Integrated Environmental Assessment and Management — Volume 19, Number 2—pp. 513–526
Received: 11 March 2022 | Revised: 19 August 2022 | Accepted: 12 September 2022

Health & Ecological Risk Assessment

Evaluation of SEAWAVE–QEX in a high agricultural intensity catchment in Belgium

Hendrik Rathjens,[1] *Maria Bettina Miguez,*[1] *Michael Winchell,*[1] *and Robin Sur*[2]
[1]*Stone Environmental Inc., Montpelier, Vermont, USA*
[2]*Bayer AG, Monheim, Germany*

Abstract

Pesticide surface water monitoring data have rarely been used as the only quantitative measure of exposure because the available monitoring data for most pesticides has not been considered robust enough for direct use in pesticide exposure assessments due to infrequent sampling. The cost of daily sample collection and analysis prohibits frequent sampling for most monitoring programs. In this context, a common question raised in assessments is how likely peak concentrations (i.e., annual maxima) may be missed if sampling intervals are longer than daily. The US Geological Survey developed the statistical model "seasonal wave with streamflow adjustment and extended capability" (SEAWAVE–QEX) to address the need to estimate infrequently occurring pesticide concentrations, such as annual maximum daily concentrations, for sites with nondaily monitoring data. This study compares the results of two postprocessing methods and evaluates the capability of SEAWAVE–QEX to estimate annual maximum concentrations of three commonly used herbicides and one metabolite in a catchment in Belgium. The study concludes that the appropriateness of using SEAWAVE–QEX to estimate annual maximum concentrations depends on pesticide characteristics and use and that the model can be particularly sensitive to nonflow correlated exposure events (e.g., point source contributions or drift). *Integr Environ Assess Manag* 2023;19:513–526. © 2022 Stone Environmental and Bayer AG Crop Science Division. *Integrated Environmental Assessment and Management* published by Wiley Periodicals LLC on behalf of Society of Environmental Toxicology & Chemistry (SETAC).

KEYWORDS: Peak concentrations; pesticide monitoring data; SEAWAVE–QEX; statistical estimation

INTRODUCTION

The use of agricultural pesticides can result in residue detections of these products and their metabolites in surface waterbodies (Baets et al., 2019). A critical component to designing and implementing mitigation practices to aid in reducing pesticide residues in surface water is a rigorous understanding of the temporal distribution of pesticide concentrations in the waterbody. There are several types of approaches to assess the concentrations of pesticides in streams and rivers, including direct measurement, the application of environmental models, and combinations of these approaches. This study focuses on the evaluation of a statistical approach developed to estimate daily average pesticide concentration data from monitoring sites with a less-than-daily sampling frequency.

According to the US Environmental Protection Agency (USEPA, 2019a), pesticide surface water monitoring data have rarely been used as a quantitative measure of exposure because the available monitoring data for most pesticides have not been considered robust enough due to infrequent sampling and lack of coverage across the landscape. To accurately characterize annual maximum pesticide concentrations in streams, daily sampling may be necessary during active pesticide runoff periods. The cost of daily sample collection and analysis prohibits sampling at such a great frequency for most monitoring programs. Even sampling frequencies of every 4 days or 10 times per month are uncommon, whereas weekly to monthly frequencies are more typical (Vecchia, 2018). In this context, a common question raised in assessments is how likely peak concentrations (e.g., annual maxima) may be missed if sampling intervals are longer than daily. Upper tail quantities and maxima are more difficult to estimate (Mosquin et al., 2017) than common target quantities such as means or medians because peaks (1) could include statistical outliers, (2) are more variable, and (3) may be driven by the timing of pesticide applications relative to storm events. Thus, the statistical model "seasonal wave with streamflow adjustment and extended capability" (SEAWAVE–QEX; Vecchia, 2018) was developed by the US Geological

This article contains online-only Supporting Information.

Address correspondence to bmiguez@stone-env.com

Published 9 October 2022 on wileyonlinelibrary.com/journal/ieam.

