Science of the Total Environment 891 (2023) 164226

Contents lists available at ScienceDirect

Science of the Total Environment

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

Methods comparison for detecting trends in herbicide monitoring time-series in streams

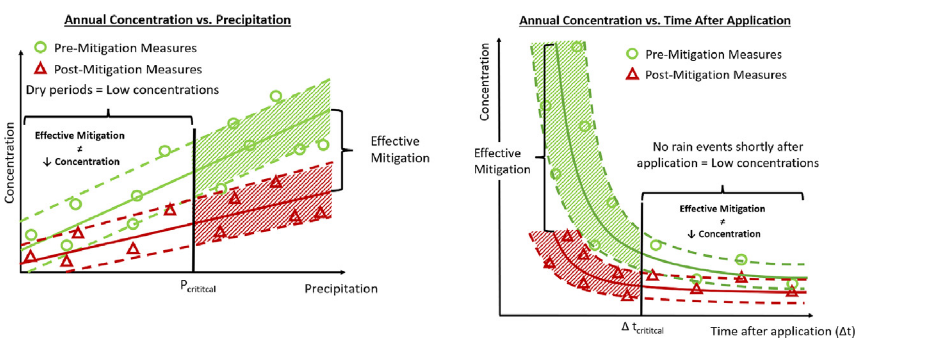
R. Chow [a,b,c,*], S. Spycher [d,e], R. Scheidegger [a,f], T. Doppler [f], A. Dietzel [f], F. Fenicia [a], C. Stamm [a]

[a] Swiss Federal Institute of Aquatic Science and Technology (Eawag), 8600 Dübendorf, Switzerland
[b] Department of Earth Sciences, Stellenbosch University, Stellenbosch, South Africa
[c] Soil Physics and Land Management Group, Wageningen University & Research, P.O. Box 47, 6700 AA Wageningen, the Netherlands
[d] EBP Schweiz AG, 8032 Zürich, Switzerland
[e] School of Agricultural, Forest and Food Sciences BFH-HAFL, 3052 Zollikofen, Switzerland
[f] VSA, Swiss Water Association, 8152 Glattbrugg, Switzerland

HIGHLIGHTS

- Variability of pesticide losses makes it hard to detect changes in water quality.
- Pesticide use data helps account for part of the temporal trends in water quality.
- Hydrological events can obscure the effects of mitigation measures.
- A strong reduction is needed to detect a change within 10 years of monitoring data.
- Sensitive methods for change detection are more prone to false-positives.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Jay Gan

Keywords:
Water quality
Agriculture
Mitigation
Pesticides
Trend detection

ABSTRACT

An inadvertent consequence of pesticide use is aquatic pesticide pollution, which has prompted the implementation of mitigation measures in many countries. Water quality monitoring programs are an important tool to evaluate the efficacy of these mitigation measures. However, large interannual variability of pesticide losses makes it challenging to detect significant improvements in water quality and to attribute these improvements to the application of specific mitigation measures. Thus, there is a gap in the literature that informs researchers and authorities regarding the number of years of aquatic pesticide monitoring or the effect size (e.g., loss reduction) that is required to detect significant trends in water quality. Our research addresses this issue by combining two exceptional empirical data sets with modelling to explore the relationships between the achieved pesticide reduction levels due to mitigation measures and the length of the observation period for establishing statistically significant trends. Our study includes both a large (Rhine at Basel, ~36,300 km^2) and small catchment (Eschibach, 1.2 km^2), which represent spatial scales at either end of the spectrum that would be realistic for monitoring programs designed to assess water quality. Our results highlight several requirements in a monitoring program to allow for trend detection. Firstly, sufficient baseline monitoring is required before implementing mitigation measures. Secondly, the availability of pesticide use data helps account for the interannual variability and temporal trends, but such data are usually lacking. Finally, the timing and magnitude of hydrological events relative to pesticide application can obscure the observable effects of mitigation measures (especially in small catchments). Our results indicate that a strong reduction (i.e., 70–90 %) is needed to detect a change within 10 years of monitoring data. The trade-off in applying a more sensitive method for change

* Corresponding author at: Swiss Federal Institute of Aquatic Science and Technology (Eawag), 8600 Dübendorf, Switzerland.
E-mail address: reynold.chow@wur.nl (R. Chow).

http://dx.doi.org/10.1016/j.scitotenv.2023.164226
Received 10 March 2023; Received in revised form 10 May 2023; Accepted 13 May 2023
Available online 24 May 2023
0048-9697/

detection is that it may be more prone to false-positives. Our results suggest that it is important to consider the trade-off between the sensitivity of trend detection and the risk of false positives when selecting an appropriate method and that applying more than one method can provide more confidence in trend detection.

1. Introduction

Agricultural pesticides are used globally to improve the efficiency of crop production to feed the growing global population (Rosling et al., 2018; Zhang, 2018). In the last decade, annual global pesticide use was estimated to be approximately 5.7 million tonnes per year (FAOSTAT, 2020). An inadvertent consequence of pesticide use is aquatic pesticide pollution, which can be detrimental to human health and the healthy functioning of aquatic ecosystems (Schäfer et al., 2012; Beketov et al., 2013; Chow et al., 2020; Fuhrimann et al., 2021). Such negative effects have triggered mitigation measures in many countries (Reichenberger et al., 2007). Examples are the national pesticide risk reduction plans across the European Union (PANE, 2013). In Switzerland a National Action Plan was developed with aims to reduce the stream length where Environmental Quality Standards are exceeded by 50 % within 10 years (Swiss Federal Council, 2017). Accordingly, there is a need to evaluate the outcome of these mitigation activities. Water quality monitoring programs are one tool for such an evaluation process (e.g., Boye et al., 2019).

While many of the long-term monitoring programs have been and still are essential to report on the status of water bodies and to identify existing water quality problems (e.g., Spycher et al., 2018), it is less clear to which degree existing monitoring programs and the subsequent statistical analyses fulfill the needs for robust trend analysis for evaluating the outcome of mitigation measures (Lloyd et al., 2014).

One key problem is the question how to detect effects of mitigation actions against the high level of inter-annual variability in pesticide losses and in-stream concentrations. Such inter-annual fluctuations have been shown to correspond to seasonal pesticide use patterns and hydrological conditions (Gilliom, 2001; Lerch et al., 2011), which can lead to the misinterpretation of effects (or lack thereof) caused by implemented mitigation measures (Vecchia et al., 2009). It was observed that discharge during the respective application period explains a substantial part of the inter-annual variability for a given catchment (Leu et al., 2010).

Chow et al. (2020) reviewed long-term pesticide monitoring studies and found that trends were only detected for large effects sizes, such as those caused by restricting pesticide use by banning a given compound. Mitigation measures, other than restricting pesticide use, rarely led to detectable long-term reductions in aquatic pesticide pollution. In the few cases where mitigation measures (e.g., buffer strips, biobeds, spray drift reduction) led to a detectable long-term reduction (Kreuger and Nilsson, 2001; Hermosin et al., 2013; Daouk et al., 2019; Budd et al., 2020), a decrease of >45 % was required to attribute the effect to the implemented mitigation measures.

There exists a long tradition in hydrology to analyse temporal trends in catchment responses to external forcings (Wolman, 1971). Accordingly, different methods have been developed and proposed. For example, the double-mass curve approach (Searcy and Hardison, 1960) used the slope of cumulative discharge against cumulative precipitation to detect long-term trends in discharge behaviour. Others have looked at the changes in concentration-discharge relationships (Choquette et al., 2019; Jarvie et al., 2017; Zhang et al., 2016) to separate between trends due to climatic forcings and land use effects.

Applying such approaches to assess pesticide trends must consider some specifics of pesticide use, transport, as well as the method of analysis. There is a pronounced seasonality of agricultural pesticide use, which additionally is very compound and crop-specific. Combined with the fact that the extent of pesticide losses depends strongly on the actual timing between application and precipitation events, traditional grab sampling may only poorly reflect the full extent of pesticide pollution. This introduces additional uncertainty that may limit the possibility to detect temporal trends. For dealing with this problem, the United States Geological Survey (USGS) has developed a statistical regression model (SEAWAVE-Q) to separate the seasonal use patterns and hydrological conditions from the underlying time trend in pesticide concentration time-series (Vecchia et al., 2009).

However, the considerable number of agricultural pesticides used can lead to relevant changes of the use of single compounds, e.g., due to substitution, and a decrease in concentration levels cannot be simply interpreted as improvement due to better practices. A useful way to avoid this problem is to express pesticide losses a loss rates by normalising observed loads in rivers by the applied pesticide use the catchment. While this approach has been used already for decades (e.g., Larson et al., 1995), its application is still hampered in many countries due to the lack of reliable pesticide application data.

These challenges are well known (e.g., Chow et al., 2020) and recommendations are provided in the literature regarding the best way to establish monitoring programs. However, there is little quantitative data available in the literature that informs researchers and authorities regarding the number of years of pesticide monitoring required to reliably detect significant trends. This paper addresses this issue by combining two exceptional empirical data sets with modelling to explore the relationships between the achieved pesticide reduction levels due to mitigation measures and the length of the observation period for establishing statistically significant trends. Firstly, we analyse a unique set of long-term (since 1995) herbicide time-series in River Rhine with daily resolution. This data set allows us to evaluate the degree to which seasonal discharge during the application period can explain compound-specific loss rates and how strong mitigation effects need to be, such that they can be detected against the background of unexplained inter-annual variability.

