Article:

Suslovaite, V., Speight, V. and Shucksmith, J.D. orcid.org/0000-0001-5497-0051 (2025) A catchment management optimization approach to mitigate rainfall induced pesticide contamination in water supply systems. Journal of Hydrology. 134123-. ISSN: 0022-1694

https://doi.org/10.1016/j.jhydrol.2025.134123

A catchment management optimization approach to mitigate rainfall induced pesticide contamination in water supply systems

Vaida Suslovaite[1], Vanessa Speight[1], James D. Shucksmith[1]

[1]School of Mechanical, Aerospace and Civil Engineering, University of Sheffield, Sheffield, S1 3JD, UK. Email: vsuslovaite@gmail.com

Keywords: water abstraction, pesticides, surface water quality. spatial optimization, catchment management, water resources

1. Introduction

Water resources are under pressure from population growth, pollutant emergence and climate change (United Nations, 2019). The effects of climate change are likely to lead to more frequent and intense extremes in rainfall and droughts (Caloiero et al., 2018, Du et al., 2019). The implications for water resources include both increased contaminant wash-off into river systems during rainfall events, and/or increasing periods of low river flows with corresponding reduced capacity for dilution of contaminants (Graydon et al., 2022). The likely increased pressure on surface raw water treatment facilities during pollutant spikes in river systems will lead to increased costs and higher risks of drinking water quality failures, which will in turn heighten risks to human health (Swinamer et al., 2024).

Alongside emerging contaminants, the management of pesticide runoff to water supply systems is a current and ongoing concern (Cooke et al., 2020). In the United Kingdom and EU water supply regulations set the maximum current legal limit for drinking water at 0.1 μg/l of one particular pesticide or not more than 0.5μg/l of all pesticides present in total (European Commission, 2024). The effectiveness of pesticide removal from surface water via treatment varies from country to country and several pesticides are commonly found in drinking water (Tröger et al. 2021). There is therefore an ongoing need for cost effective mitigation strategies to reduce contamination risks to water supply systems.

Previous studies have demonstrated that the spatial distribution of non-point sources (NPS) and/or placement of remediation approaches within a catchment can have a significant impact of pollutant loads at sensitive locations including abstraction sites (Zhang et al. 2011, Brookes et al., 2015). Hence, simulation based spatial optimization is seen as a useful method to inform catchment management strategies for the mitigation of non-point pollution of water resources (Srivastava et al., 2002; Arabi et al., 2006, Dai et al., 2024). These techniques can guide catchment best management practices (BMPs) and resources to be targeted to specific areas which are expected to be most effective in reducing pollutant impacts at specific sites (e.g. at the river basin outlet). BMP measures may include the efficient targeting of farmer subsidies to remove or reduce contaminant sources (Cooke et al., 2020), or the implementation of filter strips or riparian buffers (Schramm et al., 2024).

Simulation methodologies commonly involve an optimisation-based framework in which an objective function defined based on an aspect of a water quality model output is minimized as a function of the spatial (and/or other) properties of the pollutant sources in the catchment. An important consideration is the selection of appropriate inputs and a modelling approach which enables the optimization routine to produce faithful outputs of the water quality dynamics under investigation.

For example, Srivastava et al., (2002) showed that the use of continuous time series rainfall as a simulation input rather than combinations of design rainfall events provided a superior performance when optimizing BMP placement in a 725 ha agricultural catchment for yearly water quality improvement. In this case, the objective functions were defined as the total pollutant load simulations of the annualised AnnAGNPS model and net returns of a simple cost model. Numerous further examples of land use and/or BMP targeting optimization methodologies are available in the literature (Bodrud-Doza et al., 2023; Kaim et al., 2018), utilizing alternate optimization algorithms (Kaim, et al, 2018), remediation techniques (Lui et al., 2019), as well as coupled approaches that also consider cost or other functions (Arabi et al., 2006, Jeong et al., 2024). However, the most commonly utilized water quality evaluation criteria are total pollutant loadings which are normally seen as a general indicator of overall NPS impacts although some alternate metrics such as stream health have also been explored (Herman et al., 2016). Further, for computational efficiency under optimization, most approaches utilize lumped/or semi distributed water quality models (such as SWAT) and calculate performance metrics based on simulated water quality at daily or lower temporal resolutions. Hence existing approaches for simulation based spatial optimization are rarely tailored for specific objectives related to water supply systems.

When considering the case of impacts of specific pollutants on water resource infrastructure such as abstraction systems, objective functions which consider specific regulatory or operational target functions are likely to be more relevant when developing spatial optimization approaches. For pesticides, these targets frequently consider the duration in which concentrations remain above the regulatory threshold, such that control/abstraction/pumping decisions can be optimized to achieve water resource targets (Yassin et al., 2021, Ortiz-Lopez et al., 2022). Further, in many cases relevant water quality dynamics are highly variable at sub-daily scales, being sensitive to the spatial and temporal variability of surface rainfall runoff processes (Asfaw et al. 2018, Delpla et al., 2019, Suslovaite et al. 2024), and hence require a distributed modelling approach operating at sub-daily (i.e. event based) temporal scales to capture the necessary detail for appropriate spatial optimization.

Given the sensitivity to surface runoff dynamics, an additional challenge is to derive suitable and computationally efficient rainfall inputs to any simulation/optimization approach which account for the temporal and spatial variations in rainfall in the catchment. Whilst the use of historical rainfall time series would naturally incorporate such variations, the combination of a detailed/distributed model and extensive time series inputs within an optimization routine would be computationally prohibitive.

