, clinical severity infections csi (corresponding to an infection requiring a clinical visit) were distinguished from other infections inf (corresponding to subclinical infection or potentially Pontiac Fever endpoint). The daily probability $P_{i,A,x,P}$ was calculated for each treatment train P , distance x , activity $A \in \{passerby; resident; farmer\}$ and type of infection $i \in \{inf, csi\}$ using the exponential dose response model for *Legionella pneumophila* from Armstrong, et al. (2007)⁵⁸ defined as follows:

$$P_{i,A,x,P} = 1 - e^{-r_i D_{A,x,P} f_{CFU}}, \quad (8)$$

where $D_{A,x,P}$ is defined in equation (7) [CFU/day], r_i is the probability of the bacteria bypassing the host defenses and initiating response [unitless], and f_{CFU} denotes the relation between colony-forming units (CFU) and genome copies (GC) for *Legionella pneumophila*. The annual risk, which is the probability to be infected at least once during one irrigation day in a year was calculated as per the following equation:

$$P_{year,i,A,x,P} = 1 - (1 - P_{i,A,x,P})^{n_e d_e}, \quad (9)$$

where $n_e d_e$ is the total annual number of days of irrigation with n_e the number of episodes of irrigation over a year and d_e the duration of one irrigation episode in days.

Figure 4 represents the directed acyclic graph (DAG) of the QMRA model constructed with the equations above. The input distributions (or prior distributions in the Bayesian framework) for each parameters are defined in Table 1.

2.3 Coupling the datasets

An augmented model presented in Figure 5 (on the right, the QMRA model from Figure 4 is augmented by the data, in quantity n_m for the microbiological data and in quantity n_{day} and n_{year} for the climatic data, represented by rectangles using graphical conditional independence links) was built to account for the experimental data described in Section 2.1.

Description	Symbol X	Unit	Value	Reference
Water contamination				
Water quality in groundwater table in <i>Legionella</i>	C_{GT}	GC/L	$\log_{10}(X) \sim \mathcal{N}(10, 1)$	Expert knowledge and literature data
Water quality before treatment in <i>Leg.</i>	C_{WW}	GC/L	$\log_{10}(X) \sim \mathcal{N}(15, 1)$	Expert knowledge and literature data
Log decay in WWTP by standard treatment	k_{TP}	$\log_{10}(GC/L)$	$X \sim \mathcal{N}(8, 1)$	Project expert knowledge
Log decay by experimental pilot A	k_A	$\log_{10}(GC/L)$	$X \sim \mathcal{N}(4, 1)$	Project expert knowledge
Log decay by experimental pilot B	k_B	$\log_{10}(GC/L)$	$X \sim \mathcal{N}(3, 1)$	Project expert knowledge
Portion of <i>Legionella pneumo.</i>	f_{LP}	Unitless	$X \sim \mathcal{N}(1e-3, 2e-4)T(0,)$	From testing lab analytical techniques
Air contamination				
Flow rate of sprinkler 1	F_1	m^3/h	$X \sim \mathcal{N}(44, 1)$	Sprinklers properties ⁵⁹
Flow rate of sprinkler 2	F_2	m^3/h	$X \sim \mathcal{N}(42, 1)$	Sprinklers properties ⁵⁹
Portion of aerosols in respirable range ($< 150\mu m$)	p_{150}	Unitless	$X \sim \mathcal{U}(5e-4, 7e-4)$	Sprinklers properties ⁵⁹
Horizontal distance perpendicular to wind	y	m	$X \sim \mathcal{U}(0, 2.5)$	From literature ^{19,60}
Downwind receptor breathing zone height	z	m	$X \sim \mathcal{U}(1, 1.7)$	Height of breathing zone
Estimated aerodynamic rugosity on the field	z_0	m	$X \sim \mathcal{U}(0.1, 0.45)$	Prior knowledge on the field
Wind speed and Insolation rate	(u, E)	m/s and J/cm^2	$\log(X) \sim \mathcal{N}_2\left(\begin{pmatrix} 2 \\ 8 \end{pmatrix}, \frac{1}{8}I_2\right)$	Vague climatic prior
Microbial decay coefficient	λ	s^{-1}	$X \sim \mathcal{N}(1.32e-4, 3.44e-4)$	61,62
Exposure				
Inhalation rate	I	m^3/day	$X \sim \mathcal{N}(20, 2)$	From literature ^{63,64}
Fract. of day a passerby is exposed	t_P	Unitless	$X \sim \mathcal{N}(6.94e-4, 4.91e-4)T(0, 1)$	From literature ¹⁷
Fraction of day a resident is exposed	t_R	Unitless	$X \sim \mathcal{N}(3.15e-3, 4.91e-4)T(0, 1)$	From literature ^{17,57}
Fraction of day a farmer is exposed	t_F	Unitless	$X \sim \mathcal{N}(2.08e-2, 4.91e-4)T(0, 1)$	Knowledge on local farmer behavior
Illness				
Portion of CFU	f_{CFU}	Unitless	$X \sim \mathcal{N}(1e-3, 2e-4)T(0,)$	From testing lab analytical techniques from uncensored data
Dose response parameter for <i>Legionella pneumophila</i> infection endpoint	r_{inf}	Unitless	$\log(X) \sim \mathcal{N}(-2.934, 0.488)$	From literature ^{19,58,65}
Dose response parameter for <i>Legionella pneumophila</i> clinical severity infection endpoint	r_{csi}	Unitless	$\log(X) \sim \mathcal{N}(-9.688, 0.296)$	From literature ^{19,58,66}
Number of irrigation episodes per year	n_e	Unitless	$X \sim \text{Multi}\{3, 4, 5\}$	Process knowledge
Irrigation episodes duration	d_e	Days	$X \sim \text{Multi}\{2, 3, 4, 5, 6, 7\}$	Process knowledge

Table 1 Model input parameters and distributions for the QMRA model. $\mathcal{N}(\mu, \sigma)$ denotes the normal distribution of mean μ and standard deviation σ , eventually truncated on the interval $T(.,)$, $\mathcal{N}_2(M, \Sigma)$ denotes the bivariate Gaussian distribution of mean M and variance Σ . I_2 denotes the identity matrix of order 2. $\mathcal{U}(a, b)$ denotes the uniform distribution on the interval (a, b) . $X \sim \text{Multi}\{a_1, a_2, \dots\}$ stands for **equi**probable sampling from the values a_1, a_2, \dots (multinomial distribution).

This new model takes into account the measure uncertainty for the microbiological data and the inter and intra annual variabilities for the climatic data. In the case of the microbiological data, let $C_{Q,i}$ be the i th measured concentration of *Legionella* (in GC/L) at step $Q \in \{WW, GT, TP, A, B\}$. This random variable is linked to the concentration of *Legionella* C_Q defined in section 2.2.1. These random variables are linked through the following model:

$$\log_{10}(C_{Q,i}) \sim \mathcal{N}(\log_{10}(C_Q), \sigma_Q),$$

where $\mathcal{N}(.,)$ is the normal distribution and σ_Q represents the measure uncertainty. A Gamma distribution was assigned to the square of its inverse i.e. $1/\sigma_Q^2 \sim \Gamma(100, 1)$ which represents a vague prior confidence interval of [0.09; 0.11] at 95% for σ_Q (in GC/L).

