Research Article

Comparison of Piecewise Structural Equation Modeling and Bayesian Network for *De Novo* Construction of a Quantitative Adverse Outcome Pathway Network

Yang Cao[1,2], S. Jannicke Moe[3], Riccardo De Bin[4], Knut Erik Tollefsen[3,5,6] and You Song[3]

[1] Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, Örebro, Sweden; [2] Unit of Integrative Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; [3] Norwegian Institute for Water Research (NIVA), Oslo, Norway; [4] Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway; [5] Centre for Environmental Radioactivity (CERAD), Norwegian University of Life Sciences (NMBU), Ås, Norway; [6] Norwegian University of Life Sciences (NMBU), Faculty of Environmental Sciences and Natural Resource Management (MINA), Ås, Norway

Abstract
Quantitative adverse outcome pathway network (qAOPN) is gaining momentum due to the predictive nature, alignment with quantitative risk assessment and great potential as a computational new approach methodology (NAM) to reduce laboratory animal tests. The present work aimed to demonstrate two advanced modeling approaches, piecewise structural equation modeling (PSEM) and Bayesian network (BN), for *de novo* qAOPN model construction based on routine ecotoxicological data. A previously published AOP network comprised of four linear AOPs linking excessive reactive oxygen species production to mortality in aquatic organisms was employed as a case study. The demonstrative case study intended to answer: Which linear AOP in the network contributed the most to the AO? Can any of the upstream KEs accurately predict the AO? What are the advantages and limitations of PSEM or BN in qAOPN development? The outcomes from the two approaches showed that both PSEM and Bayesian network were suitable for constructing a complex qAOPN based on limited experimental data. Besides quantification of response-response relationships, both approaches were capable of identifying the most influencing linear AOP in a complex network and evaluating the predictive ability of the AOP, albeit some discrepancies in predictive ability were identified for the two approaches using this specific dataset. The PROs and CONs of the two approaches for qAOPN construction were discussed in detail and suggestions on optimal workflows of PSEM and BN were provided to guide future qAOPN development.

1 Introduction

The adverse outcome pathway (AOP) framework has been introduced to better utilize the existing (eco)toxicological information for mechanistically based risk assessment of environmental stressors (Ankley et al., 2010). An AOP is a pathway that links a cascade of causally related toxicological events, from a molecular initiating event (MIE), multiple key events (KE) occurring at increasing levels of biological organization, to an adverse outcome (AO) of regulatory concern. Multiple AOPs can be cross-linked to form large AOP networks (AOPN). The strength of an AOP is determined by its evidence support for causalities between different events, and the stressor and biological domains of applicability. A well-developed AOP allows better utilization of data generated from cost-efficient new approach methodologies (NAMs), such as *in vitro* high-throughput screening, high-content omics and computational models, to inform regulatory decision making, thereby reducing the number and cost for conventional toxicity assessment using laboratory animals.

A quantitative AOP network (qAOPN) is an advanced form of AOP that captures the quantitative relationships of the key events and biological complexities in a causal network, enabling the prediction of regulatory relevant AO using a mathematical/computational model. A number of mathematical and computational approaches have been proposed to facilitate the development of qAOPN models (reviewed in (Perkins et al., 2019; Spinu et al., 2020)), with regression-based deterministic approaches (Foran et al., 2019; Perkins et al., 2019; Song et al., 2020) and Bayesian network-based probabilistic approaches (Zgheib et al., 2019; Moe et al., 2021b; Song et al., 2021; Spinu et al., 2022) being extensively explored for *de novo* model

construction. Other approaches such as systems biology models have also been introduced, but may require substantial data support and mechanistic understanding of the biological processes (Zgheib et al., 2019).

Among these approaches, structural equation modeling (SEM) has been increasingly used in various types of studies as a means to infer causal processes (Grace et al., 2012). Various statistical techniques and tools have been used in the process of specifying and evaluating the models based on SEM (Kaplan, 2009). In traditional SEM, the relationships among variables (i.e., the linear coefficients) are estimated simultaneously in a single variance-covariance matrix. Although the approach is well developed, SEM can be computationally intensive (depending on the sizes of the variance-covariance matrix). Moreover, additional assumptions such as independence and normality of errors are generally violated in studies or hard to test when the sample size is small (Yuan and Bentler, 1998). Piecewise structural equation modeling (PSEM) is a more flexible and potentially more powerful technique of SEM, which was proposed as an alternate approach to traditional variance-covariance based SEM in the early 2000s (Lefcheck, 2016). The PSEM approach is based on directed acyclic graphs (DAG) and not restricted by the shortcomings of conventional SEM such as multivariate normality and independency of the observations (Pugesek et al., 2003). These key characteristics make PSEM a promising approach for network model construction. To date, no study has employed the PSEM approach for qAOPN development.

Probabilistic methods such as Bayesian network (BN) modeling are being increasingly used in ecotoxicology and risk assessment (Moe et al., 2021a; Maertens et al., 2022). Bayesian network models can also be represented by DAGs where causal/empirical relationships and the associated uncertainty/variability are quantified in conditional probability tables (CPTs). Bayesian networks inherently incorporate uncertainty and can integrate a variety of information types, including expert elicitation. A BN approach has been proposed as an alternative for quantification of AOPNs (Moe et al., 2021b). This approach is based on a combination of regression modeling and BN modeling and is typically less data-demanding than the systems biology models. The quantified AOP-BN model can be run in several directions: 1) prognostic inference, run forward from the stressor node to predict the AO level; 2) diagnostic inference, run backward from the AO node; and 3) omnidirectionally, run from the intermediate MIEs and/or KEs. The BN-qAOPN reported by Moe et al. (2021b) was able to predict the state of AO with high accuracy rate when run from intermediate KE nodes, but the model performance decreased as the starting point was further removed from the target node.

The present study used a previously published AOPN and empirical dataset to construct novel qAOPN models. Three research questions are intended to be addressed by the present study: 1) which linear AOP in the network contributed the most to the AO? 2) can any of the upstream KE (including MIE) accurately predict the AO; 3) what are the advantages and limitations of PSEM or BN in qAOPN development? The model construction workflows and outcomes of the two approaches, PSEM and BN, were compared to suggest the "best practice" for *de novo* qAOPN development.

