

Hydrol. Earth Syst. Sci. Discuss., author comment AC1
<https://doi.org/10.5194/hess-2021-477-AC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Mads Troldborg et al.

Author comment on "Probabilistic modelling of the inherent field-level pesticide pollution risk in a small drinking water catchment using spatial Bayesian belief networks" by Mads Troldborg et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2021-477-AC1>, 2021

RESPONSE TO REVIEWER 1 COMMENTS

This manuscript presents an interesting development of a hybrid Bayesian Belief Network, applied to the contamination risk of surface water by pesticides. This presents an interesting example of decision support tool with an implementation of a mix of different management options.

Though it is well written and nice to read, I have some concerns that need to be addressed before this manuscript gets published:

We thank the reviewer for a positive assessment and for constructive comments. We have responded to the comments below and highlighted any proposed changes to the manuscript text using italic.

Major concerns:

R1.1 - P7 l184-186: have you checked for time tendencies in climatology data. The mean values observed over the 1894-2019 period might be different from the 2016-2019 climate period on which the concentrations are observed.

Thank you for the comment. We apologise for the imprecise specification of what was done in the model. We propose to amend this paragraph as follows:

'Monthly rainfall and temperature data were available from a single meteorological station in the study catchment, operated by Jersey water for the period 2014-2019. Due to the short duration of this record, additional monthly total rainfall and mean monthly temperature data (from 1894-2019) were obtained from the Government of Jersey website (<https://opendata.gov.je/organization/weather>). Data for the years 1981-2019 were then used in combination with the catchment-based meteorological data to calculate monthly mean and standard deviation as priors for the model.'

R1.2: You describe your dataset as sparse. Could you please tell more about it, as

it appears that the catchment on which you are working is rather well documented with many different sources of data.

We appreciate the comment. The reviewer is correct that the modelling incorporates data from many different sources, and we believe we have already described and summarised this information in the manuscript (2.1-2.4) as well as in the appendix A. The reviewer is correct that the pesticide monitoring for some active ingredients (but not all) in the catchment by Jersey Water can be considered fairly substantial. Jersey Water currently scans for over 100 different pesticides using a risk-based water quality monitoring programme (cf. Jersey Water's Annual Water Quality reports <https://www.jerseywater.je/water-quality-report/>). As suggested in Table 1, the number of pesticide observations available over the considered time period (2016-2019) vary for the different pesticides with pendimethalin and ethoprophos roughly being sampled on a weekly basis all year round, whereas prosulfocarb generally has been sampled on a weekly basis from mid-February to end of May, bi-weekly in January and otherwise every 3 to 5 weeks. The monitoring of glyphosate and metobromuron was more sporadic and only done during spring and early summer when these pesticides are expected to be used, with no sampling taken place in the autumn and the winter.

As described in section 2.2, data on pesticide application was obtained from Jersey Royals Company and are hence representative of actual application rates from Jersey island.

However, besides from the pesticide data, the data in the catchment is sparse. As described in section 2.3, there are very limited soils data available in the catchment, so we had to identify alternative sources, using Europe-wide soil property data derived from the HYPRES database. There is also very limited information available on hydrogeology and groundwater as well as on the catchment hydrology and water balance with the hydro-climatic data used being based on Jersey wide data. We had to use considerable ingenuity to identify alternative data sources and we are pleased that we seem to have succeeded – as the impression is of a 'data rich' catchment.

We suggest that we revise the first paragraph of section 2.2 to describe the pesticide monitoring in a bit more detail as follows:

'Jersey Water, the sole water company and provider on Isle of Jersey, regularly tests and scans the quality of the raw water from the reservoir offtake for a range of different pesticides. Five active pesticide ingredients currently or recently in use in the catchment showed evidence of significant concentrations in the reservoir offtake for the drinking water supply. These included the herbicides glyphosate, metobromuron, pendimethalin and prosulfocarb, and the nematicide and insecticide ethoprophos. During 2016-2019 the sampling frequency of these pesticides was not the same (Table 1) with pendimethalin and ethoprophos being sampled on approximately a weekly basis all year round, whereas prosulfocarb generally was sampled on a weekly basis from mid-February to end of May, bi-weekly in January and otherwise every 3 to 5 weeks. The monitoring of glyphosate and metobromuron was more sporadic and only done during spring and early summer with no sampling taking place in the autumn and the winter. Metobromuron was most frequently observed above the drinking water standard, followed by ethoprophos, prosulfocarb and pendimethalin (Table 1).'

R1.3 - P10-13: the units in bracket along the text are not SI unit. Could you please check that and make the necessary changes, coherently with units used in Appendix A.

