DOI: 10.1111/risa.14030

ORIGINAL ARTICLE

Designing optimal food safety monitoring schemes using Bayesian network and integer programming: The case of monitoring dioxins and DL-PCBs

Z. Wang[1,2] | **H. J. van der Fels-Klerx**[1,3] | **A. G. J. M. Oude Lansink**[1]

[1]Business Economics, Wageningen University, Wageningen, The Netherlands

[2]College of Economics and Management, Huazhong Agricultural University, Wuhan, China

[3]Wageningen Food Safety Research, Wageningen, The Netherlands

Correspondence
H. J. van der Fels-Klerx, Business Economics, Wageningen University, 6700 EW Wageningen, The Netherlands.
Email: ine.vanderfels@wur.nl

Abstract

Efficient food safety monitoring should achieve optimal resource allocation. In this article, a methodology is presented to optimize the use of resources for food safety monitoring aimed at identifying noncompliant samples and estimating background level of hazards in food products. A Bayesian network (BN) model and an optimization model were combined in a single framework. The framework was applied to monitoring dioxins and dioxin-like polychlorinated biphenyls (DL-PCBs) in primary animal-derived food products in the Netherlands. The BN model was built using a national dataset with monitoring results of dioxins and DL-PCBs in animal-derived food products over a 10-year period (2008–2017). These data were used to estimate the probability of detecting suspect samples with dioxins and DL-PCBs levels above preset thresholds, given certain sample conditions. The results of the BN model were then inserted into the optimization model to compute an optimal monitoring scheme. Model estimates showed that the probability of dioxins and DL-PCBs exceeding threshold limits was higher in laying hen eggs and sheep meat than in other animal-derived food (except deer meat). Compared with the monitoring scheme used in the Netherlands in 2018, the optimal monitoring scheme would save around 10,000 EUR per year. This could be obtained by reallocating monitoring resources from products with lower probability of dioxin and DL-PCBs exceeding threshold limits (e.g., pig meat) to products with higher probability (e.g., bovine animal meat), and by shifting sample collection from the last quarter of the year toward the first three quarters of the year.

KEYWORDS
Bayesian network, food safety economics, food safety monitoring, optimization, sampling

1 | INTRODUCTION

Dioxins (polychlorinated dibenzo-p-dioxins and dibenzofurans, PCDD/Fs) and dioxin-like polychlorinated biphenyls (DL-PCBs) have negative impacts on human health (Hoogenboom et al., 2015; Knutsen et al., 2018; Lascano et al., 2011). These hazards are produced as by-products of industrial processes and natural phenomena (Srogi, 2008). The main route of human exposure to dioxins and DL-PCBs is via consumption of food items of animal origin (Baars et al., 2004; Hoogenboom et al., 2015).

To reduce human exposure to dioxins via food consumption and to prevent from food safety incidents, the EU has set legal limits for the presence of dioxins and DL-PCBs in food, and has stipulated a control strategy including recommended sampling and analysis procedures (European Commission, 2006a, 2014, 2017, 2020). These have resulted in the implementation of monitoring schemes in various Member States. These monitoring schemes allow to estimate the background levels of dioxins and DL-PCBs in food, to identify noncompliant agricultural products with excessive dioxins concentrations (e.g., legal limits or limits set by risk managers) and to trace back to contamination sources (European Commission, 2006a, 2014, 2017; European Food Safety Authority, 2012). In order to describe the background contamination levels of dioxins and DL-PCBs in food on the EU market, Member States should perform random monitoring with a recommended minimum number of samples collected and analyzed yearly (European Commission, 2006a). EFSA (European Food Safety Authority) merges all monitoring data from different countries to evaluate food contamination levels and to estimate the dioxin exposure of the European population (European Food Safety Authority, 2012).

Financial resources for food safety monitoring are generally limited and the analytical methods to detect dioxins and DL-PCBs are costly and time consuming (Lavric et al., 2005; Powell, 2014). In order to allocate financial resources for monitoring in an optimal way, several modeling studies have explored ways to improve the cost-effectiveness of food safety monitoring. Focker et al. (2018) reviewed methods of determining the cost-effectiveness of monitoring plans for chemical and biological hazards in the life sciences. Their results showed that most studies in food safety monitoring were based on optimization models and simulation models. For instance, Alban et al. (2016) used stochastic scenario tree modeling approach to simulate two sampling strategies in the monitoring of antimicrobial residues in Danish finishing pigs. Lascano-Alcoser et al. (2013) and Lascano-Alcoser et al. (2014) optimized monitoring schemes in milk and pork supply chains, focusing on maximizing the probability of identifying dioxins contamination given preset financial resources for sampling and analysis. Recently, some new models have been developed to optimize food safety monitoring from cost-effectiveness perspective: Focker et al. (2019b) optimized the costs and accuracy of aflatoxin monitoring in the maize supply chain by using nonlinear programming, and Wang et al. (2020) optimized the monitoring of multiple chemical hazards along the dairy supply chain to reduce potential public health impacts. Wang et al. (2021) optimized spatial sampling strategies of dioxins monitoring along the Dutch dairy supply chain by combing stochastic simulation and linear programming. However, most optimization models of food safety monitoring were developed based on simulated input data with an assumption, an approximation, or a calculation, which may not reflect reality of the historical data (Focker et al., 2018). In addition, interactions among different risk factors should be considered when optimizing food safety monitoring. The Bayesian network (BN) demonstrates a good prediction accuracy with fast response and can account for interactions of uncertainty by using causal relationships between different random variables (Namazian et al., 2019; Uusitalo, 2007). Previous studies showed the power of BN modeling in predicting types and occurrence of food safety hazards under different conditions using historical monitoring data and expert opinions (Bouzembrak & Marvin, 2019; Bouzembrak & van der Fels-Klerx, 2017). However, so far, BN modeling has rarely been applied with a mathematical programming model to solve the resource optimization problem in food safety monitoring.

The objective of this study was to develop a framework for optimizing food safety monitoring schemes to reduce monitoring costs while guaranteeing the identification of non-compliant samples. The proposed framework was illustrated using the case of the monitoring schemes for dioxins and DL-PCBs in animal-derived food products.

2 | METHODOLOGY

2.1 | The design of framework combing BN and IP models

2.1.1 | Bayesian network

The BN is a graphical model based on Bayesian statistics, decision theory, and graphical theory, which consists of nodes (i.e., random variables) connected by directed arcs. The BN can estimate causal relationships (reflected by directed arcs) between nodes or events as qualitative parameters, and inference conditional probability values between nodes or events as quantitative parameters. The Bayesian formula presents the probability of event X_i under the condition that event X_j occurs (posterior probability, $P(X_i|X_j)$):

$$P\left(X_i|X_j\right) = \frac{P\left(X_i\right) P(X_j|X_i)}{P\left(X_j\right)}, \tag{1}$$

where $P(X_i)$ is the prior probability of event X_i, $P(X_j|X_i)$ is the conditional probability of X_j under the condition of a known event X_i, and $P(X_j)$ is the probability ofX_j. A BN model contains (1) a set of variables (nodes) $U = \{X_i \ldots X_n\}$ and a set of directed arcs between variables; (2) a finite set of states for each discrete variable; and (3) a set of conditional and unconditional probabilities. If there is an arc from node X_i to node X_j, the node X_i is called the parent of node X_j and the node X_j is called the child of node X_i.

Equation 2 presents the BN formula, which is the joint probability distribution of all variables, $P(U) = P(X_i, \ldots\ldots, X_n)$, given by the product of all conditional probability tables specified in BN:

$$P\left(U\right) = \Pi^{n}_{(i=1)} P\left(X_i|pa\left(X_i\right)\right), \tag{2}$$

where $P(X_i|pa(X_i))$ is the probability of event X_i under the condition of a known parent $pa(X_i)$ of variable X_i.

