Effects of global change on the emission, fate, effects and risks of chemicals in aquatic ecosystems

Marie Curie Actions

SHORT TITLE:	ECORISK2020
COORDINATOR:	Prof Dr Paul van den Brink
ORGANISATION:	Wageningen University
TOPIC:	H2020-MSC_ITN 2018
PROJECT NUMBER:	813124

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813124

Deliverable: D.6.1 Draft risk assessment model (Bayesian network) for agricultural chemicals developed for one of the case study regions

Contributing:

Norsk institutt for Vannforskning, Sophie Mentzel (ESR11)

Norsk institutt for Vannforskning, Jannicke Moe (beneficiary, supervisor)

Norsk institutt for Vannforskning, Merete Grung (beneficiary, supervisor)

Norsk institutt for bioøkonomi, Marianne Stenrød (partner organisation, supervisor)

Norsk institutt for Vannforskning, Knut Erik Tollefsen (beneficiary, supervisor)

Norsk institutt for Vannforskning, Karina Petersen (beneficiary)

Norsk institutt for bioøkonomi, Roger Holten (partner organisation)

Norsk institutt for bioøkonomi, Randi Bolli (partner organisation)

1 Introduction

Pesticides play an important role in the food production process, but they can also lead to harmful effects in the environment. A recent European assessment for rivers and lakes over the period 2007-2017 (Mohaupt et al., 2020) report shows that 5 – 15 % of the monitoring stations showed exceedances of environmental quality standards by herbicides and 3 – 8 % by insecticides. In Europe, environmental risk assessment of pesticides is implemented to safeguard the heath and biodiversity of ecosystems (Schäfer et al., 2019). Traditionally a deterministic approach applying single values of predicted environmental concentration (PEC) (exposure assessment) and predicted no effect concentration (PNEC) (effect assessment) is used to quantify the exceedance of safe concentration that a substance poses to the environment (risk characterization) see Figure 1.

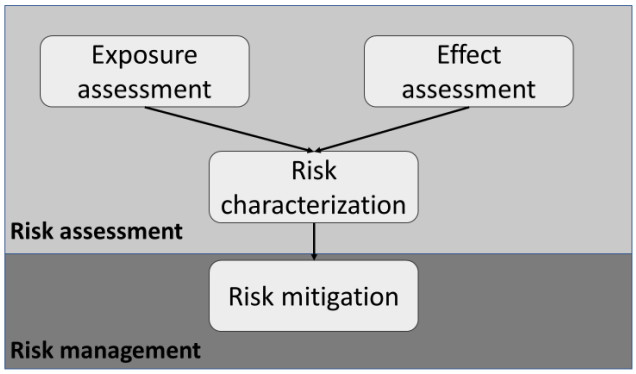


Figure 1 General environmental risk assessment process

This method has ineptitude to account for spatial and temporal variability of chemical concentrations and species sensitivity, and uncertainty (EUFRAM (2006), Verdonck (2003)). In order to overcome some of these limitations traditional risk assessment faces, ESR 11's main objective is to develop and apply a modelling tool for the assessment of the risk agricultural chemicals poses, under future scenarios. To achieve this goal a probabilistic model was developed using Bayesian network methodology which was parameterised for a northern case study region. The conceptual model underlying these networks can and will be used as a base for further study during the ECORISK2050 project.

2 Probabilistic risk assessment of pesticides

By allowing to include estimates of uncertainty and stochastic properties (Solomon et al., 2000), probabilistic risk assessment enables and expresses variability and uncertainty better than the traditional calculation of a single-value risk quotient. This is done by using probabilities or probability distributions to account for single or multiple sources of variability and uncertainty (EUFRAM, 2006) see Figure 2. In the following, underlying concept for traditional, intermediate and fully probabilistic approaches are displayed for the use of probability distributions.