In the case of the climatic data, a bivariate distribution was chosen to model the joint distribution of wind speed and insolation value to get the distribution that best suits these jointly observed data capturing the correlation between both quantities. Let $(u, E)_{j,k}^{obs}$ denote the vector of the observed wind speed (in m/s) and insolation value (in J/cm^2) at the j th day ($j = 1, \dots, n_{day}$) of the k th year ($k = 1, \dots, n_{year}$). This random vector is linked through the following model to the vector (u, E) of the QMRA

model:

$$\log \left[(u, E)_{j,k}^{obs} \right] \sim \mathcal{N}_2(M_k, \Sigma_{intra}), \quad (10)$$

$$M_k \sim \mathcal{N}_2(M_m, \Sigma_{inter}), \quad (11)$$

$$M_{year} \sim \mathcal{N}_2(M_m, \Sigma_{inter}), \quad (12)$$

$$\text{and } \log \left[(u, E) \right] \sim \mathcal{N}_2(M_{year}, \Sigma_{intra}), \quad (13)$$

where $\mathcal{N}_2(.,)$ is the bivariate normal distribution, Σ_{intra} and Σ_{inter} are the measured variability between days of each year and the measured variability between each year respectively, M_k and M_m are the mean of the k th year and the global mean over the years respectively. M_{year} is the predicted mean vector for a randomized year given by $M_{year} \sim \mathcal{N}_2(M_m, \Sigma_{inter})$. (u, E) is the vector of a predicted wind speed and insolation value in a randomized year. The following prior distributions were assigned to the variance and mean parameters described above:

$$\Sigma_{intra}^{-1} \sim W_2(I_2; 24), \quad (14)$$

$$\Sigma_{inter}^{-1} \sim W_2(I_2; 24), \quad (15)$$

$$\text{and } M_m \sim \mathcal{N}_2 \left(\begin{pmatrix} 2 \\ 8 \end{pmatrix}, \frac{1}{24}I_2 \right), \quad (16)$$

where I_2 is the identity matrix of order 2, $W_2(.,)$ is the bivariate

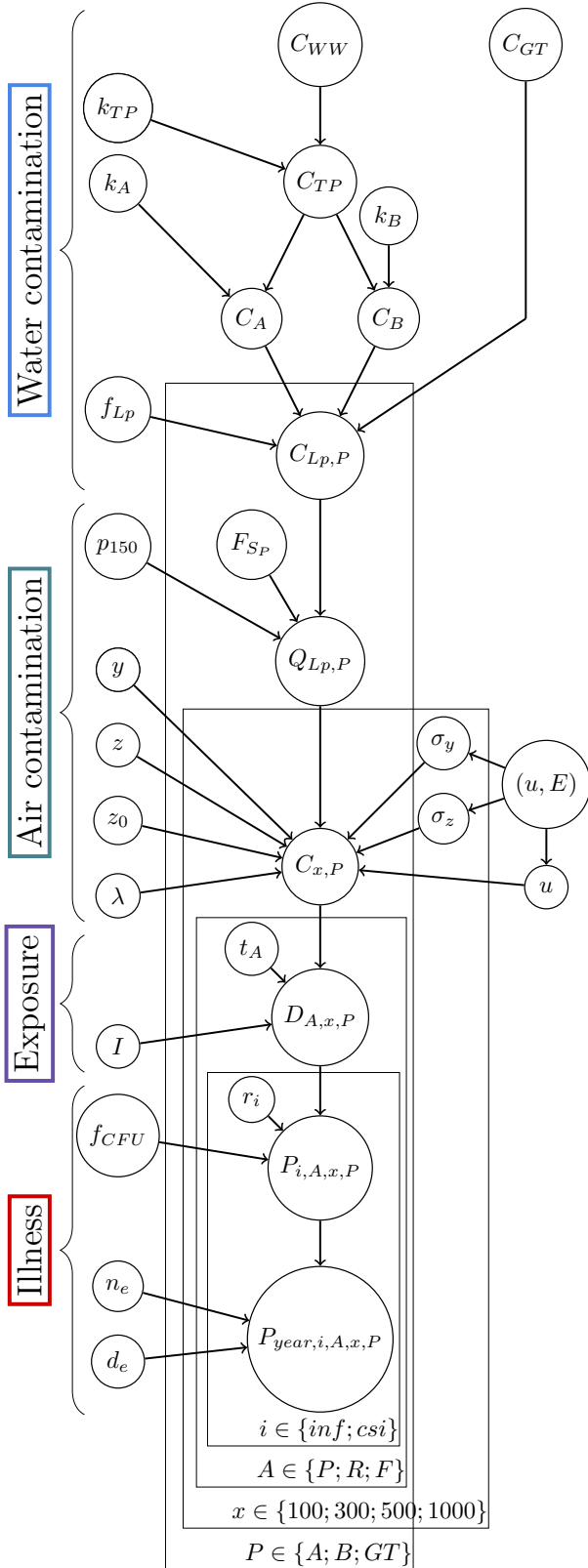


Fig. 4 Directed acyclic graph of the Bayesian network model.

Wishart distribution, and $N_2(\cdot)$ denotes the bivariate Gaussian distribution chosen to recover the vague climatic priors defined in Table 1 for (u, E) vector.

The R⁶⁷ package `rjags`⁶⁸ was used to compute the Bayesian inference of the augmented model. MCMC sampling algorithms were applied as implemented in the software JAGS 4.3.0.⁶⁹

3 Results

The QMRA model (*Prior*), and the augmented model were run at each month for which data was collected (May and September 2019 and each month from March to June 2021) (*Posteriors 1 to 6*). For each model, two independent MCMC chains (using different initial values for the parameters) of 1,400,000 simulations with a burn-in period of 1,000,000 were run. Convergence of the MCMC run was assessed by graphical inspection of the chains.

In the following, the results and the update of prior beliefs in the model are described module by module.

3.1 Water contamination

The priors described in Table 1 yielded the following 95% credible intervals for water quality concentrations expressed here in log: $k_A \in [2; 6]$, $k_B \in [1; 5]$, $k_{TP} \in [6; 10]$, $f_{LP} \in [0.06; 0.14]$, $\log_{10}(C_A) \in [-0.4; 6.5]$, $\log_{10}(C_B) \in [0.6; 7.3]$, $\log_{10}(C_{GT}) \in [8; 12]$, $\log_{10}(C_{WW}) \in [13; 17]$, $\log_{10}(C_{TP}) \in [4; 10]$, $\log_{10}(C_{LP,A}) \in [-5; 1]$, $\log_{10}(C_{LP,B}) \in [-4; 2.5]$, and $\log_{10}(C_{LP,GT}) \in [3; 7]$. These credible intervals are derived from the distributions described in Table 1 and the propagation of the uncertainty according to Figure 5 for non terminal variables (or parameters).

Figure 6 as well as Figure 7 illustrate that the groundwater ($C_{LP,GT}$) and the water after treatment by the pilot A or B ($C_{LP,A}$, $C_{LP,B}$) show similar contamination levels ($\sim 10^2$ GC/L) indicating that the new irrigation process is as safe as the previous one (with groundwater without treatment). As expected, a significant log reduction between contamination of raw and treated wastewater for the posterior distributions appears ($\sim 3 \log_{10}$ reduction). Nevertheless, one observes that the parameters k_{TP} , k_A and k_B are smaller in the posterior distributions than in the prior distribution indicating that the data introduced at the C_{TP} 's, C_A 's and C_B 's level (see Figure 5) are more contaminated than expected. This may be due to the storage of the water between treatments in large reservoirs or poor control of the treatment chain as discussed below.

From the prior distributions (in red in Figure 6) to the posterior ones, one observes reduced uncertainties (narrower distributions in posterior). The posterior distribution at June 2021 (in pink in Figure 6) shows reduction of the concentration after treatments: 1 \log_{10} reduction (from 10^6 to 10^5 GC/L) for the wastewater after treatment in treatment plant and treatment A and similarly for treatment B. The transition to *Legionella pneumophila* from *Legionella* measurements analysed by PCR reduces again the contamination by 3 \log_{10} .

