Transferability of a Bayesian Belief Network across diverse agricultural catchments using high-frequency hydrochemistry and land management data

Science of the Total Environment 949 (2024) 174926

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Transferability of a Bayesian Belief Network across diverse agricultural catchments using high-frequency hydrochemistry and land management data

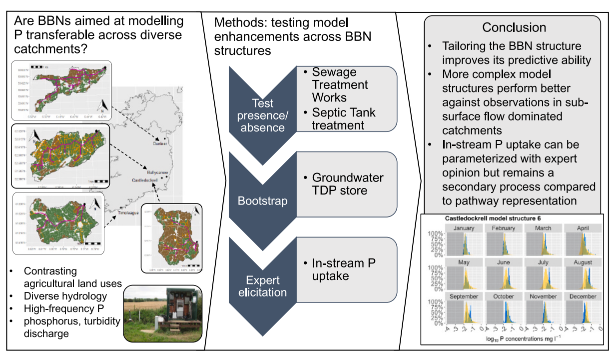
Camilla Negri [a,b,c,d,*], Nicholas Schurch [d], Andrew J. Wade [c], Per-Erik Mellander [a], Marc Stutter [b], Micheal J. Bowes [e], Chisha Chongo Mzyece [b,f], Miriam Glendell [b]

[a] *Agricultural Catchments Programme, Teagasc Environment Research Centre, Johnstown Castle, Co. Wexford Y35 Y521, Ireland*
[b] *The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, UK*
[c] *University of Reading, Department of Geography and Environmental Science, Whiteknights, Reading RG6 6DR, UK*
[d] *Biomathematics and Statistics Scotland, Craigiebuckler, Aberdeen AB15 8QH, UK*
[e] *UK Centre for Ecology & Hydrology, Wallingford OX10 8BB, UK*
[f] *Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK*

HIGHLIGHTS

- Pilot study testing transferability of Bayesian Network for P losses
- Sensitivity Analysis of a hybrid network identified redundant variables.
- Expert elicitation supports model parameterization of uncertain process.
- Models should prioritize P transfer pathways over in-stream cycling.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Ouyang Wei

Original content: A repository to test the transferability of a Bayesian Belief Network (Original data)

Keywords:
Hybrid network
Expert elicitation
Model universality
Sensitivity analysis

ABSTRACT

Biogeochemical catchment models are often developed for a single catchment and, as a result, often generalize poorly beyond this. Evaluating their transferability is an important step in improving their predictive power and application range. We assess the transferability of a recently developed Bayesian Belief Network (BBN) that simulated monthly stream phosphorus (P) concentrations in a poorly-drained grassland catchment through application to three further catchments with different hydrological regimes and agricultural land uses. In all catchments, flow and turbidity were measured sub-hourly from 2009 to 2016 and supplemented with 400–500 soil P test measurements. In addition to a previously parameterized BBN, five further model structures were implemented to incorporate in a stepwise way: in-stream P removal using expert elicitation, additional groundwater P stores and delivery, and the presence or absence of septic tank treatment, and, in one case, Sewage Treatment Works. Model performance was tested through comparison of predicted and observed total

* Corresponding author at: The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland, UK.
E-mail address: camilla.negri@hutton.ac.uk (C. Negri).

https://doi.org/10.1016/j.scitotenv.2024.174926
Received 29 March 2024; Received in revised form 31 May 2024; Accepted 19 July 2024
Available online 24 July 2024

Phosphorus
Structural uncertainty

reactive P (TRP) concentrations and percentage bias (PBIAS). The original BBN accurately simulated the absolute values of observed flow and TRP concentrations in the poorly and moderately drained catchments (albeit with poor apparent percentage bias scores; 76 % ≤ PBIAS≤94 %) irrespective of the dominant land use, but performed less well in the groundwater-dominated catchments. However, including groundwater total dissolved P (TDP) and Sewage Treatment Works (STWs) inputs, and in-stream P uptake improved model performance (−5 % ≤ PBIAS≤18 %). A sensitivity analysis identified redundant variables further helping to streamline the model applications. An enhanced BBN model capable for wider application and generalisation resulted.

1. Introduction

Generalised scientific theories are the most powerful and ideally, water quality models should be applicable to all catchments. A transferable model will likely have greater predictive power and utility, greater confidence that the model performs well for the right reasons, and an ability to help inform data collection for the model's application (Hatum et al., 2022; Mieleitner and Reichert, 2006; Schuwirth et al., 2019).

Testing model transferability is therefore important. Hatum et al. (2022) demonstrated that transferring a model across different seagrass ecosystems, using expert elicitation to support model formulation, enabled forecasting and decision-making. To date, only one Bayesian Belief Network (BBN) aimed at modelling water quality has been tested across multiple catchments. Glendell et al. (2022), tested a hybrid BBN (including both discrete and continuous variables) to predict stream P concentrations and applied the model to seven Scottish catchments. The outcomes demonstrated wider ranges in the BBN predictions than in the observations and, given that inadequate water quality data constrains the calibration and validation of P models (Drohan et al., 2019), the use of high-frequency data was suggested as a means to reduce model predictive uncertainty (Glendell et al., 2022).

Phosphorus retention in river catchments results from a combination of biological, physical, and chemical processes (Withers and Jarvie, 2008) and there is uncertainty around the retention rate in different catchments due to variations in P uptake and release by plants, adsorption to and desorption from sediment, co-precipitation, dissolution, and advection. Both biotic and abiotic in-stream P uptake can be significant, especially during low-flows and effluent exposure (Stutter et al., 2010). Its inclusion could improve process representation, and therefore transferability, in process-based semi-distributed P models (Jackson-Blake et al., 2015) and stochastic P models (Negri et al., 2024).

Mechanistic P models typically include processes such as calcite-P co-precipitation, sorption and desorption to and from suspended sediments, P exchange between the pore water and the water column, P entering the reach from upstream or Sewage Treatment Works (STWs), and epiphyte uptake (for example, the INCA-P model, (Jackson-Blake et al., 2016; Wade et al., 2002)). Similarly, stochastic P models include estimating numerous P sources, their transport through the land-water continuum, and delivery to surface waters (for example, Glendell et al., 2022; Igras and Creed, 2020; Neumann et al., 2023). However, some BBNs modelling P concentrations in the stream lack representation of processes such as stream P retention, as well as groundwater P stores (Glendell et al., 2022; Negri et al., 2024).

The observational evidence to quantify in-stream retention processes is difficult to find in a single catchment, stream, or study area, therefore gathering data and comparing across diverse catchments with different P pressures strengthens findings. Expert elicitation (acquiring experts' opinions using formal protocols, e.g., Krueger et al. (2012) is a route to help model parameterization without having to set up costly site-specific experiments and is often used to inform on model parameter uncertainty (O'Hagan, 2019).

The overall study aim was to test the transferability of a recently developed (BBN Negri et al., 2024) in a grassland catchment in Ireland, and make enhancements as necessary. The aim was addressed through three research objectives: application of the BBN to three additional catchments in Ireland with performance assessment using daily total reactive P (TRP) observations; addition of in-stream P removal processes using expert elicitation; and the assessment of whether step-wise addition of in-stream P uptake, groundwater dissolved P concentration, and the presence or absence of septic tank treatment or Sewage Treatment Works improved model performance and transferability (in terms of reduced percentage bias across all four catchments).

2. Study areas

This study focuses on four (< 1200 ha) agricultural catchments in the east and south of Ireland: Timoleague, Ballycanew, Castledockrell, and Dunleer, that are monitored by the Agricultural catchments Programme (ACP) from Teagasc; a programme created to monitor the effectiveness of Ireland's National Action Programme under the European Union Nitrates Directive (Wall et al., 2011). These catchments have different agricultural land uses and contrasting hydrology. The catchments were chosen because agriculture is the only significant anthropogenic pressure (housing density is low and domestic waste is treated with septic tanks) (Jordan et al., 2012). The location of the four catchments is shown in Fig. 1 and further information about the catchments is given in Negri and Mellander (2024).

2.1. Timoleague

The Timoleague catchment (Fig. 1, bottom left) in County Cork has an area of 758 ha, of which 85–89 % is grassland and 4–5 % tillage, the remainder being non-agricultural land use. Stocking rates are ~2 livestock units (LU) ha^{-1} (Sherriff et al., 2015) and many of the dairy farms are managing soils under derogation (i.e. deviation from the EU Nitrates Directive, with organic N loading between 170 and 250 kg ha^{-1} $year^{-1}$), (Jordan et al., 2012), which indicates the catchment has high P sources (Mellander et al., 2022). In general, the soils are well drained except for small areas neighboring the stream at the valley bottom. The catchment is mostly groundwater-fed with a large proportion (60 %) of TRP delivered via belowground pathways in winter (Mellander et al., 2016).