2 Materials and methods

2.1 AOP network

An AOPN linking reactive oxygen species (ROS) production to mortality was used as the prototype for qAOPN development in this study. This AOPN has been submitted to the AOP repository database AOPWiki as AOP # 327-330[1]. The AOPN contains four linear AOPs with a common MIE (ROS production) and AO (mortality, Fig. 1). The supporting empirical dataset for model construction was derived from a previously published laboratory experiment, where the model aquatic crustacean *Daphnia magna* (14 days old females) were exposed to a gradient (0, 0.0008, 0.05, 0.1, 0.2, 0.4 w/m^2) of a well-known oxidative stressor, ultraviolet B (UVB) radiation, for 2 and 7 days (Song et al., 2020). The AOPN (Fig. 1) serves as a conceptual model and qualitative structure for both model construction approaches (PSEM and BN).
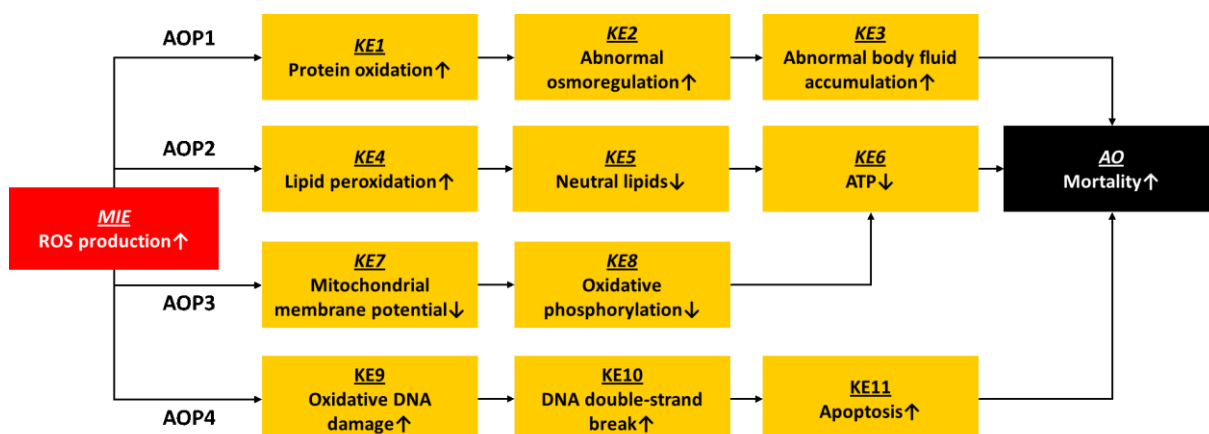

Fig. 1: An adverse outcome pathway network (qAOPN) linking excessive reactive oxygen species (ROS) production to mortality
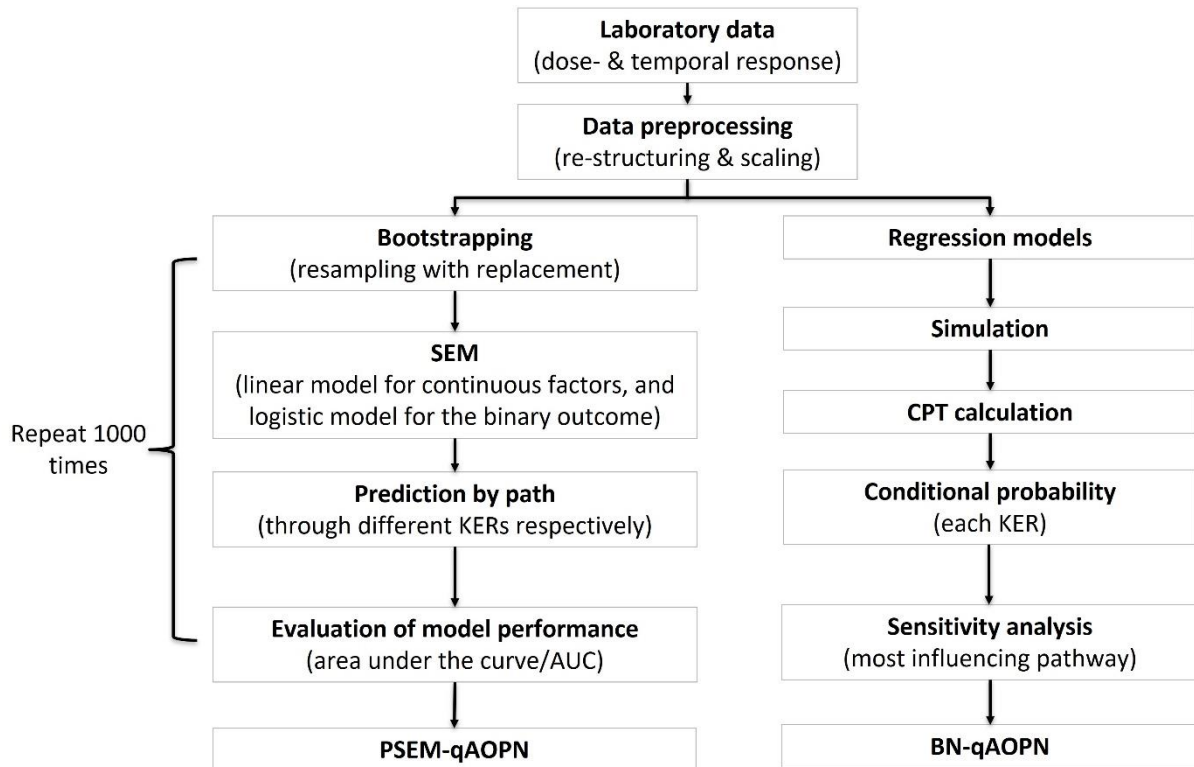MIE: molecular initiating event; KE: key event; AO: adverse outcome.

Fig. 2: The workflows for *de novo* quantitative adverse outcome pathway network (qAOPN) construction using two parallel approaches, piecewise structural equation modeling (PSEM) and Bayesian network (BN) modelling
KER: key event relationship; CPT: conditional probability table.

2.2 Piecewise structural equation modeling

In PSEM, each set of relationships is estimated independently (or locally). This process decomposes the qAOPN into the corresponding simple or multiple linear regressions for each response, each of which are evaluated separately, and then later combined to generate inferences about the entire network. Unlike traditional SEM, where the χ2 test is commonly used to compare the observed and fitted covariance matrices and evaluate the overall goodness-of-fit, the goodness-of-fit of a PSEM model is tested using the directed separation of a DAG, which tests whether the path coefficients are significantly different from zero or whether we are justified in excluding them (Shipley, 2000; Shipley and Douma, 2021).

The coefficients of the qAOPN were estimated following these steps: 1) the values of the variables were scaled by normalizing to the mean values in the control group (unexposed group) prior to analysis; 2) a random sample of the observed data was generated using the bootstrapping method by a random resampling with replacement (Varian, 2005); 3) the coefficient of each edge in the qAOPN (Fig. 1) was estimated based on the PSEM algorithm. Simple linear regression models were used for the edges connecting the continuous variables, and multivariable logistic regression models were used for the edges connecting KE3, KE6, and KE11 with AO; 4) the predictive ability in terms of area under the receiver operating characteristic (ROC) curve of the four linear AOPs (Fig. 1) were evaluated using logistic regressions models. Since a negative effect of a node on the AO could be offset by a positive effect of another node through the pathway, or transmit to the next node, both additive ($logit(AO)=\Sigma\beta_i KE_i+e$) and multiplicative ($logit(AO)=\Pi\beta_i KE_i+e$) models were adopted. These steps (Fig. 2) were repeated 1000 times and 1000 estimations were obtained for each coefficient and area under the ROC curve (AUC). Finally, distribution, mean, and 95% confidence interval of the estimates were obtained based on the 1000 estimations.