2.1.2 | Food safety monitoring costs minimized by mathematical programming

Optimization of food safety monitoring from cost-effectiveness perspective is a relevant topic of food safety economics. These problems aim to minimize monitoring costs or maximize effectiveness given some constraints and can be formulated by mathematical programming (Focker et al., 2019b; Lascano-Alcoser et al., 2013; Lascano-Alcoser et al., 2014; Wang et al., 2020). In this study, we use integer programming (IP) to minimize food safety monitoring costs by given certain constraints, and generally IP problem could be expressed as follows:

$$\text{Objective function Min}: f\,(x_1 \ldots\ldots x_n)\,, \tag{3}$$

$$\text{Constraints}\boldsymbol{G}\,(x_1 \ldots\ldots x_n) \leq \boldsymbol{b}, \tag{4}$$

$$x_{\min} \leq x_1 \ldots\ldots x_n \leq x_{\max}, \tag{5}$$

$$x_1 \ldots\ldots x_n \in Z^n, \tag{6}$$

where $f(x_1 \ldots\ldots x_n): R^n \to R$ is the objective function to be optimized; $x_1 \ldots\ldots x_n$ are decision variables and $x_1 \ldots\ldots x_n \in Z^n$ means $x_1 \ldots\ldots x_n$ are integer variables; $\boldsymbol{G}(x_1 \ldots\ldots x_n)$: $R^m \to R$ are linear or nonlinear functions and constraints vector; $\boldsymbol{b}$: $R^m \to R$ is a rational vector; $x_{min}, x_{max} \in R^n$ are the lower and the upper bound of decision variables. The optimal results of IP model highly depend on the reliability of parameters in $f(x_1 \ldots\ldots x_n)$, $\boldsymbol{G}(x_1 \ldots\ldots x_n)$, $x_{\min}$ and $x_{\max}$. These parameters are most of time subjected to uncertainties and provided by simulated data and historical data.

We combined a BN model and an IP model into a framework with the aim of minimizing monitoring costs under a set of constraints. The BN has the ability to integrate different types of data, such as expert knowledge, analytical results, and feedback of previous experiences, and to build the conditional dependencies (casual relationships) among different random variables (Buriticá & Tesfamariam, 2015; Yazdi & Kabir, 2017). The inferred results of BN would be used for parameters in IP, and IP model would be applied to the monitoring costs optimization problem. The BN could not only capture the priori probability distribution of uncertainties of parameters in IP model, but also take into account possible occurrence of some events whose observation may give new information of actual values of parameters in IP model. Figure 1 presents the proposed modeling framework in which BN and IP models are combined to optimize food safety monitoring schemes.

2.2 | Empirical application of the monitoring of dioxins and DL-PCBs

2.2.1 | Monitoring scheme of dioxins and DL-PCBs

The modeling framework was applied to 10 years of historical data (2008–2017) on dioxins and DL-PCBs in agricultural products in the Netherlands (WFSR, 2019). Table 1 presents an example record from the dataset with results of the national monitoring program of dioxins and DL-PCBs. In the monitoring scheme, samples were randomly collected each quarter at the primary production or processing stage (e.g., at farms or in slaughterhouses). We assumed that all individual samples were first screened with the DR CALUX® method without pooling and compared with threshold limits (Adamse et al, 2017). Threshold limits were set by responsible bodies and they were lower than the corresponding EC legal limits. In practice, the concentration of the reference sample was used for comparison and as threshold limit for example, 0.9 pg TEQWHO1998 g−1 for ruminant and poultry meat, 0.5 pg TEQWHO1998 g−1 for pork meat and 1.9 pg TEQWHO1998 g−1 for milk and eggs (upper bound levels) (Adamse et al., 2017). Samples suspected (samples with dioxins and DL-PCBs contamination levels exceeding threshold limits) of dioxins or DL-PCBs contamination were further examined using Gas Chromatography/High-Resolution Mass Spectrometry (GC/HRMS) to determine the concentration with high accuracy (European Commission, 2006b, 2014, 2020). If GC/HRMS results exceeded the corresponding legal limits, samples were regarded as positive samples, and if GC/HRMS results did not exceed the corresponding legal limits, samples were regarded as negative samples. Additionally, within the framework of the EU monitoring of background levels of dioxins and DL-PCBs in different food products, a certain number of samples were randomly extracted each year from all Dutch individual samples and analyzed with GC/HRMS.

2.2.2 | Bayesian network

In this study, a BN model was developed based on historical data of the monitoring schemes of dioxins and DL-PCBs to predict the conational probability of suspect samples, that is, samples with dioxins and DL-PCBs contamination levels exceeding threshold limits, under different conditions. The BN model also estimated how many samples should be analyzed yearly for each food product according to EU monitoring. The BN model could analyze the relationships between the test results on the one hand and the sample conditions on the other hand. Sample conditions considered were food product type, animal species, control point, quarter of the year, and the total number of samples collected each year. Table 2 presents the names, descriptions, and states of these variables defined in the BN. Then, the structure of the BN was established by estimating the causal relationships and probabilities between the various variables from historical monitoring data (example in the Appendix B). The probability ($P\,(sc = 1|a,\, p,\, q)$) of suspect samples given different conditions was estimated by using the BN structure and Equations (1) and (2) (Table 3). The BN model was developed using 80% of the records of the dataset, randomly selected, and it was validated using the remaining 20% of the dataset. In this (internal) model validation step, the variables of each record were used as input in the BN model to predict the screening results as retrieved from the dataset. We assumed that the prediction was correct when the screening result class, as predicted by the BN model, was similar to the screening result (suspect vs. nonsuspect) as recorded in the dataset.

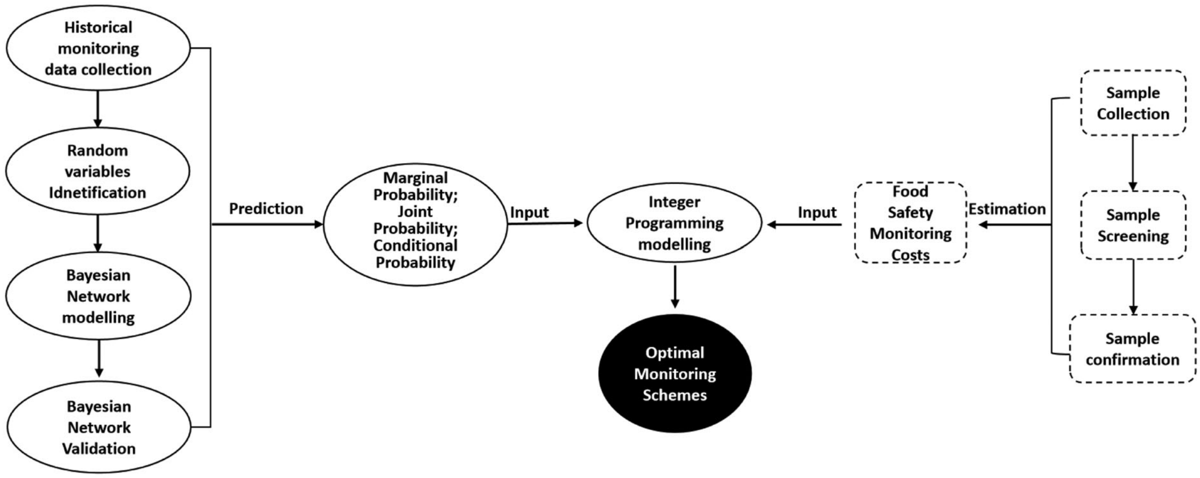

FIGURE 1 The framework combing Bayesian network and integer programming models to optimize food safety monitoring schemes