Optimization approaches which utilize more complex hydrodynamic or water quality models are often required to develop appropriate rainfall sampling strategies, such that a shorter subset of rainfall events can be derived which preserve the necessary rainfall features for simulation (Mounce et al., 2020, Eulogi et al., 2022).

This study develops and tests a spatial high-risk land management optimization approach appropriate to the reduction of acute pollutant pesticide concentrations, in this case metaldehyde, from rainfall runoff at water abstraction sites. A novel methodology to characterise rainfall inputs for optimisation-based approaches within complex/distributed catchment models is developed for this purpose, and a contaminant specific objective function based on threshold exceedance is used for the proposed application. The technique is implemented in a UK test catchment with historically high levels of observed metaldehyde concentrations, for which a validated distributed water quality model currently exists. The land use optimization methodology can be used to prioritize the mitigation of high-risk areas within the catchment, and further investigates how removal of these areas (e.g. via a subsidy scheme, as in Cooke et al., 2020) affects acute pollutant concentrations in river systems under a historical rainfall time series when compared to alternate land use mitigation strategies.

2. Methodology

This section describes the development of an inverse modelling methodology to be applied to a case study catchment. In this work the land use optimization approach is developed for the mitigation of a specific pesticide (metaldehyde), for which an event-based transport model in the case study catchment has previously been developed, calibrated, and validated. Based on the requirements of the water utility, an objective function related to concentration duration over a threshold is selected. The methodology identifies priority high-risk fields in which pesticide has been applied to crops to target for intervention, representing up to 5% of the total high-risk area. However, the optimization approach is suitable for other cases, pollutants and target area thresholds for which a distributed water quality model is required to consider water quality dynamics at appropriate scales (i.e. at high/sub-daily resolutions). Finally, the approach is then verified by the simulation of pollutant response under historical measured rainfall record for the optimised land distribution vs comparable alternate random and clustered field selection strategies.

2.1 Study site

At 300 km^2, River Leam is a moderate size sub-catchment of River Severn. Catchment elevation ranges between 46m to 232m above sea level. The abstraction site, where surface water is abstracted for drinking water supply, is maintained by the utility who carries out routine and regulatory monitoring of river water quality. A UK Environment Agency flow gauging station is present on site with data available at 15 minute intervals. Typical flow depth range is between 0.24 m and 1.16 m with mean flow of 1.57 m^3/s and mean annual catchment rainfall of 649mm.

The utility has identified pesticides as a pollutant of concern at the abstraction site through long-term routine monitoring. Within the case study catchment in 2018, 1114 ‘high risk’ fields (i.e. to which metaldehyde was applied) were identified with a total area of 88.76 km^2. Utility collaborative programmes (Farm to Tap schemes) with the farms are still ongoing to mitigate diffuse pollution from pesticides (Severn Trent Water, 2024). The schemes are carried out annually at catchment level where individual farms can apply for funding from the utility if they are eligible. Hence, in this context of cooperative land management, it would be beneficial to develop an approach to facilitate identification of priority catchment areas to inform targeted catchment management interventions.

2.2 Pesticide model

In this study we focus on the example mitigation of a specific pesticide (metaldehyde) which is a soluble molluscicide used in agriculture to control slugs and snails (Li et al., 2010). In the UK metaldehyde was historically applied to winter crops such as winter wheat, potatoes and oilseed rape, between September and December, when the conditions are most favourable for Mollusca (Asfaw, 2018). Its low sorption coefficient of active ingredient to organic carbon (KOC) value (34 to 240 L/kg) (Kay and Grayson, 2014) combined with its relatively long half-life in soil (3.17 to 223 days) allows for it to be readily leached into surface runoff during rainfall events. As such is poses significant risks for water supply systems, with frequent observations of high concentrations in arable catchments after rainfall. Metaldehyde has previously been identified to be responsible for majority of all cases of pesticide exceedances in drinking water in England and Wales. In 2016 it accounted for 87% of all pesticide exceedances recorded that year (DWI, 2017). The utility has reported exceedances at 17% of water treatment works (WTW) in 2017 and at 8% of WTWs in 2018 (Cooke et al., 2020).

Asfaw et al. (2018) presented a validated, travel time based, physically distributed model used to predict metaldehyde levels after a rainfall event accounting for variations in rainfall and distribution of land use. The model was tested/validated on the same case study catchment based on hourly metaldehyde observations at the catchment outlet following rainfall events. The model is comprised of surface runoff generation, surface runoff routing and pollutant build-up/wash-off components. The surface runoff component calculates the cumulative excess rainfall depth I^t (mm) at each timestep t based on the differential form of the Soil Conservation Service (SCS) curve number (CN) method (Mancini and Rosso, 1989). The surface runoff routing component uses a spatially distributed time variant direct runoff travel time technique to account for spatial and temporal variability of runoff generation and flow routing through overland flows and stream networks (Melesse and Graham, 2004; Du et al., 2009). The pollutant build-up/wash-off component estimates metaldehyde build-up through pesticide applications on identified high-risk areas. The model operates at 1 h time step, with input spatial rainfall data at 1 km^2, and calculates runoff at $5m^2$ resolution. The model was validated via direct measurement during independent rainfall runoff events (monitored between Oct 2014 - Feb 2017) during which metaldehyde concentration was monitored at hourly intervals at the catchment outlet. When compared to measured values, model simulations for the events all had correlation coefficients of 0.70 or more, prediction error of peak metaldehyde concentration less than 5% and time to peak concentration error of 6 or less hours. Overall, model validation returned an average coefficient of determination of 0.75 and model efficiency of 0.46. Further details of the model build, calibration and validation can be found in Asfaw et al (2018). The model has since been used in drinking water abstraction management at the utility to suspend abstraction from surface water when a peak in metaldehyde concentrations is forecasted.