Because the Rhine data set reflects the response of a large basin ($\sim$36,300 km^2), questions remain regarding the relevance of the results for monitoring sites at smaller streams. For smaller catchments, it can be expected that single precipitation events play a more important role such that seasonal metrics such as total discharge volumes have more limited explanatory power. To explore this aspect, we combine data from an intensive (i.e., high temporal resolution sampling during rain events), controlled herbicide experiment at the scale of a small catchment (Eschibach 1.2 km^2; Doppler et al., 2012) with a herbicide transport model developed for that area (Ammann et al., 2020) to investigate the long-term herbicide dynamics given the weather dynamics observed over one decade. By simulating the long-term herbicide loss dynamics with different levels of mitigation, we quantify the possibility to distinguish between natural variability due to timing between application and rainfall events and mitigation effects. Overall, the results at both scales provide insights with a dual benefit. First, they shall help to develop realistic expectations regarding monitoring programs. Second, they shall provide insights that help develop mitigation programs that are ambitious enough to cause improvements that are large enough so they can be observed against natural variability.

Finally, our study helps set realistic expectations for national water quality monitoring programs. For instance, the Swiss National Surface Water Quality Monitoring Program (NAWA TREND) consists of multiple monitoring sites draining watersheds (>21) ranging from 2.0 km^2 to $\sim$28,000 km^2 (Fabre et al., 2023). Thus, the scales of the catchments presented in our study represent scales at both ends of the spectrum. Naturally, it would be desirable to also study catchments of intermediate size, but apart from a few scientific studies with controlled applications (i.e., Eschibach; Doppler et al., 2014), most countries currently lack data on pesticide use at the small and intermediate catchment scale. Studies may be possible after 2025 when the planned georeferenced collection of all applications will be available (FOAG, 2023).

2. Methods

2.1. Empirical data sets

The Rhine River at Basel drains a basin of about 36,300 km^2 (Fig. 1), has a mean altitude of 1333 masl and an average air temperature of 4.8 °C. Precipitation generally increases with altitude and ranges between 1350 mm/yr in the lower reaches and 1930 mm/yr in the alpine regions (Uehlinger et al., 2009). Land use is mixed and includes considerable areas of arable cropping in the Swiss Plateau and the German area east of Lake Constance.

Compound selection: Since 1993, water quality at Basel has been monitored daily for an increasingly large set of compounds (Ruff et al., 2013). The introduction of high-resolution mass spectroscopy (HRMS) in 2012 drastically expanded the number of quantified compounds, which is currently >600. During this initial year, the limits of quantification also improved, which led to a consistent data set starting from 2013 that we use for our analysis (Figs. S1–S3). All concentrations evaluated in this study were determined with daily composite samples consisting of subsamples taken every 6 min at 5 locations across the river cross-section. The data is publicly available from the International Commission for the Protection of the Rhine (IKSR, 2020). Out of the 149 pesticides (76 herbicides, 35 fungicides, 26 insecticides, and 12 biocides) and 55 metabolites with quantified daily composite samples, we selected agricultural pesticides suitable for quantitative evaluation according to the following criteria: 1. availability of sales or usage data, 2. sufficiently high detection frequency, and 3. main usage for agricultural plant protection products (PPP).

The first criterion led to the exclusion of the herbicide atrazine, which has not been sold in Switzerland since 2008 (FOAG, 2021). The sale of the herbicide isoproturon was discontinued in 2020 but was included since sales stopped recently relative to this study.

The second criterion requires detection frequencies (DF) above the limit of quantification (LOQ) of >20 %, which is a recommended threshold for robust statistical analyses (Helsel, 2012). Accordingly, we considered that threshold but also examined compounds with DF between 5 and 20 % to avoid missing interesting candidate pesticides.

The third criterion was the most difficult because PPP compounds may also be authorized for use as biocide products or veterinary drugs. For such usage, no sales data were publicly available. Therefore, for these ambiguous compounds we used additional data to check the plausibility that agricultural use was the dominant source for the Rhine River. First, we only considered compounds for which the seasonal load distribution corresponded to the main application period of the respective active ingredient, e.g., for herbicides applied in spring, the highest loads are also observed in spring or early summer. For biocide usage, inputs show a different temporal distribution over the year (Wittmer et al., 2011). Second, the calculated loss rates (see below) should lie in the range of previous studies, i.e., low single-digits (cf. Burgoa and Wauchope, 1995). Some compounds that were used as PPP and biocide (e.g., carbendazim) had unrealistically large loss rates (>30 % of agricultural sales) suggesting a non-agricultural source. After excluding such critical compounds, we ended up with a set of seven herbicides that could be used for analysis (Table S1). Two of the seven (MCPA and Mecoprop) are also authorized for amenity and homeowners use for both professionals and non-professional users. The other five compounds are only registered for usage in crop production. Data on the share of amenity and homeowner use during the study period are not available. For mecoprop, a baseload from building materials can be expected with about the same load throughout the year stemming from this use (Wittmer et al., 2011).

Use data: Since Switzerland does not have regional data on pesticide use available, the national sales data (FOAG, 2021) of the corresponding year was taken as a proxy for the pesticide use in the catchment. Other

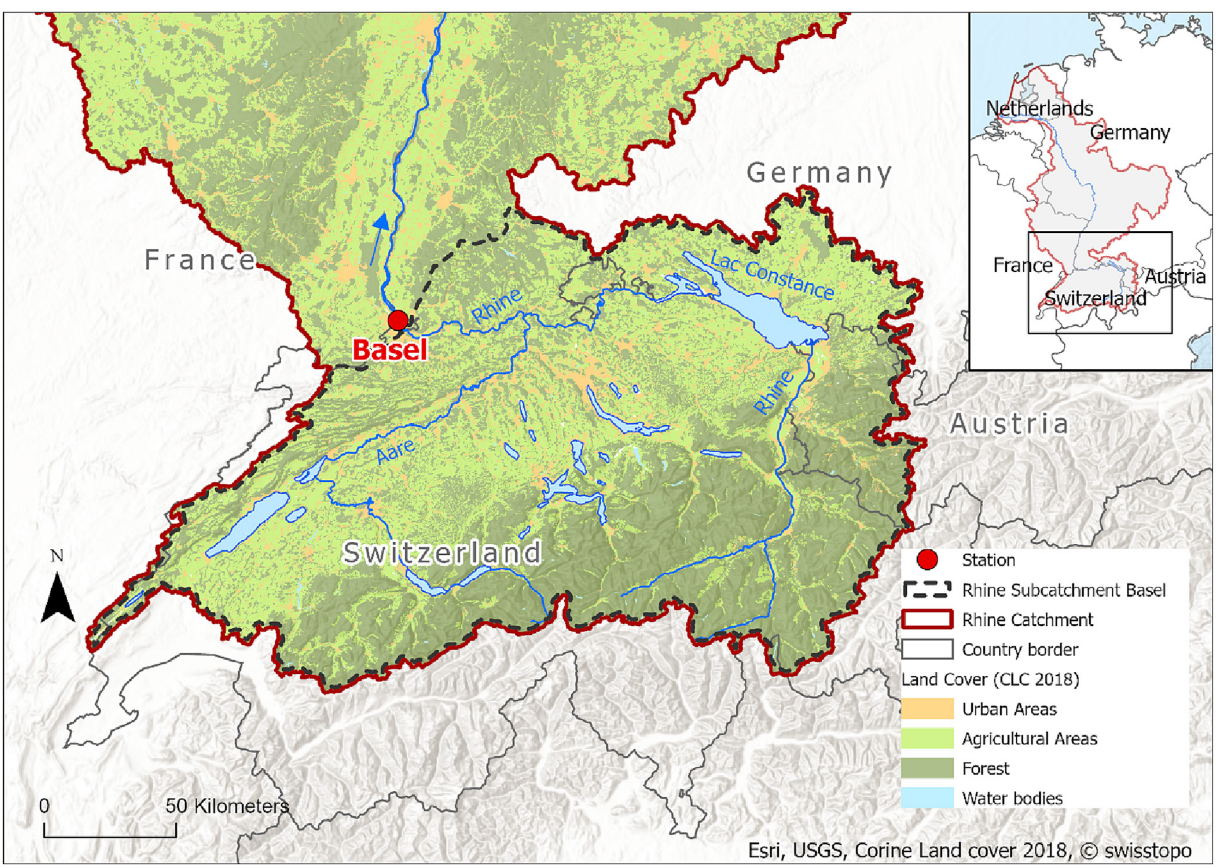


Fig. 1. The Rhine River sampling station at Basel with outlines for the Rhine River Catchment and subcatchment at Basel. Also, shown are the land cover types within the Rhine River Catchment.

researchers in Europe, the US, and South Africa, have also used PPP sales as a proxy for use when use data are not readily available (Schreder and Dickey, 2005; Dabrowski, 2015; Galimberti et al., 2020). The catchment area of the Rhine at Weil am Rhein covers about 68 % of the Swiss territory and includes almost 80 % of the total population and 87 % of the arable land (Swiss Federal Statistical Office FSO, 2018). For this reason, it seems justified to compare the national PPP sales volumes as an estimate for use in the Swiss part of the Rhine catchment with the load of the Rhine catchment area, but it should be kept in mind that the estimates tend to overestimate the true use in the Swiss part of the Rhine catchment (by a factor of ~1.15), an affect more or less compensated by additional load from tributaries of the Rhine in southern Germany (Moser et al., 2018). We acknowledge this assumption may result in biased loss calculations. However, our statistical analysis addresses the relative variation in loss rates between years, which is less susceptible to inaccuracies in the absolute loss rates themself.