TABLE 1 Example records of the past monitoring schemes of dioxins and DL-PCBs[a] used for building Bayesian network

Year	Quarter [b]	Animal species [c]	Product	Control point	Screening results [d]	Sample size [e]	Sample sizeS[f]
2008	1	Hen	Egg	Farm	Nonsuspect	226	25
2009	1	Bovine animal	Milk	Farm	Nonsuspect	254	32
2010	1	Hen	Egg	Farm	Suspect	196	27
2011	3	Bovine animal	Meat	Slaughterhouse	Nonsuspect	352	34
2012	1	Broiler	Meat	Slaughterhouse	Nonsuspect	366	48
2013	1	Sheep	Meat	Slaughterhouse	Nonsuspect	425	56
2014	1	Sheep	Meat	Slaughterhouse	Suspect	340	42
f2015	3	Sheep	Meat	Slaughterhouse	Nonsuspect	358	20
2016	3	Pig	Meat	Slaughterhouse	Nonsuspect	379	28
2017	1	Sheep	Meat	Slaughterhouse	Suspect	365	18
2018	3	Poultry (other)g	Meat	Slaughterhouse	Nonsuspect	365	31
2018	1	Bovine animal	Meat	Slaughterhouse	Nonsuspect	365	31

[a]Raw data from (WFSR, 2019); processed data in Appendix B.
[b]Quarter: the quarter of the year (1 (January–March), 2 (April–June), 3 (July–September), 4 (October–December).
[c]Animal species in this study include hen, bovine animal, broiler, sheep, pig, deer and poultry (other).
[d]Screening Results: The results of samples analyzed by DR CALUX® method; "suspect" means the contamination exceeded the threshold limit and "nonsuspect" means the contamination in samples did not exceed the threshold limit.
[e]SampleSize: Total number of randomly collected samples annually.
[f]SamplesizeS: Number of annually collected samples with dioxins and DL-PCBs concentrations exceeding threshold limits.
[g]Poultry (other): poultry excluding hen and broiler.

2.2.3 | Integer programming

The objective of the IP model was to minimize the monitoring costs for dioxins and DL-PCBs in primary animal-derived food products in the Netherlands, subject to a set of constraints reflecting the required probability of the monitoring scheme identifying noncompliant samples, and the need for background level analysis over a monitoring period of 1 year. The P ($sc = 1 \mid a,p,q$) and the number of samples needed to determine the dioxin background level (estimated from BN) were used as inputs for this optimization step. The monitoring costs during the monitoring period, as expressed in Equation (7), were the sum of costs for identifying noncompliant samples (CNS_{apq}) and costs for estimating the background level of contamination (CBL_{ap}) in animal species a and food product type p during 1 year of EU monitoring. The decision variable was the number of samples collected in animal species a and food product type p, at quarter q of the year (ns_{apq}). Equation (8) presents the model constraint for the identification of noncompliant samples, estimated as the probability of the monitoring scheme to identify the noncompliant sample in animal species a and food product type

TABLE 2 Names, descriptions, and states of nodes (random variables) used in Bayesian network model

Nodes	Description	States
Year	The monitoring year	[2008, 2009, …,2018]
Quarter	The quarter of the year	[1 (January–March), 2 (April–June), 3 (July–September), 4 (October–December)]
Place	The control points	[farm, slaughterhouse]
Product	The food product type	[milk, meat, egg, liver]
SampleSize	The number of samples collected during the monitoring period	[100−200, 200−300, 300−400, 400−500]
AnimalSpecies	The animal species monitored	[bovine animal, hen, …, poultry (other)]
ScreeningResults	The results from the screening DR CALUX® method	[suspect, nonsuspect]
GCResults	The results from the GC/MS method	[above legal limit, below legal limit][a]
SampleSizeEU	The number of samples analyzed for EU monitoring to estimate background contamination in different products	[0, 1, …, 31]

[a]Samples with positive results were classified as above legal limit, and samples with negative and N.A. results were classified as below legal limit.

TABLE 3 Conditional probability for suspect screening results according to the conditions of the product (P (ScreeningResults = 1| AnimalSpecies = *a*, Product = *p*, Quarter = *q*, Place = *c*)) as inferenced by the Bayesian network

			Quarter (*q*)			
			1	2	3	4
Place (*c*)	**Animal Species (*a*)**	**Food product types (*p*)**	**Conditional probabilities**			
Farm	Bovine[b]	Milk	3%	4%	2%	1%
Slaughterhouse	Bovine	Meat	10%	11%	9%	8%
Slaughterhouse	Broiler	Meat	3%	2%	2%	0%
Slaughterhouse	Calf[b]	Meat	2%	1%	1%	1%
Slaughterhouse	Deer	Meat	93%	97%	100%	100%
Slaughterhouse	Pig	Meat	2%	2%	4%	1%
Slaughterhouse	Sheep	Meat	42%	50%	40%	29%
Farm	Hen	Egg	26%	28%	26%	22%
Slaughterhouse	Poultry (other)	Meat	0%	0%	0%	0%

[a]ScreeningResults: the results of screening method, the suspect results with the sign of 1; AnimalSpecies: certain animal species, *a*; Product: certain food product type, *p*; Quarter: the certain quarter of the year, *q*; Place: certain control point, c. According to Equation 1 and 2 and casual relationships between variables, the conditional probability could be calculated as: $P(\mathit{ScreeningResults} = 1|\ \text{AnimalSpecies} = a, \text{Product} = p, \text{Quarter} = q, \text{Place} = c)) = \frac{P(\text{AnimalSpecies=a,|ScreeningResults=1})*(\text{Product=p|ScreeningResults=1,AnimalSpecies=a})*\text{P(Quarter=q|ScreeningResults=1)}*\text{P(Place=c,|}\mathit{ScreeningResults}\text{=1)}*P\text{(ScreeningResults=1)}}{P\text{(AnimalSpecies=a,Product=p,Quarter=q,Place=c)}}$ The detailed calculation process was presented with R codes in Appendix B. These probability values pertain to DL-PCBs.
[b]Bovine: bovine animal; Calf: calf for fattening.

p. This probability was set to be larger than or equal to the required value, which was estimated from the probability of the 2018 monitoring scheme (estimated in Equations (15) and (16)) to identify corresponding noncompliant samples. Equation (9) presents the model constraint that a sufficient number of samples needs to analyzed to estimate the background level of dioxins and DL-PCBs contamination in animal species a and food product type p. The total number of samples collected during 1 year for animal species a and food product type p was set as smaller than or equal to the number of samples collected in the actual 2018 monitoring and larger than or equal to the minimum number of yearly required samples (Ns_min_{ap}). The BN estimated the probability distribution of the number of samples analyzed for estimating the background level of contamination in animal species a and food product type p (Ns_BL_{ap}) (Appendix A Table A1), and we used the most likely values of Ns_BL_{ap} as the input of Ns_min_{ap} (Appendix Table A1). Equation (10) presents the model constraint that the value of ns_{apq} should be larger than or equal to 0, and smaller than or equal to the maximum value (Ns_max_{ap}). Equations (11) and (12) present costs for identifying noncompliant samples and costs for estimating the background levels of contamination in animal species a and food product type p during 1 year of EU monitoring. Equations (13) and (14) are used for estimating the probability of monitoring scheme identifying the noncompliant sample in animal species a and food product type p. Equations (15) and (16) are used for estimating the probability of the 2018 monitoring scheme identifying the noncompliant sample in animal species a and food product type p.