Although a significant log reduction between raw and treated wastewater is observed, the differences between treatments A and B and groundwater are very slim compared to what would have been expected by the prior knowledge of processes A and B which

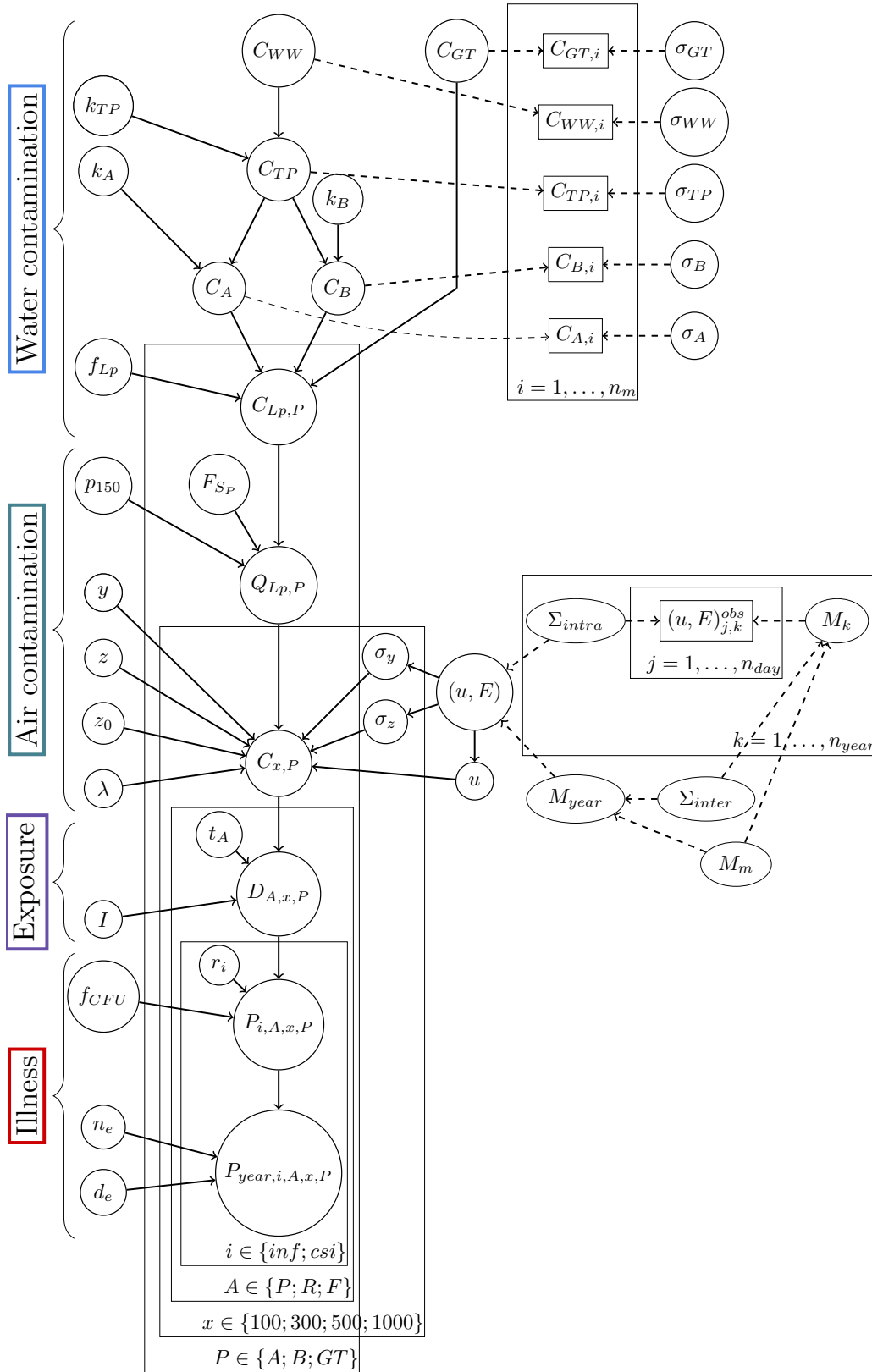


Fig. 5 Directed acyclic graph of the augmented model.

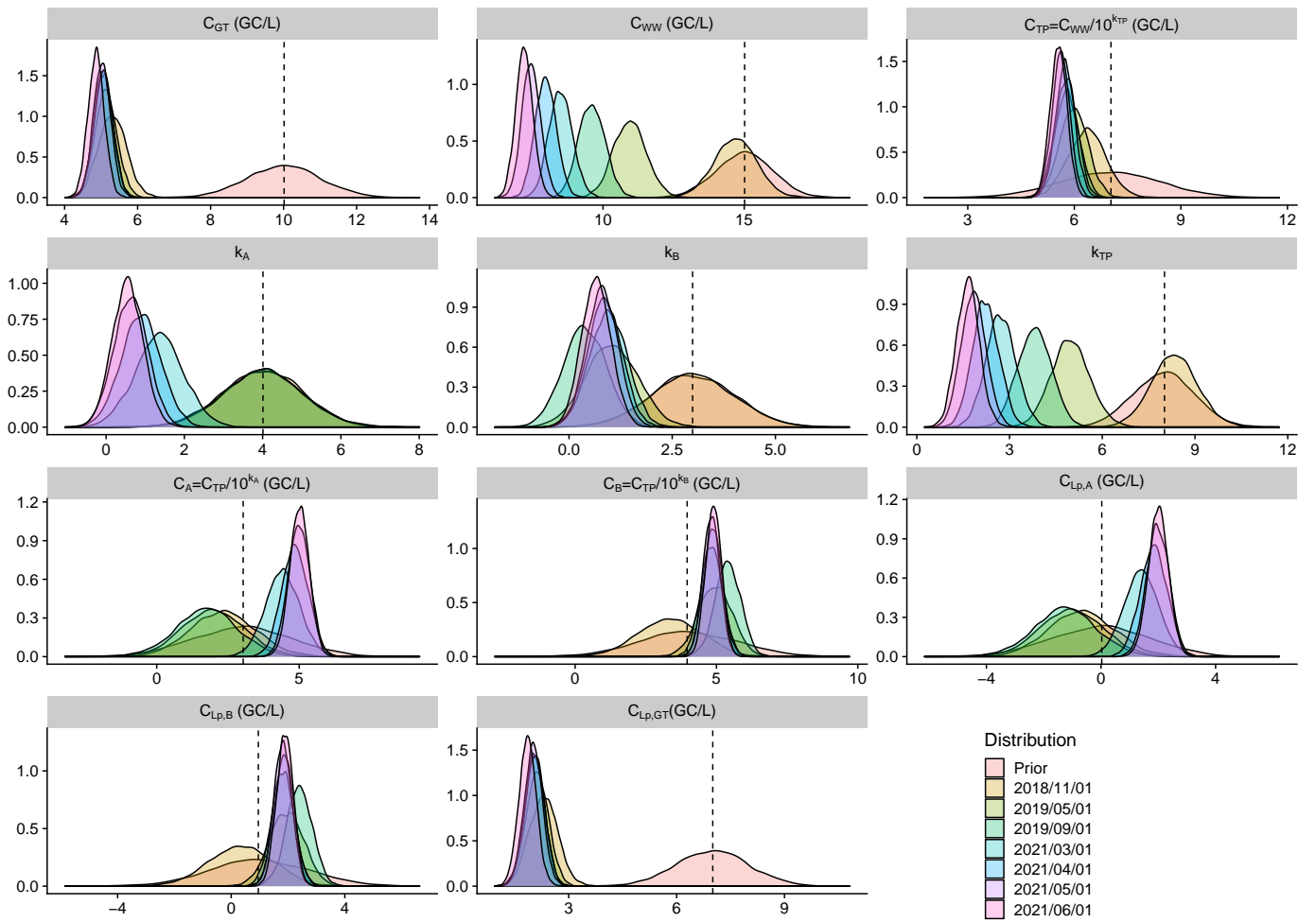


Fig. 6 Prior (red) versus posterior marginal distributions of selected parameters and variables of the water contamination module of the augmented Bayesian network. Dashed lines represent the median of the prior marginal distributions. Distributions represented are in $\log_{10}(CFU/L)$

should eliminate most of the bacteria. This may be due to storage issues or lack of control in the chain of experimentation and will be discussed in the following.