2.2. Ballycanew

The Ballycanew catchment (Fig. 1, top right) is located in County Wexford. The catchment area is 1191 ha, with 78 % grassland 14–20 % tillage. None of the farms in this catchment are tillage only, all of them have a combination of tillage plus grassland. Two-thirds of the catchments' soils (the lowlands) have poor drainage characteristics due to heavy clay deposits. However, farmers have improved the land for grass production with tile and mole drains. Due to the low soil permeability, the catchment has a flashy hydrology and a high risk of P loss to water through quick and erosive surface pathways during heavy rain events (Mellander et al., 2015).

2.3. Castledockrell

The Castledockrell catchment (Fig. 1, bottom right) is also located in County Wexford. The total area is 1117 ha, typically with 39 % grassland and 54 % tillage. Soils are generally well drained with free draining brown earths underlain by slate and shale, ideal for spring barley, which is the main crop. However, some of the lower lying areas near the stream (east- southeast of the catchment) have gley soils which are artificially drained. The catchment is mostly groundwater-fed with a large proportion (58 %) of TRP delivered via belowground pathways in winter (Mellander et al., 2016). This catchment represents an exception in that there is a single Sewage Treatment Works (STWs) managing waste for about 75 people, while the remaining population (approx. 200) uses septic tanks (Jordan et al., 2012).

2.4. Dunleer

The Dunleer catchment (Fig. 1, top left) is located in County Louth. The catchment has a total area of 948 ha with 49 % grassland and 33 % tillage. In this catchment soil drainage classes are a mixture of well, moderately, imperfectly and poorly drained soils (Thomas et al., 2016a), the latter being up to 70 % of the total area (Teagasc - Agriculture and Food Development Authority, 2018).

3. Methods

3.1. BBN parameterization

We developed a catchment-specific Bayesian Belief Network that simulates flow-weighted P concentrations and parameterized the model using high-frequency data from the Ballycanew catchment (Negri et al., 2024). The BBN was parameterized with high-frequency discharge and turbidity data (collected every 10 min and summarized at daily time-step), as well as 515 (Timoleague), 406 (Ballycanew), 408 (Castledockrell), and 392 (Dunleer) samples of soil Morgan P (McDonald et al., 2019; Thomas et al., 2016b), and calibrated against high- frequency TRP concentrations (Mellander et al., 2012). In this study, we test the BBN transferability by parametrizing the model for the first time for three further (8–12 km^2) Irish ACP catchments. The initial BBN parametrization for each catchment was identical to that presented in Negri et al. (2024) and referred to here as “Structure 1”. Structures are the graphical definitions of BBNs, also referred to as Directed Acyclic Graphs (DAGs) (Henderson and Pollino, 2010) encoding the causal (in) dependencies between variables (Aguilera et al., 2011). In this case, the structure represents the BBN’s P inputs, processes, pathways, and outputs, describing their relationships. When discussing uncertainty in environmental models, the word structure indicates the conceptual model (Refsgaard et al., 2006). Where the information was available the BBN variables were parameterized with catchment-specific datasets (these are specified in the Supplementary Information). However, catchment-specific parametrization was not possible for the following nodes (i.e., BBN variables): “Direct Discharge”, “Septic Tank Treatment”, “Sediment Water Soluble P", “Predicted Dissolved P concentration” (i.e., P that is dissolved from the soil matrix into the stream), and the total number of septic tanks in the catchment needed to calculate the total septic tank load. A detailed description of these nodes is given in Table 1. Additionally, data was not available for the “Septic Tank Treatment” node for Timoleague and Dunleer, and therefore an additional BBN structure was tested where the treatment was not implemented, and the distribution of P concentration across the catchment’s septic tanks was set up as equal to the “Unknown treatment” option (Structure 3). For the Timoleague and Castledockrell catchments, further model structure implementations (Structure 4, 5, 6) saw the addition of the node “Groundwater Dissolved P Concentration mg l^{-1}” to describe the groundwater total dissolved phosphorus (TDP) concentration contributing to the total in-stream TRP concentration at catchment outlet (the details of Structure 2 will be introduced later on in this section). This was done with the same bootstrapping methodology that was applied to observed in-stream TRP concentrations in Negri et al. (2024), here applied to monthly samples of groundwater total dissolved P (TDP, 2009–2016) monitored in multi-level wells described in Mellander et al. (2016). An assumption was made that the wells near the stream (<40 m from the stream) contribute the most to stream TRP (Mellander et al., 2016), hence we only used data from these wells for the parameterization. For all catchments, a model structure including the in-stream P uptake derived in the expert elicitation workshop was parameterized, and labelled Structure 2 for the Ballycanew and Dunleer catchments, Structure 5 for the Timoleague catchment, and Structure 5 and 6 for the Castledockrell catchment. For the Castledockrell catchment the Sewage Treatment Works loads were included in the finalized BBN labelled Structure 6. This was done by incorporating the P concentration (mg l^{-1}) after tertiary treatment (Truncated Normal ($\mu = 1.44$; $\sigma = 1.61$, (Glendell et al., 2022), in this case truncated at 0), and the design size (130 people) found through the Irish Environmental Protection Agency (EPA)

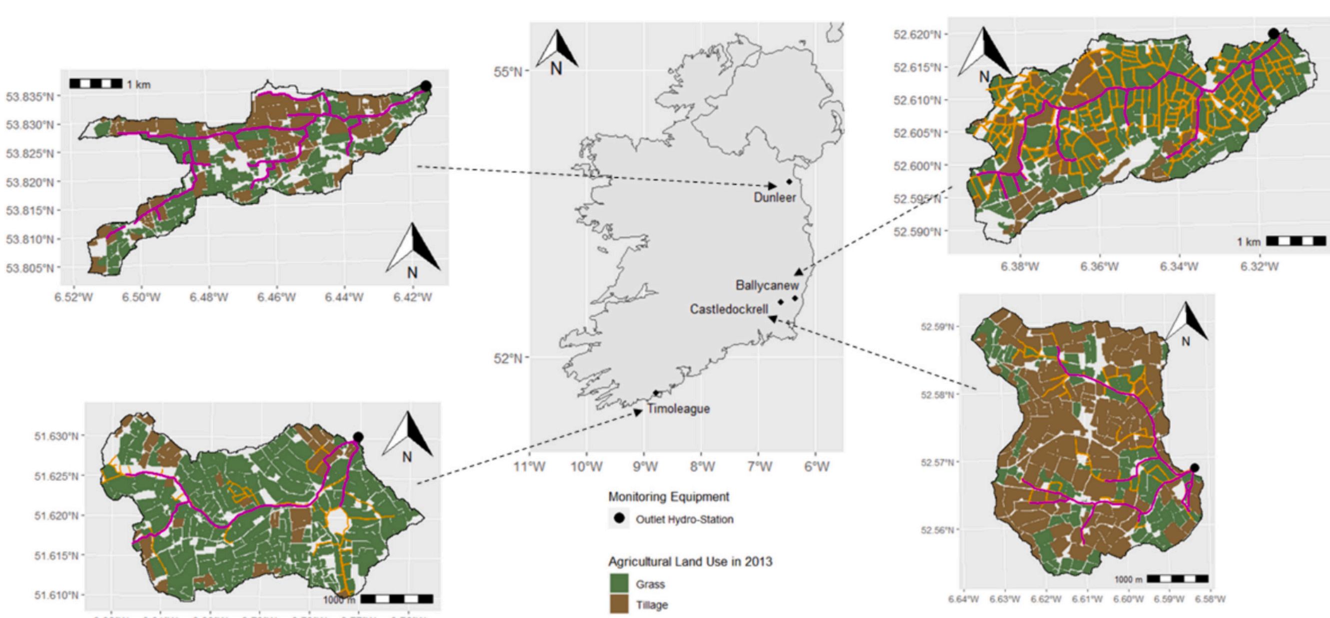


Fig. 1. Location of the four study areas within the republic of Ireland, and overview of each study area. Top left: Dunleer, top right Ballycanew, bottom left: Timoleague, bottom right: Castledockrell. Location of the high-resolution monitoring kiosk and hydrometric station at the catchment outlet is shown as a dot. Magenta lines represent streams and yellow lines represent artificial drainage (this is not available for the Dunleer catchment).

Table 1
Variables for which catchment-specific data was not unavailable in the Timoleague, Castledockrell, and Dunleer catchments. These parameters were chosen for a preliminary sensitivity analysis to understand their effects on the targeted P concentration and inform model transferability.