Objective function:

$$\text{Min} : TMC\left(ns_{apq}\right) = \sum_{apq}\left(CNS_{apq} + CBL_{ap}\right)_{apq} \tag{7}$$

Constraints:

$$PI_{ap} \geq P\left(ns\_2018_{apq}\right)_{ap2018} \tag{8}$$

$$Sum\left(ns_{ap2018}\right)_{ap2018} \geq \sum_{q} ns_{apq} \geq Ns\_minap \tag{9}$$

$$Ns\_max_{ap} \geq ns_{apq} \geq 0 \tag{10}$$

Decision variable:
ns_{apq}, integer.
With

$$CNS_{apq} = ns_{apq} * Cs + ns_{apq} * Csc + ns_{apq} * Cgc * P(sc = 1|a, p, q) \tag{11}$$

$$CBL_{ap} = ns\_BL_{ap} * Cgc \tag{12}$$

$$PI_{ap} = \left(1 - \prod_{q=1}^{q=4} P\left(0,\ ns_{apq},\ P\left(sc = 1|a,\ p, q\right)\right)_{apq}\right) * Sen\_dio \tag{13}$$

$$P\left(0, ns_{apq},\ P\left(sc = 1|a, p, q\right)\right)_{apq} = \binom{ns_{apq}}{0} * P(sc = 1|a, p, q)^{0} * (1 - P\left(sc = 1|a, p, q\right))^{ns_{apq}} \tag{14}$$

$$P\left(ns_{apq2018}\right)_{ap2018} = 1 - \left(\prod_{q=1}^{q=4} P\left(0,\ ns_{apq,2018},\ P\left(sc = 1|a,\ p,\ q, 2018\right)\right)_{apq2018}\right) * Sen\_dio \tag{15}$$

$$P\left(0, ns_{apq2018},\ P\left(sc = 1|a, p,\ s,\ 2018\right)\right)_{apq2018} = \binom{ns_{apq2018}}{0} * P(sc = 1|a, p, q,\ 2018)^{0} * (1 - P\left(sc = 1|a, p, q, 2018\right))^{ns_{apq2018}} \tag{16}$$

where TMC (ns_{apq}) are the yearly total monitoring costs; a is animal species; p represents the food product type; q is the quarter of the year (1, 2, 3 or 4); ns_{apq} is the number of samples collected from animal species a and food product type p, during quarter q; $ns_{apq2018}$ is the number of samples collected from animal species a and food product type p, during quarter q of 2018; ns_BL_{ap} is the number of samples analyzed for estimating the background level of contamination in animal species a and food product type p during the monitoring period (refers to Appendix Table A1); PI_{ap} is the probability that the monitoring scheme can identify at least one noncompliant sample in 1 year in animal species a and food product type p.

$P(0, ns_{apq},\ P(sc = 1|a,p,s))_{apq}$ is the probability that the monitoring scheme cannot identify the suspect sample in animal species a and food product type p in quarter q, following a binomial distribution; $P(sc = 1|a,p,q)$ is the conditional probability of a suspect sample collected in animal species a and food product type p in quarter q, computed by the BN; $P(ns_{apq2018})_{ap2018}$ is the probability of identifying at least one non-compliant sample in the whole year using the current sampling plan; $P(0, ns_{apq2018},\ P(sc = 1|a, p, q, 2018))_{apq2018}$ is the probability of the 2018 monitoring scheme identifying no suspect sample in animal species a and food product type p, at quarter q of 2018, following a binomial distribution; Ns_max_{ap} is the maximum number of samples collected for animal species a and food product type p during the monitoring period, estimated from 2018 monitoring results; Ns_min_{ap} is the minimum number of samples collected estimated by the BN, equal to ns_BL_{ap}; Cs are the costs of collecting one sample; Csc are the costs of screening one sample with DR CALUX; Cgc are the costs of analyzing one sample with GC/MS; CNS_{apq} are the costs of identifying a noncompliant sample in animal species a and food product type p, at quarter q of the year; CBL_{ap} are the costs of estimating background contamination levels in animal species a and food product type p during the monitoring period; Sen_dio is the assumed sensitivity of DR CALUX® combined with GC/HRMS for the identification of non-compliant samples (Table 4).

2.3 | Computing tool

The BN model was developed to obtain the parameter values for the optimization model using the program R, version 3.5.2(R Development Core Team, 2018) with the R package “bnlearn” (Scutari, 2009). The correponding R codes are displayed in Appendix B. The integer programming model was developed in Lingo software, version 17.0×64(Lindo Systems Inc, 2017). The respective codes are attached as Supporting Information.

3 | RESULTS

3.1 | Results of the BN model

After computing the best fitting BN model (network structure in the Appendix B), it is composed of (i) a set of nine

TABLE 4 Input variables in integer programming model to estimate the costs of the monitoring scheme of dioxins and DL-PCBs

Description	Variable	Value	Unit	Explanation
Costs of collecting one sample	CS	10	EUR	Reference value [a]
Costs of analyzing one sample by screening method	CsC	100	EUR	DR CALUX® [b]
Costs of analyzing one sample by confirmatory method	Cgc	350	EUR	GC/HRMS [b]
Sensitivity of dioxin analysis by combined method	Sen_dio	100	%	Assumed sensitivity of DR CALUX® combined with GC/HRMS in identifying positive samples [c]

[a] *Source*: (Focker et al., 2019a).
[b] *Source*: (Lascano-Alcoser et al., 2014).
[C] In reality the sensitivity is a bit lower than 100%, and we assumed it as 100% which could make the results give insight into the influence of sample allocation on budgets saving. The model is flexible in that this parameter can be adapted according to the requirements of different reference laboratories.

variables, namely: animal species, food product type, control point, year, quarter of the year, screening (DR CALUX®) results, confirmation (GC/MS) results, number of samples analyzed for estimating dioxin and DL-PCB background levels, and total sample size; (ii) a set of states for each variable (e.g., the node screening results has the following states: suspect and nonsuspect); and (iii) a set of direct links between the variables and an assigned conditional probability for each variable.

The BN model predicts the likelihood of excessive dioxins and DL-PCBs levels in different products given predefined conditions. Table 3 presents the conditional probabilities of positive screening results given the specific animal species, food product type, quarter of the year, preset number of samples, and control point (computed by Equation (1)). The results showed that dioxins and DL-PCBs in sheep meat and hen eggs are more likely to exceed their threshold limits than in other food product type–animal species combination (except deer meat with limited sample size) at each quarter of the year. All (other) poultry meat screening results were negative at every condition. Excessive levels of dioxins and DL-PCBs were more likely to be found in meat than in dairy milk. Except for deer meat, dioxins and DL-PCBs in most products were most likely to exceed threshold limits within the first three quarters of the year, as compared with the fourth quarter of the year.

3.2 | Results of the optimal monitoring scheme

Table 5 presents the samples collected under the 2018 monitoring scheme and applying the calculated optimal monitoring scheme, for different animal species and food product types at each quarter of the year, with their corresponding PIs. The PI2018 was estimated by the model and the optimal results were computed to minimize the total monitoring costs with the constraint that its PIs were larger than or equal to PI2018. In total, to reach the same monitoring performance, the total number of samples collected (274) using the optimal monitoring scheme was much smaller than the number actually collected (365) in the 2018 monitoring scheme. In practice, the majority of the samples (103) collected during the first quarter of 2018 and most of the samples collected in 2018 were from pig meat (98 samples in total); its PI2018 reached 91%. Even though 2 and 12 samples were collected for deer meat and sheep meat, respectively, in 2018, their PI2018 could reach 100%. Only a few samples were collected in the 2018 monitoring scheme from (other) poultry meat with four samples in total, and all of them contained subthreshold levels of dioxins and DL-PCBs. In the estimated optimal situation, most samples should be collected for bovine animal meat with 60 samples in total, thus eight samples less than with the 2018 monitoring scheme. Furthermore, more samples should be collected during the second quarter of the year, as compared to the 2018 monitoring scheme. The optimal monitoring scheme proposed to collect in total 58 samples of pig meat during the third quarter of the year, which are 40 samples fewer than with the 2018 monitoring scheme. The total number of samples for background level analysis in 2018(121) was similar to the estimated optimal number (116). However, with the optimized results, fewer samples are needed to estimate background contamination levels in bovine meat and pig meat, but more samples should be analyzed to estimate background contamination levels in milk.