3.2 Air contamination

The air contamination module output is the concentration $C_{x,P}$ (in GC/m^3) of *Legionella pneumophila* in the air at distances $x \in \{100m, 300m, 500m, 1000m\}$ from the source due to the irrigation using water of quality $P \in \{A, B, GT\}$ (see Figure 8). One observes reduced uncertainty from the prior to the posterior distributions of the concentration of *Legionella pneumophila* in the air for all treatments and distances. As expected, the distributions are globally decreasing with the distance: from 10^{-3} at 100m to 10^{-5} at 1000m for wastewater after treatment in TP and treatment A and similarly for treatment B and groundwater GT. Final predicted air concentrations are of similar intensity for irrigation using water from treatments A and B and groundwater which comforts the previous remarks on the quality of the proposed irrigation process.

3.3 Illness

Figures 9 to 11 represent the probabilities of clinical severity infection over one year, for farmers (see Figure 9), for residents (see Figure 10) and for passersby (see Figure 11). For both clinical severity infections (csi) and others (inf) a decrease in the risk of contamination is observed with the distance for each category of people at risk: from 10^{-10} at 100m to 10^{-13} at 1000m for the risk of clinical severity infection for the farmers exposed to treatment A (and after treatment TP). A decay of approximately 1 \log_{10} is observed between farmers and residents for the probability of clinical severity infection (which is again the case between residents and passersby). Treatments A and B and groundwater give similar probabilities of infection with a slightly higher risk for groundwater which confirms the previous results that the new irrigation process is slightly safer than the one under use (with groundwater, without further treatment). It is also notable that the upper bound of the 95% credibility intervals never exceeds the US EPA annual infection benchmark⁷⁰ of 10^{-4} infections per year (for drinking water) for clinical severity infections.

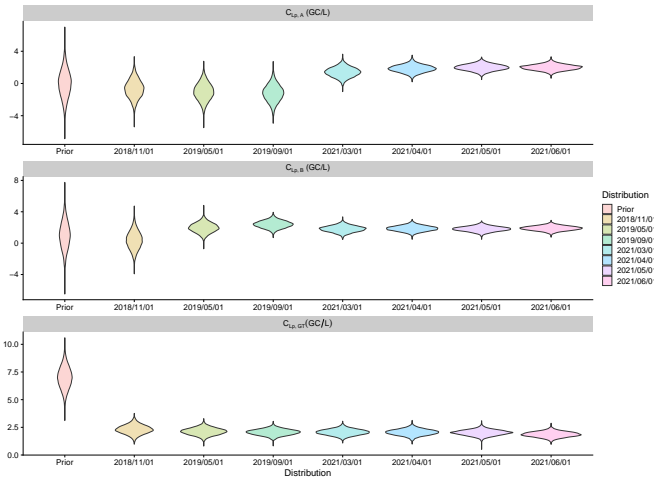


Fig. 7 Violin plots for concentrations in *Legionella pneumophila* (\log_{10} GC/L)

3.4 Scenario evaluation

Scenarios to determine how failures or very high initial contamination would affect the final risk were evaluated by computing the updated annual probabilities of infection with a new day of 10 simulated data as established by observed variable in the water contamination step of the model. In every scenario all populations were tested as well as another toy population exposed 100% of the time of irrigation to the pathogen which could relate to children playing outside all irrigation days for instance. Results of all scenarios are presented in Figure 12.

3.4.1 Scenario 1: Treatment pilot failure

For this scenario, the $\log_{10}(C_{WW,i})$ ($i = 1, \dots, 10$) data were simulated from a $\mathcal{N}(15, 1)$ as a classical contamination as established by expert knowledge and literature data (see Table 1), $\log_{10}(C_{TP,i})$ ($i = 1, \dots, 10$) data were obtained with a decay around 8 log of $\log_{10}(C_{WW,i})$ as the WWTP standard treatment is supposed to be effective (see Table 1, $k_{TP} \sim \mathcal{N}(8, 1)$ as established by expert knowledge and literature data when WWTP standard treatment is effective) and finally $\log_{10}(C_{A,i})$ ($i = 1, \dots, 10$) were obtained with around no log decay ($k_A \sim \mathcal{N}(0, 1)$) because in this scenario a failure of the experimental failure pilot A is supposed. Annual probability of infection increased by 0.8 log which gave a median of -4.89 log and an interdecile range of 1.33 log at 100m for the toy population. For all distances and all population studied the 9th decile never exceeded the US EPA threshold of -4 log annual probability of infection.

3.4.2 Scenario 2: High contamination of the source

This scenario differs from the last one in the fact that the $\log_{10}(C_{WW,i})$ ($i = 1, \dots, 10$) data were simulated from a $\mathcal{N}(20, 1)$ to simulate high contamination of source and that the pilot treatment is supposed to be effective (see Table 1, $k_A \sim \mathcal{N}(4, 1)$ as established by expert knowledge and literature data when WWTP standard treatment is effective). This resulted in an increase of 0.9 log of the annual probability of infection which is comparable to the previous scenario. Once more, the 9th decile never exceeds

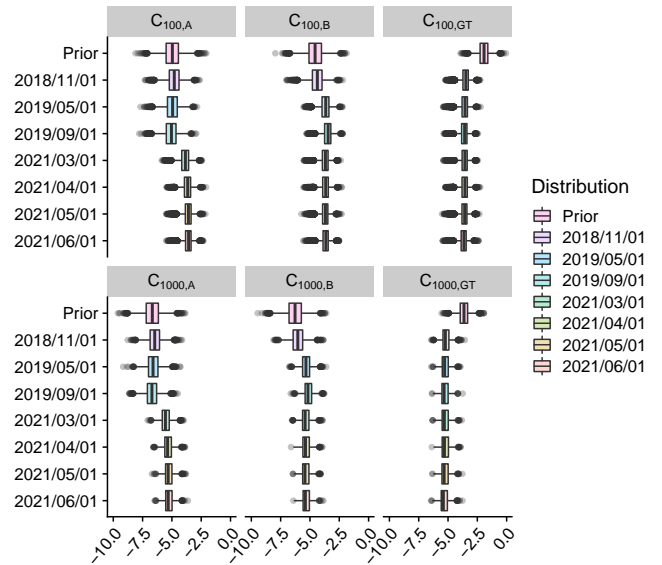


Fig. 8 Log10 of concentrations in *Legionella pneumophila* in the air due to the irrigation process for groundwater (GT), and after treatments A and B at 100m and 1000m from the source.

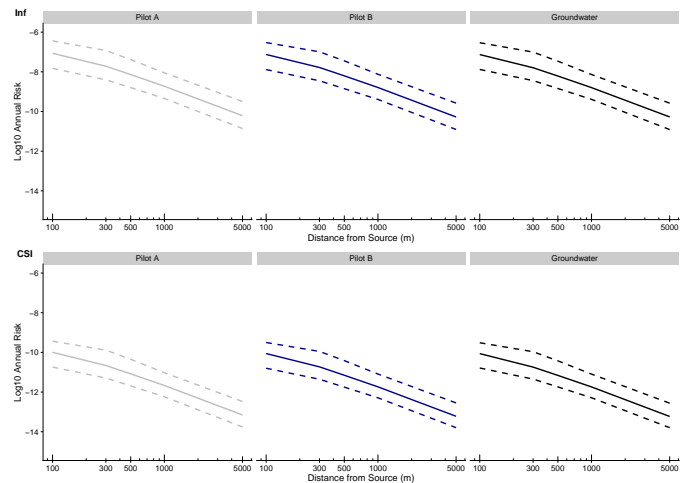


Fig. 9 Log10 annual infection risks for *L. pneumophila* for the farmers due to sprinkler exposure at downwind distances ranging from 100 to 5000 meters from the source. The median (solid line) and interdeciles confidence interval (dashed lines) are shown. Infection (Inf) or clinical severity infection (CSI) dose response model endpoints.

the US EPA threshold for annual probability of infection but gets even closer with a value of -4.2 log at 100m for the toy population.