2.3 Bayesian network modeling

The BN approach for constructing a qAOPN mainly followed the procedure described by (Moe et al., 2021b), using a combination of explorative statistical modeling, parametric regression, and simulation to quantify all components of the BN model. The general steps are as follows: 1) for each key event relationship (KER), quantify the relationship between predictor and response variables by regression modeling; 2) apply the fitted regression models with associated uncertainty to simulate a high number of response values along the predictor gradient; 3) use the simulated values to parameterize the conditional probability tables, which are the quantitative links (KERs) in the BN model (see Fig. 2 for a summary of the workflow). A general issue for the analysis of the KERs is to identify the appropriate data combinations for response-response relationships. For each KE in the dataset used here, the response variable was measured in at least three replicates (i.e., three individual organisms). However, since the test material must be processed for a given response variable, an individual organism could not be used to measure more than one response variable. Therefore, there was no obvious way to directly link the replicate of one KE to a given replicate of another KE. For example, replicate 1 of KE-4 may be linked to replicate 1, 2, or 3 of KE-5. The solution used here was to generate paired predictor-response (x-y) values for the KERs using all possible combinations of the replicates for each x-y relationship. For example, KER no. 7 (KE-4 -> KE-5) would have 3x3 = 9 data points. This approach inflated the number of data point, whereas avoided the need for a subjective or random selection of x-y pairs.

After preprocessing the data, the following procedure was used for each KER individually: 1) The functional form of the x-y relationship was explored and characterized using non-parametric regression (generalized additive models - GAM);
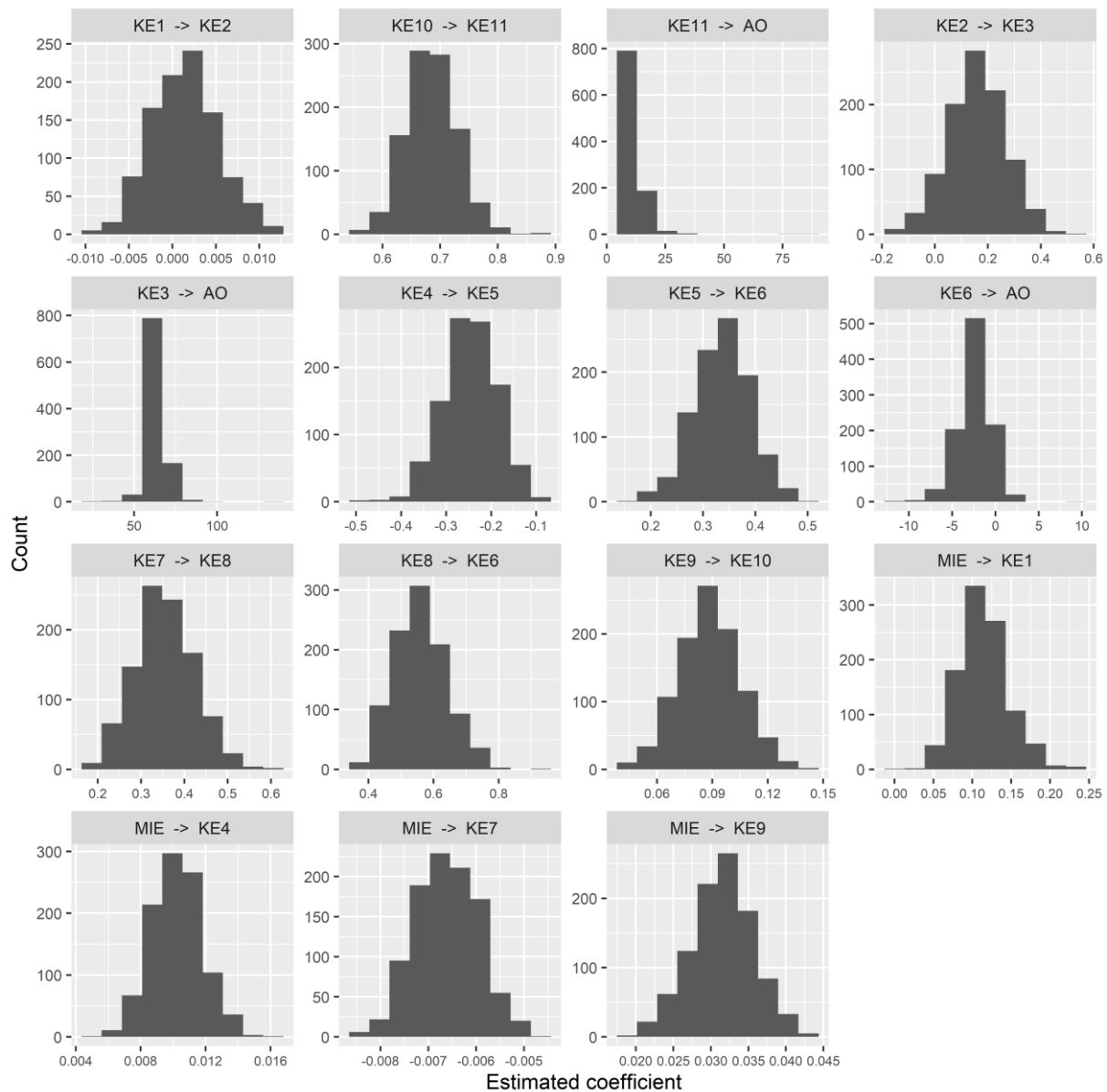
Fig. 3: Distribution of the coefficients between the adjacent key events (KEs) and adverse outcome (AO) in the quantitative adverse outcome pathway network (qAOPN)

2) An appropriate parametric regression model based on inspection of the fitted GAM curve. A ranking of the alternative models by the percent of deviance explained compared to the null model was used to guide the model selection. For the purpose of comparing the BN approach with the PSEM, which is based on linear models, we considered only linear models or linear models with transformation of the predictor variable (quadratic, logit and logarithmic), corresponding to generalized linear models (GLM); 3) The estimated coefficients and standard error of the fitted regression model was used to simulate a large number of data points (10,000) along the x-axis; 4) The x and y variables were discretized into 5-6 intervals required for implementation of BN. Regression tree analysis was used for identifying breakpoints in the x-y relationships, as a basis for selecting the appropriate number of intervals and the interval boundaries for each variable; 5) The frequency distribution of the simulated data points across the x-y grid cell were used to derive the CPT of the x-y relationship.