Survey (USGS) to address the need to estimate infrequently occurring pesticide concentrations for sites with less frequent monitoring data. The model evaluates the input, which includes limited measured pesticide concentrations and a covariate (usually streamflow), and fits coefficients in a stochastic representation of the correlations between observed chemical signals and the covariate. SEAWAVE–QEX produces several traces of daily pesticide concentration data that must be postprocessed to obtain estimates of peak concentrations. This postprocessing step is not part of SEAWAVE–QEX, and the two currently available SEAWAVE–QEX application studies used different postprocessing approaches (USEPA, 2019a; Vecchia, 2018).

A modeling approach comparable with the one used by SEAWAVE–QEX exists for estimating constituent loads. One example is the USGS Load Estimator (LOADEST model), which uses streamflow and discrete concentration measurements to develop a regression model to estimate average loads in a flowing water body (Runkel et al., 2004). Although some authors have used the LOADEST model to recreate daily time series of nutrient loads (e.g., Sharifi et al., 2017), the model was developed to produce estimates on a monthly or seasonal basis and has been applied to obtain load estimates for phosphorus, nitrogen, and carbon.

The Watershed Regressions for Pesticides (WARP) model and the WARP for multiple pesticides (WARP-MP) use a different approach to estimating pesticide concentrations in streams (Larson & Gilliom, 2001; Stone et al., 2013). The models are based on regressions between concentration statistics, use-intensity data, and watershed characteristics (Larson & Gilliom, 2001). However, WARP and WARP-MP have not been shown to reliably estimate 1-day annual maximum concentrations (e.g., Larson & Gilliom, 2001; Stone et al., 2013).

Few studies, none of them published in peer-reviewed journals, evaluate applications of SEAWAVE–QEX. Vecchia (2018) checked the model's robustness by applying SEAWAVE–QEX to four pesticide-watershed pairs in the USA (atrazine, Little Buck Creek; carbaryl, Kisco River; chlorpyrifos, Sope Creek; and fipronil, Sope Creek) and evaluated the resulting SEAWAVE–QEX estimated annual maximum concentrations and bias of approximate confidence intervals. However, the analysis was not based on in situ monitoring data. Instead, simulation results "were used to randomly generate time series of 'known' daily pesticide concentrations assuming the true model was given by the fitted model" (Vecchia, 2018, p. 17). For assessing model uncertainty, the author compiled pesticide observations for these same four chemicals from dozens of watersheds around the USA and compared the resulting SEAWAVE–QEX model coefficients (midterm flow anomaly, short-term flow anomaly, seasonal standard deviation, and correlation time scale) obtained for each chemical and concluded that model performance varied across the pesticides analyzed. SEAWAVE–QEX had lowest uncertainty when applied to fipronil and highest when applied to carbaryl, with mixed results for atrazine and chlorpyrifos. The USEPA (2019a) evaluated the accuracy of SEAWAVE–QEX estimations for atrazine, metolachlor, and simazine using the Heidelberg University's National Center for Water Quality Research and Atrazine Ecological Exposure Monitoring Program datasets. The USEPA found that SEAWAVE–QEX produces upper tails that can exceed measured data with most of the distribution lying within the range of measured concentrations for atrazine and metolachlor. In the study, they proposed a different postprocessing approach to estimating maximum concentrations from SEAWAVE–QEX results from the approach put forth by model developers described in Vecchia (2018). The USEPA, however, acknowledged that the limited range of sites and physicochemical properties used in the evaluation add uncertainty to the confidence in broadly applying SEAWAVE–QEX across pesticides and use areas and found that the simazine dataset did not result in an appropriate model fit.

Based on the literature review outlined above, it was concluded that further research on SEAWAVE–QEX model's performance is required. The present study evaluates the capability of SEAWAVE–QEX to estimate annual maximum concentrations of three commonly used herbicides (flufenacet, diflufenican, and isoproturon) and one degradation product of flufenacet (flufenacet sulfonic acid) in the predominantly agricultural catchment of Grote Kemmelbeek (GKb) in Belgium. Flufenacet is a herbicide commonly used on winter wheat in central Europe. The herbicide and its metabolite are of environmental concern because they are frequently detected in monitoring programs (e.g., Willkommen et al., 2019). The other two compounds were selected based on availability of monitoring data and to span a range of organic–carbon partitioning coefficients (K_{oc}), which is an important factor describing how mobile the chemical is in the environment. High temporal resolution (once or twice daily) sampling of flow and multiple pesticides are available for a period of 3.5 years at the outlet of the catchment (Baets et al., 2019). Concurrently, a farmer survey was carried out among all farmers cultivating fields in the watershed providing detailed data on pesticide use in the watershed (see Baets et al., 2019 for details).