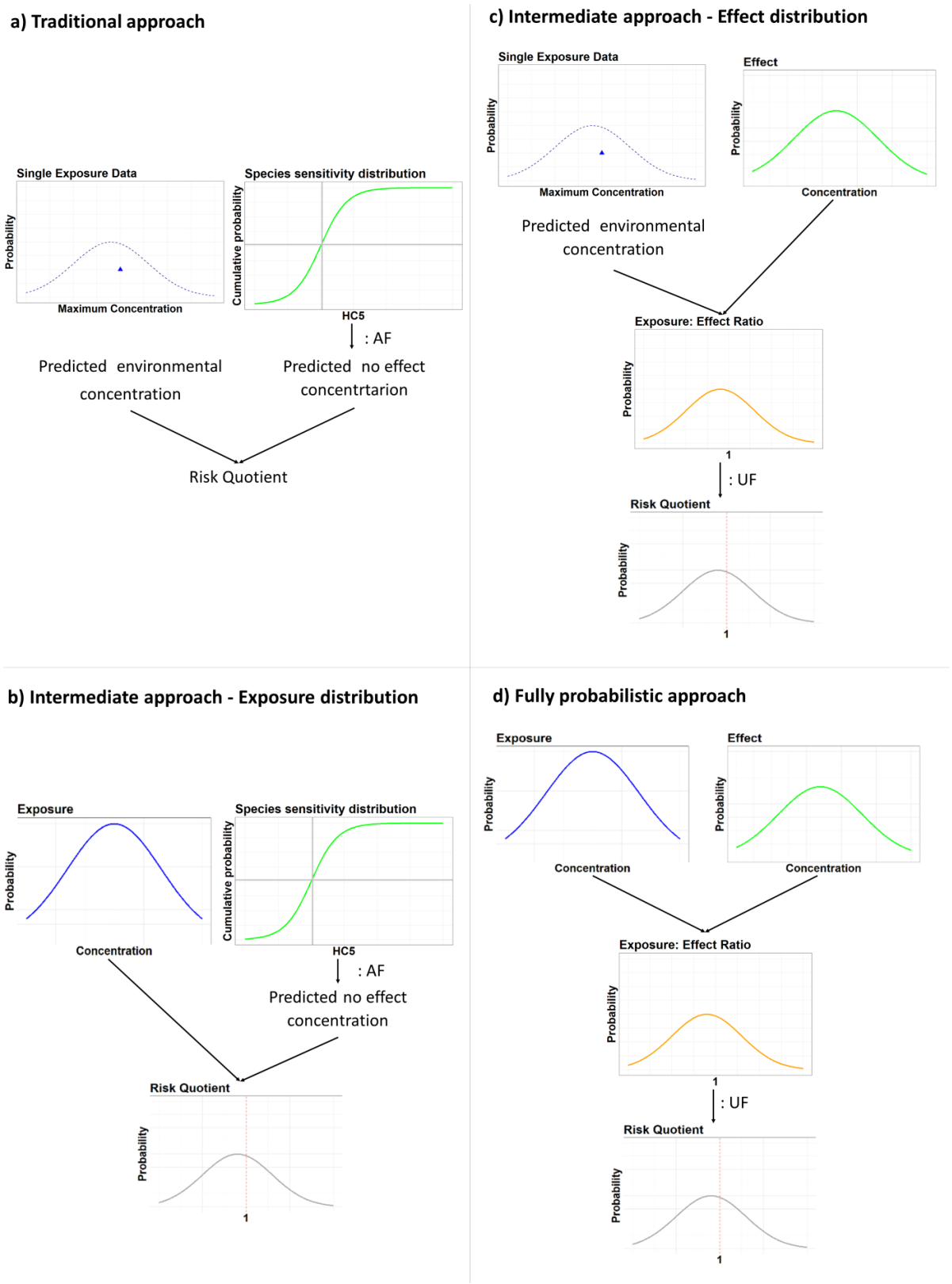


Figure 2 Systematic overview of the traditional approach to derive a risk quotient, compared to two intermediate probabilistic approaches and a fully probabilistic approach that derive a risk quotient distribution

In the traditional approach single value PEC and PNEC derived for example from a HC5 (hazardous concentration for 5% of the species) of a species sensitivity distribution (SSD) based on a cumulative distribution, together with an assessment factor (AF) are used to derive a single value risk quotient (RQ)

(see Figure 2a). Figure 2b displays an intermediate approach using an exposure distribution together with a single value PNEC to derive a risk quotient distribution. The other intermediate approach does not derive a PNEC as no HC5 can be determined, so no AF can be applied. In this case, after deriving an exposure-effect ratio distribution, an uncertainty factor (UF) is applied to account for the uncertainties in the data set (primarily the toxicity data set) in order to derive the risk quotient distribution see Figure 2c. The UF has the same role as an AF. Although, as it is not defined exactly the same a different term was used in this study. The fully probabilistic approach can be applied whenever exposure and effect distributions are available. After calculating the exposure: effect ratio distribution, an UF is applied to derive the RQ distribution see Figure 2d.

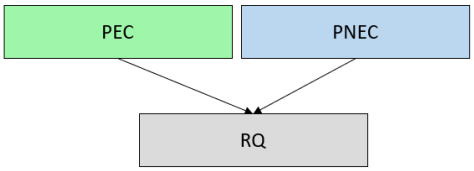
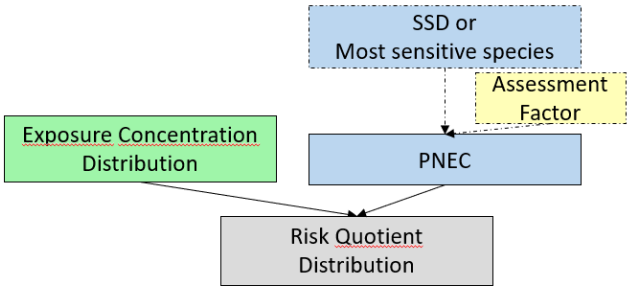
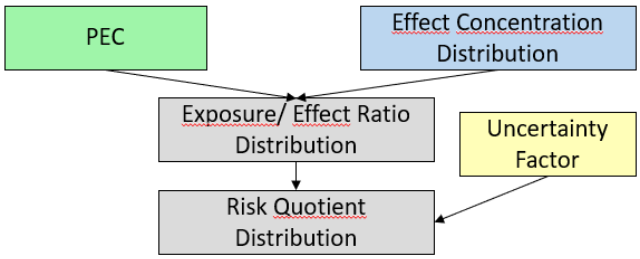
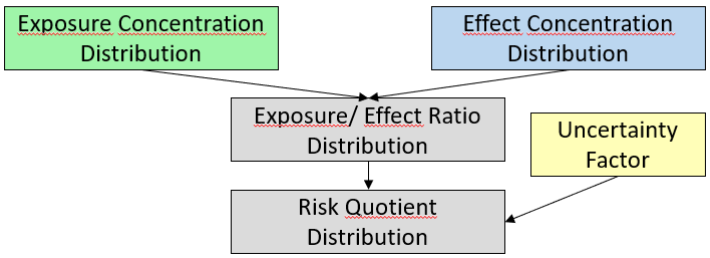
3 Developed Bayesian network approaches – traditional, intermediate and fully probabilistic approach