3.4.3 Scenario 3: High contamination of the source, and failure of treatment pilot

In this last scenario, the simultaneity of a pilot A failure and high contamination was considered to illustrate the attainable risk in this event at risk. This was the only scenario which resulted in an annual probability of infection exceeding the US EPA threshold of -4 log. This threshold was exceeded by the median for the toy population for 100m with a value of -3.79 and by the 9th decile

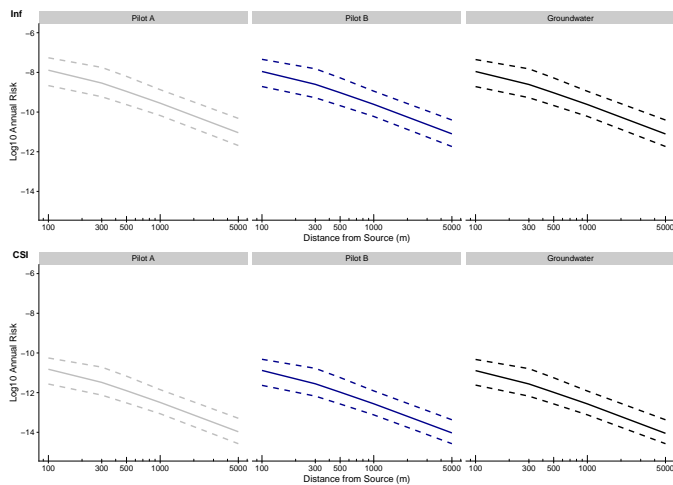


Fig. 10 Log₁₀ annual infection risks for *L. pneumophila* for the **residents** due to sprinkler exposure at downwind distances ranging from 100 to 5000 meters from the source. The median (solid line) and interdeciles confidence interval (dashed lines) are shown. Infection (Inf) or clinical severity infection (CSI) dose response model endpoints.

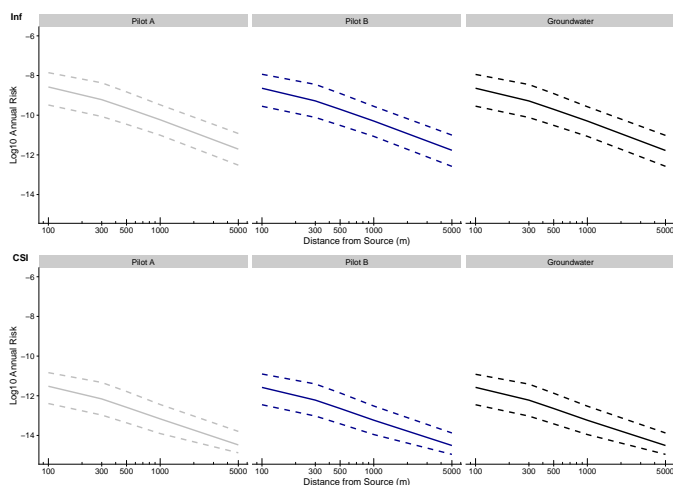


Fig. 11 Log₁₀ annual infection risks for *L. pneumophila* for the **passersby** due to sprinkler exposure at downwind distances ranging from 100 to 5000 meters from the source. The median (solid line) and interdeciles confidence interval (dashed lines) are shown. Infection (Inf) or clinical severity infection (CSI) dose response model endpoints.

at 300m with a value of -3.20 . The risk for the other populations did not exceed this threshold with a maximum value of -4.88 attained for farmers at 100m from the source.

4 Discussion

A general Bayesian network methodology was applied to a QMRA problematic to monitor the risk of *Legionella* infection in the vicinity of agricultural plots irrigated with two experimental water treatment pilots. General Bayesian networks approach allows for simple accounting for variability and uncertainty in a context of complex modelling such as QMRA models. In the developed approach a modified Gaussian plume dispersion model was used to compute health risks according different scenarios, using knowledge of meteorological conditions over long periods (> 20 years

here) and distinguishing three categories of persons at risk, two dose-response endpoints and different downwind distances from the sprinkler.

The general Bayesian network methodology used is a very powerful approach without any discretization to quantify complex phenomenon in presence of prior knowledge and scarce data. Indeed, the approach allows to take account of expert knowledge and literature data in the prior distributions of the random variables and to consider simultaneously collected data all over the phenomenon. The uncertainty and variability of each variable is evaluated through the strength of its prior knowledge combined with the strength of the data information provided by the size of data sets, which is not the case in simulated networks, just using Monte Carlo simulations, where the data are fitted marginally to establish the prior distributions, and then their provided uncertainty spread poorly in the whole network. Also, by contrast with non-parametric Bayesian network, in the highly parametric approach used in this paper, the knowledge of the functional relationships between the random variables is maximized to permit the estimation of quantities of interest in context of scarce data.

Both for annual infection and clinical severity infection, no major difference was observed between risk induced by irrigation using water treated by treatment A (ultra fine filtration, after standard treatment plant), treatment B (UV, after standard treatment plant) or groundwater. This comforts the prior belief that treatments A and B give water quality at least comparable to the quality of the groundwater. Risks were observed to decrease with distance from the source, which was expected according to the atmospheric dispersion modelling. Approximately 1 log difference was observed between each of the studied categories of people (farmers, residents and passersby), with passersby being the population least at risk and farmers the population most at risk according to the model (logical order according to the exposure time).

High values of *Legionella* in GC/L were found in the different analyzed compartments even after the two studied treatments. Impact of the storage time in large reservoirs after the two pilots can in part explain some high observed levels. The cleaning of these tank covers appears as an important point to check for operational applications. A such observation reveals the interest of working in the first steps with GC for an operational monitoring of the water quality, GC are uncensored data and PCR analysis are less time consuming and can be done routinely. The values reported in GC were first introduced in the model until the risk characterization, then converted in CFU applying a significant decrease of $10e3$ between GC and UFC, as quantified from the laboratory analytical results and as often observed in bibliography. Let us mention that if one applies this decrease factor (f_{CFU}) on the contamination after the tertiary treatment A and B, i.e. on C_A and C_B , the two treatments are class A reclaimed water because only 0.1% are above the new European limit (< 1000 CFU/L).

Note that the US EPA criterion of annual probability of infection was used throughout the study instead of disability adjusted life years (DALY) recommended by the WHO. Indeed, the current method of estimation of the DALY for *Legionella pneumophila* is still controversial to our knowledge. DALY can be computed as