Previous studies evaluated SEAWAVE–QEX performance by using large monitoring programs located in the US, aggregating data across watersheds or chemicals (USEPA, 2019a; Vecchia, 2018). This study expands the existing body of knowledge by evaluating four different chemicals in the same watershed over the same period in a different geographical region, which allows for comparison based on chemical characteristics and use. Using the high-frequency monitoring dataset and detailed knowledge of pesticide use available for the GKb catchment, the study answers the following research questions: (1) Can SEAWAVE–QEX be used to fill gaps in pesticide concentration time series to estimate concentration peaks, and how does the model perform for different chemicals in the same watershed? (2) What are the implications of using the

USGS (Vecchia, 2018) or USEPA (2019a) postprocessing approaches to estimate the 1-day annual maximum concentration from SEAWAVE–QEX conditional simulations? (3) How sensitive is SEAWAVE–QEX to individual observations and nonrunoff-driven events caused by point sources or drift of flufenacet? The study builds partly on a previously conducted Soil and Water Assessment Tool (SWAT; Arnold et al., 1998) modeling study conducted by Sur et al. (2018) that identified point source events of flufenacet in the GKb catchment.

MATERIALS AND METHODS

Study framework

The study framework is illustrated in Figure 1. Daily streamflow data (gap-filled with SWAT model results) and subsamples of daily pesticide concentration data from the GKb catchment are the SEAWAVE–QEX inputs. A subsampling method is applied mimicking typical designs of monitoring programs. The 200 traces of SEAWAVE–QEX estimated daily pesticide concentration data are processed according to two postprocessing approaches (USGS and USEPA approaches) to calculate annual statistics, in this case 1-day annual maximum concentrations. The 1-day annual maximum concentration was selected in this study because it is a standard metric used in regulatory risk assessments. Those annual statistics are then compared with the annual statistics of the observed data to assess SEAWAVE–QEX. Additionally, interannual performance variability is discussed. The same process was carried out for the dataset of flufenacet (FFA) excluding previously identified point sources, and results were used to assess SEAWAVE–QEX sensitivity toward point sources. Root mean squared error (RMSE) was used to compare the performance of the two postprocessing approaches.

SEAWAVE–QEX model

SEAWAVE–QEX was developed by the USGS (Vecchia, 2018) and is a freely available statistical model written in R (USEPA, 2019b). It uses streamflow or a different covariate (such as precipitation) and the available measured chemical concentration data, including nondetects, as input to produce several "traces," or estimated daily chemographs of chemical concentrations at a site. Each daily value will either be set at the observed concentration or, when no data are available, to an estimated concentration randomly generated by the model according to statistical parameters. For nondetects (censored samples), SEAWAVE–QEX will randomly generate a concentration below the censoring limit. The model's main drivers are the seasonal trends determined from the observed chemical concentration data, the covariate's (e.g., streamflow) short- and long-term variabilities, and correlation between the observed pesticide concentrations and the covariate. SEAWAVE–QEX can accommodate up to two seasonal waves (i.e., distinct application periods) per 365-day period, which is sufficient for the chemicals evaluated in this study with the exception of diflufenican (DFF), which is used mainly preemergence on winter grains as well as in urban areas with no clear seasonality. The program fits all possible combinations of the seasonal curve to the input data and selects the seasonal wave that best fits the observed concentrations for all years. To obtain a reasonable model fit, use patterns and application timing of the chemical of interest should remain relatively consistent over the years analyzed with SEAWAVE–QEX. The generation of traces in the model occurs in a log-transformed space and is given by the daily selection of