Discharge data: Discharge at the sampling location in Basel is measured every 5 min by the canton of Basel-Stadt and is published by the Federal Office of the Environment (FOEN, 2022). The mean daily discharge was provided with the daily composite sample concentration data by the cantonal laboratory.

Load calculations: Yearly loads were calculated as the product of the daily composite sample concentration (C_i) and the mean discharge on day i (Q_i) summed over all days of the corresponding calendar year (Eq. (1))

$$L = \sum_{i=1}^{n} C_i \cdot Q_i \tag{1}$$

Concentrations below the LOQ were set to zero, which is a simplification as the real value is likely to be between zero and the LOQ. However, tests of substituting these values with LOQ/2 showed only minor differences for all compounds with high DF, e.g., 0.5 % and 2 % higher mean load for mecoprop and metolachlor, respectively. For dimethenamid and MCPA, two compounds with mean DF < 50 %, the effect was more pronounced with 30 % and 36 % higher loads. However, given the high interannual variability of the loads of up to a factor of 5, the approach for dealing with concentrations below 0 adds very little to the total variability.

Calculation of loss rates: The loss rates (LR) for a given year k were calculated by dividing the load L_k in Basel by the sales data for Switzerland (Eq. (2)).

$$LR_k = \frac{L_k}{Sale_k} \tag{2}$$

Pesticides are not necessarily used in the year they are sold, but no compound-specific information was found about the tendency to use stocks from previous reporting years. Therefore, the sales of all compounds were weighted by taking 2/3 of the sales data from the current year and 1/3 from the previous year thereby dampening fluctuations in the sales data. Two approaches were used to calculate the loss rates: 1. calculates the load for all months of each evaluated year, and 2. calculates the load only for the months making up the main loss period (MLP). The MLP is often a very distinct period of 2–4 months in the spring or early summer that makes up the largest fraction of the annual load. E.g., for dimethenamid (a herbicide applied only to maize), the MLP is from May–July (cf. Fig. S3). For chlorotoluron and isoproturon, two herbicides applied in winter crops, the MLP ranges from October to February and from October to May, respectively. Although most farmers do not apply pesticides between November 1st and February 15th, the losses can be elevated over several months. Thus, the term main loss period is distinctly different from the application period. The prolonged period with increased load of isoproturon is due to it being first applied in the fall and then again in early spring. The load of the MLP of these two compounds was compared to the annual sales data of the year the MLP starts. As the sales data for the year 2021 were not available at the time of this study, loss rates could only be calculated from 2013 to 2020.

Eschibach catchment is a small (1.2 km^2) agricultural catchment located in northeastern Switzerland (Fig. 2). The main crops produced there are corn, sugar beet, winter wheat, and rape seed. It was the site of an experimental study, which consisted of a controlled application of two herbicide mixtures on May 19, 2009 (Doppler et al., 2012). The first mixture (containing atrazine, CAS No.: 1912-24-9) was applied to six experimental corn fields where Doppler et al. (2012) had full control over the application. Thus, spraying of the first mixture took place all on the same day with the same spraying method. The second mixture (containing terbuthylazine, CAS No.: 5915-41-3), which was applied to the rest of the corn fields in the catchment, could not be sprayed all on the same day nor with the same spraying method. Hydrological monitoring (i.e., stream discharge, precipitation, groundwater levels, and soil moisture) took place within the catchment from summer 2008 to autumn 2009. Water samples for pesticide analysis were taken from the stream and tile drains at five discharge measurement stations prior to the herbicide application and during a two-month period after application. During the two-month period, 13 rain events occurred, which triggered automatic high-resolution water sampling. This consisted of taking a water sample every 15 min for the first six hours after the start of an event, followed by hourly sampling until returning to baseflow discharge levels. Grab samples were taken periodically during baseflow periods. In total 1500 samples were taken, of which 600 were selected for chemical analysis to represent the chemograph dynamics based on seven rainfall events that occurred during the experiment (Doppler et al., 2012). Pesticide concentrations were measured with online solid-phase extraction (SPE) coupled to liquid chromatography and a triple quadrupole mass spectrometer (LC-MS/MS). The limit of detection (LOD) for all compounds was between 2 and 10 ng/L. For all compounds isotope-labelled internal standards were used for quantification. For further details see Doppler et al. (2012). This data set formed the basis for the development of a perceptual and conceptual hydrological pesticide transport model (Ammann et al., 2020, see below).

2.2. Model description and setup

A conceptual hydrological model of the Eschibach catchment was developed to simulate fast herbicide transport (Ammann et al., 2020). The Eschibach model is capable of simulating discharge and pesticide concentration time-series at the outlet of the catchment given an equal length time-series for precipitation, evapotranspiration, and the application of pesticides. The model accounts for degradation and sorption of pesticides (Table S2 for physico-chemical characteristics), as well as the spatial distribution of hydrologic properties by using hydrological response units (HRU) (Doppler et al., 2014). There are four main HRUs represented in the Eschibach model, listed in order of increasing travel times:

1. Impervious areas, such as roads (both paved and unpaved) and driveways.
2. Overland flow areas connected directly to the stream or via artificial shortcuts, such as maintenance manholes of the tile drainage system or roadside storm drains.
3. Fast vertical infiltration through macropores in the upper soil layers that reach tile drains.
4. Slower moving saturated groundwater flow.

The Eschibach model was calibrated to stream discharge and high-resolution concentration measure-ments of atrazine and terbuthylazine described in the previous section. Compared to the original model application, we adjusted the time-dependent parameters in the Eschibach model from a 15-minute time-step to a 10-minute time-step to align with the maximum time resolution obtainable from MeteoSwiss (2019) for meteorological data. We also added an additional interception bucket to the HRU representing impervious areas to overcome unrealistic concentration peaks that would otherwise occasionally occur simultaneously with the pesticide application.

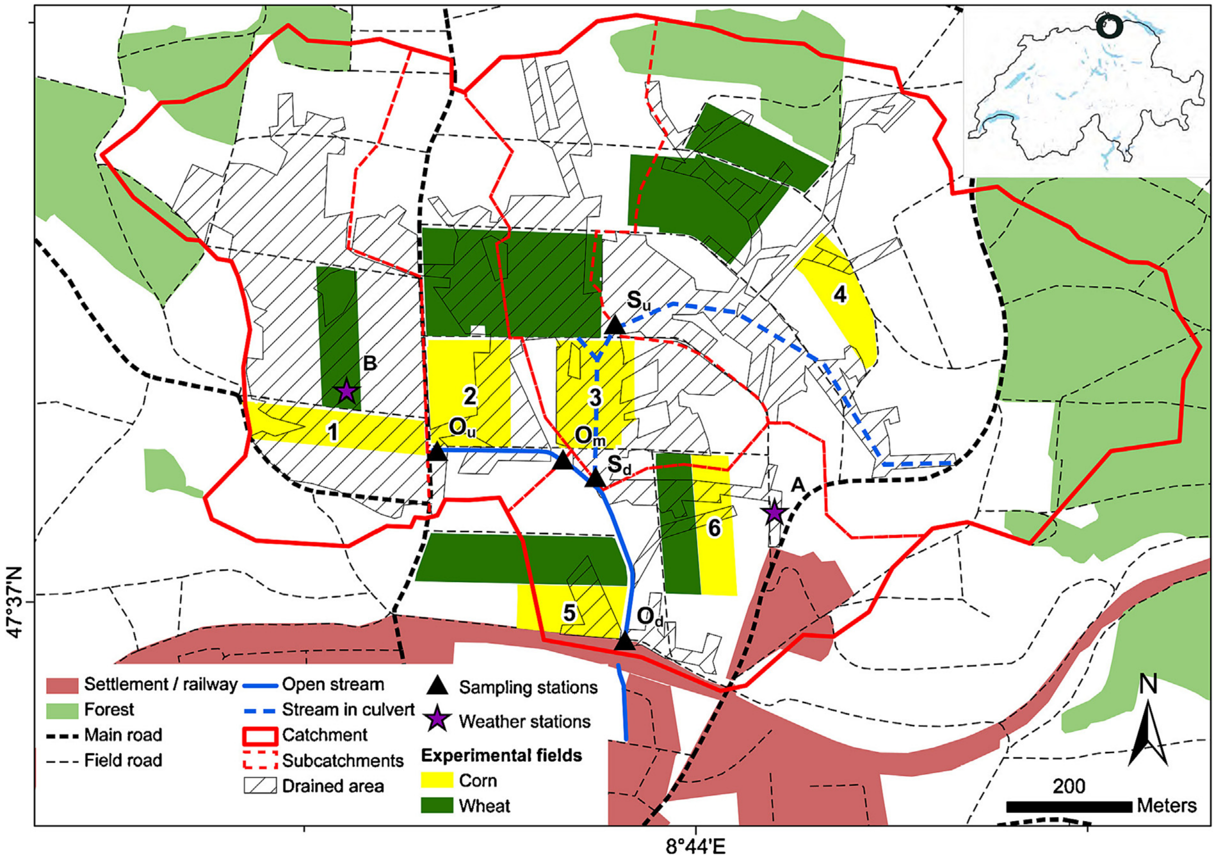


Fig. 2. The Eschibach catchment with the experimental setup. Five sampling stations are shown with their corresponding subcatchments (Su and Sd: subsurface upstream and downstream; Ou, Om, and Od: open upstream, middle, and downstream). The weather stations A and B, land use and drained areas are shown. Also, shown are the experimental wheat and corn fields where two controlled herbicide applications took place in the spring of 2009 (numbered 1–6). Inset map shows location within Switzerland (from Doppler et al., 2014).