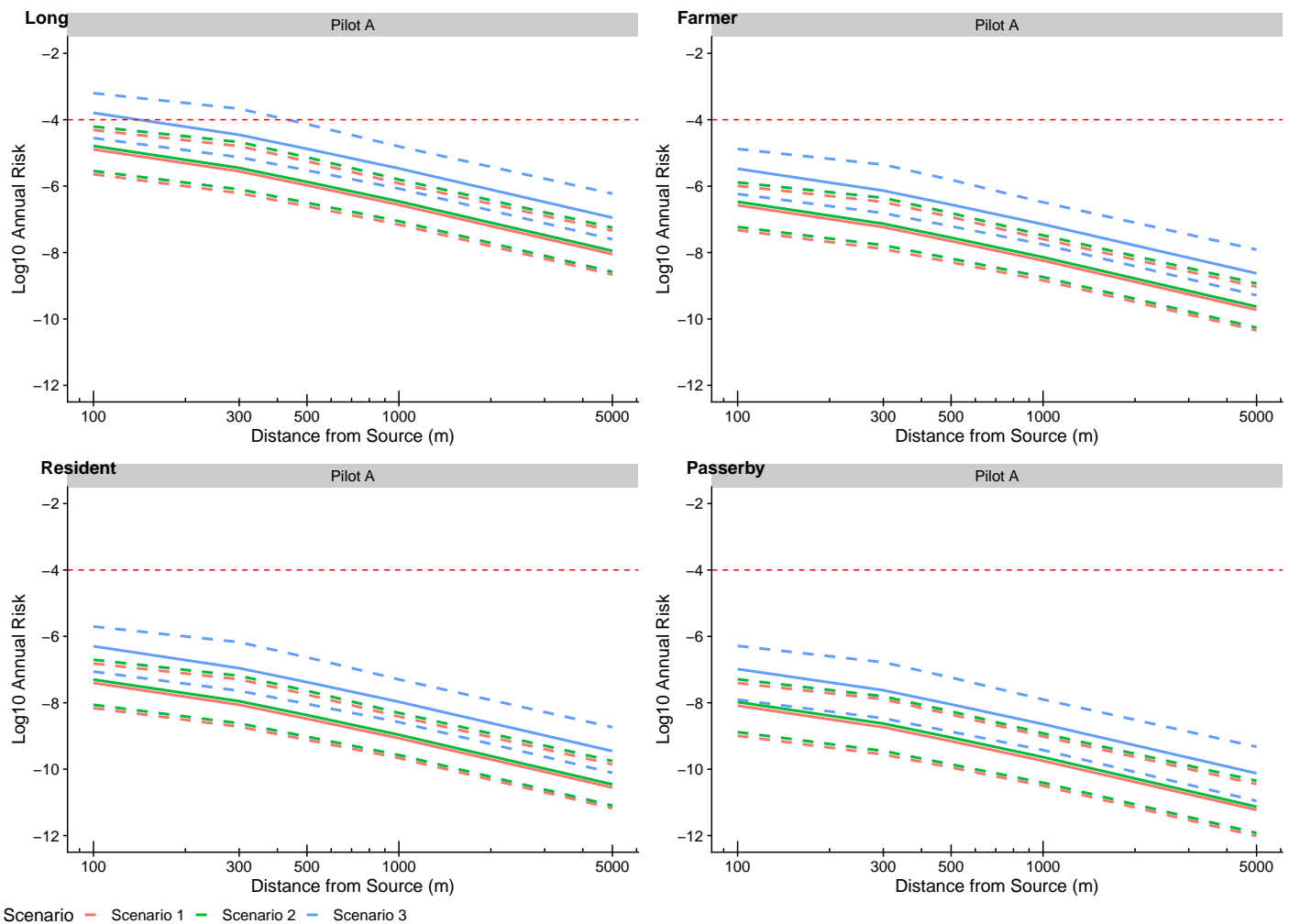


Fig. 12 Log₁₀ annual infection risks for *L. pneumophila* resulting from scenarios 1, 2 and 3. The red dotted line represents the US EPA threshold of -4 log annual probability of infection. Toy population and farmers are represented from left to right on the first line and residents and passerby on the second line.

0.97 times annual probability of infection with low precision⁷¹ and a recent study⁷² taking heterogeneity of responses to contamination into account could be used to derive a better estimation of DALY.

It is notable that the upper bound of the interdeciles credibility intervals of the posterior never exceeded the target risk values for infection (10^{-4} annual probability of infection for the US EPA criterion and 10^{-6} disability adjusted life year per person per year, i.e. DALY pppy for the WHO criterion). Although, a slow decrease of risk is observed with the distance to the source which is the consequence of the meteorological conditions in the vicinity of the agricultural plots. This observation is consistent with previous studies that have predicted long-range transport of *Legionella*.^{19,22,73} In the local climatic context of Tarbes where the decrease is slow it seems essential that the concentration have to be already low (below 10^4) at 100m from the sprinklers.

The developed tool is ready to be updated especially by air measured concentrations on the plots by the irrigations with reclaimed water that will take place starting in Summer 2021. This tool will then be used to monitor the potential risks in the vicinity

of the experimental plots and thus meet the public health demands of population protection.

In a more general QMRA context, the proposed model gives an operational tool, and theoretically stable methodology to maintain a continuous monitoring of the risks induced by the irrigation practices on the agricultural plots under study. This methodology can adapt to new sets of data and easily update the model if given new information. This model can easily be adapted to other well known pathogen through the enlightened adaptation of a few priors such as the ones on the decay parameter λ , or the dose-response parameter r , ...

A challenge would lie in the usage of the model to answer simultaneous multipathogen risk analysis. Also, one of the next main interest lies in the addition of a modelisation of the impact of the irrigation with reclaimed wastewater on the quality of the groundwater to simultaneously quantify the inhalation induced risk and the groundwater table contamination by reclaimed wastewater transport in soil. Furthermore, it could be very interesting to link the final probability of infection over the year to epidemiological data over the region as done by Albert, *et al.*

(2008)³⁶ linking an epidemiological study of campylobacteriosis cases to the probability of suffering from campylobacteriosis over one year in France due to broiler meat. In the same spirit, the link would necessitate here to introduce a random variable for the water reuse attributable fraction of *Legionella* infection to take other routes of contaminations into account. Also it would necessitate to consider only a fraction of the cases that could be due to this irrigation zone. But for the moment, the epidemiological data at disposal are only available at a large geographical scale (surface of Occitanie region is approximately 70,000 km² compared to the surface of the study zone which is approximately 6 km²) and it does not seem reasonable to link this data to the very local modeling proposed.

5 Conclusions

- A general Bayesian network as an operational tool to maintain a monitoring of the risks induced by a water reuse irrigation practice.
- First developed general Bayesian networks approach in QMRA for wastewater surveillance. The main asset is that risks are quantified with their uncertainties at each desired time taking into account new and past data.
- The QMRA model describes the exposure pathway of the water contamination from the entrance of the WWTP to the infection risk using pathogen decay models in the WWTP, a Gaussian plume model for the air contamination, an inhalation model for the population exposure and finally a dose-response model for risk characterization.
- The uncertainty of the probabilistic QMRA model is reduced by the introduction of observed data (pathogens concentrations and regional meteorological data) along the modelisation to obtain the distribution of the number of pathogens before and after treatment in WWTP, of the number of aerosolised pathogens, of the concentration of pathogens in the air at different distances, and finally the exposure and illness distributions for three different categories of population at risk and two illness endpoints.
- *Legionella* annual subclinical infection risk and annual clinical severity infection risk linked to the agricultural irrigation using the groundwater table or two experimental water treatment pilots are inferred below 10⁻⁴ tolerable limit defined by the US EPA for farmers, passersby and residents at distances between 100 m and 1000 m away from sprinklers.
- Such dynamic approach can be applied to various pathogens in the context of wastewater reuse

Conflicts of interest

There are no conflicts to declare.

Acknowledgment

The partners of the project are Polymem, Bio-UV, Ecofilae, VERI, Veolia Eau, Sede, and Inrae-Transfert Narbonne.

Established with the authorization of the Tarbes-Lourdes-Pyrénées Syndicate, owner of the treatment plant, this project involves farmers as well as the Departmental Federation of Farmers' Unions of Hautes-Pyrénées (FDSEA 65) and the Chamber of farming Hautes-Pyrénées.

The project is FUI funding, co-funded by the Sud-West region and Occitania region (France).