Figure 2 presents the optimal allocation of monitoring costs and the estimated monitoring costs in 2018 in terms of animal species and food product types during each quarter of the year. It would require fewer resources to reach the same performance as the 2018 monitoring scheme. More than 10,000 EUR could be saved each year by implementing the optimized monitoring schemes. In the optimal situation, budget should be allocated to collect more samples in the first three quarters of the year to identify noncompliant samples. In accordance with Table 5, costs could mostly be saved in sampling pig meat as in the optimal monitoring compared with the 2018 monitoring scheme. Except deer meat, the optimal monitoring costs for each food products were allocated at earlier quarters (1, 2, and 3) of the year as compared with the monitoring costs spent in 2018.

TABLE 5 Sample allocation at different conditions in the optimal monitoring scheme of dioxins and DL-PCBs with a comparison of the sample allocation following the 2018 Dutch national monitoring scheme

		Number of samples in 2018 monitoring scheme [a]							Optimal number of samples						
		Quarter [b]							Quarter						
Animal species	**Food product types**	**1**	**2**	**3**	**4**	**Total [d]**	**Ns_bl2018 [e]**	**PI2018 [c]**	**1**	**2**	**3**	**4**	**TotalO [d]**	**Ns_blO [e]**	**PI [c]**
Hen	Egg	2	35	4	0	41	16	100%	2	32	5	0	39	16	100%
Bovine e	Meat	20	9	25	14	68	22	100%	7	51	1	1	60	16	100%
Broiler	Meat	25	22	19	14	80	26	80%	54	0	0	0	54	22	80%
Calf e	Meat	16	15	10	8	49	15	49%	38	1	1	0	40	14	49%
Deer	Meat	2	0	0	0	2	2	100%	0	0	0	1	1	1	100%
Pig	Meat	24	10	36	28	98	28	91%	0	0	58	0	58	20	91%
Sheep	Meat	8	0	2	2	12	6	100%	0	8	1	0	9	9	100%
Bovine	Milk	5	5	1	0	11	3	31%	5	5	1	0	11	16	31%
Poultry (other) e	Meat	1	0	2	1	4	3	0%	0	0	2	0	2	2	0%
Total		103	96	99	67	365	121		106	97	69	2	274	116	

[a] *Source*: (WFSR, 2019).
[b] Quarter: the quarter of the year.
[c] PI2018/PI: the probability that the monitoring scheme can identify at least one noncompliant sample through the monitoring year. The probabilities from both sides were set as equal for each animal species and food product type.
[d] "Total/TotalO": the total number of samples collected and analyzed yearly for both estimating background level of chemicals and identifying noncompliant products in 2018 or in the optimal monitoring design.
[e] Ns_bl2018/ns_blO: the number of samples analyzed yearly for estimating background level of chemicals in different food products in 2018 or in the optimal monitoring design.
[f] Bovine: bovine animal; Calf: calf for fattening; Poultry (other): excluding hen and broiler.

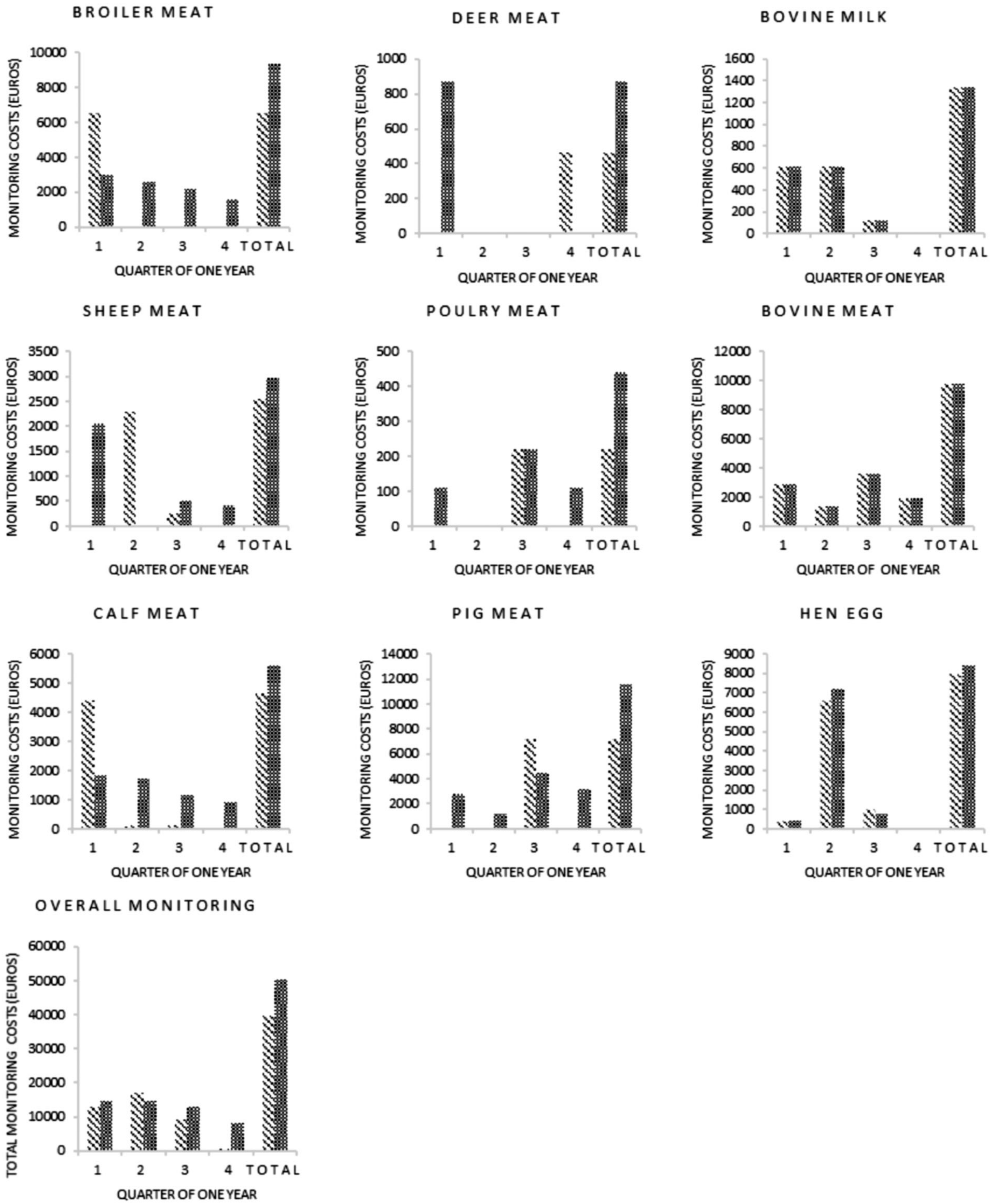

FIGURE 2 The optimal monitoring costs compared with monitoring costs in four quarters of the year 2018 with equal food safety monitoring performance

4 | DISCUSSION

This article presents a modeling approach to design optimal food safety monitoring schemes. With our approach, noncompliant food products from animal origin can be identified, given required conditions, while minimizing monitoring costs. The framework consists of a BN model and an IP model. The BN model estimates the states of each variable and the conditional probability of suspect samples given the required conditions. The optimization model uses the results of the BN as input to compute optimal monitoring schemes.