Bayesian networks (BN) are probabilistic graphical models implementing Bayes' rule for updating probability distributions based on evidence (Carriger et al., 2016). Bayesian networks are increasingly used in environmental risk assessment and management (Moe et al., 2021). The nodes (variables) have states (intervals), quantified by probability distributions. The causal links between nodes are shown as arrows and represent conditional probability tables (CPTs). Following construction guidelines provided by Marcot et al. (2006) and Pollino and Henderson (2010)the BNs used in this project are set up.

3.1 Conceptual models for Bayesian networks and data preparation

The Bayesian networks developed for this research are set up according to the conceptual models in Table 1.

Table 1 Conceptual models for Bayesian Network including intermediate and fully probabilistic approach

Approach	Bayesian Network approach	Data input
Traditional approach (a)		PNEC value derived from NOEC based SSD Exposure concentration max value from exposure data set
Intermediate approach – exposure distribution (b)		PNEC value derived from NOEC based SSD Exposure distribution derived from exposure data set
Intermediate approach – effect distribution (c)		Effect distribution from collected NOEC values Exposure concentration max, min and mean value from exposure data set
Fully probabilistic Approach (d)		Effect distribution from collected NOEC values Exposure distribution derived from exposure data set

The collected data for the exposure distribution should include an abundance of measured or predicted environmental exposure concentrations for the selected chemical and study/ representative area. For the effect distribution as many NOECs (aquatic species, adverse effect toxicity tests) should be collected as possible (following guidelines for SSD modelling in the *Technical Guidance For Deriving Environmental Quality Standards* (SCHEER, 2017)). In case of multiple endpoints and multiple values of the same species, the mean should be used to derive the effect distribution.

Effect and exposure data was provided by The NIVA Risk Assessment database (RAdb, www.niva.no/radb) for the case study area in the south east Norway and the selected pesticide Azoxystrobin. RAdb was used as a toolbox that aggregates exposure and effect data, identifies assessment criteria on basis of international recommendations and provides transparent and harmonized cumulative risk estimates. The exposure data was collected through the *Norwegian Agricultural Environmental Monitoring Program (JOVA)* via RAdb database from the period of 2011 to 2016. Effects data for several freshwater species representing various taxonomic groups was collected from RAdb. The data set consist of NOECs (no observed effect concentration) for adverse effects such as growth, reproduction, and population.

Further processing of the data sets included the fitting of a log-normal distribution to the exposure and effects data. For the exposure distribution, also the recorded limits of quantification were considered. The *R* version 4.0.2 (Team, 2020) package "MASS" version 7.3-51.6 (Venables & Ripley, 2002) was used to fit the distribution to the data. In case of exposure data, the values below limit of quantification (LOQ) were simulated from the fitted lognormal distribution. The mean and standard deviation are estimated from the derived distribution.

3.2 Network assumptions and equations

For this research, the software *Netica* (Norsys Software Corp., www.norsys.com) was used to construct the BN. The software offers shortcuts and equations to be entered for generating conditional probability tables. For the input nodes, the following equation is used to train the network: *NormalDist (x,μ,σ)*. With "*x*" indicating the node, "*μ*" being the mean of the normal distribution, and "*σ*" being the standard deviation of the distribution. For the seasonal BN using the fully distributed approach, another equation is used to implement conditions (for the different seasons).

P(X|B) =
(B == Spring) ? NormalDist (x,μ,σ):
(B == Summer) ? NormalDist (x,μ,σ):
(B == Autumn) ? NormalDist (x,μ,σ): 0

The parent of "*X*" (e.g. Exp_Log) in this network is "*B*" (e.g. Sea = Season). The formula describes the way to condition states of a discrete node and each condition has a distribution assigned. Other common mathematical operators used in the model were *exp (x)* for exponential (e ^ x). The network assumptions and connections are as described in Table 2.