2.3 Development of inverse modelling method for designing catchment management options

The inverse modelling approach searches for model input, in this case a distribution of catchment high-risk fields (i.e. those which act as a significant source for pesticide, based on seasonal distribution of crops as identified by land cover maps), that result in desired model output (defined pesticide levels in river water). The objective function is set to be the number of predicted hours that pesticide levels exceed the specified EU and UK threshold of 0.1 μg L^{-1} in drinking water at the potable water abstraction site situated at catchment outlet. This objective function is of specific relevance for water supply systems with limited raw water storage or ability to blend with other sources, in which abstracted water is directly treated and distributed to consumers.

There are vast amounts of possible catchment high-risk field distributions and so a guided search algorithm is needed. In this case, genetic algorithm (GA) was selected to carry out land use optimisation. GA is widely used to solve optimization problems in water resources planning and management (Nicklow et al., 2010, Eulogi et al., 2022). GA is an evolutionary search algorithm based on natural selection. It works with parameter sets of a model while checking the outcome of the model as its objective function. The parameter values that produce the most optimal model outcome are then selected to produce the next set of parameters ('offspring') through crossover and mutation. In many applications, there is a need to have several near optimal solutions as alternatives because not all solutions can be implemented for practical reasons. Hence, GA is especially suited to mitigation measure allocation searches (Srivastava et al., 2003; Srivastava et al., 2002; Arabi et al., 2006; Perez-Pedini et al., 2005) because it searches from populations rather than a single point and can provide more than one solution. A more detailed summary of Genetic Algorithms and their applications is detailed in Tang et al. (1996).

2.3.1 Zero-one integer programming

Combinatorial GA problems require an input as a list of values that can be presented in different combinations which the algorithm can optimise. In this study, zero-one integer programming is used to represent source fields within the catchment where the pesticide is present (1) or not present (0). The technique has been previously used in solving similar allocation problems where the method (or land use) is either implemented or not implemented (Wang et al., 2019, Aerts et al., 2002).

Since fields have a non-uniform area, maintaining the same total pesticide use area for each iteration is not always possible. A targeted land use mitigation approach was assumed in which up to 5% of the land area can be considered for intervention. In practice this is likely to represent measures such as such as farmer subsidies paid by the water utility to use alternate, less harmful alternatives (such as ferric phosphate) at these locations (Cooke et al., 2020). Therefore, in this case the GA aims to find a combination of high-risk (pesticide present) fields which constitutes a reduction of high-risk field area by 5%±1% and minimises the number of hours that forecasted total pesticide levels exceed the threshold of 0.1 μg L^{-1}. Hence, every new solution created in GA contains at least 95%±1% of the total original (2018) high-risk field area.

To begin, an initial solution is created, based on the known 2018 distribution of fields containing crops to which the target pesticide is applied. This is a list of 1's where the total number of digits, represent all the fields present in the original 2018 high-risk shapefile. Then, presence of pesticides is removed from a number of randomly selected fields equalling 5%±1% of total area, with corresponding 1's in the list replaced by 0's. This list is then used to create a shapefile of high-risk fields were pesticides are present. The Asfaw et al. (2018) pesticide model is first run with the initial solution land use shapefile and the outcome forms the initial objective function that GA uses to compare to its subsequent objective functions.

All new solutions form a list of same length as the initial solution but where fields are selected to be removed, the 1s are replaced with 0s. For each iteration, a new high-risk shapefile is created and the resulting objective function value is determined based on the water quality model and associated variables.

Genetic algorithm parameters are set out in table 1. Figure 1 shows a flow chart of the GA method. The initial solution represents the current distribution of high risk land use. Checking the objective function runs the model with the new high risk shapefile and checks the resulting forecasted total hours pesticide levels are above threshold, and redefines objective function.

Table 1 Parameter settings used in the genetic algorithm.

Parameter name	**Parameter setting**
Probability of crossover	1
Probability of mutation	0.3
Tournament size	3
Population size	100
Number of generations	100
Stopping criteria	4.5 days runtime

2.3.2 Rainfall Inputs

A significant complicating factor is that the spatial and temporal variations in rainfall have a significant effect on the dynamics of pesticide concentrations at the abstraction site (Asfaw et al. 2018). Therefore, an appropriate model rainfall input for the inverse modelling/optimisation approach needs to be carefully considered to account for these processes when considering the spatial distribution of high-risk fields. Utilising observed long-term rainfall datasets within the optimisation would be a valid approach, as this would inherently capture the variations of temporal and spatial rainfall over the catchment area. However, the use of long time series datasets is infeasible in practice due to the computationally intensive nature of the optimisation routines. Hence, to account for the influence of spatial variability of rainfall patterns and intensities, a shorter compilation of rainfall events representative of the historic catchment rainfall is required. In this study, a rainfall event ‘mashup’ was created which contains a selected subset of the historical catchment rainfall spanning several years (2015-2019).