Node	Septic tank treatment	Direct discharge	Number of septic tanks	Sediment Water Soluble P (WSP) [mg kg^{-1}]	Predicted dissolved P concentration [mg l^{-1}]
Description	Probability of having "unknown", "primary" or "secondary" treatment of the effluent in a septic tank.	Probability of ST discharging directly into the stream.	Septic tanks within catchment boundary.	Describes the phosphorus released from the sediments into the stream. Defined as the best fitting distribution, fitted with the *SHELF* R package version 1.8.0 (Oakley, 2020a) to observed Water Extractable P in the catchment sediments (Shore et al., 2016) when data was available.	Describes the Water Extractable Phosphorus (WEP) dissolved from the soil matrix into the stream. Defined with the linear model: Predicted Dissolved P = β(WEP)-α, where β =0.08, α =0.158, derived from (Thomas et al., 2016b), whereby WEP stands for Water Extractable P. An assumption is made that when the linear model yields a negative value, that is resampled as a zero. This equation is not catchment specific.
Timoleague	Unavailable.	As Ballycanew.	As Ballycanew.	As Ballycanew.	Same everywhere.
Ballycanew	Probabilities based on a survey conducted within WaterProtect, a research project supported by the European Union research and innovation funding programme Horizon 2020 [grant no. 727450]. Probabilities are reported in Negri et al. (2024)	Assumed.	Available from data (88 tanks).	Defined as a Lognormal distribution (μ = −0.9, σ = 1), fitted with the *SHELF* R package (version 1.8.0, Oakley, 2020) to observed Water Extractable P in the catchment sediments (Shore et al., 2016).	Same everywhere.
Castledockrell	Probabilities based on a survey conducted within WaterProtect, a research project supported by the European Union research and innovation funding programme Horizon 2020 [grant no. 727450].	As Ballycanew.	As Ballycanew.	Defined as a Gamma distribution (k = 1.03, ɵ = 0.44).fitted with the *SHELF* R package (version 1.8.0, Oakley, 2020) to observed Water Extractable P in the catchment sediments (Shore et al., 2016).	Same everywhere.
Dunleer	Unavailable.	As Ballycanew.	As Ballycanew.	As Ballycanew.	Same everywhere.

data. This is consistent with the fact that STWs with tertiary treatment are required to keep the orthophosphate concentration below 2 mg l^{-1} for their effluent discharge (Fitzsimons et al., 2016). The model structure and the datasets used for the finalized parametrization are specified for each catchment in the Supplementary Information, and a summary of each Structure's specifications is given in Table 3.

3.2. In-stream P uptake

P uptake is complex with multiple components based on physical, chemical, and biological processes (Withers and Jarvie, 2008). Whilst uptake rates might be available for specific components (e.g., algal uptake, large plant uptake, sediment uptake), this study focusses on providing an overall effect from the combined processes in each catchment for each season via expert elicitation because the necessary data to quantify uptake rates was not available for the catchments. Here P uptake was framed as the percentage (%) of in-stream P that is removed by both biotic and abiotic uptake. A simplified version of the methodology presented in Mzyece et al. (2024) for expert elicitation was used to determine P uptake for the four catchments for four seasons. We selected 6 key papers describing UK-led experiments on this topic (Bowes et al., 2016; Jarvie et al., 2002; Stutter et al., 2021, 2010; Wade et al., 2001; Withers and Jarvie, 2008), and invited four authors of these papers to contribute to our elicitation exercise as experts who have published on the topic of P uptake in rivers. Three accepted, one declined. The elicitation process then comprised of three steps: 1) The Sheffield Elicitation Framework (SHELF) e-learning course for experts (Gosling, 2018), which the experts took on their own, 2) a preliminary exercise where the experts were asked to complete an elicitation table per catchment, providing their personal judgement on the two tertiles, T1 and T2, (33th and 66th quantiles) and the median M (50th quantile) percentage in-stream P uptake for each season. Initially, the upper limit of the distribution was fixed at 100 % removal and the lower limit at 0 % removal. To aid the experts with their judgements, supporting documentation containing both a summary of the literature on in-stream P removal and information on the four catchments, was provided to the experts ahead of time (published in an evidence dossier in Negri and Mellander, 2024). For the scope of this elicitation, we aimed to quantify global uptake (see, for example, the quantities in bold under the column 'P retained' in Table 1 Negri and Mellander (2024)) and asked the experts to provide judgement on what was the likely P uptake based on their experience of other river systems. 3) Preliminary prior Normal distributions were fitted with the *SHELF* R package version 1.8.0 (Oakley, 2020) to the elicited distributions at Step 2 and presented to the experts during the workshop. In the workshop, the experts were asked to discuss the preliminary distributions and to agree on a single consensus distribution per season per catchment. Based on what emerged during the discussion, and to facilitate consensus, distributions were re-fitted and plotted in real time for the experts to examine. The final consensus distributions were then used to parameterize the "In-stream P uptake" node in the BBNs, and the updated BBN was subsequently tested against in-stream TRP observations.

3.3. Sensitivity analysis

Sensitivity analysis was done to understand the effect of using non-catchment specific data on model transferability for the variables listed in Table 1. For direct discharge presence (0–100 %) or absence (0–100 %), the relative fraction of direct discharge presence/absence was varied in 5 % steps, with the probabilities for the two categories summing to 100 %. To assess the impact of number of septic tanks within each catchment boundary on in-stream P concentration, increases of two septic tanks per step were applied, ranging from 30 to 150 septic tanks. This range assumed 2.4 people per household (and therefore per tank) for these scarcely populated catchments. To understand the effects of varying the Water Soluble P (WSP, described in Table 1) we applied a stepwise variation (0.1 increments) on the parameters of the Lognormal distribution used in the Ballycanew catchment: the mean ($-2 \leq \mu \leq 2$) and, separately, the same variation on the standard deviation ($0 \leq \sigma \leq 2$). The Gamma distribution has two parameters (shape, k, and scale, ɵ) that together control the shape of the distribution. These parameters do not correspond directly to physical values (unlike, for example, the mean value of a Normal or Poisson distribution) and are always >0. Here we stepped through these parameters in increments of 0.1 over the range $0.1 \leq k \leq 2$ and 0.005 increments over the range $0.05 \leq \theta \leq 1$ for the WSP node parameterized for Castledockrell. For the "Predicted Dissolved P Concentration, 0.02 stepwise increases were applied to the β parameter ($-1 \leq \beta \leq 1$), and 0.005 stepwise increases were applied to the α ($0 \leq \alpha \leq 0.2$). The sensitivity analysis was conducted independently for each parameter in each catchment, (no nodes were varied simultaneously) and carried out only for the finalized and best performing model structure in each catchment (specifically, Structure 5 for Timoleague, Structure 6 for Castledockrell, and Structure 2 for Dunleer). The analysis was performed using *rSMILE* version 2.0.10 (BayesFusion, 2019a), an API engine available in R which can perform the same Bayesian inference operations performed by GeNIe Modeler (BayesFusion, 2019b), the software used to design the BBNs. In each catchment, the parameter variations were applied to predict the TRP concentration (in the model, the target node is called "In-stream P concentration [mg l^{-1}]" and it describes the variable of interest). The effects of changing the input parameters on the full distributions was assessed visually by comparison against the distribution from "simulation 0″ (the initial BBN parameterization). The effects of the presence of the nodes "Septic Tank Treatment", "Groundwater Dissolved P Concentration mg l^{-1}", the nodes relative to in-stream P uptake, and those pertaining to the STWs in Castledockrell were tested by comparing distributions derived from different model structures to those obtained from the original BBN.

3.4. Model evaluation

The model structures were evaluated by comparing the predicted TRP concentrations with the available observed TRP concentrations (available as daily mean, mg l^{-1}) (2009-10-01 to 2016-12-31) by: calculating percentage bias (PBIAS) in the R package *hydroGOF* version 0.4–0 (Zambrano-Bigiarini, 2020), plotting and visually comparing the full posterior distributions, and comparing median concentrations. PBIAS calculation and visual assessment are recommended when modelling P in catchments with a prevalence of diffuse sources, as in these instances models struggle to produce good Nash-Sutcliffe statistics (Jackson-Blake et al., 2015). In addition, for the model version containing the "In-stream P removal", the Normal distribution allows for negative concentrations due to the potential for release of in-stream P (considered to be plausible by the expert elicitation). For the purposes of model evaluation these were resampled into zeros prior to analysis.

4. Results and discussion

4.1. BBN parameterization

The results of the preliminary BBN parametrization are shown in Fig. 3, where the density plots from all model structures are shown against the distribution fitted to the observations. When comparing the TRP distributions and boxplots of Structure 1 (in green) against the observed (light brown), the figure shows that the model performs well for Ballycanew and Dunleer, and less well for the Timoleague and Castledockrell catchments. For Timoleague and Castledockrell, the initial parameterization (Structure 1) overpredicts the stream P concentration by 65–82 % (data not shown), which is a consequence of the parameterization being tailored for a surface-driven catchment instead of a groundwater-driven one. Specific details of the each of the models' performances are discussed in Section 4.4. The state-of-the-art high-resolution and long-term monitoring data available in these catchments could also facilitate other model structures besides the ones considered in this study. For example, soil chemistry data could be leveraged to improve process representation for the groundwater-driven catchments, because the presence of aluminium-rich or iron-rich soils is known to impact P solubilization and transfer to the groundwater table (Mellander et al., 2016).