Table 2 Node overview containing type, number of states and equations used for the intermediate approach b, c and the fully distributed approach

Node / Variable	Type	States	Netica equation
Intermediate approach (b)			
Exposure concentration distribution Log	DC	10	P(Exp_Log|) = NormalDist (Exp_Log, μ, σ)
Exposure concentration distribution (ug/L)	DC	10	Exp_Norm (Exp_Log) = exp(Exp_Log)
PNEC (ug/L)	Constant	1	-
Risk quotient distribution	DC	7	RQ (Exp_Norm, PNEC) = (Exp_Norm/PNEC)
Intermediate approach (c)			
PEC (ug/L)	Constant	2	-
Effect Concentration distribution Log	DC	10	P(Eff_Log|) = NormalDist (Eff_Log, μ, σ)
Effect concentration distribution (ug/L)	DC	10	Eff_Norm (Eff_Log) = exp(Eff_Log)
Exposure - effect - ratio distribution (ug/L)	DC	7	Exp_Eff_Ra (Exp_Norm, Eff_Norm) = (Exp_Norm/Eff_Norm)
Uncertainty factor (UF)	Constant	5	-
Risk quotient distribution	DC	7	RQ (UF, Exp_Eff_Ra) = (UF*Exp_Eff_Ra)

Fully probabilistic approach (d)			
Exposure concentration distribution log	DC	10	P(Exp_Log) = NormalDist(Exp_Log, μ, σ)
Effect concentration distribution log	DC	10	P(Eff_Log|) = NormalDist(Eff_Log, μ, σ)
Exposure concentration distribution (ug/L)	DC	10	Exp_Norm (Exp_Log) = exp(Exp_Log)
Effect concentration distribution (ug/L)	DC	10	Eff_Norm (Eff_Log) = exp(Eff_Log)
Exposure - effect - ratio distribution	DC	7	Exp_Eff_Ra (Exp_Norm, Eff_Norm) = (Exp_Norm/Eff_Norm)
Uncertainty factor	Constant	5	-
Risk quotient distribution	DC	7	RQ (UF, Exp_Eff_Ra) = (UF*Exp_Eff_Ra)

*DC -> discretized continuous; continuous variables were binned into the states

3.3 Example for case study area and a pesticide

The following example displays a fully probabilistic approach of Azoxystrobin for the case study area Heia in the south east of Norway (for detailed model input see Appendix, Table A.I). The fully probabilistic approach (d) uses distributions for both exposure and effect (Figure 3). For the event of using an UF or 10, the probability of the risk quotient (RQ) exceeding 1 (to be in the interval of "1 to 5") is 0.57%.

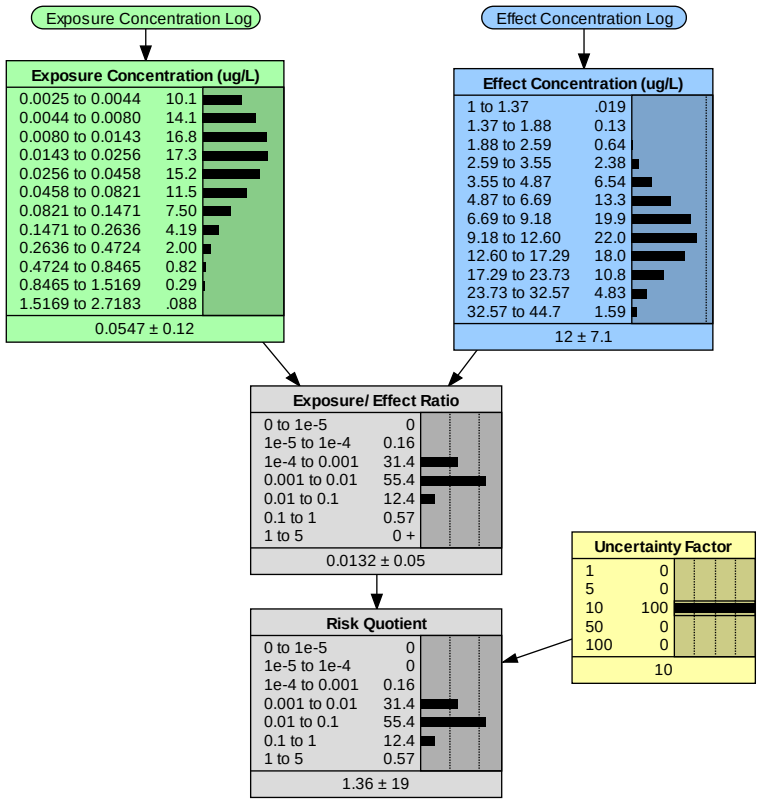


Figure 3 Example Bayesian network for the fully probabilistic approach for Azoxystrobin for the study area 'Heia'. Using distributed exposure and effect concentration, and uncertainty factor of 10.

The most likely interval for RQ when applying an UF of 1 and 5 is "0.001 to 0.01". As the probability of the RQ to be below 1 is 99.43 %, it can be assumed that Azoxystrobin does not pose a risk to the environment in this case. If an UF of 10 is applied, the RQ is most likely in the interval of "0.01 to 0.1". When using a UF of 50 the RQ is most likely in the interval of "0.1 to 1" (Figure 4).