Similar approaches have been used previously for computationally intensive optimisation based methods which are sensitive to temporal variation in rainfall inputs (Mounce et al. 2020). In this study, due to the nature of the rainfall runoff, it is important to retain elements in the subset that capture both the temporal and spatial distribution of rainfall within the catchment. To achieve this, the statistical characteristics (temporal and spatial variability) of the historic catchment rainfall patterns were analysed and recreated as closely as possible in the rainfall ‘mashup’ input file.

A routine was first developed to analyse the 1 km^2 spatial and 5 min temporal resolution spatial rainfall radar data from the Met Office Nimrod System (Met Office, 2003). Initially every 5-minute time step data point was averaged over the catchment area to produce a single value, producing a time series of 5-minute catchment averaged values for several years (2015-2019). As the metaldehyde model is used for September-December months, data for these months only was taken forward for analysis (Figure 2). For every 5-minute time step, the spatial standard deviation, (defined as standard deviation of all the values within the original spatial rainfall file for that 5 minute time step), was calculated as a metric of rainfall variability.

A routine was then developed to loop through the resulting averaged rainfall time series to automate the recognition of storm events. The routine identifies a gap in rainfall, takes the next nonzero value as a start of a storm event, and the subsequent next gap in rainfall as the end of a storm event. It assigns a storm ID to each event, records its start and end date/time, calculates its length (time in hours), total event rainfall depth, average spatial standard deviation, and antecedent moisture condition for 15 days prior to the start of the rainfall event (AMC15). The identified events were further refined so that each included event was at least 1 hour long and be expected to produce an increase in pesticide levels at the abstraction site (based on running the Asfaw et al. 2018 model for each event). Over the full time series (months September to December in the years 2015-2019) this resulted in 188 identified storm events in the catchment expected to influence pesticide levels at the catchment outlet.

A multivariate stratified sampling method (Speight et al., 2004) was then used to select a subset of these events which efficiently characterised the overall temporal and spatial variability of all identified rainfall events in the catchment without bias by taking into account the clustered nature of the data. The rainfall events were assigned into strata by spatial standard deviation, each strata was then stratified by temporal standard deviation. At sub-strata level, random numbers were assigned to elements and sorted largest to smallest. A single element at the top of each sorted sub-strata was then selected. Based on this analysis, sixteen rainfall events were selected to create the mashup subset. Summary statistics of the full and sampled dataset are shown in table 2.

Table 2 summary statistics of catchment averaged hourly rainfall for the full dataset of 188 events and the sampled dataset of 16 events

.Statistic	Full dataset	Sampled dataset
Mean	0.52	0.52
Standard Error	0.02	0.07
Median	0.32	0.31
Mode	1.48×10^{-5}	N/A
Standard Deviation	0.6	0.67
Sample Variance	0.36	0.44
Kurtosis	4.39	6.16
Skewness	1.92	2.31
Range	4.09	3.36
Minimum	4.93×10^{-6}	7.40×10^{-5}
Maximum	4.09	3.36
Sum	668.38	43.55
Count	1287	83

A two-sample Kolmogorov-Smirnov test was used to check if the full rainfall dataset and rainfall mashup dataset obtained through multivariate stratified sampling have the same distribution of temporal and spatial standard deviation. If the Kolmogorov–Smirnov test statistic exceeds critical D (D_α, equation 1) the null hypothesis of both samples come from a population with the same distribution can be rejected.

$$D_\alpha = c(\alpha)\sqrt{\frac{m+n}{mn}} \qquad \text{(Equation 1)}$$

Where *c(α)* is the inverse of the Kolmogorov distribution at significance level *α*, *m* is the first sample size and *n* is the second sample size. As the test statistic was lower than critical D at α = *0.05*, a failure to reject the null hypothesis is implied. Therefore, the two datasets can be assumed to be from the same distribution when checked both by spatial and temporal standard deviation distributions. The histograms below show the distributions for spatial standard deviation (figure 3) and temporal standard deviation (figure 4).

Based on this analysis, the mashup event was taken forward and used as an input to the GA simulation/optimisation routine. When considering the rainfall mashup as an input, the optimisation algorithm for metaldehyde land use was evaluated to have reached the best solution (considering a maximum of 5%±1% land mitigation) after 3023 runs. On a Windows10 computer with Intel I9-10900x processor and 64GB of RAM this optimisation analysis required approximately 4.5 days of simulation time.

2.3.3. Simulation and verification of river pesticide concentrations

To define the performance of the GA approach (including the use of the simplified representation of rainfall inputs via the mashup event), the resulting optimised high-risk mitigation solutions were evaluated by running the pesticide model with the full 188 rainfall event record identified for September-December 2015-2019. The outcome of total duration pesticide levels were above regulatory threshold was compared to two alternative methods for verification: 1) a random 5% removal of high-risk area, and 2) removal of the 5% of fields closest to the abstraction site, as selected by travel time.

3. Results and discussion

3.1 Land-use optimisation

Model outputs for the range of simulated high-risk removal areas resulted in pesticide levels at the catchment outlet above the regulatory threshold for between 307 and 322 hours for the duration of the mashup event. After 3023 runs the GA routine produced a reduction of 39 hours in duration over the regulatory threshold from the initial solution (figure 5).

The location of high-risk fields removed in the initial solution as well as fields removed for the optimal solution as identified by the GA are plotted in figure 6. In this case there is no discernible clustering in the identified fields or identified proximity to the river network. However it is noted that the shape of the temporal distribution of the removed high-risk areas (in terms of travel time from the abstraction site under a uniform 1 mm/hr rainfall event, calculated as in Asfaw et al., 2018) is of similar form to the overall initial 2018 distribution of identified high-risk fields in the catchment (figure 7).