4.2. In-stream P uptake

During elicitation, consensus was reached by initially focusing on the first catchment (Timoleague), first comparing summer and winter, then spring and autumn. Consensus about the other three catchments was then reached by comparison with the first. For wintertime, the experts agreed to use −100 % as the lower limit which represents complete sediment P release into the water column and biotic uptake close to zero. For wintertime in Timoleague, averaged values (tertiles and median) and fitted a Normal distribution were used (Fig. 2). For summertime in Timoleague, expert consensus had the probability density centred on a 43 % removal rate, and this was the same for autumn and spring. For Ballycanew, the experts decided to reduce P removal by 30 %, due to the high flashiness of the catchment. The experts considered the catchment P saturation and loading to be the most influential factors in determining in-stream uptake, catchment flashiness was considered less important. As a result, the mean P uptake was similar across catchments (Table 2). The consensus also reflected that, even though Castledockrell is less flashy than Ballycanew, the two catchments have similar P uptake because high loading in Castledockrell due to a Sewage Treatment Works and septic tanks. An exception is made for Castledockrell in the wintertime, with tertiles and median values similar to Timoleague, and therefore the same parameterization (Table 2). For Dunleer, the wintertime uptake was considered very low, then the rest of the seasons were considered comparable to Timoleague. Overall, the experts had greater confidence estimating P removal in the colder seasons (winter and autumn), than in the warmer ones, where the distributions are wider and more uncertain (Fig. 2). Furthermore, the experts suggested that visual aids such as photos of the river corridor could assist in estimating uptake, allowing the approximate width and depth of ditches and rivers to be estimated, as well as the presence of submerged and emergent vegetation and algae to be assessed. This is especially important because increased riparian vegetation and algae can lead to decreased dissolved P concentrations (Bowes et al., 2016; Chase et al., 2016). The distributions obtained were used in each catchment model to calculate the in-stream P load reduction (Eq. (1)):

$$r = (1 - Normal\,(\mu_s; \sigma_s)\,)*L \tag{1}$$

where r is the in-stream reduced load, L the total catchment load, and N $(\mu_s; \sigma_s)$ is a Normal distribution with a seasonal dependent mean and standard deviation (specified in Table 2), and the loads are expressed in T $month^{-1}$. In the BBN, the seasonal monthly distributions are child nodes of a deterministic node termed "Season", which indicates the meteorological seasons.

4.3. Sensitivity analysis

The sensitivity analysis showed that the three tested BBNs are not sensitive to changes in the variables representing septic tank "Direct Discharge" (% of tanks that discharge the effluent directly into the stream), and "Sediment Water Soluble P" (that is, P released into the stream by sediments). One BBN showed sensitivity to changes in the β parameter used for the node "Predicted Dissolved P concentration". Details of the sensitivity to the Predicted Dissolved P concentration" node are shown for one catchment (Dunleer) in the Appendix. This shows the $\log_{10}$(TRP) concentration boxplot for each parameter value against the "simulation 0″ (in light green) overlayed with a sample of the full distribution plotted as dots. The equation in the node "Predicted Dissolved P concentration" was derived from Thomas et al. (2016b), and is an aggregated result of catchment-specific regression models, which were not available at the time of model parametrization. It would be instructive to reparametrize the BBN if/when these individual models become available, and to compare the results of a corresponding sensitivity analysis on this new model structure with these results.

Sensitivity analysis is a pivotal component of model calibration and

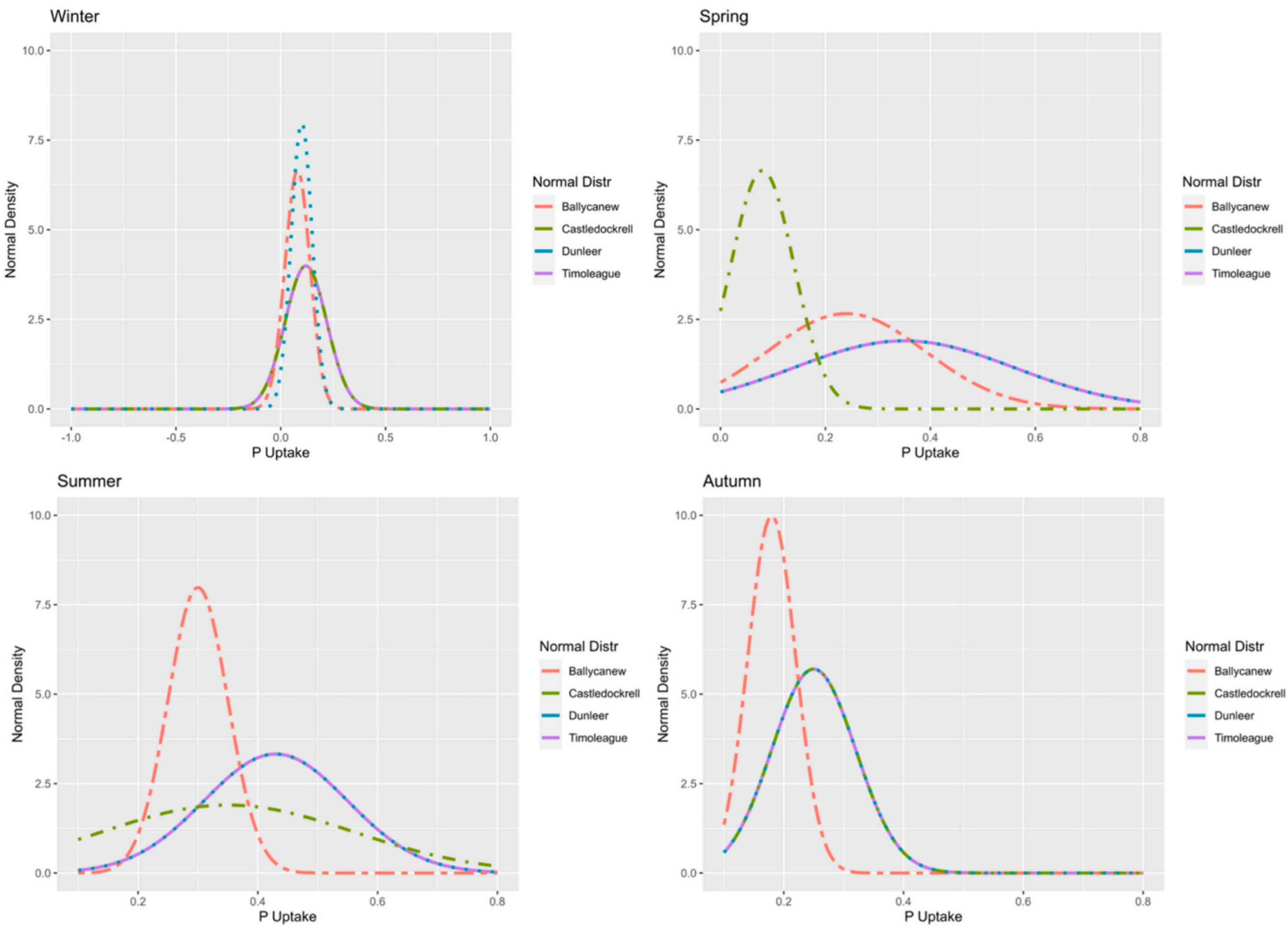


Fig. 2. Consensus Normal distributions grouped by season. The y axis shows the probability density function, the x axis is the agreed upon plausible range for in-stream P uptake (%). Different colours show the distributions for each catchment. For the winter season, Castledockrell and Timoleague are overlapping; for spring and summer, Timoleague and Dunleer are overlapping; and for the autumn, Timoleague, Castledockrell, and Dunleer are overlapping.

Table 2
Characteristics of seasonal P uptake as discussed by the experts during the workshop, including re-defined lower and upper limits of uptake, and the elicited parameters for the Normal distributions. A mean (μ) of 0.10 corresponds to 10 % mean uptake.

	% P uptake		Justification	Normal distributions parameters fitted from consensus							
				Timoleague		Ballycanew		Castledockrell		Dunleer	
	Lower limit consensus	Upper limit consensus		μ	σ	μ	σ	μ	σ	μ	σ
Winter	−100	+100	To describe the fact that there can be release of P (−100 %) rather than uptake (+100 %)	0.12	0.10	0.08	0.06	0.12	0.10	0.10	0.05
Spring	0	80	Uptake can never be 100 %, but the experts agree on absent or negligible P release	0.35	0.21	0.24	0.15	0.08	0.06	0.35	0.21
Summer	10	80	Biological uptake always present, so lower limit cannot be 0 %	0.43	0.12	0.30	0.05	0.35	0.21	0.43	0.12
Autumn	0	65	Uptake can never be 100 % and is lower than in spring, but the experts agree on absent or negligible P release	0.25	0.07	0.18	0.04	0.25	0.07	0.25	0.07

design, however, methodologies for conducting it for hybrid BBNs aren't readily accessible in the software used for BBN parameterization or in R packages, and therefore require bespoke coding for implementation. For example, Glendell et al. (2022), conducted a sensitivity analysis on a discretized version of their hybrid network, which causes loss of information (Uusitalo, 2007), and makes the BBN sensitive to the chosen discretization. Similarly, Piffady et al. (2021) tested the sensitivity of a discretized BBN by varying nodes deemed important across a reasonable range. Here we provided a preliminary approach to the sensitivity analysis of a hybrid BBN without triggering discretization.