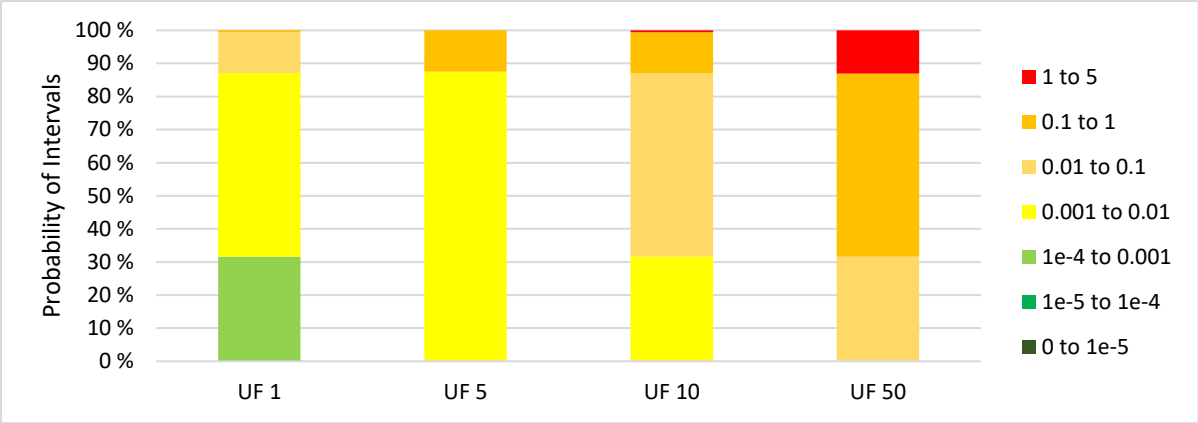


Figure 4 Risk quotient distribution output from the Bayesian network for fully probabilistic approach. This used exposure and effect distribution to derive the risk quotient distribution for an uncertainty factors of 1, 5, 10, and 50.

The following example displays a seasonal BN (for detailed model input see Appendix, Table A.II). This network offers an additional temporal variable (node) as it uses exposure distributions derived for the individual seasons. In general, when an UF of 1 is applied there is no predicted risk to the environment as there is 0% probability of the RQ to be above 1. The RQ for summer has high probability of being in the interval of "0.01 to 0.1" see Figure 5.

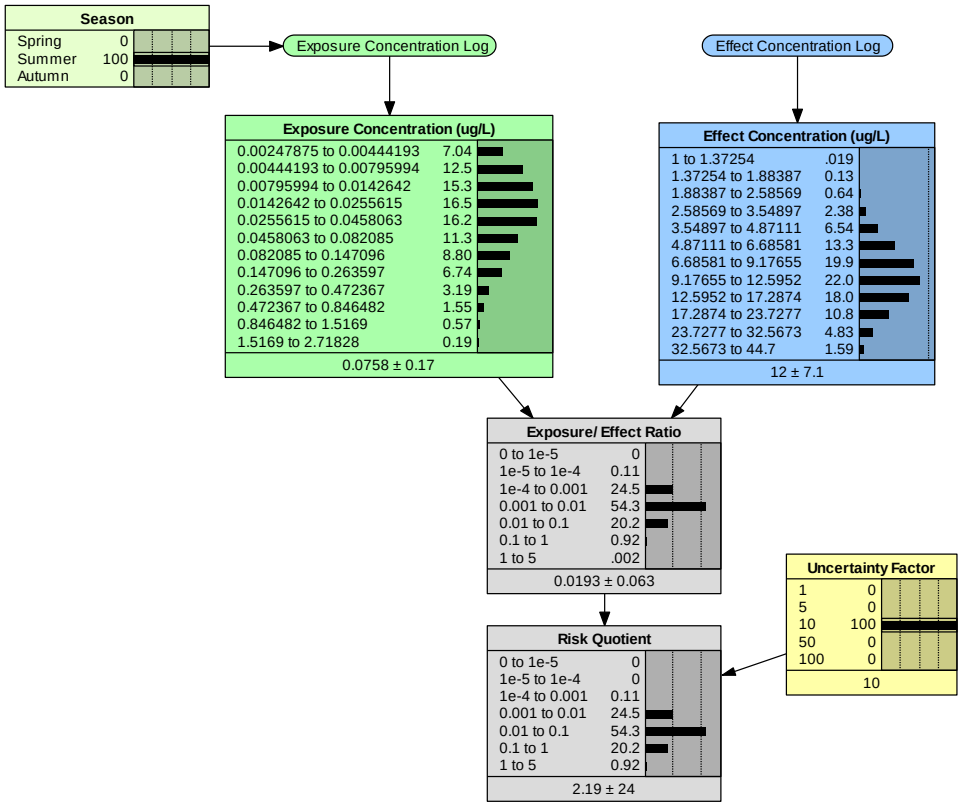


Figure 5 Example Bayesian network model with calculation of a Risk Quotient distribution from the effect and exposure concentration distributions of Azoxystrobin, with application of an uncertainty factors of 10 for summer season.