We used the Eschibach model to simulate herbicide concentrations at a 10-minute time resolution from 2008 to 2018 using meteorological forcings (i.e., temperature, ET, precipitation) from a meteorological station located roughly 11 km north of Eschibach in Schaffhausen (MeteoSwiss, 2019). The quantity of herbicides applied each year in the model was the same amount applied in the 2009 experiment for the first mixture (Doppler et al., 2012). The timing of pesticide application occurred on a single day each year based on a growing degree-day of 5.9 °C for corn (Neild and Seeley, 1977), which represents conditions similar to those on the day of the controlled pesticide application in the 2009 experiment (Doppler et al., 2012). Additionally, the date of pesticide application was chosen based on the prevailing hydrological conditions (i.e., herbicide was not applied on a rainy day). We use this scenario as a control, representing the inter-annual variability in herbicide concentrations purely caused by the inter-annual variability in weather conditions.

2.3. Simulation of mitigation effects

For both study areas, we represented the effect of mitigation measures by reducing either the observed loss rates (The Rhine at Basel) or the amount of herbicide applied (Eschibach catchment). Next, we assumed the monitoring period to consist firstly of a control period without mitigation measures implemented and secondly a mitigation period of the same duration.

The Rhine River at Basel: For each of the herbicides considered, the effects of the mitigation measures were expressed as the Reduction Level (RL) in percent of the control period. Accordingly, we calculated for each year of observed loss rates (LR_{obs}) the corresponding loss rates for the different RLs as follows (Eq. (3)):

$$LR_{RL_i} = \frac{100 - RL_i}{100} LR_{obs} \tag{3}$$

We varied the RLs between 0 and 90 % in steps of 5 %.

Eschibach catchment: Mitigation effects were simulated by reducing the herbicide application rate corresponding to the RLs-values described above. To evaluate the effects of application timing on our ability for trend detection, we also simulate the same reduction levels with application split across a 1-month period (−10 days and + 20 days from the date of single-day application) on days without rainfall. To reflect the sampling scheme currently implemented in the Swiss NAWA monitoring program (Kunz et al., 2016) we converted the simulated 10-minute concentration time-series to 14-day average composite values (see S2.2 in Supporting Information for simulated concentration time-series).

2.4. Statistical approaches

The Rhine River at Basel: To investigate the impact of the duration of the observation period, we combined the original and the reduced LRs of variables length between 6 and 16 years. For each duration, the control and the mitigation period had the same duration (between 3 and 8 years each). The 8 years (2013 – 2020) of monitoring data served as an empirical control period since it did not contain any significant trends. Through random sampling of individual years, we generated sets of shorter control periods as well. By reducing the loss rates as described above, we also derived sets

Table 1
Overview of statistical approaches.

Method	Data used		Study area
	Concentration	Discharge	
t-Test	All data	None	Eschibach, Rhine
Seasonal Mann-Kendall	All data	None	Eschibach
C-Q relationships	All data	All data	Eschibach, Rhine
Double-mass curve	Seasonal data	Seasonal data	Eschibach

of mitigation periods. We fully exploited the combinatorial possibilities and analysed 12,021 sets of control-mitigation years (3136 combinations of six and 10 years, 4900 of eight years, 784 of 12, 64 of 14, and one of 16 years).

For the assessment of mitigation effects in the Rhine basin, we started from the observation that loss rates LR_{obs} of each herbicide were well correlated with the discharge volume Q_{MLP} during the respective main loss period (see Result section below). Mitigation measures reducing the loss rates will cause the Q_{MLP} - LR_{obs} relationship to decrease. We tested whether the observed slope was significant between the control and the mitigation period (Table 1). To that end, we used a linear regression model for the loss rate of each compound with the average discharge during MLP (Q_{MLP}) and the indicator variable "period" (control or mitigation) as an indicator variable. We tested whether there was a significant interaction between these two explanatory variables implying the slopes to differ between the two periods. We tested this interaction for all year combinations and all reduction levels yielding a distribution of p-values over 64 (for a duration 14 years) up to 4900 year-combinations (for a duration of eight years). In the result section, we mainly refer to the 10 % quantile and the median of these distributions. An exception was of course the 16-year period where we rely on the single, complete set of years for the control period and the respective set for the mitigation period.

Eschibach catchment: The simulated atrazine time-series for the control and treatment scenarios were analysed for statistically significant trends or changes using a variety of statistical approaches (Table 1), which include the *t*-test (Gosset, 1908), Seasonal Mann-Kendall test (Hirsch et al., 1982), the Welch's *t*-test (Welch, 1947), and the double-mass curve analysis (Searcy et al., 1960).

We also tested relationships between the yearly maximum concentration, C_{max}, in a composite sample and the respective rainfall sum P_{sum} for the period of the respective sample. The C_{max}-P_{sum} slope difference was tested by the Welch's t-test implemented in R. We also cumulated these annual maximum concentrations and the respective rainfall amounts as double-mass curves. Such curves have been used in hydrology for decades (e.g., Searcy & Hardison, 1960) to check whether the relationship between two variables change over time. Here, we tested for a change in the slope of the double-mass curve and how probable the cumulated C_{max} values at the end of the treatment period were under assumption that the C_{max}-P_{sum} slope of the control period prevailed during the treatment period.

The slope difference was tested based on Monte Carlo simulations. To formulate a proper null hypothesis, we randomly sample from the C_{max}-P_{sum} slope correlation (n = 10,000) for the precipitation values of the treatment period. For each realisation, we calculated the resulting slope and the cumulative C_{max} sum after the treatment period under the null hypothesis (H_o). This resulted in a distribution for the C_{max}-P_{sum} slopes and the cumulative C_{max} sum. The slope and concentration sum for the treatment were then compared with the respective distributions based on H_o resulting in an empirical exceedance probability.

To account for the effect of the specific sequence of years attributed to the control or treatment period, we applied the procedure described in the previous paragraph to all combinations of years (252 combinations for 10 years split into 5 years of control and treatment, respectively).

3. Results

3.1. Herbicide dynamics in the Rhine River at Basel

The loss rates calculated from the observed loads and weighted sales data ranged on average between 0.33 % and 1.6 % of the amounts applied (Table 2). They revealed considerable interannual variability with the minimum and maximum loss rates for each compound differing by factors between 2.0 and 12.8 (median = 5.3) between years. For the seven study compounds, average discharge during the main loss period (MLP) explained between 60 and 96 % of the variance (Table 2, Figs. 3, S4–S6).

If mitigation measures are effective in reducing loss rates, the slope of the observed loss rate for each compound would decrease. We tested the necessary length of control-mitigation observation period and strength of reduction level to observe a significant change in the slope of the Q_{MLP} - LR_{obs} relationship ($p < 0.05$). Despite decent relationships for most compounds, the analysis revealed that very effective or very long time-series are needed to detect significant changes against the observed, unexplained variance in the system. With 14 years of observations (seven years of control and mitigation, respectively), for four compounds a 50 % reduction level will only lead to a significant change in 10 % of all year combinations (see Method section for details). Only three compounds are less demanding (required reduction between 17 and 42 %). If one expects 50 % of the year combinations to cause a significant change with 14 years of observations, reduction levels must exceed 50 % except for one compound (metolachlor).

The possibility to detect changes depends on how well discharge describes the loss rate (Fig. 4). The smaller the unexplained variability the smaller the necessary reduction level or duration of the observation period (i.e., metolachlor has an R^2 = 0.96 and with 14 years of observations requires a 25 % reduction level to observe a significant change, while chlortoluron with an R^2 = 0.60 and the same observation period requires a 85 % reduction level). Roughly, 1 % of explained variance reduced the required reduction level by 1 % as well.

However, even with a good model, the required reduction level increases strongly if the observation period gets short. While a 22.5 % reduction level was sufficient for metolachlor based on 16 years of observation, with only six years (three years of control and mitigation, respectively) of observation, a reduction level of almost 70 % is required to achieve a median p-value <0.05 (Fig. 5).

Table 2
Regression models between loss rates and discharge volumes between application periods. Mean LR: average loss rate over the eight-year monitoring period, intercept and slope: refer to regression between average discharge and loss rate, sd: standard error of the slope, $RL_{10\%}$: minimal reduction level necessary for 10 % of year-combinations to observe a significant change based on 14 years of observations, $RL_{50\%}$: minimal reduction level necessary for 50 % of year-combinations to observe a significant change based on 14 years of observations.