4.4. Model evaluation

Results of the model evaluation are shown in Fig. 3, which shows boxplots with the median, interquartile range with the whiskers extending to the highest and lowest datapoints, and a representative selection of datapoints, from ten-thousand simulated realizations of each BBN structure tested. These are summarized in Table 3, where

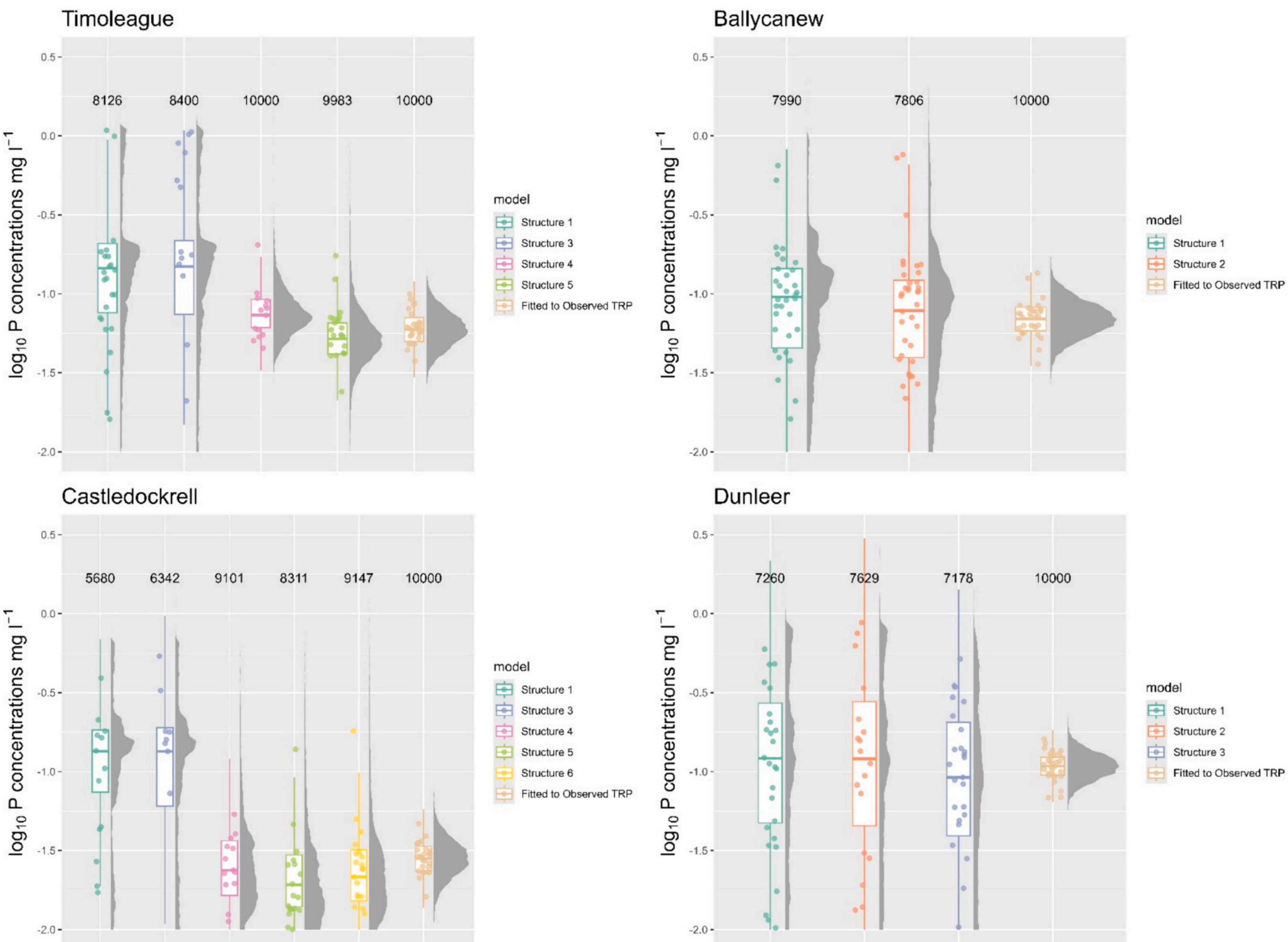


Fig. 3. Predicted and observed $\log_{10}$(TRP) concentrations for each of the four catchments. The grey density shows the distribution obtained by simulated realizations from the BBN (all plots except the rightmost of each panel), filled points the scatter of the realizations (100 samples per catchment), coloured boxplots show the median (central line), interquartile range (box) and highest and lowest datapoints (shown by the whiskers). Observations are shown in the rightmost plot in each panel, where the grey density shows the distribution fitted to the full suite of observations, filled points the scatter of the realizations, the light brown boxplots show the median (central line), interquartile range (box) and the 95 % quantile range for the distribution. Data outside the instrument's limit of detection (0.01–5.00 mg l^{-1}) were excluded from the plot, and the text shows the number of valid samples for each model (with 10,000 being the maximum number of available samples generated by the model). This plot was produced with the ggdist R package version 3.3.0 (Kay, 2023). A complete description of the finalized model structures is given in the Supplementary Information for the Timoleague, Dunleer, and Castledockrell catchments, a description of Structure 1 is given in Negri et al. (2024).

predicted $\log_{10}$ TRP concentrations are compared to the observations (daily time-step, data from 2009 to 2016). In the surface-drained catchments (Fig. 3, Ballycanew and Dunleer, right-hand side), the distribution of $\log_{10}$(TRP) concentrations predicted by the BBN models is not sensitive to the structure of the BBN. The BBN parameterized in Negri et al. (2024) (Structure 1) can reproduce the mean and median observed P concentrations in the Ballycanew and Dunleer catchments. For Ballycanew, the percentage bias is within acceptable ranges (close to the 50 % departure from observations or less, shown in Table 3). For Dunleer, a bias of 94 % is still considered acceptable because the mean predicted concentration was 0.11, whilst the observed was 0.10 mg l^{-1}. The addition of an in-stream P removal node improved the ability of the model to replicate the mean and median in-stream P concentration in these two catchments (Table 3, comparing Structure 1 and 2), by introducing a linear scaling factor. Further, the percentage bias in Dunleer went from 94 % to 45 % with the addition of removal, however, because the concentrations being predicted are small, small changes in their absolute values represent large changes in bias, therefore bias values should be looked at critically in context with mean TRP concentrations, as shown in Table 3. For the two groundwater-dominated catchments (Timoleague and Castledockrell), the introduction of groundwater TDP concentration (Structures 4, 5, and 6) improved the simulated TRP concentrations: in the final structure, the predicted median was the same as the observed, 0.05 (Timoleague) and 0.02 mg l^{-1} (Castledockrell). This could not be achieved in the Castledockrell catchment with a process-based model such as SimplyP (Hawtree et al., 2023), even though the BBN and SimplyP deploy similar strategies to represent below-ground processes. An improvement in percentage bias (from 40 % to −5 %) is provided by the addition of in-stream P removal in the Timoleague catchment (also in Table 3, comparing Structure 4 and 5), however, the bias was already within the 50 % departure from observations, which indicates that this remains a secondary process, at least if compared to correctly representing groundwater concentrations (Structure 4).

Knowledge of the type of septic tank treatment adopted (i.e., comparing Structure 1 to Structure 3), provides little to no advantage (concentrations remain unvaried), except for better representing the available datasets. Increasing the structural complexity of the BBN had the most impact in the Castledockrell catchment, where the percentage bias of posterior simulations has decreased more than twenty-fold

Table 3
Overall results of the different BBN versions for the four catchments, concentrations (mg l^{-1}) outside the instrument limit of detection (0.01–5.00 mg l^{-1}) have been excluded from the analysis. Both observed and predicted TRP concentrations were log-transformed before calculating the statistics, and then converted back to normal values. A positive bias indicates overestimation.