Thank you for the comment. In the model development sections, we were describing and defining all model variables using SI dimensional units (L=length, M=mass, T=time etc.), a common way of describing generic units for inputs and parameters in models and equations, which avoids the need to introduce unit conversions directly in the equations.

However, to avoid any confusion we will revise the method section 2.5 to use SI-units so that this section is consistent with the appendix and the results section as suggested by the reviewer.

R1.4 - P10/11: there is an incoherence between equation (3) and Csw unit in microg.L-1. According to equation (3) which refers to fluxes, Csw should be a quantity per volume per year. Please check carefully your formulas and units to make it coherent.

The reviewer is correct, the units in Eq. 3 are inconsistent. The reason for this inconsistency was that we were referring to annual pesticide application rates or fluxes (i.e., mass per area per time), when in fact we were actually considering the amount of pesticide mass applied per area in the model. We will change the term flux to load throughout the revised manuscript (and appendix A) and will make it clear that we refer to applied mass per area.

R1.5: I have a concern about the way Csw is computed. Why did you divide the fluxes by the total volume of the reservoir. Doing so, you consider that a field is at risk only if its contribution to the reservoir contamination is enough to make it higher than the standard drinking limitation. This appears rather limiting the potential risk. And what about the sum of fields contribution to assess the catchment risk?

Thank you for the comment. It is true that when calculating Csw, the flux (load) calculated from a given field (L_{gw}+L_{of}) is divided by the total volume of the reservoir, but please note that in Eq. 3 we also multiply the load by the total area of ALL fields in the catchment (and not by the area of the given field). As stated after Eq. 3, this calculation essentially corresponds to assuming that the load from every field in the catchment is the same as the load computed for the given field. Please note that this calculation was primarily done to help define the discretisation boundaries (low, medium, high) for the combined surface water risk node. We suggest rephrasing how Csw is calculated to make it clearer that we use the total area of all fields in the catchment for the calculations as follows:

'This combined pesticide load was converted to a surface water concentration C_{sw} [$\mu\text{g L}^{-1}$] to evaluate the risk to surface water as follows:

$$C_{sw} = (L_{gw} + L_{of}) * A_f / V_{res} * 10^6 \quad (3)$$

where A_c is the total area of all fields in the catchment (192 ha) and V_{res} is the water volume in the reservoir (938,700 m³).

As the reviewer points out, the loads/contributions from each field could be calculated and summed to assess the overall catchment risk. As discussed in section 3.5 (we propose to move this to section 3.4 in the revised version of the manuscript as stated in response to **R1.8** below), the spatial application in the R package `bnspatial` allows to simulate expected quantities, based on the median value of each discretisation interval. Hence, by multiplying the combined loads from each field and summing the resulting pesticide masses over all fields in the catchment and dividing by the reservoir volume, a concentration in the reservoir water could be estimated for each pesticide and month, which could be compared to measured concentrations if available for further model validation. However, we found that this deterministic calculation was heavily reliant on the discretisation of the target node in question, and, coupled with rare extreme high values generated by the stochastic model, made such "validation" uncertain. We therefore did not follow this approach. We will revise section 3.4 on model validation to make this aspect

clearer as follows:

'A qualitative 'reasonable fit' visual inspection has been shown to be an effective means of assessing model performance using diverse incomplete data sets (Ghahramani et al., 2020). Ghahramani et al. (2020) found that the ranking of confidence in model predictions between determinands was related to data availability as much as to the model itself, with pesticide simulations performing less well than those for hydrology, sediments and phosphorus.'

*Further validation approaches could be explored in future implementations. The spatial application in the R package *bnspatial* allows to simulate expected quantities, based on the median value of each discretisation interval. Hence, by multiplying the expected loads from each field with the probability of the field falling into each discretised interval, then summing the resulting pesticide masses over all fields in the catchment and dividing by the reservoir volume, a concentration in the reservoir water could be estimated for each pesticide and month. This could then be compared to measured concentrations if available for further model validation. However, this deterministic calculation would be heavily reliant on the discretisation of the target node in question and, coupled with rare extreme high values generated by the stochastic model, would make the validation uncertain. Hence, it would be best applied in combination with dynamic discretisation, if available, and with further model development constraining upper simulated values.'*

R.1.6 - P15I416-417: what volume did you use as a *dmix* for the overland flow flux (comparatively to eq 11)?