The RQ for spring, summer and autumn have high probabilities to be in intervals of "0.001 to 0.01" and "0.001 to 0.01" when an UF of 10 is applied Figure 6. Though for this example, based on Azoxystrobin

exposure in the Heia catchment there is a small probability that the RQ is in the intervals of "0.1 to 1" and "1 to 5" during summer and autumn.

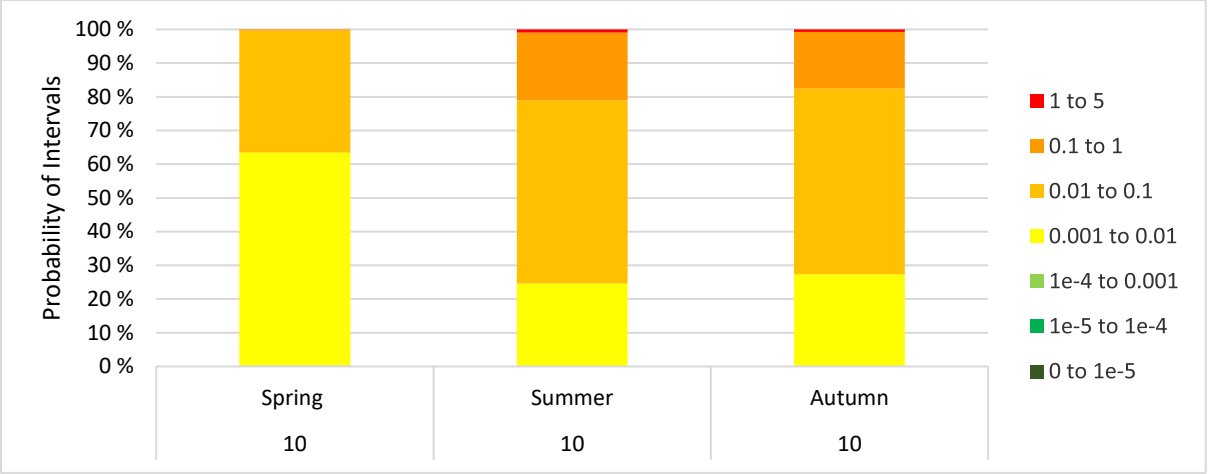


Figure 6 Bayesian network model with calculation of a probabilistic risk quotient, for spring, summer, autumn and uncertainty factors 10 for Azoxystrobin

4 Conceptual model for the Bayesian network incorporating future change scenarios

We also developed a conceptual model for a BN to assess environmental risk of pesticides under future scenarios, which incorporates various types of information from climate projections, pesticide exposure models and toxicity tests.

4.1 Future change scenario Bayesian network input data - Exposure prediction model assumptions, input and output

Environmental factors such as soil and site parameters together with chemical properties and climate scenarios are linked to the exposure of a pesticide in the selected study area by using pesticide fate and transport models. For the case study area "Skuterud" in the south east of Norway, exposure modelling is carried out with the *World Integrated System for Pesticide Exposure (WISPE)* (Bolli et al., 2013). In general, several realistic scenarios are developed and used as the input for the WISPE model. The aim is to:

- Simulate chemicals to specify unique pesticide application conditions for different scenarios.
- Run the model with current and predicted climate data for a representative field site.
- Determine application scenarios by using the assumed worst case application rate and creating two additional inputs for 25 % above and below this worst case.

The WISPE model output are of 26 predicted concentrations, corresponding to the 26 years over which the model runs. The concentrations are predicted for instantaneous, 24 h, 96 h, 21, 60 and 90 days. These concentrations can be used as input for the BN. A log-linear equation was fitted to the time series concentrations generated by the WISPE and used to generate the conditional probability table for the Time-specific concentration node.

Effect concentration distribution was again derived by collecting toxicity test data from RAdb to derive acute and chronic distributions. Acute toxicity distribution together with a time since application of 1 and 2 days derive RQ for acute exposure situations, whereas chronic toxicity distributions together with longer time since application 'events' derives a chronic RQ.

4.2 Future change scenario Bayesian network model - assumptions and equations

Hereinafter, the conceptual model for the BN connecting exposure modelling with risk estimation is displayed (Figure 7). The instantaneous concentration together with time since application determines the time-specific concentration. The effect concentration distribution node is determined by the endpoint node to differentiate between acute and chronic exposure.