Notes and references

- 1 Vo PT, Ngo HH, Guo W, Zhou JL, Nguyen PD, Listowski A, et al. A mini-review on the impacts of climate change on wastewater reclamation and reuse. *Science of The Total Environment*. 2014;494-495:9-17. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969714009632>.
- 2 Elbasiouny H, El-Ramady H, Elbehiry F. Sustainable and Green Management of Wastewater Under Climate Change Conditions. *The Handbook of Environmental Chemistry*. 2021:1-19. Available from: <https://doi.org/10.1007/978-2021-787>.
- 3 Guest JS, Skerlos SJ, Barnard JL, Beck MB, Daigger GT, Hilger H, et al. A New Planning and Design Paradigm to Achieve Sustainable Resource Recovery from Wastewater. *Environmental Science & Technology*. 2009;43(16):6126-30. Available from: <https://doi.org/10.1021/es9010515>.
- 4 Fito J, Van Hulle SW. Wastewater reclamation and reuse potentials in agriculture: towards environmental sustainability. *Environment, Development and Sustainability*. 2021;23(3):2949-72.
- 5 Hristov J, Barreiro-Hurle J, Salputra G, Blanco M, Witzke P. Reuse of treated water in European agriculture: Potential to address water scarcity under climate change. *Agricultural Water Management*. 2021;251:106872. Available from: <https://www.sciencedirect.com/science/article/pii/S0378377421001372>.
- 6 EP. Water Reuse - Setting Minimum Requirements. European Parliament. 2020.
- 7 Food and Agriculture Organization of the United Nations. The state of the world's land and water resources for food and agriculture: Managing systems at risk. Earthscan; 2011.
- 8 Adegoke AA, Amoah ID, Stenström TA, Verbyla ME, Mihelcic JR. Epidemiological evidence and health risks associated with agricultural reuse of partially treated and untreated wastewater: a review. *Frontiers in Public Health*. 2018;6:337.
- 9 Bartram J, Chartier Y, Lee JV, Pond K, Surman-Lee S. *Legionella and the prevention of legionellosis*. World Health Organization; 2007.
- 10 Organization WH, WHO. Guidelines for drinking-water quality. vol. 1. world health organization; 2004.
- 11 Organization WH. Guidelines for safe recreational water environment. vol. 1. World Health Organization; 2006.
- 12 Organization WH, et al. WHO guide to ship sanitation. World Health Organization; 2011.
- 13 EC. Water Reuse. European Commission. 2021.
- 14 Courault D, Albert I, Perelle S, Fraisse A, Renault P, Salemk-

- our A, et al. Assessment and risk modeling of airborne enteric viruses emitted from wastewater reused for irrigation. *Science of the Total Environment*. 2017;592:512-26.
- 15 Paez-Rubio T, Viau E, Romero-Hernandez S, Peccia J. Source bioaerosol concentration and rRNA gene-based identification of microorganisms aerosolized at a flood irrigation wastewater reuse site. *Applied and Environmental Microbiology*. 2005;71(2):804-10.
 - 16 Teltsch B, Katzenelson E. Airborne enteric bacteria and viruses from spray irrigation with wastewater. *Applied and Environmental Microbiology*. 1978;35(2):290-6.
 - 17 ANSES. Avis et rapport d'expertise : Réutilisation des eaux usées traitées pour l'irrigation des cultures, l'arrosage des espaces verts par aspersion et le lavage des voiries. Saisine n°2009-SA-0329. 2012.
 - 18 Blanky M, Sharaby Y, Rodríguez-Martínez S, Halpern M, Friedler E. Greywater reuse - Assessment of the health risk induced by *Legionella pneumophila*. *Water Research*. 2017;125:410-7. Available from: <https://www.sciencedirect.com/science/article/pii/S0043135417307212>.
 - 19 Hamilton KA, Hamilton MT, Johnson W, Jjemba P, Bukhari Z, LeChevallier M, et al. Health risks from exposure to *Legionella* in reclaimed water aerosols: Toilet flushing, spray irrigation, and cooling towers. *Water Research*. 2018;134:261-79. Available from: <https://www.sciencedirect.com/science/article/pii/S0043135417310175>.
 - 20 Hamilton KA, Haas CN. Critical review of mathematical approaches for quantitative microbial risk assessment (QMRA) of *Legionella* in engineered water systems: research gaps and a new framework. *Environ Sci: Water Res Technol*. 2016;2:599-613.
 - 21 Diederer B. *Legionella* spp. and Legionnaires' disease. *J Infect*. 2008;56(1):1-12.
 - 22 Walser SM, Gerstner DG, Brenner B, Höller C, Liebl B, Herr CE. Assessing the environmental health relevance of cooling towers—a systematic review of legionellosis outbreaks. *International journal of hygiene and environmental health*. 2014;217(2-3):145-54.
 - 23 Xu P, Zhang C, Mou X, Wang XC. Bioaerosol in a typical municipal wastewater treatment plant: concentration, size distribution, and health risk assessment. *Water Science and Technology*. 2020;82:1547-59.
 - 24 Loenenbach AD, Beulens C, Euser SM, van Leuken JP, Bom B, van der Hoek W, et al. Two community clusters of Legionnaires' Disease directly linked to a biologic wastewater treatment plant, the Netherlands. *Emerging infectious diseases*. 2018;24(10):1914.
 - 25 Tang JW. The effect of environmental parameters on the survival of airborne infectious agents. *Journal of the Royal Society Interface*. 2009;6(suppl_6):S737-46.
 - 26 Haas CN, Rose JB, Gerba CP. Quantitative microbial risk assessment. John Wiley & Sons; 2014.
 - 27 Janevska DP, Gospavic R, Pacholewicz E, Popov V. Application of a HACCP-QMRA approach for managing the impact of climate change on food quality and safety. *Food Research International*. 2010;43(7):1915-24. Available from: <https://www.sciencedirect.com/science/article/pii/S0963996910000438>.
 - 28 Petterson S, Ashbolt N. QMRA and water safety management: review of application in drinking water systems. *Journal of Water and Health*. 2016;14(4):571-89.
 - 29 Owens CEL, Angles ML, Cox PT, Byleveld PM, Osborne NJ, Rahman MB. Implementation of quantitative microbial risk assessment (QMRA) for public drinking water supplies: Systematic review. *Water Research*. 2020;174:115614. Available from: <https://www.sciencedirect.com/science/article/pii/S0043135420301500>.
 - 30 Mara D, Sleigh P, Blumenthal U, Carr R. Health risks in wastewater irrigation: comparing estimates from quantitative microbial risk analyses and epidemiological studies. *Journal of water and health*. 2007;5(1):39-50.
 - 31 Wang Rn, Li X, Yan C. Seasonal fluctuation of aerosolization ratio of bioaerosols and quantitative microbial risk assessment in a wastewater treatment plant. *Environmental Science and Pollution Research*. 2021:1-18.
 - 32 Simhon A, Pileggi V, Flemming CA, Lai G, Manoharan M. Norovirus risk at a golf course irrigated with reclaimed water: Should QMRA doses be adjusted for infectiousness? *Water Research*. 2020;183:116121.
 - 33 Carducci A, Donzelli G, Cioni L, Verani M. Quantitative microbial risk assessment in occupational settings applied to the airborne human adenovirus infection. *International journal of environmental research and public health*. 2016;13(7):733.
 - 34 Yan C, Gui Zc, Wu Jt. Quantitative microbial risk assessment of bioaerosols in a wastewater treatment plant by using two aeration modes. *Environmental Science and Pollution Research*. 2021;28(7):8140-50.
 - 35 Zhiteneva V, Hübner U, Medema GJ, Drewes JE. Trends in conducting quantitative microbial risk assessments for water reuse systems: A review. *Microbial Risk Analysis*. 2020;16:100132. Available from: <https://www.sciencedirect.com/science/article/pii/S2352352220300384>.
 - 36 Albert I, Grenier E, Denis JB, Rousseau J. Quantitative risk assessment from farm to fork and beyond: A global Bayesian approach concerning food-borne diseases. *Risk Analysis: An International Journal*. 2008;28(2):557-71.
 - 37 Smid J, Verloo D, Barker G, Havelaar A. Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment. *International Journal of Food Microbiology*. 2010;139:S57-63.
 - 38 Nielsen TD, Jensen FV. Bayesian networks and decision graphs. Springer Science & Business Media; 2009.
 - 39 Scutari M, Denis JB. Bayesian Networks: With Examples in R (2nd ed.). Chapman and Hall/CRC; 2022.
 - 40 Rigaux C, Ancelet S, Carlin F, Nguyen-thé C, Albert. Inferring an augmented Bayesian network to confront a complex quantitative microbial risk assessment model with durability