We assumed the same mixing volume for the overland flow load as used for the groundwater load. We did this purely to allow a comparison of the relative contribution of the two components to the combined risk. Defining the discretisation boundaries (low, medium, high) for the groundwater load was based on an assumption that the pesticide load L_{gw} reaching the groundwater was mixed in the upper $dmix=10$ cm of the groundwater aquifer; the resulting concentration ($C_{gw} = L_{gw}/dmix$) can then be compared to regulatory standards based on which the discretisation boundaries were defined. It is therefore relatively straightforward to define the boundaries for the groundwater load; all that is needed is a mixing depth assumption. It is harder to do the same for the overland flow load, as it is not easy to define the depth (or volume) that the load (or mass) should be mixed in. For simplicity, we therefore opted for using the same boundaries as for groundwater load. We propose to revise the end of section 2.5.2 as follows to make this aspect clearer:

'It is not straightforward to assess when the calculated overland flow load from a given field to the reservoir should be considered high, medium or low, as it is not easy to define the water depth (or volume) that this pesticide load (or mass) should be mixed in. For simplicity, the overland flow load was therefore evaluated similarly to the leaching risk (Eq. 11) to allow a comparison of the relative contribution of the two components to the combined risk.'

R1.7: Have you performed any inference according to observed concentrations data. If so, please give details about it, if not, please state clearly your model remains an expert based model.

No, we did not perform any inference according to observed concentration data. We will emphasise this in section 3.4 of the revised manuscript that our model is an 'expert-based' model.

R1.8: The results/discussion sections are merged. This could be acceptable, but the subsections 3.2/3.3/3.4 are only results with no confrontation to literature or prior hypothesis. This would be great to add discussion elements to these paragraphs or to move them to an independent result section.

Thank you for this point. We will put our findings in a wider context by discussing them in relation to literature and propose to amend the text as follows.

In section 3.2 we proposed to include additional text:

'While the above influential variables have been previously identified among important factors for rapid overland flow risk (Bereswill et al., 2014; Tang et al., 2012), our modelling approach allows to distinguish between generic risk factors and variables with the greatest influence on pesticide pollution risk in a local catchment setting. Coupled with the ability to identify critical source areas, and how they change at a monthly time step, our approach offers a major advancement as compared to static GIS-based risk index assessment approaches (e.g. Quaglia et al., 2019).'

In section 3.3 we propose to revise the text to:

'Reduction in application rates unsurprisingly results in a reduced risk, particularly in the groundwater leaching risk (Figure 9), which is reduced to Low even after 10% reduction. This is in line with findings of Reichenberger et al. (2007) who concluded that application rate reduction was one of a few mitigation measures that could address pesticide leaching to groundwater with some confidence. However, while reduction in application rates is an easy measure to implement and may lead to cost savings, it can only be implemented to a point whereby it remains effective, thus potentially limiting its acceptability as a mitigation measure to farmers (Bereswill et al., 2014). Introduction of buffers reduces the runoff risk to a similar extent as a 50% reduction of pesticide application rates for overland flow risk. Buffers, particularly edge of field buffers, have been found to be effective mitigation measures for the overland flow pathway (Bereswill et al., 2014; Reichenberger et al., 2007). Although buffer effectiveness is variable, most studies report efficiencies over 60% (Bereswill et al., 2014). In addition, permanent buffers are seen as an acceptable mitigation intervention, as long as farmers can be compensated for potential losses of growing area (Bereswill et al., 2014). Managing and removing potential plough pans increases the amount of infiltration into soils, thus reducing the pesticide runoff risk to an extent that is comparable to 10% reduction in application rates (Fig. 8). By combining all available maximum interventions of 50% reduction in pesticide application rate, management of plough pan, delayed application timing and installing additional field buffers, the probability of all types of risk is notably reduced (Figs. 4-6, 8-9). Whilst some of these measures may be deemed less acceptable or difficult to implement (e.g. delayed application timing (Bereswill et al., 2014)), an ambitious mitigation approach may be required to achieve real water quality improvements (Villamizar et al., 2020).'

In section 3.4 we propose to revise the text (and move some of the text from 3.5) to:

'We constructed a causal network, where the model structure was informed by expert knowledge. However, Bayesian Networks can also be used as machine-learning associative tools that are suitable for deriving patterns in datasets without a specific response variable. It could be argued that pesticide risk, expressed as load or concentration of pesticide in different potential loss pathways (overland flow or groundwater leaching) is a latent variable without available observational data (Piffady et al., 2020), making it difficult to calibrate or validate a risk model. Hence, model credibility and salience (Cash et al., 2005) need to be evaluated by experts and stakeholders. Here, we implemented a simple validation approach to confirm that the model predictions fall within the realms of

credibility, using the limited observational data to validate the expert-based model.