The optimization showed that it is possible to reduce the total monitoring costs by monitoring samples differently between food products and at different quarters of the year. For example, in the 2018 monitoring scheme, most samples were collected from pig meat, but in the optimal scheme most samples should be collected from bovine meat. For pig meat monitoring, collecting all 58 samples during the third quarter of the year would be more cost-effective (Table 5). The probability of a pig meat sample being suspect was 4% during the third quarter of the year, which is twice as high as during the first quarter of the year and four times higher than during the fourth quarter of the year. Thus, suspect pig samples could be more easily identified during the third quarter of the year.

Humans are exposed to dioxins and DL-PCBs mainly through food consumption of animal derived products with potential health effects (Baars et al., 2004; Hoogenboom et al., 2015), but the potency of the congeners among these two classes of chemicals are different. The International Agency for Research on Cancer (IARC) considered that DL-PCBs (e.g., PCB 126) and some dioxins (e.g., 2,3,7,8-TCDD) are carcinogenic to humans (Group 1) (Knutsen et al., 2018). Acute exposure of humans to high levels of dioxins may result into skin lesions and may damage liver functions, and chronic exposure is associated with impairment of the immune system, the endocrine system, and reproductive functions (Marinković et al., 2010). In order to compare the toxicity of a mixture of dioxins and DL-PCBs in facilitating risk assessment, the concept of toxic equivalents (TEQs) based on different toxic equivalency factors (TEFs) for the different dioxin and DL-PCB congeners was introduced, with TEFs expressing the toxicity of individual dioxins and DL-PCB congeners relative to the most toxic dioxin congener (being2,3,7,8-TCDD) (EFSA Panel on Contaminants in the Food Chain (CONTAM), 2012). The WHO foodborne disease burden epidemiology reference group (FERG) estimated the global burden of foodborne diseases, and in this context, the disease burden (with symptoms like hypothyroidism and male infertility) related to dioxins and DL-PCBs exposure was measured as long-time exposure for chronic toxicity risk (WHO, 2015). According to FERG reports, the Southeast Asia Region had the highest foodborne disease burden caused by dioxins with 10–41 DALYs (disability-adjusted life years)/100,000 population, and 0.9–19 DALYs/100,000 population were estimated for western Europe countries (Gibb et al., 2015). Although the background dietary exposure to dioxins in the Netherland currently does not pose risks to public health (Boon et al., 2014), elevated dioxin levels should be prevented to protect human health and to prevent from economic losses. In order to guarantee the safety of the food supply and protect human health, many national authorities monitor dioxins and DL-PCBs to estimate their background levels in food, and to identify noncompliant agricultural and food products with excessive dioxins concentrations. Consistent with these monitoring tasks, we set these two monitoring functions as main constraints in our optimization model with the objective to minimize related monitoring costs.

Mathematical programming models can solve optimization problems of food safety monitoring from the cost-effectiveness perspective (Focker et al., 2019b; Wang et al., 2020), but the model parameters are subject to uncertainties of some random events (e.g., the positive rate of samples, collecting places, product types, and so on). Combing a stochastic simulation process or using stochastic programming may deal with the randomness of (some) model parameters (Powell, 2013; Rijpkema et al., 2016; Wang et al., 2021); however, these methods depend on a prior probability distribution for each of the random parameters without considering possible new evidence that can provide other information on these parameters. Compared with the above-mentioned methods, BN is able to translate uncertainties into numerical values, combing new evidence to estimate posterior probabilities of certain events and, also to show good prediction accuracy with fast response. Regarding monitoring dioxins in agricultural production chains, Lascano-Alcoser et al. (2013, 2014) and Wang et al. (2021) optimized dioxin monitoring schemes in a specific food product type to identify noncompliant samples. Our study did not only identify noncompliant samples among different food products, but also guaranteed sufficient sample collection to estimate background levels of dioxins and DL-PCBs in these food products, which is consistent with the objective of the current EU dioxin monitoring recommendation (European Food Safety Authority, 2012; European Commission, 2017). For instance, four samples were collected for (other) poultry meat in the 2018 monitoring scheme and two in the optimal monitoring scheme, although the probability of a suspect sample is zero. These samples need to be collected and analyzed to estimate background dioxin levels in this type of meat. Therefore, in the case of dioxin monitoring, the BN model is able to capture the interaction of uncertainty in historical monitoring data, and to estimate the number of samples required to get a reliable estimate of the background contamination level, described by a distribution of possible results, and the probability of detecting suspect samples under different conditions.

This study shows that the modeling framework, combining a BN model and an optimization model, is useful for selection of the optimal monitoring scheme. This modeling framework, has previously been used to build real-time decision-making systems in other domains (Efe et al., 2018; Tchangani, 2004), but rarely been applied to the domain of food safety monitoring. Monitoring data collected going forward can be used to update the occurrence of contamination over time. The methodology could be extended to any food safety hazard for which historical monitoring data is available on different conditions. However, the results of this study cannot directly be applied as optimal monitoring scheme by industry or Food Safety Authorities to monitoring food safety hazards, because when food producers apply new technologies to reduce risk or unpredictable climate happens, the proposed framework cannot capture these uncertainties. Therefore, food safety risk managers should also take into account any other potential risk factor and use the proposed framework as a basis to optimize the final food safety monitoring scheme.

The BN model correctly predicted 93% of the validation cases (Appendix Table A2), which showed a high accuracy of the predictions of parameters in the IP model compared with deterministically assumed parameters and simulated parameters only based on their prior probabilities. It should be noted that the remaining 7% of the cases is wrongly predicted and would cause economic losses and negative impacts on public health. Even though the current level and human exposure of dioxin decrease, the health impacts of monitoring dioxin are still important when risk managers design monitoring schemes (Adams et al., 2017; Ali et al., 2022). The economic impacts of dioxin monitoring in this study only covered costs for sampling and analysis procedures, but the loss costs due to wrong decisions were not included in the model, because the focus of the study was on how to allocate monitoring budgets in monitoring procedures. An optimal design from a social welfare point of view should also include the negative public health effects and other economic losses due to prediction errors. Future research could expand the current framework along the lines sketched above by addressing the wider economic losses and public health consequences. For instance, experts could weight different hazards based on their disease burdens and economics losses, and the model could allocate monitoring resources more accurately to the target food product to reduce adverse impacts due to prediction errors.

The framework was built based on monitoring results from the Dutch official control program for dioxins and DL-PCBs in foods of animal origin and its results were therefore limited by the information available in these reports. For instance, geographical information was only reported for 2008 in the historical dataset. With more relevant information, the accuracy of the BN results and the performance of the optimal monitoring scheme could be further improved. For future studies, we also recommend collecting and storing all possible relevant information (for instance geographical information, frequency of monitoring, types of hazards, etc.). In addition, a larger input database, such as the monitoring results collected by EFSA, would help to further optimize monitoring schemes for multiple food safety hazards across Europe.

5 | CONCLUSION

We combined BN and IP models into a single framework to optimize the use of resources for food safety monitoring aimed at identifying noncompliant samples and estimating background level of hazards in food products. The methodology was applied to dioxins and DL-PCBs in primary animal derived food products in the Netherlands. The optimization results were compared to the monitoring scheme used in 2018 in the Netherlands. By estimating the prevalence of suspect samples, sampling pig meat could be done in a more cost-effective manner. In addition, conducting sample collection earlier in the year could increase the performance of the dioxin monitoring program. This framework can help risk managers from both the government and the industry to design and implement resource efficient food monitoring schemes.