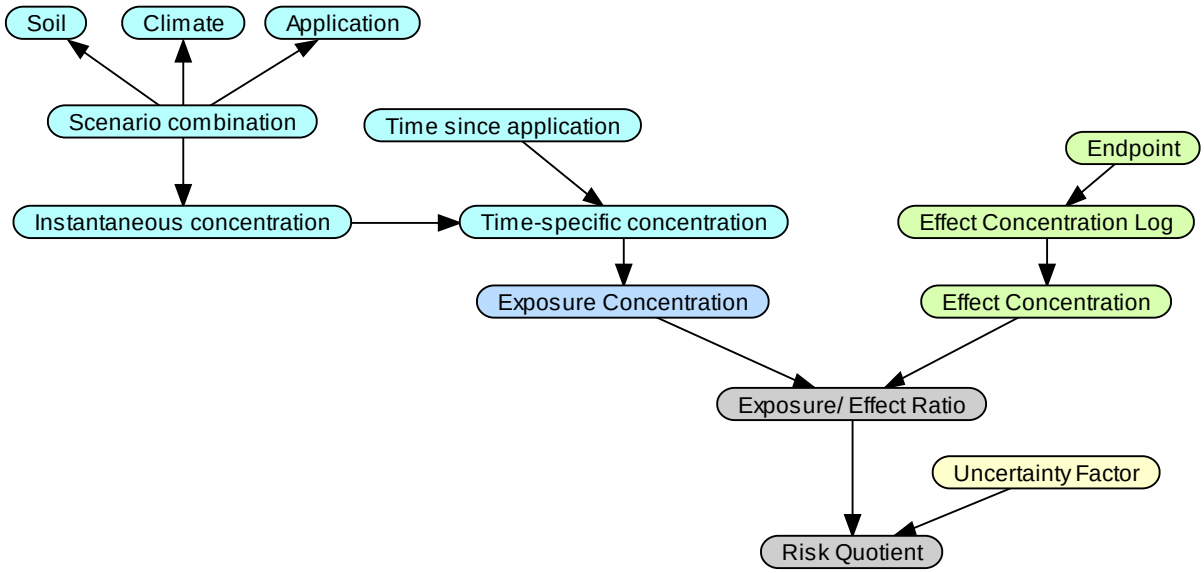


Figure 7 Conceptual model for the Bayesian network for the calculation of a Risk quotient distribution for various future change scenarios.

Exposure and effect together determine the exposure/ effect ratio node. An UF can be applied to derive the RQ see Table 3. The finalized BN model can be run by selecting a set of scenarios as evidence. Given this evidence, probability distributions will be updated throughout the BN.

Table 3 Node overview of Bayesian network containing future change scenarios

Node / Variable	Type	States	Netica equation and node input
Soil	Constant	3	Scenarios
Climate	Constant	3	Scenarios
Application	Constant	3	Scenarios
Scenario combination	Constant	?	-
Instantaneous concentration	DC	5	Probabilities for each interval
Time since application	Constant	5	-
Time specific concentration	DC	10	Conc_time (time, Conc_inst) = (Conc_inst + (slope* time))
Endpoint	Constant	2	-
Effect concentration distribution Log	DC	10	P(Eff_Log | Endpoint) = Endpoint == Acute? NormalDist(Eff_Log, μ, σ): Endpoint == Chronic? NormalDist(Eff_Log, μ, σ): 0
Exposure concentration distribution	DC	10	Exp_Norm (Exp_Log) = exp(Exp_Log)
Effect concentration distribution	DC	10	Eff_Norm (Eff_Log) = exp(Eff_Log)
Exposure – effect – ratio distribution	DC	7	Exp_Eff_Ra (Exp_Norm, Eff_Norm) = (Exp_Norm/Eff_Norm)
Uncertainty factor	Constant	5	-
Risk quotient distribution	DC	7	RQ (UF, Exp_Eff_Ra) = (UF*Exp_Eff_Ra)

5 Future outlooks

These examples for intermediate and fully probabilistic approaches demonstrate that BN modelling is a promising tool in for calculating the RQ as a full probability distribution instead of a single value. Also, these BNs can predict the probability of several risk levels and facilitate the communication of risk estimates and uncertainties. The examples shown here also demonstrate the importance of the uncertainty factor (UF) in the calculation of the RQ. The UF value should be selected by a risk assessor based on an evaluation of the data availability, representativity and other factors contibuting to uncertainty.

Furthermore, BNs have the potential to assess the environmental risk of pesticides under future scenarios, by integrating different types of information from e.g. climate, pesticide exposure models and toxicity testing. Also, it is possible to quantify uncertainty of all components in the BN model as they are propagated and incorporated in the probabilistic risk calculation. In addition, it gives a transparent way to display the risk quotient and the selected uncertainty factor. Further plans for the Bayesian network are to integrate other crop scenarios. We are also planning to run the model with two predicted climate scenarios for Norway and other representative field sites that grow other crop types.

References

Bolli, R. I., Eklo, O. M., Holten, R., & Mulder, P. (2013). *Development of WISPE for surface- and groundwater modelling of pesticides in major crops* (National Scenarios - Norway, Issue. https://core.ac.uk/download/pdf/285987623.pdf

Carriger, J. F., Barron, M. G., & Newman, M. C. (2016). Bayesian Networks Improve Causal Environmental Assessments for Evidence-Based Policy. *Environ Sci Technol*, *50*(24), 13195-13205. https://doi.org/10.1021/acs.est.6b03220

EUFRAM. (2006). *Concerted action to develop a europea framework for probabilistic risk assessment of the environmental impacts of pesticides* (Detailed reports on role, emthods, reporting & validation, Issue.