Figure 10 shows a comparison between the probability density distributions based on 10,000 surface water risk simulations for each active ingredient in the hybrid network and the limited observational data (in $\mu\text{g L}^{-1}$) available for the months January – March between 2016 and 2019. The model typically over-estimates the simulated risk for glyphosate and pendimethalin, albeit with low probability of high values. The simulated and observed distributions for prosulfocarb are comparable, whilst the model seems to under-estimate the risk from metobromuron. However, it has to be noted that the very few observations available for metobromuron (N=8) seem to be higher and less accurate than for the other pesticides. It should also be noted that the developed model was never intended to represent the complex transport and fate processes in the catchment in detail or to accurately simulate the pesticide concentration levels in the reservoir, so the comparison in Figure 10 was mainly carried out as a sense check of the model predictions. Overall, model simulations appear conservative, which is helpful in terms of informing a precautionary management approach. *Further model refinement could focus on constraining the upper simulation values throughout the model. A qualitative 'reasonable fit' visual inspection has been shown to be an effective means of assessing model performance using diverse incomplete data sets (Ghahramani et al., 2020). Ghahramani et al. (2020) found that the ranking of confidence in model predictions between determinands was related to data availability as much as to the model itself, with pesticide simulations performing less well than those for hydrology, sediments and phosphorus.*

*Further validation approaches could be explored in future implementations. The spatial application in the R package *bnspatial* allows to simulate expected quantities, based on the median value of each discretisation interval. Hence, by multiplying the expected loads from each field with the probability of the field falling into each discretised interval, then summing the resulting pesticide masses over all fields in the catchment and dividing by the reservoir volume, a concentration in the reservoir water could be estimated for each pesticide and month. This could then be compared to measured concentrations if available for further model validation. However, this deterministic calculation would be heavily reliant on the discretisation of the target node in question and, coupled with rare extreme high values generated by the stochastic model, would make the validation uncertain. Hence, it would be best applied in combination with dynamic discretisation, if available, and with further model development constraining upper simulated values.'*

R1.9 - P16 I456-458 and Section 3.4: have you simulated 10.000 results for each field? It is unclear here. And once again, I have a concern with the direct comparison of Csw and concentrations as concentrations are the sums of all field contributions. Furthermore, it might be interesting to also consider / discuss the fact that ground fluxes are probably slower than surface ones.

Thank you for the comment. For the validation, we did not run 10.000 simulations for each field but instead tested the simulations derived from the hybrid model against observational data, before discretisation and spatial application was implemented. We will clarify this in the text (L637) as:

'Figure 10 shows a comparison between the probability density distributions based on 10,000 surface water risk simulations for each active ingredient in *the hybrid network* and the limited observational data'.

Please refer to our response regarding the calculation of Csw in R1.4 above. The reviewer is correct, the groundwater fluxes will be slower than the runoff ones. This is something

the model is currently not accounting for. We will highlight this in the limitation and outlook section 3.5 as:

'Although quite detailed in process representation that is based on established mechanistic approaches, the modelling of pesticide leaching and runoff in the BBN is still simplified and could be extended to e.g., consider preferential flow pathways and/or capture the slower rate of groundwater pesticide leaching as compared to fluxes via surface runoff.'

Minor comments and typos:

R1.10 - P5 I142: why is fluopyram underlined?

Thank you for pointing out this error, which will be corrected in the revised manuscript

R1.11 - P11 I320: check reference brackets.

We will correct this formatting error.

R1.12 - P15 I419: typo on 'asses'.

We will correct this typographical error.

R1.13 - P16 I471: Piffady (2020) approach was clearly not mechanistic. This would be more helpful to cite here an example of mechanistic approach.

Thank you for the comment. We will remove the reference to Piffady and replace the reference with the studies by Vilamizar et al. 2020 (using SWAT to model pesticide and influence of management interventions) as well as the review paper by Kohne et al. (2009) focussing on mechanistic modelling of pesticide transport at various scales.

R1.14 - P24 I578-580: a reference would be needed.

Thank you for spotting the missing reference. We will add the reference to the study by Reichenberger 2007 in the revised manuscript.

R1.15 - P24 I581: typo on 'metabrouron'. Please check metabromuron spelling through the document, you sometimes write it with an h, sometimes without.

Thank you for spotting this inconsistency. We will check the spelling and ensure that we write 'metobromuron' throughout the revised manuscript

R1.16 - P25 I584: please precise which figure you refer to.

Apologies for the omission, we will amend the Figure number in the revised manuscript.

R1.17 - P29 I666: '..' typo.

We will replace the ... with (...) to mark the gap in the text quoted from the Schuwirth et al. (2019) page 2 and add the page number in the manuscript

R1.18 - P32 actual evapotranspiration: there are overlapping boundaries between low and medium classes.

Thank you for spotting this error. This will be corrected as Low <35, Medium 35-40, High >40

R1.19 - P35 'groundwater fluxes': I would expect an 'f-leach' factor in the Leach formula, coherently with eq (1).

The reviewer is correct. We will correct this in a revised manuscript

R1.20 - P36 'surface water risk' : according to my units checking, there should be a time factor in the unit (see major concerns).

Please see the response to major comment in **R1.4**.