Compound	Main loss period	Mean LR	Intercept	Slope	sd	R^2	p-value	$RL_{10\%}$	$RL_{50\%}$
Chlortoluron	Oct–Feb	0.43	−0.56	0.0012	0.18	0.60	0.02	50	85
Dimethenamid	May–July	0.33	−0.61	0.0007	0.14	0.70	0.01	53	70
Isoproturon	Oct–May	0.44	−0.65	0.0012	0.13	0.69	0.02	53	70
MCPA	Apr–Sept	0.82	−1.85	0.0024	0.22	0.82	0.002	35	52
Mecoprop	Apr–Aug	1.6	−0.34	0.0017	0.19	0.78	0.004	42	57
(S-)Metolachlor	Apr–June	0.46	−0.53	0.0008	0.06	0.96	0.00002	17	25
Terbuthylazine	May–Aug	0.39	−0.51	0.0007	0.13	0.69	0.01	56	68

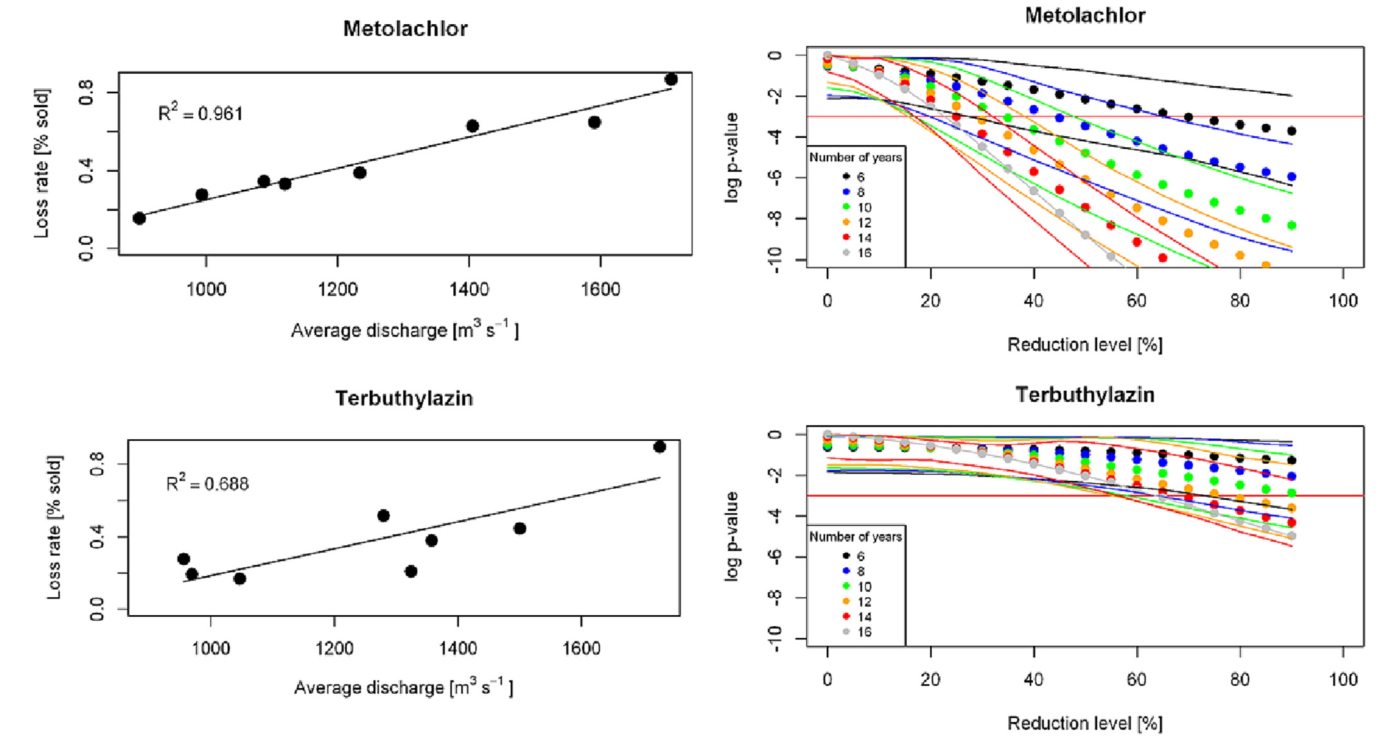


Fig. 3. Correlations between average discharge and the loss rate (% of combined sales of the previous and current year, see methods for details) for metolachlor (top left) and terbuthylazine (bottom left); relationships between the median p-value (in natural log units; dots) for the slope differences between control and mitigation period as a function of reduction level and duration of the observation period (years) for metolachlor (top right) and terbuthylazine (bottom right). The years indicate the full duration ranging between 6 and 16 years split equally between control and mitigation period. The red horizontal line indicates a p-value of 0.05.

3.2. Herbicide dynamics in the Eschibach catchment

The average simulated loss rate for atrazine was 0.42 % and ranged between 0.008 % and 1.25 %. Fig. 6 shows the simulated annual atrazine concentration distributions from 2008 to 2018. The distribution for each year contains concentration values within a 100-day application period. The blue box plots represent the annual concentration distributions for the control scenario with the same quantity of herbicide applied as in the 2009 experiment (Doppler et al., 2012). The red box plots represent the mitigation scenario where half the atrazine is applied. Fig. 6 shows the high level of inter-annual variability of atrazine concentrations and that some years with half application have concentrations distributions that are higher than those with full application (e.g., comparing 2016 to 2011 concentrations). The variability is solely due to the annual differences in the timing and intensity of rainfall events relative to pesticide application (see Figs. S10–S15 in Supporting Information). We can see that the variability in hydrological conditions alone can make it difficult to evaluate the effectiveness of mitigation measures (see S2.1 of Supporting Information for further discussion).

3.2.1. Statistical tests

Based on the modelled time-series described above, we have derived simulated monitoring data consisting of 14-day composite samples. Subsequently, we used different statistical tests to observe their power to detect

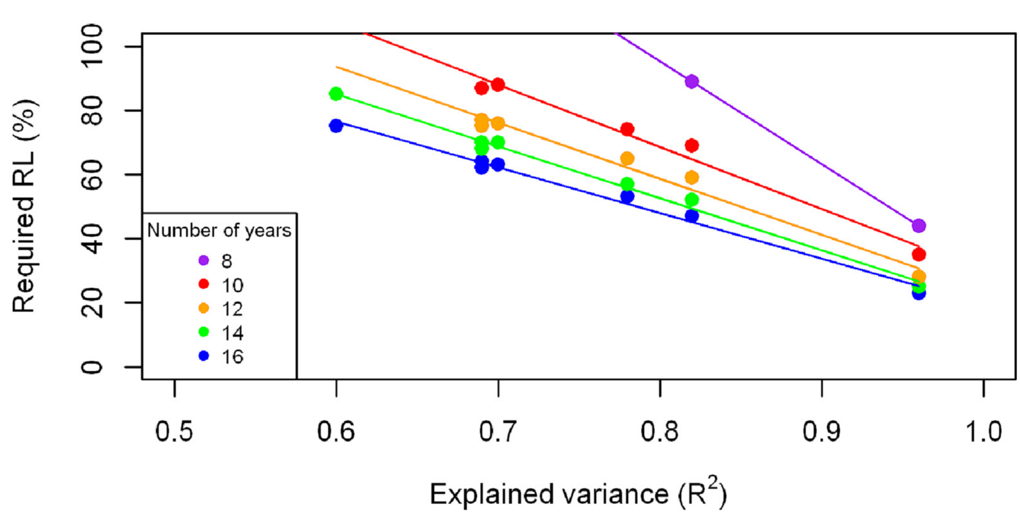


Fig. 4. Relationship between explained variance and the reduction level needed to get a median p-value <0.05 for differing years of observations.

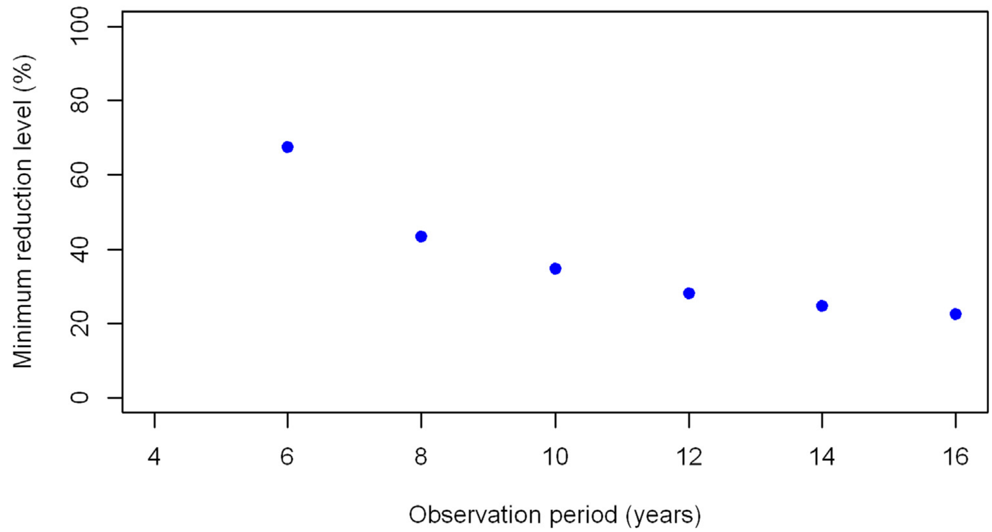


Fig. 5. Relationship between duration and reduction level for metolachlor (median p-value <0.05).