ACKNOWLEDGMENTS

This research was financially supported by Chinese Scholarship Council, Wageningen University and the Fundamental Research Funds for the Central Universities (Program No. 2662022ZKQD003). We kindly thank Wageningen Food Safety Research for making the data available, and Dr. Bouzembrak (Wageningen Food Safety Research) for his valuable input to this study.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Z. Wang https://orcid.org/0000-0001-5109-5335
H. J. van der Fels-Klerx https://orcid.org/0000-0002-7801-394X

REFERENCES

Adamse, P., Schoss, S., Theelen, R. M., & Hoogenboom, R. L. (2017). Levels of dioxins and dioxin-like PCBs in food of animal origin in the Netherlands during the period 2001–2011. *Food Additives and Contaminants: Part A, Chemistry, Analysis, Control, Exposure and Risk Assessment*, *34*(1), 78–92. https://doi.org/10.1080/19440049.2016.1252065

Alban, L., Rugbjerg, H., Petersen, J. V., & Nielsen, L. R. (2016). Comparison of risk-based versus random sampling in the monitoring of antimicrobial residues in Danish finishing pigs. *Preventive Veterinary Medicine*, *128*, 87–94. https://doi.org/10.1016/j.prevetmed.2016.04.007

Ali, B. M., Andersson, M. G., van den Borne, B. H. P., Focker, M., & van der Fels-Klerx, H. J. (2022). Multi-criteria decision analysis in food safety risk management: the case of dioxins in baltic fish. *Foods*, *11*(7), 1059. https://doi.org/10.3390/foods11071059

Baars, A., Bakker, M. I., Baumann, R. A., Boon, P. E., Freijer, J. I., Hoogenboom, L. A., Hoogerbrugge, R., van Klaveren, J. D., Liem, A. K., Traag, W. A., & De Vries, J. (2004). Dioxins, dioxin-like PCBs and non-dioxin-like PCBs in foodstuffs: Occurrence and dietary intake in The Netherlands. *Toxicology Letters*, *151*(1), 51–61. DOI:https://doi.org/10.1016/j.toxlet.2004.01.028

Boon, P. E., te Biesebeek, J. D., de Wit-Bos, L., & van Donkersgoed, G. (2014). *Dietary exposure to dioxins in the Netherlands*. National Institute of Public Health and the Environment (RIVM). https://rivm.openrepository.com/handle/10029/557043

Bouzembrak, Y., & Marvin, H. J. (2019). Impact of drivers of change, including climatic factors, on the occurrence of chemical food safety hazards in fruits and vegetables: A Bayesian Network approach. *Food Control*, *97*, 67–76. DOI: https://doi.org/10.1016/j.foodcont.2018.10.021

Bouzembrak, Y., & van der Fels-Klerx, H. (2017). Effective sampling strategy to detect food and feed contamination: Herbs and spices case. *Food Control*, *83*, 28–37. https://doi.org/10.1016/j.foodcont.2017.04.038

Buriticá, J. A., & Tesfamariam, S. (2015). Consequence-based framework for electric power providers using Bayesian belief network. *International Journal of Electrical Power & Energy Systems*, *64*, 233–241. https://doi.org/10.1016/j.ijepes.2014.07.034

Efe, B., Kurt, M., & Efe, Ö. F. (2018). Hazard analysis using a Bayesian network and linear programming. *International Journal of Occupational Safety and Ergonomics*, *26*, 573–588. https://doi.org/10.1080/10803548.2018.1505805

European Commission. (2006a). Commission Recommendation of 16 November 2006 on the monitoring of background levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in foodstuffs (notified under document number C(2006) 5425). *Official Journal of the European Union* L 322/324.

European Commission. (2006b). Commission Regulation (EC) No.1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs. *Official Journal of the European Union L*, *364*, 5–24.

European Commission. (2014). Commission Recommendation of 3 December 2013 on the reduction of the presence of dioxins, furans and PCBs in feed and food (consolidated version). Official Journal of the European Union L323/37.

European Commission. (2017). Commission Regualtion (EU) 2017/644 of 5 April 2017 laying down methods of sampling and analysis for the control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation (EU) No. 589/2014. Official Journal of the European Union L 92/9.

European Commission. (2020). Commission Regulation (EC) No.1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs (consolidated version). Official Journal of the European Union L 364/5.

European Food Safety Authority. (2012). Update of the monitoring of levels of dioxins and PCBs in food and feed. *EFSA Journal*, *10*(7), 2832. https://doi.org/10.2903/j.efsa.2012.2832

EFSA Panel on Contaminants in the Food Chain (CONTAM). (2012). Scientific Opinion on the presence of dioxins (PCDD/Fs) and dioxin-like PCBs (DL-PCBs) in commercially available foods for infants and young children. *EFSA Journal*, *10*(12), 2983. https://doi.org/10.2903/j.efsa.2012.2983

Focker, M., Van der Fels-Klerx, H. J., & Oude Lansink, A. G. J. M. (2018). Systematic review of methods to determine the cost-effectiveness of monitoring plans for chemical and biological hazards in the life sciences. *Comprehensive Reviews in Food Science and Food Safety*, *17*(3), 633–645. https://doi.org/10.1111/1541–4337.12340

Focker, M., Van der Fels-Klerx, H. J., & Oude Lansink, A. G. J. M. (2019a). Cost-effective sampling and analysis for mycotoxins in a cereal batch. *Risk Analysis*, *39*(4), 926–939. https://doi.org/10.1111/risa.13201

Focker, M., Van der Fels-Klerx, H. J., & Oude Lansink, A. G. J. M. (2019b). Optimization of the aflatoxin monitoring costs along the maize supply chain. *Risk Analysis*, *DOI*, https://doi.org/10.1111/risa.13364

Gibb, H., Devleesschauwer, B., Bolger, P. M., Wu, F., Ezendam, J., Cliff, J., Zeilmaker, M., Verger, P., Pitt, J., Baines, J., Adegoke, G., Afshari, R., Liu, Y., Bokkers, B., van Loveren, H., Mengelers, M., Brandon, E., Havelaar, A. H., & Bellinger, D. (2015). World health organization estimates of the global and regional disease burden of four foodborne chemical toxins, 2010: A data synthesis. *F1000Research*, *4*, (1393). https://doi.org/10.12688/f1000research.7340.1

Hoogenboom, R., Traag, W., Fernandes, A., & Rose, M. (2015). European developments following incidents with dioxins and PCBs in the food and feed chain. *Food Control*, *50*, 670–683. https://doi.org/10.1016/j.foodcont.2014.10.010

Knutsen, H. K., Alexander, J., Barregård, L., Bignami, M., Brüschweiler, B., Ceccatelli, S., Cottrill, B., Dinovi, M., Edler, L., Grasl-Kraupp, B., Hogstrand, C., Nebbia, C., Oswald, I. P., Petersen, A., Rose, M., Roudot, A.-C., Schwerdtie, T., Vleminckx, C., Vollmer, G., & Hoogenboom, L. (2018). Risk for animal and human health related to the presence of dioxins and dioxin-like PCBs in feed and food. *EFSA Journal*, *16*(11), e05333. https://doi.org/10.2903/j.efsa.2018.5333

Lascano-Alcoser, V., Mourits, M., Van der Fels-Klerx, H. J., Heres, L., Velthuis, A., Hoogenboom, L., & Oude Lansink, A. G. J. M. (2014). Cost-effective allocation of resources for monitoring dioxins along the pork production chain. *Food Research International*, *62*, 618–627. https://doi.org/10.1016/j.foodres.2014.04.011