In addition to the method for quantification of the BN based on simulations from linear models (labelled SIM-LIN as described above), two simpler methods were applied to quantify the BN for comparison. For both alternatives, the CPT of each KER was quantified with an equation specified directly in the BN analytical software Netica version 6.04[2] (Norsys Software Corp., Vancouver, Canada). The equation was based on the coefficients estimated by the GLM (as above), but without the estimated standard deviation. Instead, the uncertainty of each KER was generated by the built-in sampling uncertainty function in Netica. The first alternative BN, labelled EQU-LIN, had all KERs estimated as linear models, except for those with mortality as the response (logistic models). This set of models corresponds to the models used for SIM-LIN and for PSEM. For the second alternative BN, labelled EQU_SEL, a more comprehensive model selection was performed for each KER. The candidate regression models were linear, log-linear, quadratic and logistic (sigmoid). For each KER, the optimal model was
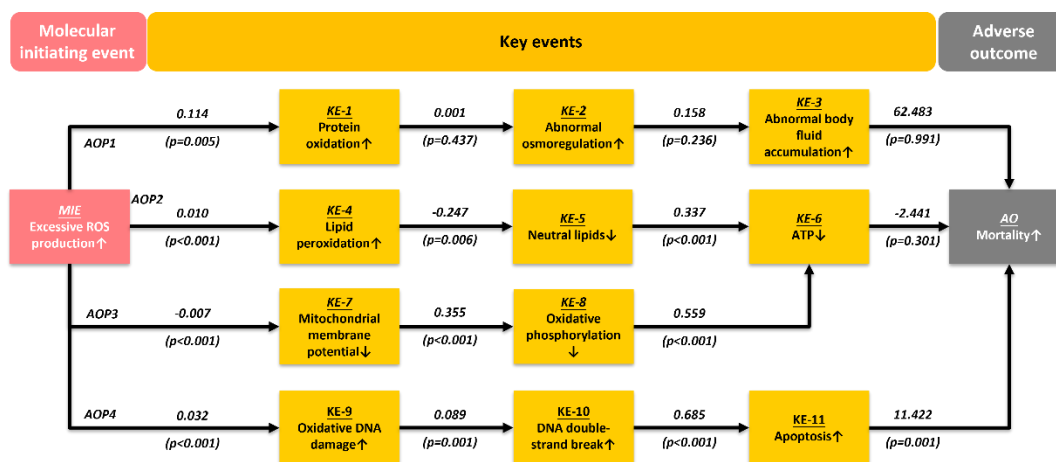
---

[2] http://www.norsys.com

Fig. 4: Average coefficients over the 1000 bootstrap iterations for the key event relationships
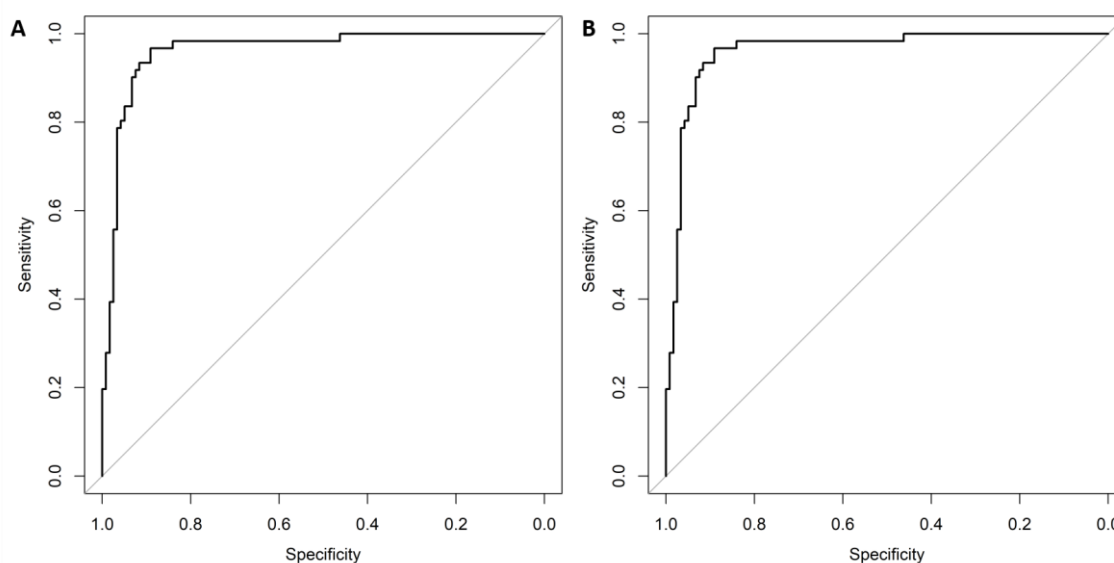

Fig. 5: One of the 1000 evaluations of predictive ability of the quantitative adverse outcome pathway #4 (AOP4) in the network using multiplicative (A) and additive multivariable (B) logistic regression models

selected based on the percentage of deviance explained, supported by visual inspection of the scatter plots and fitted GAM curves.

For nodes with more than one parent nodes, such as KE-6 and the AO, the parent nodes were assigned equal weights in the BN model. A differentiated weighting of multiple parent nodes, and thereby the different pathways of the qAOPN, can be further refined later.

Sensitivity analysis was applied to characterize the influence of individual parent nodes on the target node (adverse outcome - mortality). The sensitivity is measured as mutual information between the target node and the parent node, which corresponds to the reduction in entropy of the target node (measured in bits) due to a finding (evidence) at the parent node.

3 Results

3.1 PSEM-qAOPN

By using the PSEM approach, the regression coefficients were estimated for all adjacent KEs, with the KEs in AOP3 and 4 being highly significant (p<0.001) compared to that in AOP1 and 2 (Fig. 3). As all the coefficients between any two adjacent KEs in AOP4 were statistically significant, the prediction performance was further evaluated for this specific qAOP. Multiplicative and additive multivariable logistic regression models were fitted, and the AUCs of both models were larger than 0.9, indicating strong predictive ability of the qAOP.

The distributions and average values of the coefficients of the edges between the adjacent KEs and AO along the pathways in the qAOPN are shown in Fig. 3.

Eleven of the 15 coefficients were statistically significant (p < 0.01), and all coefficients in AOP4 (from MIE, KE9, KE10, and KE11 to AO) were statistically significant (Fig. 4).

Both multiplicative and additive multivariable logistic regression models were further fitted, and the AUCs of both models were larger than 0.9 (Fig. 5), indicating strong predictive ability of AOP4.