3.2 Performance under historical rainfall record

To evaluate the performance of the methodology in reducing pesticide risks to the drinking water abstraction site, the proposed GA solution (spatial distribution of mitigated high-risk areas) was simulated under the full historical rainfall record (i.e. 188 identified storm events from 2015 - 2019). This effectively tests the impacts of the uncertainties introduced by the simplification of the rainfall record into a mashup event during the optimization process. Water quality outputs were compared to a simulations using the original high-risk (HR) land use with no fields removed, a randomly selected removal of 5% of fields by area (m^2), and a removal of 5% of area from fields with the shortest travel time (identified after running the Asfaw et al., 2018 catchment hydrological model with 1mm/hr uniform rainfall). The total duration that pesticide remained over the regulatory threshold for each of these scenarios is presented in Table 3.

Table 3. Performance of optimization methodology under full 2015-2019 historical rainfall dataset

Scenario	Pesticide hours-above-0.1 μg/l threshold	Reduction in pesticide hours-above-threshold compared to original H.R. field distribution
Original HR fields	3248	N/A
5% random HR fields removed (initial solution)	3157	91 (2.8 %)
5% HR fields removed by best GA solution	2997	251 (7.7%)
5% shortest travel time HR fields removed	3169	79 (2.4%)

The GA solution for high-risk area mitigation reduced modelled pesticide time above the regulatory threshold at the abstraction site by 251 hours or 7.7% in the test catchment. The GA best solution performed significantly better than removing an equivalent area of high-risk fields closest to the abstraction point, or if the field distribution was selected randomly. This demonstrates the effectiveness of the proposed methodology, including the rainfall subsampling procedure in effectively and efficiently targeting high-risk (HR) fields for mitigating pesticide impacts on water supply systems. In this case clustering interventions close to the abstraction site resulted in similar levels of performance to a random chosen distribution. To enable a comparison, a simple sensitivity analysis was conducted utilising the original Asfaw et al. (2018) model by varying the uniform catchment washoff parameter. Based on an analysis over the full rainfall record, to achieve the same reduction in time above threshold using a non optimated field distribution (i.e. randomly removed fields) would require an estimated 9.5 % reduction in pollutant wash-off from catchment high risk fields.

When considering the full record, it is noted that the GA solution is effective in reducing the duration over threshold for the majority of simulated events (although to a minor degree in some cases). In contrast, due to spatial positioning and interaction with the runoff processes the shortest travel time selected H.R. field approach is observed to reduce concentrations during the rising limb of some of the larger events only. The randomly selected H.R. removal field method also performs inconstantly, reducing duration over threshold for a more limited subset of events than the GA solution.As examples, Figures 8 and 9 plot the modelled pesticide concentration for two selected rainfall events from the historical record under the different land use scenarios. Fig 8 is an example of a low intensity, frequently occurring event in which the GA solution reduced the time above threshold by 3 hours, with almost no reduction for the alternate approaches. Fig 9 displays a larger (and hence less frequent) rainfall event where the shortest travel time approach reduced the time above threshold by 1 hour by delaying the first arrival time of the pesticide. Overall, the increased performance of the GA solution is due to a more consistent ability to reduce the duration over the threshold over a greater range of rainfall events than the alternate approaches. This further indicates the effectiveness of the proposed event sub-sampling approach in characterising the full rainfall record.

It is important to note that, given model simulation and optimisation uncertainties it is not possible to claim that the GA solution represents the best possible land use distribution/targeted mitigation strategy. However the use of a catchment validated model together with the verification results provide confidence that the proposed methodology provides a feasible approach to effectively target land use interventions which provides superior performance over random or simple clustered groupings of interventions.

3.3 Future studies

Considering observed interactions between modelled water quality dynamics, the objective function and the defined regulatory threshold for pesticides, it is evident that the choice of objective function has significance when defining the optimal positioning of catchment interventions such as BMPs, and hence when considering the performance of simulation/optimisation algorithms for spatial targeting of interventions. Whilst many previous studies have focused on total pollutant load over a given time period (e.g. year) as an objective function, it is recommended for future studies that further consideration is given to appropriate objective functions based on the proposed application/environmental problem under consideration. Whilst total pollutant load may be appropriate for many environmental problems, there are a range of different environmental performance metrics, regulatory targets and practical considerations which may be appropriate under different situations. Examples include chemical frequency-concentration-duration standards for intermittent pollutant discharges based on environmental toxicology in surface waters (e.g. FWR 2012), bacterial percentile standards for bathing waters (e.g. EU. 2006), and various utility defined management techniques for raw potable water supply, often locally dependent on the availability of bankside storage options to enable blending of raw water as well as site specific treatment efficiency and local drinking water regulations.

A further notable consideration is that in this case, the properties/shape of the travel time distribution of the optimum GA solution resemble that of the initial distribution of HR fields. This suggests that targeting/removing high-risk field areas in proportion to the calculated travel time distribution within catchment may be an effective transferable strategy to other catchments when considering the targeted mitigation of rainfall driven water quality impacts, however this finding requires further verification at other sites. In this specific case model outputs suggest that this strategy is effective in producing solutions which are effective over a wider range of potential catchment rainfall/runoff events than clustering interventions in specific catchment areas. However, it should be noted that interactions between rainfall runoff / water quality dynamics / and the objective function are notably complex, and the degree that this result is a function of the water quality parameter under investigation (a highly soluble pesticide), specific catchment characteristics and nature of the objective function should be further explored in future work.