Lascano-Alcoser, V., Velthuis, A., Van der Fels-Klerx, H. J., Hoogenboom, L., & Oude Lansink, A. G. J. M. (2013). Optimizing bulk milk dioxin monitoring based on costs and effectiveness. *Journal of Dairy Science*, *96*(7), 4125–4141. https://doi.org/10.3168/jds.2012-5898

Lascano Alcoser, V., Velthuis, A., Hoogenboom, L., & Van der Fels-Klerx, H. J. (2011). Financial impact of a dioxin incident in the Dutch dairy chain. *Journal of Food Protection*, *74*(6), 967–979. https://doi.org/10.4315/0362-028X.JFP-10-350

Lavric, E. D., Konnov, A. A., & De Ruyck, J. (2005). Surrogate compounds for dioxins in incineration. A review. *Waste Management*, *25*(7), 755–765. https://doi.org/10.1016/j.wasman.2004.12.026

Lindo Systems Inc. (2017). https://lindo.com/index.php/ls-downloads?id=112:lingo-documentation&catid=82

Marinković, N., Pašalić, D., Ferenčak, G., Gršković, B., & Stavljenić Rukavina, A. (2010). Dioxins and human toxicity. *Arhiv za higijenu rada i toksikologiju*, *61*(4), 445–452. https://doi.org/10.2478/10004-1254-61-2010–2024

Namazian, A., Yakhchali Yakhchali, S. H., Yousefi Yousefi, V., & Tamošaitienė Tamošaitienė, J. (2019). Combining Monte Carlo simulation and Bayesian networks methods for assessing completion time of projects under risk. *International Journal of Environmental Research and Public Health*, *16*(24), 5024. https://doi.org/10.3390/ijerph16245024

Powell, M. R. (2013). The economic efficiency of sampling size: The case of beef trim revisited. *Risk Analysis*, *33*(3), 385–396. https://doi.org/10.3390/ijerph16245024

Powell, M. R. (2014). Optimal food safety sampling under a budget constraint. *Risk Analysis*, *34*(1), 93–100. https://doi.org/10.1111/risa.12054

R Development Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org

Rijpkema, W. A., Hendrix, E. M., Rossi, R., & van der Vorst, J. G. (2016). Application of stochastic programming to reduce uncertainty in quality-based supply planning of slaughterhouses. *Annals of Operations Research*, *239*(2), 613–624. https://doi.org/10.1007/s10479-013-1460-y

Scutari, M. (2009). Learning Bayesian networks with the bnlearn R package. arXiv preprint arXiv:0908.3817. https://doi.org/10.18637/jss.v035.i03

Srogi, K. (2008). Levels and congener distributions of PCDDs, PCDFs and dioxin-like PCBs in environmental and human samples: A review. *Environmental Chemistry Letter*, *6*, 1–28. https://doi.org/10.1007/s10311-007-0105-2

Tchangani, A. P. (2004). Decision-making with uncertain data: Bayesian linear programming approach. *Journal of Intelligent Manufacturing*, *15*(1), 17–27. https://doi.org/10.1023/B:JIMS.0000010072.16604.04

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, *203*(3–4), 312–318. https://doi.org/10.1016/j.ecolmodel.2006.11.033

Wang, Z., van der Fels-Klerx, H. J., & Oude Lansink, A. G. J. M. (2020). Optimization of sampling for monitoring chemicals in the food supply chain using a risk-based approach: The case of aflatoxins and dioxins in the Dutch dairy chain. *Risk Analysis*, *40*(12), 2539–2560. https://doi.org/10.1111/risa.13605

Wang, Z., van der Fels-Klerx, H., & Lansink, A. O. (2021). Modeling cost-effective monitoring schemes for food safety contaminants: Case study for dioxins in the dairy supply chain. *Food Research International*, *141*, 110110.

WFSR. (2019). *Monitoring dioxins, PCBs and flame retardants in agricultural products*. https://www.wur.nl/en/Research-Results/Research-Institutes/rikilt/Expertise-areas/Measuring-and-detecting-substances/Contaminants/Dioxin-analysis/Monitoring-dioxins-PCBs-and-flame-retardants-in-agricultural-products.htm

World Health Organization. (2015). *WHO estimates of the global burden of foodborne diseases: Executive summary*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/200046/WHO_FOS_15.02_eng.pdf

Yazdi, M., & Kabir, S. (2017). A fuzzy Bayesian network approach for risk analysis in process industries. *Process Safety and Environmental Protection*, *111*, 507–519. https://doi.org/10.1016/j.psep.2017.08.015

How to cite this article: Wang, Z., van der Fels-Klerx, H. J., & Oude Lansink, A. G. J. M. (2022). Designing optimal food safety monitoring schemes using Bayesian network and integer programming: the case of monitoring dioxins and DL-PCBs. *Risk Analysis*, 1–14. https://doi.org/10.1111/risa.14030

APPENDIX A

The tables below present the probability distribution of sample size estimating background levels of contaminants and the validation results of Bayesian network

APPENDIX B

The corresponding monitoring data, R codes for building BN model, and computed Bayesian Network are available from https://github.com/puhmli/BNDioxinMonitroing.git

TABLE A1 The probability distribution for estimating background levels of dioxins and DL-PCBs based on the total number of samples analyzed for different animal species and food product types

Animal species	Food product type	Probability distribution of Ns_BLapa																											Most likely Value of Ns_BLapb	95%CI of Ns_BLapc
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	24	27	28	29	31		
Hen	egg	1%	1%	1%	2%	0%	0%	0%	0%	0%	6%	0%	0%	0%	0%	12%	0%	32%	8%	17%	12%	0%	0%	8%	0%	0%	0%	0%	16	(15.1, 16.9)
Bovine	meat	0%	1%	0%	1%	2%	2%	9%	12%	0%	5%	10%	0%	5%	7%	6%	0%	23%	0%	9%	8%	0%	0%	0%	0%	0%	0%	0%	16	(11.2, 13.1)
Broiler	meat	0%	0%	0%	0%	0%	8%	4%	0%	0%	0%	5%	5%	0%	0%	0%	0%	0%	0%	11%	0%	16%	26%	13%	13%	0%	0%	0%	22	(17.7, 20.3)
Calf	meat	0%	0%	0%	0%	0%	0%	13%	7%	0%	0%	0%	10%	12%	15%	16%	15%	0%	0%	13%	0%	0%	0%	0%	0%	0%	0%	0%	14	(11.7, 13.1)
Deer	meat	38%	44%	19%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1	(0.6, 0.9)
Pig	meat	0%	0%	0%	0%	0%	0%	4%	0%	0%	6%	5%	0%	0%	0%	0%	12%	8%	0%	0%	0%	15%	0%	0%	13%	12%	11%	15%	20	(20.2, 23.4)
Sheep	meat	0%	0%	0%	8%	0%	6%	0%	8%	15%	41%	0%	15%	0%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	9	(8.1, 9.0)
Bovine	milk	0%	1%	0%	2%	1%	2%	10%	12%	0%	5%	10%	0%	4%	6%	7%	0%	25%	0%	8%	8%	0%	0%	0%	0%	0%	0%	0%	16	(11.2, 13.1)
Poultry	meat	16%	51%	33%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2	(1.0, 1.3)

Note: Ns_BLap: the number of samples analyzed dioxins and DL-PCBs background level monitoring in animal species *a* and food product type *p*. We chose the most likely value of Ns_BL as the lower limit of number of samples collected during 1 year for animal species a and food product type *p*. 95%CI: 95% confidence interval of Ns_BL.

TABLE A2 The validation results of Bayesian network

Compliance with pre-set threshold limits	Prediction results: Nonsuspect	Prediction results: Suspect	Total	Correctness (%)
Nonsuspect	569	28	597	95%
Suspect	16	39	55	71%
Overall	–	–	652	93%