Tab. 1: Predictive ability of the quantitative adverse outcome pathways (qAOPs) as indicated by area under the curve (AUC)
The values in the table are mean and 95% confidence interval of the 1000 evaluations.

qAOP	Multiplicative model	Additive model
qAOP1	0.94 (0.91, 0.96)	0.93 (0.90, 0.96)
qAOP2	0.94 (0.91, 0.97)	0.94 (0.91, 0.97)
qAOP3	0.93 (0.89, 0.96)	0.93 (0.89, 0.96)
qAOP4	0.94 (0.91, 0.97)	0.94 (0.90, 0.97)

Tab. 2: Sensitivity analysis for the Bayesian network model with the three alternative parameterizations of conditional probability tables
The prefixes SIM- and EQU- refer to the source of uncertainty: SIM = from simulations based on regression model estimates; EQU = from the built-in sampling uncertainty function in Netica. The suffices -LIN and -SEL refer to the model selection: LIN = linear; SEL = selected as best fitting.

SIM-LIN		EQU-LIN		EQU-SEL	
AOPs ranked by influence	Mutual information (%)	AOPs ranked by influence	Mutual information (%)	AOPs ranked by influence	Mutual information (%)
AOP1	29.9	AOP1	15.6	AOP2+3	31.7
AOP2+3	14.9	AOP4	13.5	AOP1	16.1
AOP4	4.61	AOP2+3	12.6	AOP4	0.399
KEs ranked by influence	Mutual information (%)	KEs ranked by influence	Mutual information (%)	KEs ranked by influence	Mutual information (%)
KE-6	10.9	KE-3	2.56	KE-6	16.9
KE-3	1.23	KE-2	0.536	KE-11	0.404
KE-11	0.51	KE-11	0.201	KE-10	0.159

Although one or two coefficients in the AOP1, AOP2, and AOP3 were not statistically significant, they showed excellent predictive ability regarding the AO and all the AUCs were larger than 0.9 (Table 1).

3.2 BN-qAOPN

The curves estimated by GAM together with the estimated standard errors (Fig. 6) were used to characterize the shape of the KERs. Thresholds estimated by regression tree modeling were applied to identify the most important breakpoints in the KER shape.

The selected node intervals guided by the regression tree analysis are shown as gridlines in Fig. 6. with the selected interval boundaries guided by the regression tree analysis. Most of the KE nodes were discretized into 5 intervals, which was considered sufficient for capturing the KERs. The stressor node (UVB) and the MIE node (ROS production) were given a higher resolution (8 and 6 intervals, respectively), to better capture changes in the start of the AOP. The AO node was discretized into 5 equidistant intervals, for easier interpretation of the outcome.

For the BN presented here, we attempted to align the structure with the SEM as closely as possible to facilitate comparison. Therefore, the conditional probability tables (CPTs) of all key events were derived from linear models ($y = b0 + b1* x$), while the CPTs of the AO nodes were derived from logistic models $y = \exp(b0 + b1* x)/(1 + \exp(b0 + b1* x))$. The estimated parameters are reported in Table 2.

The resulting structure of the BN model with the initial settings is shown in Fig. 7(a-c). The circles represent intermediate versions of the subsequent nodes (KE-6 and AO) belonging to different pathways. In both cases, the intermediate nodes are combined into their respective child nodes by equal weighting. Fig. 7a shows the uninstantiated (default) version with uniform probability distribution for the stressor node and prior probability distributions for the subsequent nodes. For each child node, the prior probability distribution was determined by the conditional probability table, represented by an arrow from its parent node(s). When the BN was instantiated for the highest UV level (Fig. 7b), the probability distribution of the MIE (ROS production) shifted towards higher intervals (average increased from 19 to 30). The probability distributions of the subsequent key events correspondingly shifted towards higher or lower values; the signal was fading for each step. In consequence, the predicted AO (mortality) was practically unaltered (average from 0.406 to 0.41). However, the BN can also be instantiated from one of the key events, which are fewer steps away from the AO and therefore more of the signal is retained. For example, a situation with the highest interval of KE-6 (ATP, representing intracellular energy transfer) predicts a lower probability of mortality (average 0.35). Running the BN model from an intermediate node can also provide a prediction from the parent nodes (diagnostic inference). For example, the highest interval of ATP corresponds to a lower stressor level than the initial setting: the probability of UVB exposure being below 0.075 increased from the initial 37.5% to 68%.

Two alternative versions of the BN model were also constructed (EQU-LIN and EQU-SEL) to explore the performance of the BN with different assumptions underlying the quantification of the CPTs (Table 2). For all alternative parameterization, the sensitivity of the target node (the final AO node) to the stressor and to the MIE node were close to zero. An analysis of the model performance in terms of accuracy (correct predictions) when the model is run from these nodes is therefore not meaningful. Instead, the sensitivities of the different model components have been inspected more closely.

For each pathway, the sensitivity score listed for an AOP represents the percentage mutual information between the intermediate AO node of the pathway (shown as grey circles in Fig. 7) and the target node (the final AO node). Likewise, the sensitivity listed for individual KE nodes is the mutual information between the KE node and the target node. In general, the target node had very low sensitivity to changes in the stressor node, for all of the alternative parameterizations (Table 2). This reflect the low percentage of deviance explained for some of the KERs in the network even when the relationships are estimated
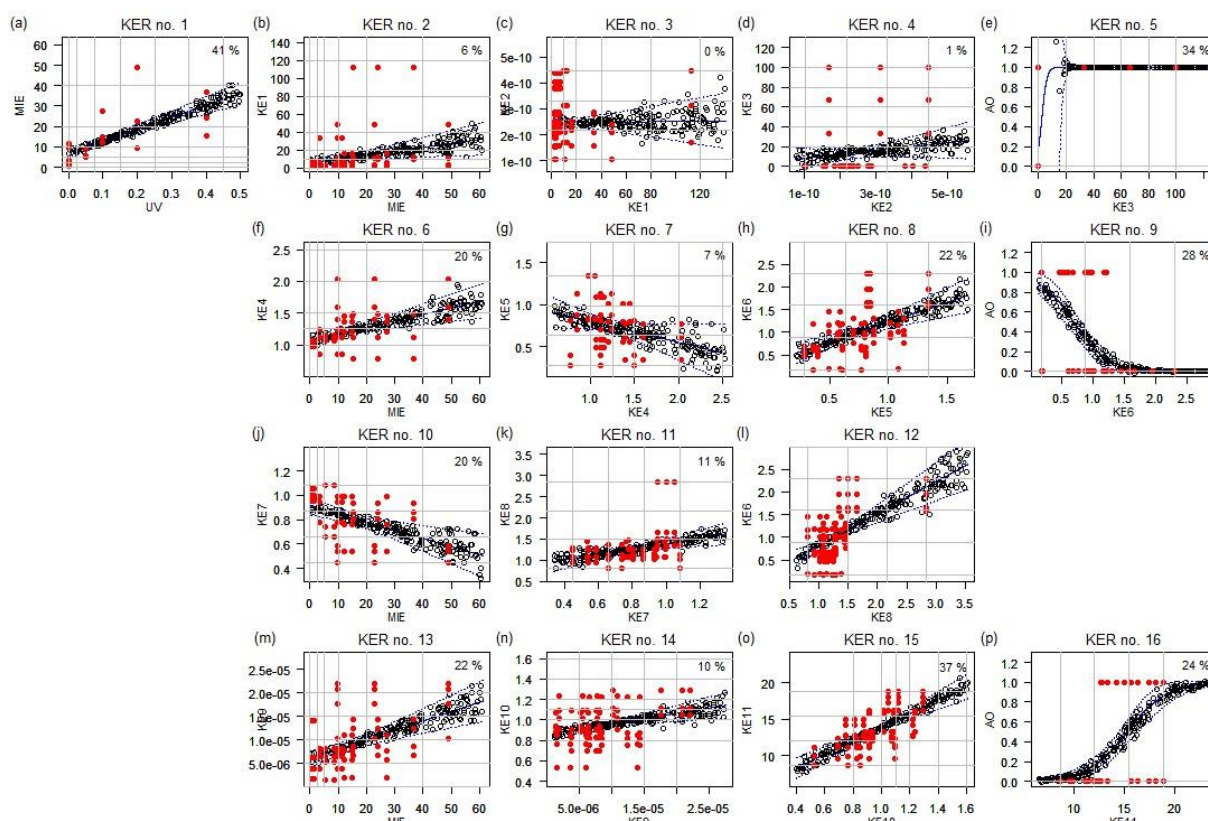
AOP for UV - Daphnia: glm plots with simulations