			1	2	3	4	5	6	Observations
			Negri et al. (2024)	Negri et al. (2024) + in-stream removal (no ST treatment in Dunleer)	Negri et al. (2024), no ST treatment	No ST treatment + GW TDP	No ST treatment + GW TDP + in-stream P removal	No ST treatment + GW TDP + in-stream P removal + STWs 130 p.e.	
Timoleague	mean	mg l^{-1}	0.14	–	0.14	0.08	0.05	–	0.05
	lower limit (μ-1σ)		0.05	–	0.05	0.05	0.03	–	0.03
	upper limit (μ + 1σ)		0.40	–	0.41	0.11	0.08	–	0.09
	median		0.14	–	0.15	0.07	0.05	–	0.05
	5th quantile		0.02	–	0.02	0.05	0.03	–	0.02
	25th quantile		0.08	–	0.07	0.06	0.04	–	0.04
	75th quantile		0.21	–	0.21	0.09	0.08	–	0.08
	PBIAS	%	285	–	291	40	−5	–	–
Ballycanew	mean	mg l^{-1}	0.08	0.07	–	–	–	–	0.06
	lower limit (μ-1σ)		0.03	0.03	–	–	–	–	0.03
	upper limit (μ + 1σ)		0.21	0.17	–	–	–	–	0.11
	median		0.10	0.08	–	–	–	–	0.06
	5th quantile		0.02	0.02	–	–	–	–	0.01
	25th quantile		0.05	0.04	–	–	–	–	0.04
	75th quantile		0.14	0.12	–	–	–	–	0.14
	PBIAS	%	80	49	–	–	–	–	–
Castledockrell	mean	mg l^{-1}	0.11	–	0.10	0.03	0.02	0.02	0.02
	lower limit (μ-1σ)		0.04	–	0.04	0.01	0.01	0.01	0.01
	upper limit (μ + 1σ)		0.29	–	0.29	0.05	0.04	0.05	0.04
	median		0.13	–	0.13	0.02	0.02	0.02	0.02
	5th quantile		0.02	–	0.02	0.01	0.01	0.01	0.01
	25th quantile		0.07	–	0.06	0.02	0.01	0.02	0.02
	75th quantile		0.18	–	0.19	0.04	0.03	0.03	0.04
	PBIAS	%	445	–	453	34	12	18	–
Dunleer	mean	mg l^{-1}	0.11	0.09	0.11	–	–	–	0.10
	lower limit (μ-1σ)		0.03	0.03	0.03	–	–	–	0.06
	upper limit (μ + 1σ)		0.38	0.28	0.39	–	–	–	0.16
	median		0.12	0.09	0.12	–	–	–	0.09
	5th quantile		0.01	0.01	0.01	–	–	–	0.05
	25th quantile		0.05	0.04	0.05	–	–	–	0.06
	75th quantile		0.27	0.20	0.28	–	–	–	0.14
	PBIAS	%	94	45	97	–	–	–	–

Abbreviations: ST septic tanks; GW TDP groundwater total dissolved phosphorus; STWs sewage treatment works; p.e. people equivalent.

(Table 3, comparing Structure 1 with Structure 6). To further demonstrate this, monthly predicted $\log_{10}$(TRP) concentrations (yellow bars) are plotted as histograms against daily observed $\log_{10}$(TRP) concentrations (blue bars, grouped by month) across all model structures developed for the Castledockrell catchment in Fig. 4. This shows the progress made in adapting Structure 1 in this catchment (top histograms), where yellow and blue are not overlapping, up to the last model structure (Structure 6, bottom panel), which shows good correspondence between predicted and observed TRP concentrations. The addition of P removal had the added benefit of improving seasonality in the BBN predictions, which was not a behaviour that emerged in the first parameterization; however, the observations still show stronger seasonality patterns than the simulations. A summary table of these results is reported in the Supplementary Information, where, for each catchment, monthly predictions from the first versus the final model version are compared against the observations. Percentage bias shows that the final and best performing model in each catchment performs best in dry conditions (summer months). However, in Dunleer and Ballycanew, the model predicts the mean concentration better in winter than in summer. This is notable, as predicting P concentrations correctly in summer may be more relevant from the point of view of assessing ecological impacts in running waters than predicting them during the ecologically less active winter period. In the groundwater-dominated catchments, the final model is better constrained than in the runoff-dominated catchments (Ballycanew and Dunleer), as evident when comparing the predicted upper (μ + σ) concentrations versus the observed in Table 5 of Supplementary Information. Table 4 shows both the observed and the marginal probabilities of Environmental Quality Standard of 0.035 mg

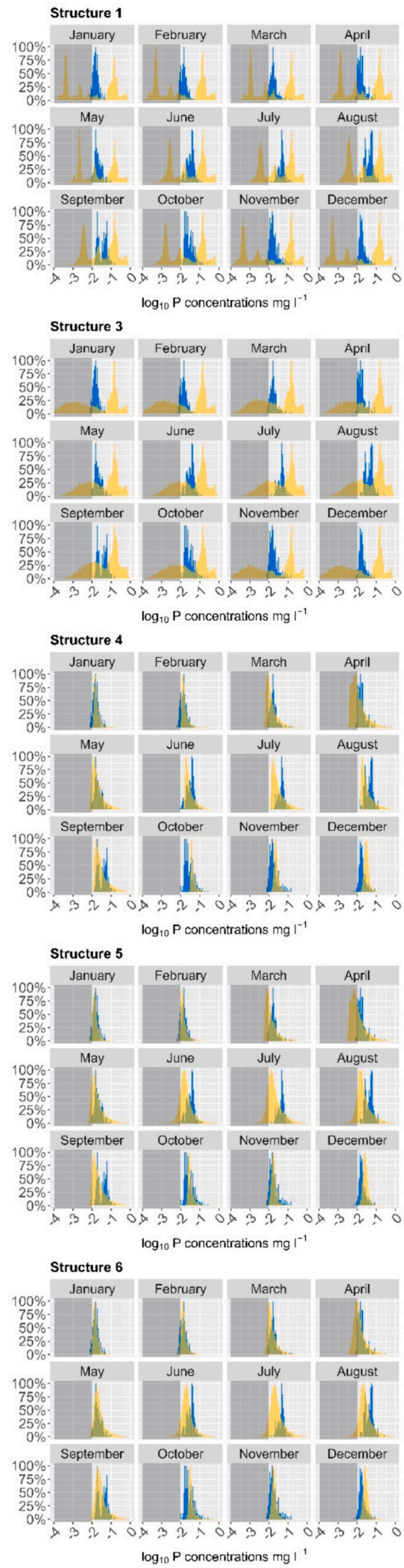


Fig. 4. Histograms of monthly $\log_{10}$(TRP) concentrations (mg l^{-1}). Observations are shown in blue, predictions obtained from each model structure adapted for the Castledockrell catchment are shown in yellow. The dark grey box indicates concentration values below the limit of detection (0.01 mg l^{-1}).

Table 4
Marginal probability of exceeding EQS limits in the four catchments.

	Probability to exceed EQS limits			
	2010–2020 data (Mellander et al., 2022)	2009–2016 data	Model in Negri et al. (2024)	Model with in-stream P removal
	Hourly mean concentration	Daily mean concentration	Structure 1	Final structure (a different one for each catchment)
Timoleague	81 %	80 %	72 %	88 %
Ballycanew	94 %	88 %	65 %	61 %
Castledockrell	29 %	28 %	46 %	18 %
Dunleer	99 %	99 %	58 %	55 %

l^{-1} (EQS) exceedance in each catchment and across two model structures. The Table shows that even though the models can work for two catchments and is improved by the inclusion of P removal, the model predicts a lower probability of exceeding the EQS than the observational data in the two P risky catchments (Ballycanew and Dunleer). Meanwhile, the prediction of EQS exceedance for the Timoleague catchment is either under- or over- predicting by 8 %, depending on BBN model structure, while at Castledockrell, the prediction of exceedance for the final model is 10 % lower than the observed. These findings suggest that the BBN may be best used as Decision Support Tool by calculating the quantiles of monthly predicted concentrations as seen in Negri et al. (2024) or the monthly mean and upper and lower limits ($\mu \pm \sigma$, as shown in the Supplementary) rather than as a discrete probability of EQS exceedance, due to the predicted distributions being wider and more skewed than the observations, also seen in Negri et al. (2024).

5. Conclusions

This study is the first application of a BBN aimed at predicting stream P concentrations in four Irish agricultural catchments. We set out to test the transferability of a hybrid BBN targeting P pollution across agricultural catchments with diverse dominant hydrological processes. The initial BBN proved to be transferrable between catchments dominated by surface or mixed hydrological pathways, irrespective of land use, but less so between catchments dominated by sub-surface delivery. Inclusion of groundwater total dissolved P (TDP), Sewage Treatment Works (STWs) inputs, and in-stream P uptake improved model performance in all four catchments and made the BBN more transferable, though at the cost of increased complexity and data requirements.

In this work, we explored two strategies to improve model structure: bootstrapping to estimate the groundwater TDP concentration, and expert elicitation to assess in-stream P removal. The addition of groundwater TDP loads improved the predictions in sub-surface-driven catchments. Expert elicitation aided the P uptake parameterization, which lacked generalizable data, highlighting a research gap. However, we found that in-stream P uptake remained a secondary process compared to the representation of P transfers via both surface and subsurface pathways when simulating daily P concentrations.

To avoid discretizing the continuous distributions that form critical components of the BBN nodes prior to sensitivity analysis, we implemented a method to evaluate the effects of parameter variation on the full posterior distribution of the target node, by varying the parameters of interest while holding the others fixed. This demonstrated the transferability of non-catchment specific data to further catchments and found redundant parameters in the sediments and septic tanks components of the model.

Testing BBN applicability also revealed constraints in this study related to the limited presence of BBN studies conducted in catchments comparable to those examined in this research, and the fact that few modelling studies have been performed in our study catchments. Therefore, future work should involve the use of other modelling approaches in these catchments, allowing the intercomparison of models parameterized with high-frequency datasets. Given the scope of the Agricultural Catchments Programme, in the future, the BBNs developed here present an effective tool for modelling of catchment-scale effects of water quality mitigation measures.