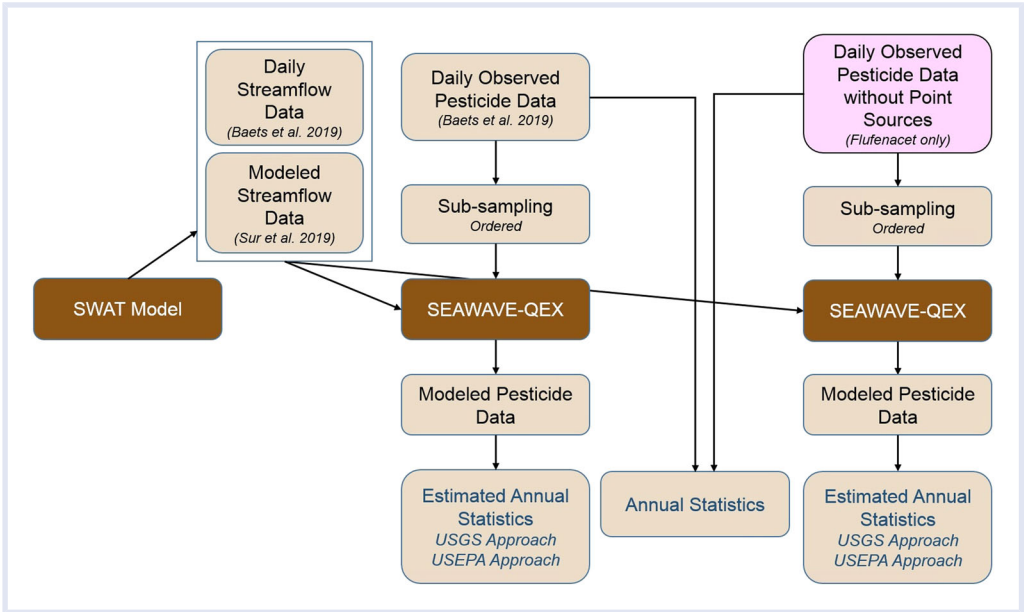

FIGURE 1 Study design framework and references to input data sources

 DOI: 10.1002/ieam.4688

TABLE 1 Pesticide data summary

Chemical	LOQ (µg/L)	K_{oc} (ml/g)	DT50 (days)	Number of samples	Percent censored samples	Maximum observed concentration (µg/L)	Date of maximum observed concentration
Flufenacet	0.01	221[a]	12[a]	1268	20%	5.1	2010-05-25
Flufenacet sulfonic acid	0.01	11.1[a]	31.6[a]	1052	2%	1.07	2011-12-15
Diflufenican	0.01	3417[b]	143.2[b]	1268	32%	0.97	2012-03-30
Isoproturon	0.01	122[c]	12.6[c]	1052	1%	13.55	2011-10-26

Abbreviations: DT50, soil half-life at 20 °C and field capacity; censored samples are samples with concentrations equal to or less than LOQ/2; K_{oc}, organic–carbon partitioning coefficient; LOQ, limit of quantitation.
[a]Values obtained from European Food Safety Authority (EFSA, 2022).
[b]Values obtained from EFSA (2007), deviation of DT50 as a result of normalization to field capacity with a Q10 of 2.58 instead of 2.2.
[c]Values obtained from European Union (EU) (2002).

a random value from the population of normalized residuals, which is then multiplied by the seasonal standard deviation (in turn determined by the seasonal wave) to get an estimated concentration. The model's minimum data requirements are three years of daily streamflow observations with concurrent pesticide sample data. At least 10 to 12 samples per year and a censoring rate (nondetects) lower than 70% are required of the pesticide sample data (see Vecchia, 2018). The three herbicides and one metabolite selected for this study meet the minimum data requirements. Uncensored detection rates range from 68% to 99% (Table 1), and the subsampling of the almost daily monitoring program met or exceeded the minimum number of samples per year.