Fig. 6: Observations (red dots) and fitted generalized additive models (GAM, black curves) with standard errors (SE, dashed blue lines)
The numbers in the upper right corner are the percentage of deviance explained by the GAM model.

with non-parametric regression by GAM (Fig. 6), which in general provided a better fit than any of the parametric models. The comparison of the three alternative parameterizations shows that the rank of influence of the AOPs differs among the three alternatives. For example, when the KERs are forced to be linear, the pathway AOP1 has the highest influence; but when the best-fitting regression model is selected, the pathway AOP2+3 has higher influence. This means that the relative influence of the different AOPs in the network depends on the method used for estimation of coefficients and quantification of uncertainty for the CPTs. Likewise, the influence of the individual KERs on the target node also depends on the method of BN parameterization.

Considering the sum of percent mutual information for all the nodes combined (not shown), the BN based on CPTs from simulations had higher sensitivity (69%) than the BN based on linear equation with built-in sampling uncertainty (EQU-LIN, 45%), but slightly lower than the BN based on selected best-fitting regression models (EQU-SEL, 73).

4 Discussion

The present study has employed two directed acyclic graph (DAG) types of model construction approaches, piecewise structural equation modeling (PSEM) and Bayesian network (BN) for *de novo* construction of a complex quantitative adverse outcome pathway network (qAOPN). Both approaches were able to quantify the response-response relationships and assist the identification of the most influencing AOP in a network. The advantages and limitations of these approaches will be discussed below.

4.1 Piecewise structural equation modeling

Although PSEM has been successfully applied in some studies such as determination of the links between the environmental conditions and diatom diazotroph associations (Stenegren et al., 2017) and evaluation of the role of quantity implicature in child language (Grinstead et al., 2022), it is still new in the field of qAOP development. Traditional SEM requires multivariate normality and independence of the variables, like many other statistical tests. The new extension of the PSEM relaxes these assumptions, so it is now possible to fit models to a variety of non-normal distributions and model non-independence through specification of fixed correlation structures or even random effects (Lefcheck, 2016). Unlike traditional modeling approaches, PSEM is often preferred in a system where a number of predictors and outcomes are connected. It incorporates hypothesized causalities between the nodes in the network, which is based on previous knowledge of the investigated system, and it facilitates identification of the direct, indirect, and cascading effects in the system. Although PSEM is built on assumed causalities, whether the causalities are correct is difficult to test when the predictors and outcomes were observed at the same time. Fortunately, in the current experimental study, the predictors and outcome occurred sequentially, indicating that the causalities were not inversed. It should be noted that the PSEM also needs to meet all the assumptions of the underlying generalized linear
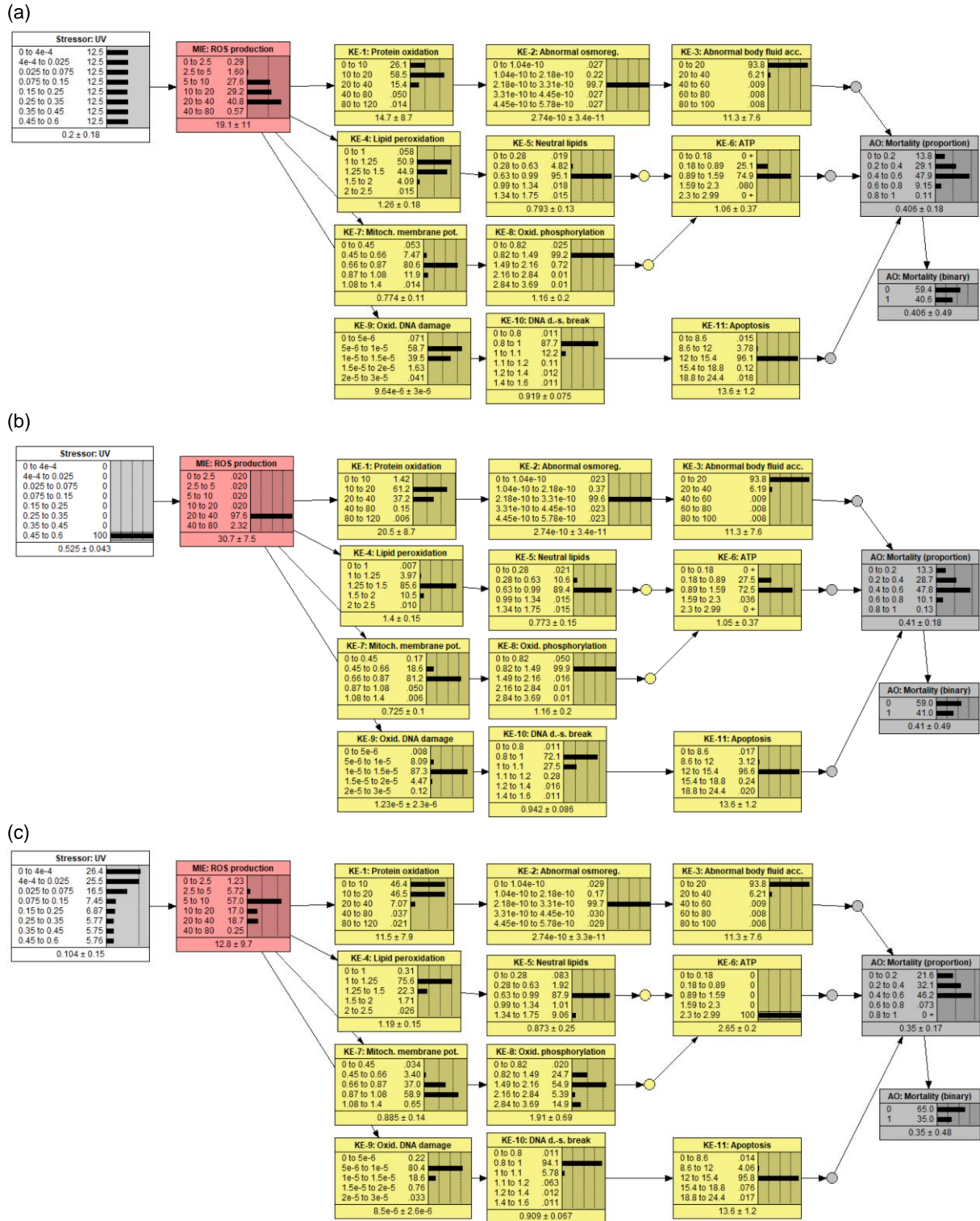
Fig. 7: A Bayesian network (BN) representation of a quantitative adverse outcome pathway network (qAOPN)
(a) The BN with prior probabilities and uniform probability distribution for the stressor node, UV. (b) The BN instantiated from the stressor node. (c) The BN instantiated from key event (KE) no. 6, ATP.