It is also acknowledged that modelling uncertainties (e.g. those associated with pollutant build up and wash off quantification), within the simulation approach will affect the identification of optimal field distributions using the GA methodology developed herein. Due to computational requirements, adopting formal uncertainty assessments such as Monte Carlo are infeasible when applied to optimisation problems at this scale. However, in this specific case, a site validated simulation model has been used, and performance verified using a comprehensive rainfall record of 188 district events to consider the effectiveness of the model input subsampling approach. Ideally, further studies would consider measuring in river pollutant response before and after GA targeted land use mitigations have been adopted, however this would require a significant multi-season study with the cooperation and resources of water utilities and landowners/farmers. The transferability of this approach to other catchments and contaminants with different characteristics is also worth further study, however when dealing with alternate and/or more complex catchments or contaminants a recalibrated or alternate/more complex water quality simulation model should be used within the optimization framework, and confidence in the validity of model outputs should be validated by further field observations of the contaminant under study during rainfall-runoff events. It should also be noted that practical adoption of mitigation strategies is dependent on the cooperation and agreement with local landowners/farmers (Cooke et al. 2020), and that precise specification of specific areas may not always be feasible to local practicalities such as land access issues.

4. Conclusions

This study has developed a new methodology for targeting catchment mitigation options for the reduction of impacts from pesticides following acute rainfall events on water abstraction systems. In this case impact is defined as duration the pesticide remains above the UK regulatory target of 0.1 μg/l, which is the relevant metric for the water supply utility. Whilst land use optimisation to address water quality problems is a common topic in the literature, to date such methods have focused on an evaluation of long-term water pollutant loads based on lumped or semi distributed catchment models. When considering the mitigation of acute impacts, a key challenge is the required representation of the effects of spatial and temporal variations in rainfall within the optimisation framework. To enable the required computational efficiency required for such an optimisation with a distributed water quality model, this work has proposed the selection of a subset of rainfall events based on statistical interrogation of the historic rainfall record (based on statistical properties related to the spatial and temporal variability of rainfall events) for use within the optimisation routine. The use of this 'mashup' event within a GA optimisation algorithm was able to identify priority areas for catchment intervention which resulted in a significant simulated reduction in pesticide risk to the water supply system. Pending successful transferability, the approach can be feasibly applied using a desktop computer for catchments of moderate scale (300 km^2 in this case). The application of the GA defined solution to the historic rainfall record demonstrated a significant improvement when compared to a simple selection/prioritisation of fields closed to the abstraction site (7.8% in duration target pesticide remained over the regulatory limit, based on a 5% removal of high-risk field area in the catchment). Findings from this case suggest that superior intervention strategies are those which can be effective over the widest possible range of rainfall events within the target catchment. Here, this can be considered by proportionally applying measures which reference to the distribution of high-risk areas according to travel time from the site of concern.

The method of optimisation and rainfall subsampling is potentially transferable to other catchments/pollutant types assuming the availability of appropriate, validated catchment specific models for the pollutant and application of concern (e.g. bacterial models such as in Suslovaite et al., 2024). The methodology is expected to be of most relevance in cases where fully distributed/complex modelling approaches are required to capture the necessary temporal dynamics of the water quality parameter, most notably for water supply applications. Results also demonstrated the complexity of the interaction between rainfall runoff process, water quality dynamics and the objective function used within the optimization problem, which further highlights the requirement to carefully consider an appropriate choice of objective function, required model structure and input variables for the required environmental objective at appropriate scales.

Based on identified priority areas for interventions, utilities may invest in catchment measures such as subsidies to use alternative pesticides, cultural controls or investment in local drainage solutions. The use of such optimization methods is anticipated to enable utilities to obtain the best return for the money invested by replacing blanket mitigation measures with a more targeted catchment intervention approach.

Acknowledgement

This work was funded by an EPSRC CASE award (EP/R513313/1, studentship no. 2282601) together with Severn Trent Water Ltd. We would like to thank the Catchment team at Severn Trent Water for providing the Metaldehyde model and associated data. Special thanks to Helen Pickett and Laura Flower for continued support throughout the project.

References

Aerts, J. C. J. H., Eisinger, E., Heuvelink, G. B. M. and Stewart T. J. (2002) ‘Using Linear Integer Programming for Multi-Site Land-Use Allocation’, Geographical Analysis, 35(2), pp. 148-169.

Arabi, M., Govindaraju, R. S. and Hantush, M. M. (2006) ‘Cost-effective allocation of watershed management practices using a genetic algorithm’, Water Resources Research, 42, pp. 1-14.

Asfaw, A. , Maher, K., Shucksmith, J.D. (2018) ‘Modelling of metaldehyde concentrations in surface waters: A travel time based approach’, Journal of Hydrology, 562, pp. 397-410.

Bodrud-Doza, Md., Yang, W., de Queiroga Miranda, R., Martin, A., DeVries, B., Fraser, E.D.G. (2023) ‘Towards implementing precision conservation practices in agricultural watersheds: A review of the use and prospects of spatial decision support systems and tools’, *Science of The Total Environment*, 905, 167118, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2023.167118.

Brooks, E.S., Saia, S.M., Boll, J., Wetzel, L., Easton, Z.M., Steenhuis T.S. (2015) ‘Assessing Bmp effectiveness and guiding bmp planning using process-based modeling’, *J Am. Water Resour. As.,* 51, pp. 343-358, 10.1111/1752-1688.12296.