CRediT authorship contribution statement

Camilla Negri: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Nicholas Schurch:** Writing – review & editing, Supervision, Methodology. **Andrew J. Wade:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Per-Erik Mellander:** Writing – review & editing, Supervision, Funding acquisition, Data curation, Conceptualization. **Marc Stutter:** Methodology. **Micheal J. Bowes:** Methodology. **Chisha Chongo Mzyece:** Methodology. **Miriam Glendell:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

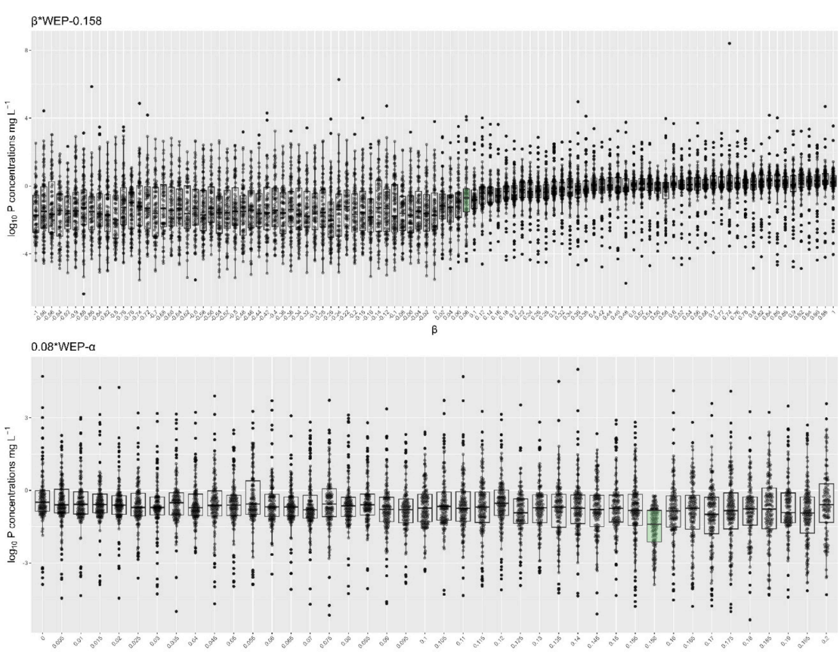
Supplementary information including the R code for the sensitivity analysis conducted in this study and the supplementary results is publicly available at the shared link under the MIT license.

A repository to test the transferability of a Bayesian Belief Network (Original data) (Github)

Acknowledgements

We acknowledge the Teagasc Walsh Fellowship Programme for providing the funding (Reference Number 2019021). We thank the experts for their time and expertise during the workshop and providing valuable insights not only regarding in-stream phosphorus uptake, but also into the topic at large and future research paths. We wish to thank the team at BayesFusion (https://www.bayesfusion.com/) for providing us with the necessary academic licensing and software support. We wish to thank the reviewers and Dr. Daniel Hawtree for their suggestions that greatly improved this manuscript.

Appendix A



Appendix. Results of the sensitivity analysis on the two parameters for the "Predicted Dissolved P concentration" node, β (slope, top plot) and α (intercept, bottom plot) displayed as boxplots showing the median (central line), interquartile range (box) for the $\log_{10}$(TRP) concentration (mg l^{-1}) distribution of each simulation, filled black points show the scatter of the realizations. Values assumed for each parameter in each simulation are shown on the x axis, the boxplots of the "simulation 0" are shown in light green. Results are shown for the model Structure 2 for the Dunleer catchment.

Appendix B. Supplementary data

The R code for the sensitivity analysis conducted in this study as well as the supplementary results are publicly available at the link https://github.com/CamillaNegri/Transferability_Ptool under the MIT license. The figures in this study have been made by adapting the code published by Negri et al. (2024) and available at https://github.com/CamillaNegri/Ballycanew_Ptool under the MIT license (https://github.com/git/git-scm.com/blob/main/MIT-LICENSE.txt). Supplementary data to this article can also be found online at https://doi.org/10.1016/j.scitotenv.2024.174926.

References

Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. Environ. Model Softw. 26, 1376–1388. https://doi.org/10.1016/j.envsoft.2011.06.004.

BayesFusion, 2019a. SMILE Engine [WWW Document]. URL. https://www.bayesfusion.com/smile/ (accessed 7.26.23).

BayesFusion, 2019b. GeNIe 2.4 [WWW Document]. URL. https://www.bayesfusion.com/ (accessed 5.6.20).

Bowes, M.J., Loewenthal, M., Read, D.S., Hutchins, M.G., Prudhomme, C., Armstrong, L. K., Harman, S.A., Wickham, H.D., Gozzard, E., Carvalho, L., 2016. Identifying multiple stressor controls on phytoplankton dynamics in the River Thames (UK) using high-frequency water quality data. Sci. Total Environ. 569–570, 1489–1499. https://doi.org/10.1016/j.scitotenv.2016.06.239.

Chase, J.W., Benoy, G.A., Hann, S.W.R., Culp, J.M., 2016. Small differences in riparian vegetation significantly reduce land use impacts on stream flow and water quality in small agricultural watersheds. J. Soil Water Conserv. 71, 194–205. https://doi.org/10.2489/jswc.71.3.194.

Drohan, P.J., Bechmann, M., Buda, A., Djodjic, F., Doody, D., Duncan, J.M., Iho, A., Jordan, P., Kleinman, P.J., McDowell, R., Mellander, P.-E., Thomas, I.A., Withers, P. J.A., 2019. A global perspective on phosphorus management decision support in agriculture: lessons learned and future directions. J. Environ. Qual. 48, 1218–1233. https://doi.org/10.2134/jeq2019.03.0107.

Fitzsimons, L., Clifford, E., McNamara, G., Doherty, E., Phelan, T., Horrigan, M., Delauré, Y., Corcoran, B., 2016. Increasing Resource Efficiency in Wastewater Treatment Plants, No. 168. Environmental Protection Agency, Ireland.

Glendell, M., Gagkas, Z., Stutter, M., Richards, S., Lilly, A., Vinten, A., Coull, M., 2022. A systems approach to modelling phosphorus pollution risk in Scottish rivers using a spatial Bayesian Belief Network helps targeting effective mitigation measures. Front. Environ, Sci, p. 10.

Gosling, J.P., 2018. SHELF: The Sheffield elicitation framework. In: Dias, L.C., Morton, A., Quigley, J. (Eds.), Elicitation: The Science and Art of Structuring Judgement. Springer International Publishing, Cham, pp. 61–93. https://doi.org/10.1007/978-3-319-65052-4_4. International Series in Operations Research & Management Science.

Hatum, P.S., McMahon, K., Mengersen, K., Wu, P.P.-Y., 2022. Guidelines for model adaptation: a study of the transferability of a general seagrass ecosystem dynamic Bayesian networks model. Ecol. Evol. 12, e9172 https://doi.org/10.1002/ece3.9172.

Hawtree, D., Galloway, J., Zurovec, O., Jackson-Blake, L., Norling, M., Mellander, P.-E., 2023. Performance of a Parsimonious Phosphorus Model (SimplyP) in Two Contrasting Agricultural Catchments in Ireland (No. EGU23-7759). Copernicus Meetings. https://doi.org/10.5194/egusphere-egu23-7759.

Henderson, C.M., Pollino, C.A., 2010. Bayesian Networks : A Guide for their Application in Natural Resource Management and Policy. (No. Landscape Logic, Technical Report 14). Australian Government – Department of the Environment, Water, Heritage and the Arts.

Igras, J.D., Creed, I.F., 2020. Uncertainty analysis of the performance of a management system for achieving phosphorus load reduction to surface waters. J. Environ. Manag. 276, 111217 https://doi.org/10.1016/j.jenvman.2020.111217.

Jackson-Blake, L., Wade, A., Futter, M., Butterfield, D., Couture, R.-M., Cox, B., Crossman, J., Ekholm, P., Halliday, S., Jin, L., Lawrence, D.S.L., Lepistö, A., Lin, Y., Rankinen, K., Whitehead, P., 2016. The INtegrated CAtchment model of phosphorus dynamics (INCA-P): description and demonstration of new model structure and equations. Environ. Model Softw. 83, 356–386. https://doi.org/10.1016/j.envsoft.2016.05.022.

Jackson-Blake, L.A., Dunn, S.M., Helliwell, R.C., Skeffington, R.A., Stutter, M.I., Wade, A. J., 2015. How well can we model stream phosphorus concentrations in agricultural catchments? Environ. Model Softw. 64, 31–46. https://doi.org/10.1016/j.envsoft.2014.11.002.

Jarvie, H.P., Neal, C., Williams, R.J., Neal, M., Wickham, H.D., Hill, L.K., Wade, A.J., Warwick, A., White, J., 2002. Phosphorus sources, speciation and dynamics in the lowland eutrophic River Kennet, UK. Sci. Total Environ., Water quality functioning of lowland permeable catchments:inferences from an intensive study of the River Kennet and upper River Thames, 282–283, pp. 175–203. https://doi.org/10.1016/S0048-9697(01)00951-2.