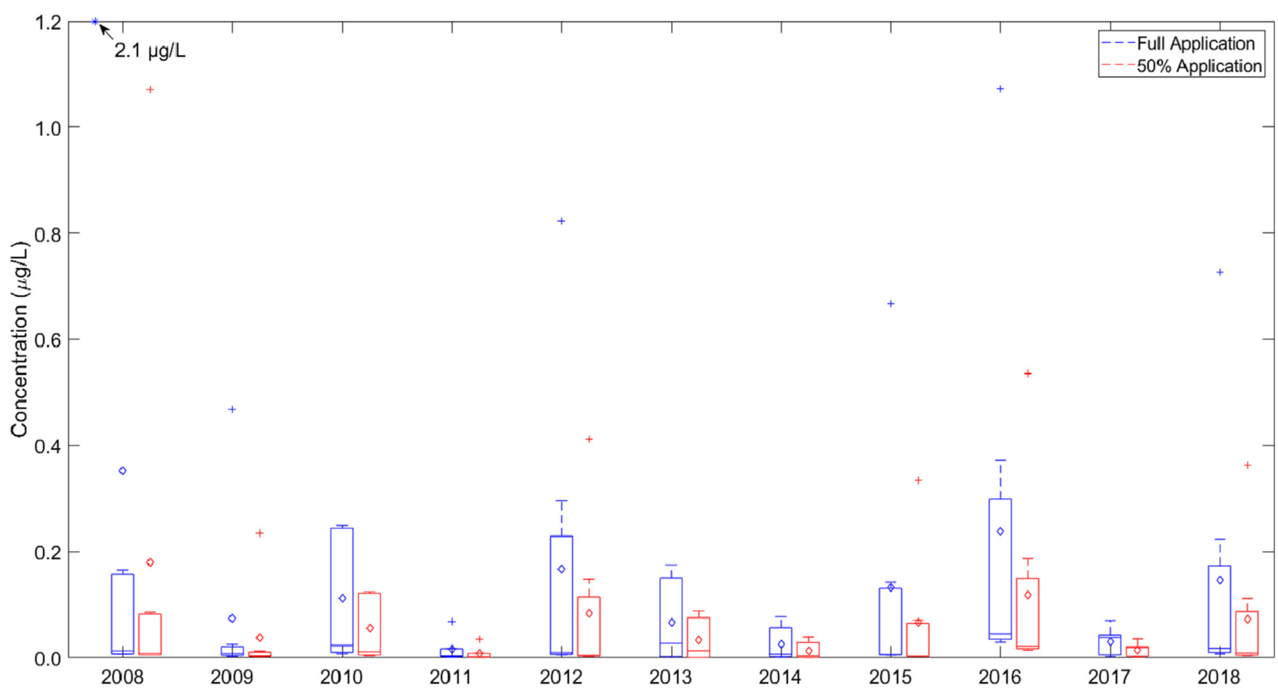


Fig. 6. Simulated annual atrazine concentrations over application period (14-day average composite values). Each box shows outliers (crosses), upper and lower adjacent (whiskers), 25th and 75th percentiles values (box limits), median values (bar within box), and mean values (diamonds). One outlier of 2.1 μg/L in 2008 plots off the graph. Time-series of annual simulated atrazine concentrations from 2008 to 2018 are available in Supporting Information (Figs. S10–S15).

the simulated trends over time. We have treated each single year as independent and tested all combinations of years assigned to either the control or mitigation period.

The different statistical methods that we applied use different levels of information. The *t*-test for example pools all data irrespective of season, C-Q relationships consider the potential influence of hydrological conditions on concentration levels and Seasonal Mann-Kendall test adds seasonal information. Accordingly, one can expect that trends were easier to detect when more information was considered.

Even though we analysed the output of a completely deterministic model, the statistical tests revealed a very limited power to detect the existing trends against the interannual variability (Table 3). The Seasonal

Table 3
Summary of test efficiencies for simulated Eschibach atrazine concentrations. Percent chance of rejecting H_o (no change, p-value = 0.05) given 10 years of monitoring data (5 years control, 5 years mitigation) at different reduction levels of applied atrazine.

Application reduction level		0 %	10 %	30 %	50 %	70 %	90 %
T-test	Annual mean conc.	5 %	6 %	10 %	21 %	43 %	72 %
	Annual median conc.	5 %	6 %	9 %	21 %	47 %	71 %
	Annual 90th percentile	4 %	6 %	8 %	17 %	35 %	58 %
SMK-test		14 %	19 %	33 %	52 %	80 %	99 %
C-Q relationship		2 %	2 %	6 %	16 %	33 %	89 %
Double mass curve		22 %	25 %	36 %	54 %	80 %	100 %

Mann-Kendall test and the double mass curve analysis performed best. Yet, even with a 50 % reduction of the herbicide input, only in about half of the cases (i.e., combinations of year considered control or treatment) a significant trend was observed. This implies a high rate of false negative results.

Conversely, even for the control treatment (i.e., no mitigation measures applied), depending on the combination of years, an apparent negative trend was observed (i.e., false positive). The chance of false positives is shown in the column with '0% Application Reduction Level' in Table 3. Thus, there is a chance (2–22 %) the test would suggest that there was a significant change even if no reduction in application was made.

We suspected that our results might be influenced strongly by the distribution of the herbicide input over time. Because the calibration of the model was performed with data from a controlled experiment where all the atrazine was applied on the same day, the standard model also assumed such an application timing. To be closer to conditions resembling normal farm practice, we also split the application over several days during a one-month period. This change in the model set-up hardly affected our analyses.

We also hypothesized that herbicides with shorter half-lives than atrazine might be even more sensitive to timing effect between time of application and the rainfall triggering the major loss events. This might imply that it would even more difficult to detect trends over time. However, our findings do not support these expectations.

This was different for the influence of concentrations reported below limits of quantification (LOQ). To test for this factor, we subjected the time-series for atrazine and sulcotrione to stipulated LOQs of 6 and 15 ng/L for the 14-day composite samples, respectively. We then analysed the results of these censored time-series for the two cases where the concentrations were either set equal to LOQ or to zero. For both compounds, the statistical power to detect significant trends was further diminished.

Based on these findings that time averaged samples and the sequence of years limit the possibility to reliably detect existing trends, one may ask how long monitoring must last to ensure such trend detection. In the absence of longer measured time-series of weather data to run the model, we created synthetic input time-series by repeatedly sampling from the existing data set with full years as units. These simulations revealed that depending on the mitigation effort unrealistic long monitoring time-series would be required. Even with a 50 % reduction, it would take a 40-years time-series for 80 % of all possible year combinations to detect a significant trend using the Seasonal Mann-Kendall test (SMK-test).

4. Discussion

4.1. Trend detection in long-term monitoring data

In many countries, governmental programs monitoring water quality aim to evaluate the success of implemented policies for improving water quality, which can include the reduction of aquatic pesticide risks. From a political and environmental perspective, there is generally the desire to achieve improvements within short periods of time. The analysis of the high-resolution, long-term data sets of several important herbicides in the Rhine River at Basel provides several important lessons regarding the possibility to detect improvements with confidence. As known from many studies (e.g., Chow et al., 2020; La Cecilia et al., 2021), the observed concentrations and loads may fluctuate strongly at daily, seasonal, and inter-annual scale. This variability though can be rather well explained if pesticide use or sales data are available, one accounts for the seasonal application and loss patterns (i.e., major loss period, MLP) while factoring in the hydrological conditions during the MLP. However, even with such a solid base of data (i.e., pesticide use, in-stream concentrations, hydrology), our results demonstrate that effects of mitigation measures must be substantial (most cases load reductions >50 %) and monitoring periods must span many years (>10 years) to be able to detect significant trends or improvements. Conversely, it would be challenging or nearly impossible to detect statistically significant improvements from mitigation measures if their effectiveness is small to moderate (e.g., 10–30 % load reduction) or if monitoring periods for control and mitigation phases are limited to only a few years. Although mitigation measures have been shown to be effective in controlled experiments (Reichenberger et al., 2007), there is little empirical evidence of strong reductions in aquatic pesticide pollution by individual mitigation measures (e.g., buffer strips) beyond banning or limiting pesticide use (Chow et al., 2020). Therefore, we encourage the implementation of multiple mitigation measures.

4.2. Signal to noise ratio in the synthetic example

Our analysis was based on the output of a deterministic hydrological transport model of the Eschibach catchment, which revealed substantial problems to detect implemented temporal trends against the inter-annual variability. Obviously, part of the deterministic signal (i.e., model output) cannot be captured by the metrics applied to the concentration time-series but appears as noise in the subsequent data analysis. There are three main factors that contribute to this effect, which are also relevant for the analysis of real monitoring data. First, model outputs are aggregated over time to result in time-integrated samples reflecting a realistic sampling strategy (i.e., same used in Swiss monitoring program). Second, the information of the exact timing between herbicide application and rainfall events is ignored. This also corresponds to the real-world situation. Finally, there is indeed a random component to the analysis in that the sequence of years that are either attributed to the control or mitigation period have a substantial influence on the outcome. Even with strong herbicide reduction implemented, depending on the sequence of years, an unlucky situation may prevent the detection of a significant trend or improvement.