Study area

The GKb catchment is a 1030 ha catchment located in the Flanders region of Belgium (Figure 2). The catchment is predominantly agricultural (>85%), with some forest, farmsteads, and noncultivated grassland. Main crops grown include corn, potato, and winter wheat (Rathjens et al., 2022). Agricultural fields are located close to surface water bodies (i.e., no riparian buffers), which heightens the risk of off-field drift depositions into the surface waterbody. Community gardens are located in the watershed where herbicides are occasionally applied (Dirk Baets, Bayer, personal communication, 10 June 2017; see also Baets et al., 2019). The mean elevation of the catchment is 53 m, and ranges from a minimum of 24 m to a maximum of 159 m (Rathjens et al., 2022). Based on the years from 2009 to 2013, the average annual precipitation for the GKb catchment is 797 mm/year, with 18 mm/year (liquid equivalent) falling as frozen precipitation (data from four weather stations close to the watershed provided by the Flemish government and the INAGRO research institution, personal communication, 19 September 2016; see also Rathjens et al., 2022). The watershed soils are mostly poor to imperfectly drained loams, silts, and silt loams, and approximately 50% of the watershed area is tile drained (Rathjens et al., 2022).

Streamflow data

Flow data for the GKb catchment were available from the monitoring station located at the outlet of the catchment from 17 May 2010 through 31 December 2013. This monitoring station was outfitted with a programmed automatic sampler that measured water level and flow velocity at a 5-min time step. Based on an estimated cross-sectional area associated with the measured depth, and the measured velocity, a flow rate (m^3/s) was calculated for each time step. There was one notable period of missing data (from 5 March 2012 to 29 March 2012) when a high flow event damaged the flow monitoring equipment. In this study, the streamflow results from a SWAT model developed by Sur et al. (2018) were used to fill the first 5 months of 2010 and the 25-day period in 2012 when there was no measured streamflow data. As SEAWAVE–QEX requires a complete daily time series, filling the period from January to May 2010 and the 25-day period in 2012 was necessary for running the model. The gap only represents 12% of the available monitoring data and the SWAT model used for gap filling was judged to be very good (Sur et al., 2018) based on standard SWAT model calibration evaluation guidelines (Moriasi et al., 2007). The mean daily flows used as input to SEAWAVE–QEX are shown in Figure A.1 (Supporting Information). The year 2010 was included in the study despite its partial data availability to take advantage of all available daily pesticide monitoring data. However, the results from 2010 should be interpreted with caution.

Pesticide data

SEAWAVE–QEX evaluation was conducted for three commonly used herbicides (FFA, DFF, and isoproturon [IPU]) and one metabolite of FFA (flufenacet sulfonic acid [FFA-SA]). Monitoring data for two of the herbicides in this study (FFA and DFF) were available beginning on 17 May 2010 and lasting through 31 December 2013. Monitoring data for the metabolite and IPU were available from 1 January 2011 to 31 December 2013. A summary of the pesticide monitoring data is presented in Table 1.