models, such as homogeneity of variance, which was obtained in this study by scaling the variables during preprocessing of raw data.

While the PSEM is a considerable leap forward of traditional SEM, there are also limitations in the implementation of PSEM in the current study. First, only direct effects between the nodes were investigated due to the complexity of the network, Second, we did not incorporate latent or composite variables in the current study, which is indeed one of the great advantages of conventional SEM. Nevertheless, the qAOPs constructed by PSEM showed excellent predictive ability, suggesting that the KEs acted directly and in concert to influence the AO. In addition, no bidirectional relationships caused by

Tab. 3: Parameter estimates from regression models from the piecewise structural equation modeling (PSEM) and Bayesian network (BN) workflows
Note that the PSEM used variables in normalized scale while the BN workflow used variables in original scale (raw data), which resulted in different coefficient of different magnitudes for some of the key event relationships (KERs). MIE: molecular initiating event; KE: key event; AO: adverse outcome.

KER no.	Predictor (x)	Response (y)	Model	PSEM Coefficient (β)	PSEM Standard error	PSEM p-value	BN Coefficient (β)	BN Standard error	BN p-value
2	MIE	KE_1	linear	0.114	0.030	0.005	0.437	0.184	0.020
3	KE_1	KE_2	linear	0.001	0.003	0.437	<0.001	<0.001	0.874
4	KE_2	KE_3	linear	0.158	0.102	0.236	>1000	2.87E+10	0.167
5	KE_3	AO	logistic	62.483	76.011	0.991	0.559	41.40	0.989
6	MIE	KE_4	linear	0.010	0.002	<0.001	0.010	0.003	0.001
7	KE_4	KE_5	linear	-0.247	0.069	0.006	-0.246	0.127	0.058
8	KE_5	KE_6	linear	0.337	0.062	<0.001	0.864	0.195	<0.001
9	KE_6	AO	logistic	-2.411	1.981	0.301	-3.647	0.689	<0.001
10	MIE	KE_7	linear	-0.007	0.001	<0.001	-0.007	0.002	0.001
11	KE_7	KE_8	linear	0.355	0.076	<0.001	0.648	0.224	0.005
12	KE_8	KE_6	linear	0.559	0.085	<0.001	0.705	0.119	<0.001
13	MIE	KE_9	linear	0.032	0.004	<0.001	<0.001	<0.001	<0.001
14	KE_9	KE_10	linear	0.089	0.020	0.001	>1000	3158	0.001
15	KE_10	KE_11	linear	0.685	0.067	<0.001	9.27	1.25	<0.001
16	KE_11	AO	logistic	11.422	2.805	0.001	0.509	0.090	<0.001

a shared underlying driver have been considered, only crude approximate correlated errors are implemented, and cyclic relationships cannot be evaluated (Lefcheck, 2016).

4.2 Bayesian network

An inherent strength of BN models is the explicit and transparent handling of uncertainty in each step (node) of the model. In this study, for each KER, the predictor-response relationships as well as the variability was first explored by GAM and then quantified by generalized linear models (including linear regression). The uncertainty estimated by the selected GLM could then be incorporated into each KER of the BN via the CPT. In the AOP context, the BN represents a causal model, in a similar way as the PSEM. An additional benefit of the BN compared to the PSEM is the flexibility of the CPTs (causal links) to allow for more flexible relationships. Although the CPTs in this study were limited to GLMs for easier comparison with PSEM, a CPT can in principle represent any functional form. Moreover, a CPT can easily capture interactions (synergisms or antagonisms) between multiple parent nodes, although this aspect was not explored by the present study.

Another unique feature of BN models is the ability to run the model in multiple directions: prognostic (forward) from the stressor node or from the MIE; diagnostic (backward) from the AO, or omnidirectional from one or more intermediate key events. This feature allows for more types of inferences from a qAOP(N), if the model has sufficiently high predictive ability.

The BN model developed here, however, did not obtain sufficient sensitivity to have predictive ability. There are several explanations for the lack of sensitivity. One reason is the low correlation of x-y values for some the KERs (Figure 6), resulting in a weak signal in the corresponding CPTs. A different handling of the replicates could have resulted in lower uncertainty in the BN. For example, selecting only one replicate for each x and each y by random would potentially lead to lower variability and thereby better model fit. Alternatives for handling of response-response relationships was beyond the scope of this paper but will be addressed in later investigations.

Another obvious reason for low sensitivity of the BN model is the discretization of continuous variables, which inevitably results in loss of information and precision. The discretization of variables is a standard process and inherent shortcoming of traditional BN modeling. However, more advanced BN methods allow for continuous nodes (e.g., Jackson-Blake et al., 2022) or a combination of discrete and continuous (e.g., Moe et al. (2020). The propagation of quantified uncertainty throughout the model is also a unique property of BN models, which provides transparency and reflects the modeler's best knowledge of the system, but also contributes to low sensitivity of the target node (adverse outcome) to the earlier components of the network.

The comparison of the three methods for parameterization of CPTs (Table 2) showed that the method with estimation of linear models with uncertainty and subsequent simulations (SIM-LIN) had overall higher sensitivity than the alternative with built-in uncertainty (EQU-LIN). This comparison suggests that the combined estimation and simulation method (SIM-LIN), which provides a precise characterization of the uncertainty in the linear KERs, has the potential to generate qAOPNs with high sensitivity, even though the resulting qAOPN in this example had low predictability. However, the alternative method where the best-fitting regression models were selected for generating CPTs (EQU-SEL) resulted in even higher sensitivity, suggesting that selecting appropriate non-linear parametric model for the KERs can be important overall model sensitivity of the qAOPN. A repetition of this exercise on a new dataset with more observation and/or with less variation among replicates might result in a BN-AOP which also has higher predictability.