Our results demonstrate that these factors may prevent the detection of trends even with substantial real mitigation success and lead to false negative results. On the other hand, the analyses of the control treatment shows that the randomness of the sequence of years may also lead to apparent trends despite a lack of change (false positive). Table 3 shows that tests that have higher chance of change detection also have a higher chance of delivering a false positive. Although the double mass curve analysis has the greatest chance of detecting a change with a given application reduction level, it comes at the price of having the highest chance of a false positive (Table 3). Thus, there appears to be a trade-off when selecting the change detection method between the sensitivity of the method and the risk of a false positive.

The effects are not restricted to such synthetic examples but can also affect the analyses of real monitoring time-series. However, in reality, additional factors may further increase the risk for false negative or positive findings. For instance, other pesticide types (e.g., insecticide, fungicides) can be more varied in their physiochemical properties, making it harder to predict their fate and transport within the watershed. Furthermore, non-agricultural sources can potentially contribute to aquatic pesticide pollution (e.g., urban), which can obfuscate trend detection. Mitigation measures targeted against agricultural sources could be misjudged as ineffective (false negative) when non-agricultural sources are overlooked.

Furthermore, our analysis represents the simple scenario where mitigation measures lead to an abrupt change in water quality (i.e., immediate reduction in loss rate or application). Whereas, in reality, there could be more gradual improvements in water quality from the implementation of mitigation measures. However, it is likely that the detection of significant gradual improvements (against the high levels of interannual variability) is even more challenging than detecting abrupt changes.

5. Conclusions

Our findings have implications firstly for scientists analyzing aquatic pesticide pollution data for trends and secondly for authorities and policy makers regarding their expectations of detectable trends when implementing policies to improve water quality.

Our results indicate that different types of statistical tests result in different p-values, varying in their sensitivity to change detection. Furthermore, our results suggest that a strong reduction (i.e., 70–90 %) is needed to detect a change within 10 years of aquatic pesticide monitoring data. The double-mass curve endpoint test may have the appropriate level of sensitivity to detect the subtle changes in aquatic pesticide pollution when constrained by the length of monitoring time-series; however, the trade-off in applying a more sensitive method for change detection is that it may be more prone to false-positives. Our results suggest that it is important to consider the trade-off between detection sensitivity and the risk of false positives when selecting an appropriate trend detection method and that applying more than one method could provide more confidence in trend detection.

Still, even with the most powerful statistical tests, there are several important requirements from the monitoring program to allow for trend detection. For example, enough baseline monitoring is required before implementing mitigation measures. When selecting compounds for evaluating mitigation policies, it is important to consider the difficulty to detect significant trends for pesticides with concentrations close to or below the limits of quantification. Additionally, in small catchments, effects caused by changes in pesticide use or the implementation of mitigation measures can be obscured by the timing and magnitude of hydrological events relative to pesticide application.

In the context of using pesticide monitoring programs for evaluating the success of mitigation policies, our findings demonstrate that the existence of a high-quality monitoring program per se may not be sufficient to demonstrate the success of such policies. In fact, most European countries use monthly or quarterly grab samples for their pesticide monitoring. Our findings suggest that such low-resolution monitoring would make the detection of trends due to mitigation measures, such as sustainable pesticide use policies (Möhring et al., 2020), nearly impossible. Comprehensive monitoring studies have revealed that measured environmental concentrations of several compounds are more than a factor of 10 above their water quality objectives (e.g., Spycher et al., 2018; Liess et al., 2021). Thus, compliance with the water quality regulations would require a corresponding reduction of those concentrations, which if achieved should lead to significant observable trends. It is important that researchers communicate to policy makers and authorities that data from monitoring programs will only be able to demonstrate positive developments if the effects of mitigation measures are strong and the monitoring is long enough (generally longer than the typical 4-year policy and funding cycle).

CRediT authorship contribution statement

R. Chow: Writing – original draft, Writing - review and editing, Data curation, Formal Analysis, Visualization.

S. Spycher: Investigation, Data curation, Formal Analysis, Visualization, Writing- Reviewing and Editing.

R. Scheidegger: Investigation, Data curation.

T. Doppler: Investigation, Writing- Reviewing and Editing.

A. Dietzel: Investigation, Writing- Reviewing and Editing.

F. Fenicia: Conceptualization, Writing- Reviewing and Editing, Funding acquisition, Resources, Supervision.

C. Stamm: Conceptualization, Writing – original draft, Formal Analysis, Visualization, Funding acquisition, Resources, Supervision.

Data availability

Data will be made available on request.

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.

Acknowledgments

Funded by the Swiss Federal Office for the Environment (FOEN). Thanks to Lorenz Ammann for providing us with a calibrated hydrological pesticide transport model of the Ossingen catchment, along with his technical support. Thanks to Jan Mazacek for providing us discharge and concentration data of the Rheinüberwachungsstation (RÜS). We would also like to thank Rosi Sieber for her GIS support.

Supplementary information

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.164226.

References

Ammann, L., Doppler, T., Stamm, C., Reichert, P., Fenicia, F., 2020. Characterizing fast herbicide transport in a small agricultural catchment with conceptual models. J. Hydrol. 586 124812 (124815 pp.).

Beketov, M.A., Kefford, B.J., Schäfer, R.B., Liess, M., 2013. Pesticides reduce regional biodiversity of stream invertebrates. Proc. Natl. Acad. Sci. 110 (27), 11039–11043.

Boye, K., Lindström, B., Boström, G., Kreuger, J., 2019. Long-term data from the Swedish national environmental monitoring program of pesticides in surface waters. J. Environ. Qual. 48 (4), 1109–1119.

Budd, R., Wang, D., Ensminger, M., Phillips, B., 2020. An evaluation of temporal and spatial trends of pyrethroid concentrations in California surface waters. Sci. Total Environ. 718, 137402.

Burgoa, B., Wauchope, R.D., 1995. Pesticides in run-off and surface waters. In: Roberts, T.R., Kearney, P.C. (Eds.), Environmental Behaviour of Agrochemicals. Wiley & Sons Ltd., John, pp. 221–255.

Choquette, A.F., Hirsch, R.M., Murphy, J.C., Johnson, L.T., Confesor, R.B., 2019. Tracking changes in nutrient delivery to western Lake Erie: approaches to compensate for variability and trends in streamflow. J. Great Lakes Res. 45 (1), 21–39.

Chow, R., Scheidegger, R., Doppler, T., Dietzel, A., Fenicia, F., Stamm, C., 2020. A review of long-term pesticide monitoring studies to assess surface water quality trends. Water Res. X 9, 1–13 100064.

Dabrowski, J.M., 2015. Development of pesticide use maps for South Africa. S. Afr. J. Sci. 111 (1–2) pp.07-07.

Daouk, S., Doppler, T., Wittmer, I., Junghans, M., Coster, M., Stamm, C., 2019. Pesticides dans les eaux de surface : Mesures de réduction et monitoring – synthese des apprentissages lies aux projects. Aqua Gas N° 1.

Doppler, T., Camenzuli, L., Hirzel, G., Krauss, M., Lück, A., Stamm, C.H., 2012. Spatial variability of herbicide mobilisation and transport at catchment scale: insights from a field experiment. Hydrol. Earth Syst. Sci. 16 (7), 1947–1967.

Doppler, T., Honti, M., Zihlmann, U., Weisskopf, P., Stamm, C., 2014. Validating a spatially distributed hydrological model with soil morphology data. Hydrol. Earth Syst. Sci. 18, 3481–3498. https://doi.org/10.5194/hess-18-3481-2014.

Fabre, C., Doppler, T., Chow, R., Fenicia, F., Scheidegger, R., Dietzel, A., Stamm, C., 2023. Challenges of spatially extrapolating aquatic pesticide pollution for policy evaluation. Sci. Total Environ. 875, 162639.

FAOSTAT Statistical Database, 2020. Food and Agriculture Organization of the United Nations, Rome Accessed on June 3, 2020.

FOAG, 2021. Federal Office of Agriculture, Verkaufsmengen je Pflanzenschutzmittel-Wirkstoff Stand 23.11.2021 Accessed on January 7, 2022.

FOAG, 2023. Federal Office of Agriculture, digiFLUX - Information zur Mitteilungspflicht für Pflanzenschutzmittel und Nährstoffe. https://www.blw.admin.ch/blw/de/home/politik/datenmanag-ement/digiflux.html accessed on 2023-04-24.

FOEN, 2022. Federal Office of the Environment - Hydrological Data and Forecasts. https://www.hydrodaten.admin.ch/en/2613.html.

Fuhrimann, S., Wan, C., Blouzard, E., Veludo, A., Holtman, Z., Chetty-Mhlanga, S., Dalvie, M.A., Atuhaire, A., Kromhout, H., Röösli, M., Rother, H.A., 2021. Pesticide research on environmental and human exposure and risks in sub-saharan africa: a systematic literature review. Int. J. Environ. Res. Public Health 19 (1), 259.