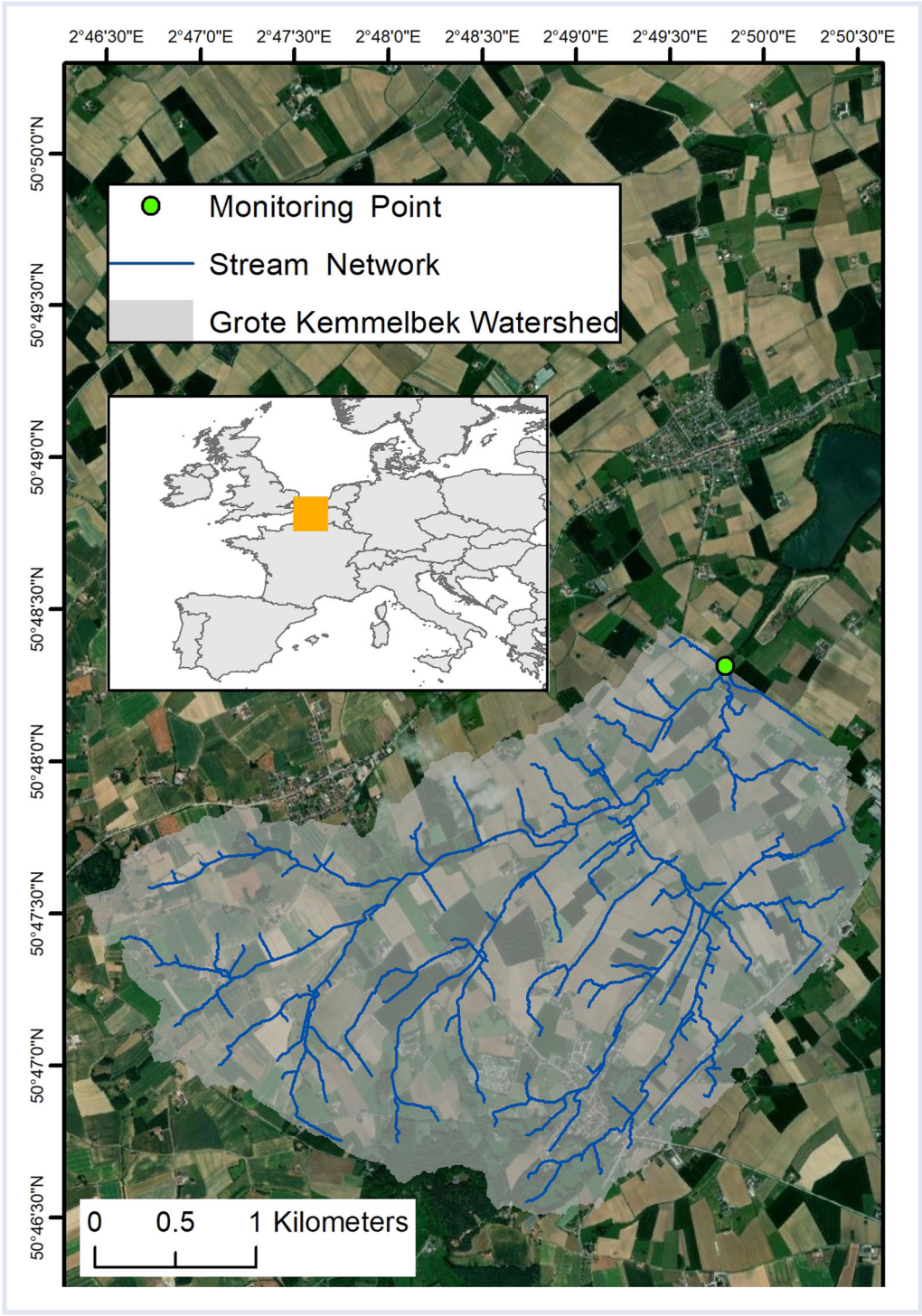

FIGURE 2 Site map of GKb catchment, stream network, and monitoring point location

In the GKb catchment, FFA is a herbicide used primarily on corn (including corn silage), potatoes, and grains with typical applications occurring between April and June. Elevated concentrations were observed in May 2010 (5.10 µg/L) and June–July 2012 (3.02 and 3.85 µg/L). Although FFA, with moderate mobility, tends to peak within its application period (April–June) or shortly after and disappears within 1 month of the application period, FFA-SA peaks occur approximately 6 months after the FFA application period and only 2% of the samples are below the detection limit (Table 1). Flufenacet sulfonic acid had its maximum observed concentrations in December 2011 (1.07 µg/L) and October 2013 (1.04 µg/L). Diflufenican is moderately persistent in the environment and the herbicide with the highest K_{oc} studied (3417 ml/g). It is used on agricultural crops (e.g., winter wheat, winter barley) in October and

 DOI: 10.1002/ieam.4688

November as well as in urban areas with no clear seasonality (e.g., community gardens), resulting in low concentrations with sustained detections throughout the study period. The only exception is the highly unusual DFF peak of 0.97 µg/L on 30 March 2012, occurring 2 weeks after a high streamflow event. Isoproturon is used to control broadleaf weeds in winter wheat and cereal with applications in early postemergence and has a low carbon–water partitioning coefficient. Isoproturon exhibited the highest concentrations in late October to early November 2011, in March 2012, and in May 2013, with a maximum concentration of 13.55 µg/L. Chemographs of these chemicals are presented in Figures A.2–A.5 (Supporting Information).