4.3   Recommendations for future direction

Comparing the two approaches used in this paper, PSEM and BN, the PSEM obtained higher predictive ability than then BN for this given dataset. However, in spite of the low overall sensitivity of the BN model, some important properties have been identified for the purpose of qAOPN development. The exploration of alternative methods for parameterization of the CPTs (Table 3) showed that the selection of the parametric model underlying each CPT is also of importance, both for the overall sensitivity of the BN, and for the relative sensitivity of the individual AOPs in the network. An optimal workflow for qAOPN development would combine the more flexible modeling of KERs allowed by the BN, with the higher predictive ability displayed by the PSEM. Apart from these, additional experimentation that captures the dose-response patterns across a time scale (i.e., temporal response) may be beneficial for constructing more robust causal models using these approaches. Although not achieved in the present study, validation of the qAOPN models by follow-up experimental studies would be highly appreciated in the future.

5   Conclusion

The present study has demonstrated that both piecewise structural equation modeling and Bayesian network are suitable for *de novo* construction of a complex qAOP network. Besides quantification of response-response relationships, these two approaches were also capable of identifying the most influencing linear AOP in a complex AOPN and evaluating the predictive ability of the qAOP. Future studies will also identify the most critical (influencing) KE(s) in the qAOPN and how accurately this KE can be used to predict the final adverse outcome that is relevant for hazard and risk assessment. The workflow described herein can be easily repeated by other laboratories for efficient qAOP model construction utilizing routine ecotoxity data.

References

Ankley, G. T., Bennett, R. S., Erickson, R. J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem 29*, 730-741. doi:10.1002/etc.34

Foran, C. M., Rycroft, T., Keisler, J. et al. (2019). A modular approach for assembly of quantitative adverse outcome pathways. *ALTEX 36*, 353-362. doi:10.14573/altex.1810181

Grace, J. B., Schoolmaster Jr., D. R., Guntenspergen, G. R. et al. (2012). Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere 3*, art73. doi:10.1890/ES12-00048.1

Grinstead, J., Ortiz-Ramirez, P., Carreto-Guadarrama, X. et al. (2022). Piecewise structural equation modeling of the quantity implicature in child language. *Lang Speech* 238309211066086. doi:10.1177/00238309211066086

Jackson-Blake, L., Clayer, F., Haande, S. et al. (2022). Seasonal forecasting of lake water quality and algal bloom risk using a continuous Gaussian Bayesian network. *Hydrol Earth Syst Sci Discuss 2022*, 1-34. doi:10.5194/hess-2021-621

Kaplan, D. (2009). *Structural equation modeling (2nd ed.): Foundations and extensions*. Vol. 2nd. doi:10.4135/9781452226576

Lefcheck, J. S. (2016). Piecewisesem: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Meth Ecol Evol 7*, 573-579. doi:10.1111/2041-210X.12512

Maertens, A., Golden, E., Luechtefeld, T. H. et al. (2022). Probabilistic risk assessment - the keystone for the future of toxicology. *ALTEX 39*, 3-29. doi:10.14573/altex.2201081

Moe, S. J., Madsen, A. L., Connors, K. A. et al. (2020). Development of a hybrid Bayesian network model for predicting acute fish toxicity using multiple lines of evidence. *Environmental Modelling & Software 126*, 104655. doi:10.1016/j.envsoft.2020.104655

Moe, S. J., Carriger, J. F. and Glendell, M. (2021a). Increased use of Bayesian network models has improved environmental risk assessments. *Integr Environ Assess Manag 17*, 53-61. doi:10.1002/ieam.4369

Moe, S. J., Wolf, R., Xie, L. et al. (2021b). Quantification of an adverse outcome pathway network by Bayesian regression and Bayesian network modeling. *Integr Environ Assess Manag 17*, 147-164. doi:10.1002/ieam.4348

Perkins, E. J., Ashauer, R., Burgoon, L. et al. (2019). Building and applying quantitative adverse outcome pathway models for chemical hazard and risk assessment. *Environ Toxicol Chem 38*, 1850-1865. doi:10.1002/etc.4505

Pugesek, B. H., Tomer, A. and von Eye, A. (2003). *Structural equation modeling: Applications in ecological and evolutionary biology*. Vol. Cambridge, MA: Cambridge University Press. http://pubs.er.usgs.gov/publication/70006389

Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling 7*, 206-218. doi:10.1207/S15328007SEM0702_4

Shipley, B. and Douma, J. C. (2021). Testing piecewise structural equations models in the presence of latent variables and including correlated errors. *Structural Equation Modeling 28*, 582-589. doi:10.1080/10705511.2020.1871355

Song, Y., Xie, L., Lee, Y. et al. (2020). De novo development of a quantitative adverse outcome pathway (qAOP) network for ultraviolet b (UVB) radiation using targeted laboratory tests and automated data mining. *Environ Sci Technol 54*, 13147-13156. doi:10.1021/acs.est.0c03794

Song, Y., Kamstra, J. H., Cao, Y. et al. (2021). High-throughput analyses and Bayesian network modeling highlight novel epigenetic adverse outcome pathway networks of DNA methyltransferase inhibitor mediated transgenerational effects. *Journal of Hazardous Materials 408*, 124490. doi:10.1016/j.jhazmat.2020.124490

Spinu, N., Cronin, M. T. D., Enoch, S. J. et al. (2020). Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. *Arch Toxicol 94*, 1497-1510. doi:10.1007/s00204-020-02774-7

Spinu, N., Cronin, M. T. D., Lao, J. et al. (2022). Probabilistic modelling of developmental neurotoxicity based on a simplified adverse outcome pathway network. *Comput Toxicol 21*, 100206. doi:10.1016/j.comtox.2021.100206

Stenegren, M., Berg, C., Padilla, C. C. et al. (2017). Piecewise structural equation model (SEM) disentangles the environmental conditions favoring diatom diazotroph associations (DDAS) in the Western Tropical North Atlantic (WTNA). *Front Microbiol 8*, 810. doi:10.3389/fmicb.2017.00810

Varian, H. (2005). Bootstrap tutorial. *Mathematica Journal 9*, 768-775.

Yuan, K.-H. and Bentler, P. M. (1998). Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology 51*, 289-309. doi:10.1111/j.2044-8317.1998.tb00682.x

Zgheib, E., Gao, W., Limonciel, A. et al. (2019). Application of three approaches for quantitative AOP development to renal toxicity. *Computational Toxicology 11*, 1-13. doi:10.1016/j.comtox.2019.02.001

Data availability statement

The demonstration data and concept codes for PSEM and BN are accessible upon request via Github[3].

Conflict of interest

The authors declare that they have no conflict of interest.

---

[3] https://github.com/RiskAOP/PSEM_BN_qAOPN

[4] https://www.niva.no/en/projectweb/riskaop

[5] http://www.niva.no/nctp