Jordan, P., Melland, A.R., Mellander, P.-E., Shortle, G., Wall, D., 2012. The seasonality of phosphorus transfers from land to water: Implications for trophic impacts and policy evaluation. Sci. Total Environ., Climate Change and Macronutrient Cycling along the Atmospheric, Terrestrial, Freshwater and Estuarine Continuum - A Special Issue dedicated to Professor Colin Neal, 434, pp. 101–109. https://doi.org/10.1016/j.scitotenv.2011.12.070.

Kay, M., 2023. {ggdist}: Visualizations of Distributions and Uncertainty. https://cran.r-project.org/web/packages/ggdist/index.html.

Krueger, T., Page, T., Hubacek, K., Smith, L., Hiscock, K., 2012. The role of expert opinion in environmental modelling. Environ. Model Softw. 36, 4–18. https://doi.org/10.1016/j.envsoft.2012.01.011.

McDonald, N.T., Wall, D.P., Mellander, P.E., Buckley, C., Shore, M., Shortle, G., Leach, S., Burgess, E., O'Connell, T., Jordan, P., 2019. Field scale phosphorus balances and legacy soil pressures in mixed-land use catchments. Agric. Ecosyst. Environ. 274, 14–23. https://doi.org/10.1016/j.agee.2018.12.014.

Mellander, P.-E., Melland, A.R., Jordan, P., Wall, D.P., Murphy, P.N.C., Shortle, G., 2012. Quantifying nutrient transfer pathways in agricultural catchments using high temporal resolution data. Environ. Sci. Policy, CATCHMENT SCIENCE AND POLICY EVALUATION FOR AGRICULTURE AND WATER QUALITY 24, 44–57. https://doi.org/10.1016/j.envsci.2012.06.004.

Mellander, P.-E., Jordan, P., Shore, M., Melland, A.R., Shortle, G., 2015. Flow paths and phosphorus transfer pathways in two agricultural streams with contrasting flow controls. Hydrol. Process. 29, 3504–3518. https://doi.org/10.1002/hyp.10415.

Mellander, P.-E., Jordan, P., Shore, M., McDonald, N.T., Wall, D.P., Shortle, G., Daly, K., 2016. Identifying contrasting influences and surface water signals for specific groundwater phosphorus vulnerability. Sci. Total Environ. 541, 292–302. https://doi.org/10.1016/j.scitotenv.2015.09.082.

Mellander, P.-E., Galloway, J., Hawtree, D., 2022. Phosphorus mobilization and delivery estimated from long-term high frequency water quality and discharge data. Front. Water 4. https://doi.org/10.3389/frwa.2022.917813.

Mieleitner, J., Reichert, P., 2006. Analysis of the transferability of a biogeochemical lake model to lakes of different trophic state. Ecol. Model. Special Issue on the Fourth European Conference on Ecological Modelling 194, 49–61. https://doi.org/10.1016/j.ecolmodel.2005.10.039.

Mzyece, C.C., Glendell, M., Gagkas, Z., Quilliam, R.S., Jones, I., Pagaling, E., Akoumianaki, I., Newman, C., Oliver, D.M., 2024. Eliciting expert judgements to underpin our understanding of faecal indicator organism loss from septic tank systems. Sci. Total Environ. 171074 https://doi.org/10.1016/j.scitotenv.2024.171074.

Negri, C., Mellander, P.-E., 2024. Evidence Dossier for the Workshop Titled "in-Stream Phosphorus Cycling in Agricultural Catchments: A Workshop to Quantify Abiotic and Biotic P Uptake" https://doi.org/10.6084/m9.figshare.25055165.v1.

Negri, C., Mellander, P.-E., Schurch, N.J., Wade, A.J., Gagkas, Z., Wardell-Johnson, D.H., Adams, K., Glendell, M., 2024. Bayesian network modelling of phosphorus pollution in agricultural catchments with high-resolution data. Environ. Model Softw. 106073 https://doi.org/10.1016/j.envsoft.2024.106073.

Neumann, A., Blukacz-Richards, E.A., Saha, R., Alberto Arnillas, C., Arhonditsis, G.B., 2023. A Bayesian hierarchical spatially explicit modelling framework to examine phosphorus export between contrasting flow regimes. J. Gt. Lakes Res. 49, 190–208. https://doi.org/10.1016/j.jglr.2022.10.003.

Oakley, J., 2020. Getting Started with SHELF.

O'Hagan, A., 2019. Expert knowledge elicitation: subjective but scientific. Am. Stat. 73, 69–81. https://doi.org/10.1080/00031305.2018.1518265.

Piffady, J., Carluer, N., Gouy, V., le Henaff, G., Tormos, T., Bougon, N., Adoir, E., Mellac, K., 2021. ARPEGES: a Bayesian belief network to assess the risk of pesticide contamination for the river network of France. Integr. Environ. Assess. Manag. 17, 188–201. https://doi.org/10.1002/ieam.4343.

Refsgaard, J.C., van der Sluijs, J.P., Brown, J., van der Keur, P., 2006. A framework for dealing with uncertainty due to model structure error. Adv. Water Resour. 29, 1586–1597. https://doi.org/10.1016/j.advwatres.2005.11.013.

Schuwirth, N., Borgwardt, F., Domisch, S., Friedrichs, M., Kattwinkel, M., Kneis, D., Kuemmerlen, M., Langhans, S.D., Martínez-López, J., Vermeiren, P., 2019. How to make ecological models useful for environmental management. Ecol. Model. 411, 108784 https://doi.org/10.1016/j.ecolmodel.2019.108784.

Sherriff, S., Rowan, J.S., Melland, A.R., Jordan, P., Fenton, O., Ó hUallacháin, D., 2015. Investigating suspended sediment dynamics in contrasting agricultural catchments using ex situ turbidity-based suspended sediment monitoring. Hydrol. Earth Syst. Sci. 19, 3349–3363. https://doi.org/10.5194/hess-19-3349-2015.

Shore, M., Jordan, P., Mellander, P.-E., Kelly-Quinn, M., Daly, K., Sims, J.T., Wall, D.P., Melland, A.R., 2016. Characterisation of agricultural drainage ditch sediments along the phosphorus transfer continuum in two contrasting headwater catchments. J. Soils Sediments 16, 1643–1654. https://doi.org/10.1007/s11368-015-1330-0.

Stutter, M., Richards, S., Ibiyemi, A., Watson, H., 2021. Spatial representation of in-stream sediment phosphorus release combining channel network approaches and in-situ experiments. Sci. Total Environ. 795, 148790 https://doi.org/10.1016/j.scitotenv.2021.148790.

Stutter, M.I., Demars, B.O.L., Langan, S.J., 2010. River phosphorus cycling: separating biotic and abiotic uptake during short-term changes in sewage effluent loading. Water Res. 44, 4425–4436. https://doi.org/10.1016/j.watres.2010.06.014.

Teagasc - Agriculture and Food Development Authority, 2018. Agricultural Catchments Programme - Phase 2 Report.

Thomas, I.A., Jordan, P., Mellander, P.-E., Fenton, O., Shine, O., Ó hUallacháin, D., Creamer, R., McDonald, N.T., Dunlop, P., Murphy, P.N.C., 2016a. Improving the identification of hydrologically sensitive areas using LiDAR DEMs for the delineation and mitigation of critical source areas of diffuse pollution. Sci. Total Environ. 556, 276–290. https://doi.org/10.1016/j.scitotenv.2016.02.183.

Thomas, I.A., Mellander, P.-E., Murphy, P.N.C., Fenton, O., Shine, O., Djodjic, F., Dunlop, P., Jordan, P., 2016b. A sub-field scale critical source area index for legacy phosphorus management using high resolution data. Agric. Ecosyst. Environ. 233, 238–252. https://doi.org/10.1016/j.agee.2016.09.012.

Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203, 312–318. https://doi.org/10.1016/j.ecolmodel.2006.11.033.

Wade, A.J., Hornberger, G.M., Whitehead, P.G., Jarvie, H.P., Flynn, N., 2001. On modeling the mechanisms that control in-stream phosphorus, macrophyte, and epiphyte dynamics: an assessment of a new model using general sensitivity analysis. Water Resour. Res. 37, 2777–2792. https://doi.org/10.1029/2000WR000115.

Wade, A.J., Whitehead, P.G., Butterfield, D., 2002. The Integrated Catchments model of Phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogeneous river systems: model structure and equations. Hydrol. Earth Syst. Sci. 6, 583–606. https://doi.org/10.5194/hess-6-583-2002.

Wall, D., Jordan, P., Melland, A.R., Mellander, P.-E., Buckley, C., Reaney, S.M., Shortle, G., 2011. Using the nutrient transfer continuum concept to evaluate the

European Union Nitrates Directive National Action Programme. Environ. Sci. Pol. 14, 664–674. https://doi.org/10.1016/j.envsci.2011.05.003.
Withers, P.J.A., Jarvie, H.P., 2008. Delivery and cycling of phosphorus in rivers: a review. Sci. Total Environ. 400, 379–395. https://doi.org/10.1016/j.scitotenv.2008.08.002.
Zambrano-Bigiarini, M., 2020. hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series. https://doi.org/10.5281/zenodo.839854.