- studies: application to *Bacillus cereus* on a courgette purée production chain. *Risk Analysis*. 2013;33(5):877-92.
- 41 Beaudouin D, Harden F, Roiko A, Mengersen K. Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. *Science of The Total Environment*. 2016;541:1393-409. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969715308469>.
 - 42 D B, F H, Roiko A SH, C L, K M. Modelling microbial health risk of wastewater reuse: A systems perspective. *Environ Int*. 2015;84:131-41.
 - 43 Aguilera PA, Fernández A, Fernández R, Rumí R, Salmerón A. Bayesian networks in environmental modelling. *Environmental Modelling & Software*. 2011;26(12):1376-88.
 - 44 Korb KB, Nicholson AE. *Bayesian artificial intelligence*. CRC press; 2010.
 - 45 Beaudouin D, Harden F, Roiko A, Stratton H, Lemckert C, Mengersen K. Beyond QMRA: Modelling microbial health risk as a complex system using Bayesian networks. *Environment International*. 2015;80:8-18. Available from: <https://www.sciencedirect.com/science/article/pii/S0160412015000719>.
 - 46 Beaudouin D, Harden F, Roiko A, Mengersen K. Potential of Bayesian networks for adaptive management in water recycling. *Environmental Modelling & Software*. 2017;91:251-70. Available from: <https://www.sciencedirect.com/science/article/pii/S1364815217300919>.
 - 47 Herrera-Murillo J, Mora-Campos D, Salas-Jimenez P, Hidalgo-Gutierrez M, Soto-Murillo T, Vargas-Calderon J, et al. Wastewater Discharge and Reuse Regulation in Costa Rica: An Opportunity for Improvement. *WATER*. 2021 Oct;13(19).
 - 48 Zhiteneva V, Carvajal G, Shehata O, Hübner U, Drewes JE. Quantitative microbial risk assessment of a non-membrane based indirect potable water reuse system using Bayesian networks. *Science of The Total Environment*. 2021;780:146462. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969721015308>.
 - 49 Pouillot R, Albert I, Cornu M, Denis J. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. *Int J Food Microbiol*. 2003;81(2).
 - 50 Højsgaard S. Graphical independence networks with the gRain package for R. *Journal of Statistical Software*. 2012;46(10):1-26.
 - 51 Albert I, Espié E, de Valk H, Denis JB. A Bayesian evidence synthesis for estimating campylobacteriosis prevalence. *Risk Analysis: An International Journal*. 2011;31(7):1141-55.
 - 52 Gilks WR, Richardson S, Spiegelhalter D. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Interdisciplinary Statistics. Taylor & Francis; 1995.
 - 53 Verbyla ME, Symonds EM, Kafle RC, Cairns MR, Iriarte M, Mercado Guzman A, et al. Managing microbial risks from indirect wastewater reuse for irrigation in urbanizing watersheds. *Environmental science & technology*. 2016;50(13):6803-13.
 - 54 Palusińska-Szyszk M, Cendrowska-Pinkosz M. Pathogenicity of the family Legionellaceae. *Archivum immunologiae et therapeuticae experimentalis*. 2009;57(4):279-90.
 - 55 Cianciotto NP. Pathogenicity of *Legionella pneumophila*. *International Journal of Medical Microbiology*. 2001;291(5):331-43.
 - 56 Seinfeld JH, Pandis SN. *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons; 2016.
 - 57 US EPA. *Exposure Handbook*. US Environmental Protection Agency. 2011.
 - 58 Armstrong TW, Haas CN. A Quantitative Microbial Risk Assessment Model for Legionnaires' Disease: Animal Model Selection and Dose-Response Modeling. *Risk Analysis*. 2007;27(6):1581-96. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2007.00990.x>.
 - 59 Molle B, Huet L, Tomas S, Granier JM, Dimaiolo P, Rosa C. Caractérisation du risque de dérive et d'évaporation d'une gamme d'aspenseurs d'irrigation. Application à la définition des limites d'utilisation de l'aspersion en réutilisation d'eaux usées traitées [Ph.D. thesis]. Irstea; 2009.
 - 60 Paez-Rubio T, Ramarui A, Sommer J, Xin H, Anderson J, Pecchia J. Emission rates and characterization of aerosols produced during the spreading of dewatered class B biosolids. *Environmental science & technology*. 2007;41(10):3537-44.
 - 61 Hambleton P, Broster M, Dennis P, Henstridge R, Fitzgeorge R, Conlan J. Survival of virulent *Legionella pneumophila* in aerosols. *Epidemiology & Infection*. 1983;90(3):451-60.
 - 62 Dennis P, Lee J. Differences in aerosol survival between pathogenic and non-pathogenic strains of *Legionella pneumophila* serogroup 1. *Journal of Applied Bacteriology*. 1988;65(2):135-41.
 - 63 Stellacci P, Liberti L, Notarnicola M, Haas CN. Hygienic sustainability of site location of wastewater treatment plants: A case study. I. Estimating odour emission impact. *Desalination*. 2010;253(1-3):51-6.
 - 64 Brooks JP, McLaughlin MR, Gerba CP, Pepper IL. Land application of manure and class B biosolids: An occupational and public quantitative microbial risk assessment. *Journal of Environmental Quality*. 2012;41(6):2009-23.
 - 65 Muller D, Edwards ML, Smith DW. Changes in iron and transferrin levels and body temperature in experimental airborne legionellosis. *Journal of Infectious Diseases*. 1983;147(2):302-7.
 - 66 Fitzgeorge R, Baskerville A, Broster M, Hambleton P, Dennis P. Aerosol infection of animals with strains of *Legionella pneumophila* of different virulence: comparison with intraperitoneal and intranasal routes of infection. *Epidemiology & Infection*. 1983;90(1):81-9.
 - 67 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria; 2020. Available from: <https://www.R-project.org/>.
 - 68 Plummer M. rjags: Bayesian Graphical Models using MCMC;

2019. R package version 4-10. Available from: <https://CRAN.R-project.org/package=rjags>.
- 69 Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd international workshop on distributed statistical computing. vol. 124. Vienna, Austria.; 2003. p. 1-10.
- 70 EPA U. Potable reuse compendium. EPA 810-R-17-002 2017. Office of Ground Water and Drinking Water: Washington ...; 2017.
- 71 Hamilton KA, Hamilton MT, Johnson W, Jjemba P, Bukhari Z, LeChevallier M, et al. Risk-Based Critical Concentrations of *Legionella pneumophila* for Indoor Residential Water Uses. *Environmental Science & Technology*. 2019 Apr;53(8):4528-41. Publisher: American Chemical Society. Available from: <https://doi.org/10.1021/acs.est.8b03000>.
- 72 Weir MH, Mraz AL, Mitchell J. An Advanced Risk Modeling Method to Estimate Legionellosis Risks Within a Diverse Population. *Water*. 2020;12(1). Available from: <https://www.mdpi.com/2073-4441/12/1/43>.
- 73 Borgen K, Aaberge I, Werner-Johansen Ø, Gjørund K, Størnsrud B, Haugsten S, et al. A cluster of Legionnaires' disease linked to an industrial plant in southeast Norway, June-July 2008. *Eurosurveillance*. 2008;13(38):18985.