Marcot, B. G., Steventon, J. D., Sutherland, G. D., & McCann, R. K. (2006). Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research*, *36*(12), 3063-3074. https://doi.org/10.1139/x06-135

Moe, S. J., Carriger, J. F., & Glendell, M. (2021). Increased Use of Bayesian Network Models Has Improved Environmental Risk Assessments. *Integrated Environmental Assessment and Management*, *17*(1), 53-61. https://doi.org/https://doi.org/10.1002/ieam.4369

Mohaupt, V., Völker, J., Altenburger, R., Kirst, I., Kühnel, D., Küster, E., Semeradova, S., & Subelj, G. (2020). *Pesticides in European rivers, lakes and groundwaters - Data assessment* (ETC/ICM Technical Report 1/2020: European Topic Centre on Inland, Coastal and Marine waters, Issue. https://www.eionet.europa.eu/etcs/etc-icm/products/etc-icm-report-1-2020-pesticides-in-european-rivers-lakes-and-groundwaters-data-assessment

Pollino, C. A., & Henderson, C. (2010). Bayesian networks: A guide for their application in natural resource management and policy. *Landscape Logic, Technical Report*, *14*. https://www.utas.edu.au/__data/assets/pdf_file/0009/588474/TR_14_BNs_a_resource_guide.pdf

SCHEER. (2017). Scientific Advice on Environmental Quality Standards. *Technical Guidance for Deriving Environmental Quality Standards*(27). https://doi.org/10.2875/018826

Schäfer, R. B., Liess, M., Altenburger, R., Filser, J., Hollert, H., Roß-Nickoll, M., Schäffer, A., & Scheringer, M. (2019). Future pesticide risk assessment: narrowing the gap between intention and reality. *Environmental Sciences Europe*, *31*(1). https://doi.org/10.1186/s12302-019-0203-3

Solomon, K., Giesy, J., & Jones, P. (2000). Probabilistic risk assessment of agrochemicals in the environment. *Crop protection (Guildford, Surrey)*, *19*(8/10), 649-655. https://doi.org/10.1016/s0261-2194(00)00086-7

Team, R. C. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4 ed.). Springer. http://www.stats.ox.ac.uk/pub/MASS4

Verdonck, F. A. M. (2003). *Geo-referenced probabilistic ecological risk assessment* (Publication Number 257) Ghent University]. Gent. https://biblio.ugent.be/publication/521906

Appendix – Node input for the example of Azoxystrobin for the case study area Heia, Norway

Table A.I Node overview for the fully probabilistic approach – example of Azoxystrobin in the case study catchment Heia, Norway.

Node / Variable	Type	States	Netica equation
Exposure concentration distribution log	DC	10	P(Exp_Log) = NormalDist(Exp_Log, -4.148, 1.484)
Effect concentration distribution log	DC	10	P(Eff_Log|) = NormalDist(Eff_Log, 2.3224782, 0.5680065)
Exposure concentration distribution (ug/L)	DC	10	Exp_Norm (Exp_Log) = exp(Exp_Log)
Effect concentration distribution (ug/L)	DC	10	Eff_Norm (Eff_Log) = exp(Eff_Log)
Exposure - effect - ratio distribution	DC	7	Exp_Eff_Ra (Exp_Norm, Eff_Norm) = (Exp_Norm/Eff_Norm)
Uncertainty factor	Constant	5	-
Risk quotient distribution	DC	7	RQ (UF, Exp_Eff_Ra) = (UF*Exp_Eff_Ra)

Table A.II Node overview for the seasonal fully probabilistic approach – example of Azoxystrobin in the case study catchment Heia, Norway.

Node / Variable	Type	States	Netica equation and node input
Season	Constant	3	-
Exposure concentration distribution Log	DC	10	P(Exp_Log| Sea) = Sea == Spring? NormalDist(Exp_Log, -5.0293471, 0.7116285): Sea == Summer? NormalDist(Exp_Log, -3.939194, 1.528558): Sea == Autumn? NormalDist(Exp_Log, -4.017832, 1.540513): 0
Effect concentration distribution Log	DC	10	P(Eff_Log|) = NormalDist(Eff_Log, 2.3224782, 0.5680065)
Exposure concentration distribution	DC	10	Exp_Norm (Exp_Log) = exp(Exp_Log)
Effect concentration distribution	DC	10	Eff_Norm (Eff_Log) = exp(Eff_Log)
Exposure - effect - ratio distribution	DC	7	Exp_Eff_Ra (Exp_Norm, Eff_Norm) = (Exp_Norm/Eff_Norm)
Uncertainty factor	Constant	5	-
Risk quotient distribution	DC	7	RQ (UF, Exp_Eff_Ra) = (UF*Exp_Eff_Ra)