Cooke, A., Rettino, J., Flower, L., Filby, K. and Freer, A. (2020) ‘Farming for water; catchment management initiatives for reducing pesticides’, Water and Environment Journal, 34, pp. 679-691. https://doi-org.sheffield.idm.oclc.org/10.1111/wej.12609

Caloiero, T.; Veltri, S.; Caloiero, P.; Frustaci, F. Drought Analysis in Europe and in the Mediterranean Basin Using the Standardized Precipitation Index. Water 2018, 10, 1043. https://doi.org/10.3390/w10081043

Cui, H., Guo, P., Li, M., Guo, S., Zhang, F. (2019) 'A multi-risk assessment framework for agricultural land use optimization', Stochastic Environmental Research and Risk Assessment, 33, pp. 563–579, https://doi.org/10.1007/s00477-018-1610-5

Dai, C., Zhang, X., Tan, X., Hu, M., Sun, W. (2024) 'A stochastic simulation-based chance-constrained programming model for optimizing watershed best management practices for nonpoint source pollution control under uncertainty', *Journal of Hydrology*, 632, 130882, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2024.130882.

Delpla, I., Florea, M., Rodriguez, M.J. (2019) 'Drinking Water Source Monitoring Using Early Warning Systems Based on Data Mining Techniques', *Water Resour Manage*, 33, pp. 129–140, https://doi.org/10.1007/s11269-018-2092-4.

Du, H., Alexander, L. V., Donat, M. G., Lippmann, T., Srivastava, A., Salinger, J., et al. (2019). Precipitation from persistent extremes is increasing in most regions and globally. Geophysical Research Letters, 46, 6041–6049. https://doi.org/10.1029/2019GL081898

Du, J., Xie, H., Hu, Y., Xu, Y., Xu, C-Y. (2009) 'Development and testing of a new storm runoff routing approach based on time variant spatially distributed travel time method' J. Hydrol., 369 (1–2), pp. 44-54.

DWI (2023) 'Drinking Water 2022 – Summary of the Chief Inspector's report for drinking water in England', ISBN 978-1-911087-50-2, https://dwi-content.s3.eu-west-

2.amazonaws.com/wp-content/uploads/2023/07/11131751/E02864254_DWI-Public-water-supplies-in-England-2022_Accessible.pdf.

DWI (2017) Annual Report Drinking Water 2016. London, UK: Drinking Water Inspectorate.

Eulogi M, Ostojin S, Skipworth P, Kroll S, Shucksmith JD, Schellart A. (2022) 'Optimal Positioning of RTC Actuators and SuDS for Sewer Overflow Mitigation in Urban Drainage Systems', *Water*, 14(23):3839. https://doi.org/10.3390/w14233839

Eulogi M, Ostojin S, Skipworth P, Kroll S, Shucksmith JD, Schellart A. (2022) 'Comparing methods to place adaptive local RTC actuators for spill volume reduction from multiple CSOs', *Journal of Hydroinformatics*, 24 (1), pp. 78–92, doi: https://doi.org/10.2166/hydro.2021.085.

European Commission (2024) 'Drinking water', Drinking water - European Commission (europa.eu), Website, Accessed on 08/09/2024.

EU (2006) 'Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC', https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32006L0007.

FWR (2012) 'Urban pollution management manual 3rd edition, Retrieved from http://www.fwr.org/UPM3.

Graydon, R.C., Mezzacapo, M., Boehme, J., Foldy, S., Edge, T.A., Brubacher, J., Chan, H.M., Dellinger, M., Faustman, E.M., Rose, J.B., Takaro; T.K. (2022) 'Associations between extreme precipitation, drinking water, and protozoan acute gastrointestinal illnesses in four North American Great Lakes cities (2009–2014)', J Water Health, 20 (5), pp. 849–862. doi: https://doi.org/10.2166/wh.2022.018.

Herman, M.R., Pouyan Nejadhashemi, A., Daneshvar, F., Abouali, M., Ross, D.M., Woznicki, S.A., Zhang, Z. (2016) 'Optimization of bioenergy crop selection and placement based on a stream health indicator using an evolutionary algorithm', *Journal of Environmental Management*, 181, pp. 413-424, ISSN 0301-4797, https://doi.org/10.1016/j.jenvman.2016.07.005.

Jeong, D.S.; Kim, J.H.; Kim, J.H.; Park, Y. (2024) 'Evaluation of the Best Management Practices for Reducing Phosphorus Load in a Watershed in Terms of Cost and Greenhouse Gas Emissions', *Agronomy*, 14, 906. https://doi.org/10.3390/agronomy14050906.

Kaim, A., Cord, A.F., Volk, M. (2018) 'A review of multi-criteria optimization techniques for agricultural land use allocation', *Environmental Modelling & Software*, 105, pp. 79-93, ISSN 1364-8152, https://doi.org/10.1016/j.envsoft.2018.03.031.

Kay, P. & Grayson, R. (2014) 'Using water industry data to assess the metaldehyde pollution problem', Water and Environment Journal, 28(3), pp.410–417.

Li, C., Wu Y.L., Yang, T., Zhang, Y. (2010) 'Determination of Metaldehyde in Water by SPE and UPLC – MS – MS', Chromatographia, 72(9-10), pp.987–991.

Li, M., Li, J., Singhd, V.P., Fua, Q., Liu, D., Yange, G. (2019) 'Efficient allocation of agricultural land and water resources for soil environment protection using a mixed optimization-simulation approach under uncertainty', Geoderma, 353, pp. 55-69, https://doi.org/10.1016/j.geoderma.2019.06.023.