Galimberti, F., Dorati, C., Udias, A., Pistocchi, A., 2020. Estimating Pesticide Use Across the EU. Accessible Data and Gap-filling. Publication Office of the European Union.

Gilliom, R.J., 2001. Pesticides in the hydrologic system-what do we know and what's next? Hydrol. Process. 15 (16), 3197–3201.

Gosset, W.S. (Student), 1908. The probable error of a mean. Biometricka 6 (1), 1–25 (JSTOR 2331554).

Helsel, D., 2012. Statistics for Censored Environmental Data Using Minitab and R. John Wiley & Sons, Hoboken, New Jersey.

Hermosin, M.C., Calderon, M.J., Real, M., Cornejo, J., 2013. Impact of herbicides used in olive groves on waters of the Guadalquivir river basin (southern Spain). Agric. Ecosyst. Environ. 164, 229–243.

Hirsch, R.M., Slack, J.R., Smith, R.A., 1982. Techniques of trend analysis for monthly water quality data. Water Resour. Res. 18 (1), 107–121.

Internationale Kommission zum Schutz des Rheins, 2020. Assessment Rhine 2020. Accessed from: https://www.iksr.org/fileadmin/user_upload/DKDM/Dokumente/Broschueren/EN/bro_En_Assessment_%E2%80%9CRhine_2020%E2%80%9D.pdf.

Jarvie, H.P., Johnson, L.T., Sharpley, A.N., Smith, D.R., Baker, D.B., Bruulsema, T.W., Confesor, R., 2017. Increased soluble phosphorus loads to Lake Erie: unintended consequences of conservation practices? J. Environ. Qual. 46 (1), 123–132. https://doi.org/10.2134/jeq2016.07.0248.

Kreuger, J., Nilsson, E., 2001. Catchment scale risk-mitigation experiences- key issues for reducing pesticide transport to surface waters. British Crop Protection Council Symposium Proceedings, pp. 319–324 No. 78.

Kunz, M., Schindler, Wildhaber Y., Dietzel, A., 2016. Zustand der Schweizer Fliessgewässer, Ergebnisse der Nationalen Beobachtung Oberflächengewässerqualität (NAWA) 2011–2014. Bundesamt für Umwelt (BAFU), Schweizerische Eidgenossenschaft.

La Cecilia, D., Dax, A., Ehmann, H., Koster, M., Singer, H., Stamm, C., 2021. Continuous high-frequency pesticide monitoring to observe the unexpected and the overlooked. Water Res. X 13, 100125.

Larson, B.T., Capel, P.D., Goolsby, D.A., Zaugg, S.D., Sandstrom, M.W., 1995. Relations between pesticide use and riverine flux in the Mississippi river basin. Chemosphere 31 (5), 3305–3321.

Lerch, R.N., Sadler, E.J., Sudduth, K.A., Baffaut, C., Kitchen, N.R., 2011. Herbicide transport in Goodwater Creek ExperimentalWatershed: I. Long-term research on atrazine 1. J. Am. Water Resour. Assoc. 47 (2), 209–223.

Leu, C., Schneider, M.K., Stamm, C., 2010. Estimating catchment vulnerability to diffuse herbicide losses from hydrograph statistics. J. Environ. Qual. 39, 1441–1450. https://doi.org/10.2134/jeq2009.0323.

Liess, M., Liebmann, L., Vormeier, P., Weisner, O., Altenburger, R., Borchardt, D., Brack, W., Chatzinotas, A., Escher, B., Foit, K., Gunold, R., Henz, S., Hitzfeld, K.L., Schmitt-Jansen, M., Kamjunke, N., Kaske, O., Knillmann, S., Krauss, M., Küster, E., Link, M., Lück, M., Möder, M., Müller, A., Paschke, A., Schäfer, R.B., Schneeweiss, A., Schreiner, V.C., Schulze, T., Schüürmann, G., von Tümpling, W., Weitere, M., Wogram, J., Reemtsma, T., 2021. Pesticides are the dominant stressors for vulnerable insects in lowland streams. Water Res. 201, 117262.

Lloyd, C.E., Freer, J.E., Collins, A.L., Johnes, P.J., Jones, J.I., 2014. Methods for detecting change in hydrochemical time series in response to targeted pollutant mitigation in river catchments. J. Hydrol. 514, 297–312.

MeteoSwiss, 2019. Climate Data Switzerland, Federal Office of Meteorology and Climatology MeteoSwiss. Data downloaded May 2019. Accessed from: https://www.meteoschweiz.admin.ch/home/service-undpublikationen/beratung-und-service/datenportal-fuer-experten.html.

Möhring, N., Ingold, K., Kudsk, P., Martin-Laurent, F., Niggli, U., Siegrist, M., Studer, B., Walter, A., Finger, R., 2020. Pathways for advancing pesticide policies. Nat. Food 1 (9), 535–540.

Moser, A., Wemyss, D., Scheidegger, R., Fenicia, F., Honti, M., Stamm, C., 2018. Modelling biocide and herbicide concentrations in catchments of the Rhine basin. Hydrol. Earth Syst. Sci. 22 (8), 4229–4249.

Neild, R.E., Seeley, M.W., 1977. Growing degree days predictions for corn and sorghum development and some applications to crop production in Nebraska. Research Bulletin No. 280.

Pesticide Action Network Europe (PANE), 2013. Reducing Pesticide Use across the European Union: Implementation of the EU Directive on Sustainable Use of Pesticides. European Union.

Reichenberger, S., Bach, M., Skitschak, A., Frede, H.G., 2007. Mitigation strategies to reduce pesticide inputs into ground-and surface water and their effectiveness; a review. Sci. Total Environ. 384 (1–3), 1–35.

Rosling, H., Rosling, O., Rönnlund, A.R., 2018. Factfulness: Ten Reasons We're Wrong About the World–And Why Things Are Better Than You Think: Sceptre. Great Britain.

Ruff, M., Singer, H., Ruppe, S., Mazacek, J., Dolf, R., Leu, C., 2013. 20 Jahre Rheinüberwachungsstation -Erfolge und analytische Neuausrichtung in Weil am Rhein. Aqua Gas 5, 16–25.

Schäfer, R.B., von der Ohe, P.C., Rasmussen, J., Kefford, B.J., Beketov, M.A., Schulz, R., Liess, M., 2012. Thresholds for the effects of pesticides on invertebrate communities and leaf breakdown in stream ecosystems. Environ. Sci. Technol. 46 (9), 5134–5142.

Schreder, E., Dickey, P., 2005. Toxic Tradeoff: Exit Diazinon Enter Carbaryl. Clean Water for Salmon Campaign Northwest Coalition for Alternatives to Pesticides. Washington Toxics Coalition (2005).

Searcy, J.K. and Hardison, C.H. (1960): Double-mass curves. Man. Hydrology: Part, 1. https://pubs.usgs.gov/wsp/1541b/report.pdf.

Searcy, J.K., Hardison, C.H., Langbein, W.B., 1960. Double-mass Curves: Manual of Hydrology: Part 1. General Surface-water Techniques. US Government Printing Office.

Spycher, S., Mangold, S., Doppler, T., Junghans, M., Wittmer, I., Stamm, C., Singer, H., 2018. Pesticide risks in small streams—how to get as close as possible to the stress imposed on aquatic organisms. Environ. Sci. Technol. 52 (8), 4526–4535.

Swiss Federal Council, 2017. Aktionsplan zur Risikoreduktion und nach-haltigen Anwendung von Pflanzenschutz-mitteln. Bericht des Bundesrates, Schweizerische Eidgenossenschaft.

Swiss Federal Statistical Office FSO, 2018. Land Use Statistics (Arealstatistik) Based on the Standard Nomenclature NOAS04 (2018). BFS/OFS GEOSTAT, Neuchâtel, Switzerland.

Uehlinger, U., Wantzen, K.M., Leuven, R.S.E.W., Arndt, H., 2009. The Rhine River Basin. In: Tockner, K., Robinson, C.T., Uehlinger, U. (Eds.), Rivers of Europe. Elsevier Academic Press, pp. 199–245.

Vecchia, A.V., Gilliom, R.J., Sullivan, D.J., Lorenz, D.L., Martin, J.D., 2009. Trends in Concentrations and Use of Agricultural Herbicides for Corn Belt Rivers, 1996–2006. Environ. Sci. Technol. 43 (24), 9096–9102.

Welch, B.L., 1947. The generalization of "Student's" problem when several different population vairances are involved. Biometrika 34 (1–2), 28–35 (PMID 20287819).

Wittmer, I.K., Scheidegger, R., Bader, H.-P., Singer, H., Stamm, C., 2011. Loss rates of urban biocides can exceed those of agricultural pesticides. Sci. Total Environ. 409, 920–932.

Wolman, M.G., 1971. The nation's rivers. Science 174 (4012), 905–918. https://doi.org/10.1126/science.174.4012.905.

Zhang, Q., Hirsch, R.M., Ball, W.P., 2016. Long-term changes in sediment and nutrient delivery from Conowingo Dam to Chesapeake Bay: effects of reservoir sedimentation. Environ. Sci. Technol. 50 (4), 1877–1886. https://doi.org/10.1021/acs.est.5b04073.

Zhang, W., 2018. Global pesticide use: profile, trend, cost/benefit and more. Proc. Int. Acad. Ecol. Environ. Sci. 8 (1), 1–27.