Liu, G., Chen, L., Wei, G., Shen, Z. (2019) 'New framework for optimizing best management practices at multiple scales', *J. Hydrol.*, 578, pp. 124-133, 10.1016/j.jhydrol.2019.124133.

Mancini, M., Rosso, R. (1989) 'Using GIS to assess spatial variability of SCS Curve Number at the basin scale. In In New directions for surface water modelling', Proceedines of the Baltimore Symposium. IAHS, 181, pp. 435–444.

Melesse, A.M., Graham, W.D. (2004) 'Storm runoff prediction based on a spatially distributed travel time method utilizing remote sensing and GIS',. J. Am. Water Resour. Assoc. 40 (4), 863–879.

Met Office (2003) '1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System', NCAS British Atmospheric Data Centre. Available at: http:// catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350 (Accessed 11.07.2019).

Mounce, S. R., Shepherd, W. , Ostojin, S. , Abdel-Aal, M. , Schellart, A. N. A. , Shucksmith, J. D. , Tait, S. J. (2020) 'Optimisation of a fuzzy logic-based local real-time control system for mitigation of sewer flooding using genetic algorithms', *Journal of Hydroinformatics*, 22 (2), pp. 281–295. doi: https://doi.org/10.2166/hydro.2019.058

Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A. and Zechman, E. (2010) 'State of the art for genetic algorithms and beyond in water resources planning and management. Journal of Water Resources Planning and Management', 136(4), pp.412-432.

Ortiz-Lopez, C., Bouchard, C., & Rodriguez, M. (2022) 'Machine learning models with potential application to predict source water quality for treatment purposes: a critical review', *Environmental Technology Reviews*, 11(1), pp. 118–147, https://doi.org/10.1080/21622515.2022.2118084.

Perez-Pendini, C., Limbrunner, J. F. an Vogel, R.M. (2005) 'Optimal Location of Infiltration-Based Bst Management Practices for Storm Water Management', Journal of Water Resources Planning and Management, 131(6), pp 441-448.

Sadeghi, S.H.R., Jalili, Kh., Nikkami, D. (2009) 'Land use optimization in watershed scale, Land Use Policy', 26(2), pp. 186-193. https://doi.org/10.1016/j.landusepol.2008.02.007.

Schramm M, Kikoyo D, Wright J, Jain S (2024) 'A meta-analysis of the impacts of best management practices on nonpoint source pollutant concentration', *Front. Water,* 6,1397615, doi: 10.3389/frwa.2024.1397615

Severn Trent Water (2024) 'Farm to Tap', Website, https://www.stwater.co.uk/about-us/environment/working-with-farmers-and-land-managers/farming-for-water/farm-to-tap/#:~:text=Farm%20to%20Tap%20is%20currently,our%20Farming%20for%20Water%20team. (Accessed on 07/10/2024).

Speight, V. L., Kalsbeek, W. D., DiGiano, F.A. (2004) 'Randomized Stratified Sampling Methodology for Water Quality in Distribution Systems', Journal of Water Resources Planning and Management, 130(4), pp. 330-338.

Srivastava, P., Hamlett, J. M., Robillard, P.D. and Day, R.L. (2002) 'Watershed Optimization of best management practices using AnnAGNPS and a genetic algorithm', Water Resources Research, 38(3), pp. 3.1-3.14.

Srivastava, P., Hamlett, J. M., and Robillard, P.D. (2003) 'Watershed Optimization of agricultural best management practices: Continuous simulation versus design storms', Journal of American Water Resources Association, 39(5), pp. 1043-1054. https://doi.org/10.1111/j.1752-1688.2003.tb03691.x

Suslovaite, Vaida (2023) 'Development of a risk based approach to surface water abstraction', PhD thesis, University of Sheffield. https://etheses.whiterose.ac.uk/34483/.

Swinamer R., Anderson L.E., Redden D., Bjorndahl P., Campbell J., Krkošek W.H., Gagnon G.A. Environ. Sci. Technol. 2024, 58, 27, 11958–11969.

Tang, K.S., Man, K.F., Kwong, S., He, Q. (1996) 'Genetic Algorithms and their Applications', IEEE Signal Processing Magazine, 13(6), pp. 22-37. DOI:10.1109/79.543973

Tröger, R., Ren, H., Yin, D., Postigo, C., Nguyen, P.D., Baduel, C., Golovko, O., Been, F., Joerss, H, Boleda, M.R., Polesello, S., Roncoroni, M., Taniyasu, S., Menger, F., Ahrens, L., Lai, F., Wiberg, K. (2021) 'What's in the water? – Target and suspect screening of contaminants of emerging concern in raw water and drinking water from Europe and Asia', Water Research, 198, 117099.

United Nations, Department of Economic and Social Affairs, Population Division (2019) 'World Population Prospects 2019: Highlights', ST/ESA/SER.A/423.

Wang, M., Wang, Y., Gao, X. and Sweetapple, C. (2019) 'Combination and placement of sustainable drainage system devices based on zero-one integer programming and schemes sampling', Journal of Environmental Management, 238, pp. 59-63.

Zhang, P., Liu, Y., Pan, Y., Yu, Z. (2011) ‘Land use pattern optimization based on CLUE-S and SWAT models for agricultural non-point source pollution control’, Mathematical and Computer Modelling, 58, pp. 588–595.