Point sources

Previous work in the GKb catchment (Sur et al., 2018) used FFA field-level application data obtained from a farmer survey and a Monte Carlo simulation approach to identify point source events. Examples of point source events are spillage on hard surfaces during filling or wash-off from rinsing spraying equipment and unreported use or misuse of the product. In the context of this study, drift (i.e., the transport of chemicals by wind to a nearby water body) was not considered a point source. The analysis found that 46% (34 of 74) of elevated FFA concentration detections were not likely to be caused by diffuse sources and were therefore likely to be a result of point source contributions (see Figure A.2; the methodology of the point source classification is explained in Sur et al., 2018). Although the previous work (Sur et al., 2018) focused on identifying potential FFA point events, similar point source events could have also occurred for the other chemicals in this study. The results from the point source analysis conducted by Sur et al. (2018) were used to analyze SEAWAVE–QEX sensitivity on point source events. The days classified as FFA point source events were removed from the monitoring input SEAWAVE–QEX data (and not replaced by any value) when analyzing SEAWAVE–QEX sensitivity on point sources (see study design framework in Figure 1).

Subsampling

To assess SEAWAVE–QEX performance, subsampling method and intervals were chosen to be representative of potential sampling program design. Two subsampling intervals were selected that picked observed concentrations from the available almost daily monitoring dataset: 7 and 30 days, and the subsampling itself was performed with an ordered method (e.g., sample every Monday or the third day of each month). One subsample was obtained per each day in that interval and thus the number of subsamples generated for each sampling interval is equal to the number of days in that interval (e.g., monthly subsampling yields 30 distinct subsamples).

Results evaluation

Each subsample input to SEAWAVE–QEX generated 200 traces. For each of those subsamples, the estimated value of the statistic of concern, 1-day annual maximum concentrations, was calculated as suggested by the model developers (Vecchia, 2018) by taking the average of 1-day annual maximum concentration of each of the 200 traces (referred to subsequently as the "USGS's approach"). A second postprocessing approach was used as suggested by USEPA (2019a) where the maximum of the 99th percentile of the 1-day annual maximum concentration of the 200 traces was calculated (referred to subsequently as the "USEPA's approach").

Because the streamflow and pesticide concentration values varied significantly from year to year, results were interpreted on an annual basis so that interannual differences could be recognized. All estimates of USGS's and USEPA's postprocessing approaches were then compared with the observed annual statistic of concern (Figure 1). The RMSE was calculated for each interval and postprocessing approach, both per year and an average across the study period, for each pesticide. Each subsample that yielded a final estimate of the annual statistic that was 100 times higher than the annual maximum observed concentration in any of the years was flagged as unrealistic. Two orders of magnitude higher than the maximum observed concentration was selected as a conservative (but subjective) threshold that a modeler would use to not include the results in a study. Careful examination of the model output for subsamples with unrealistic results according to the procedure described in USEPA (2019b) confirmed a poor model fit. Unrealistic subsamples were excluded from tables comparing results between FFA and FFA without point sources because they introduce noise to the comparison of results and would be discarded in any actual application of the SEAWAVE–QEX model. This approach is confirmed by USEPA (2019a), which stated that analyzing SEAWAVE–QEX output and verifying a proper model fit is essential to ensure that the resulting concentrations are trustworthy for use in risk assessment.

RESULTS AND DISCUSSION

Flufenacet results

Estimated annual statistics for FFA demonstrated great variability across years, subsampling frequency, and postprocessing approach (Figure 3). The 1-day annual maximum concentration estimates of FFA using USEPA's postprocessing approach have higher minimum estimates across all sampling intervals than USGS's approach. Although the highest estimates using USEPA's approach are still at least an order of magnitude higher than the observed concentrations, application of the USGS's approach resulted in annual maximum concentrations that were from less than 1 to almost 4 orders of magnitude higher than the observed value. The extremely high values calculated with the USGS's approach are caused by traces in three subsamples with unrealistic (i.e., two orders magnitude greater than the observed values) results (two 7-day interval and one 30-day interval subsamples). These

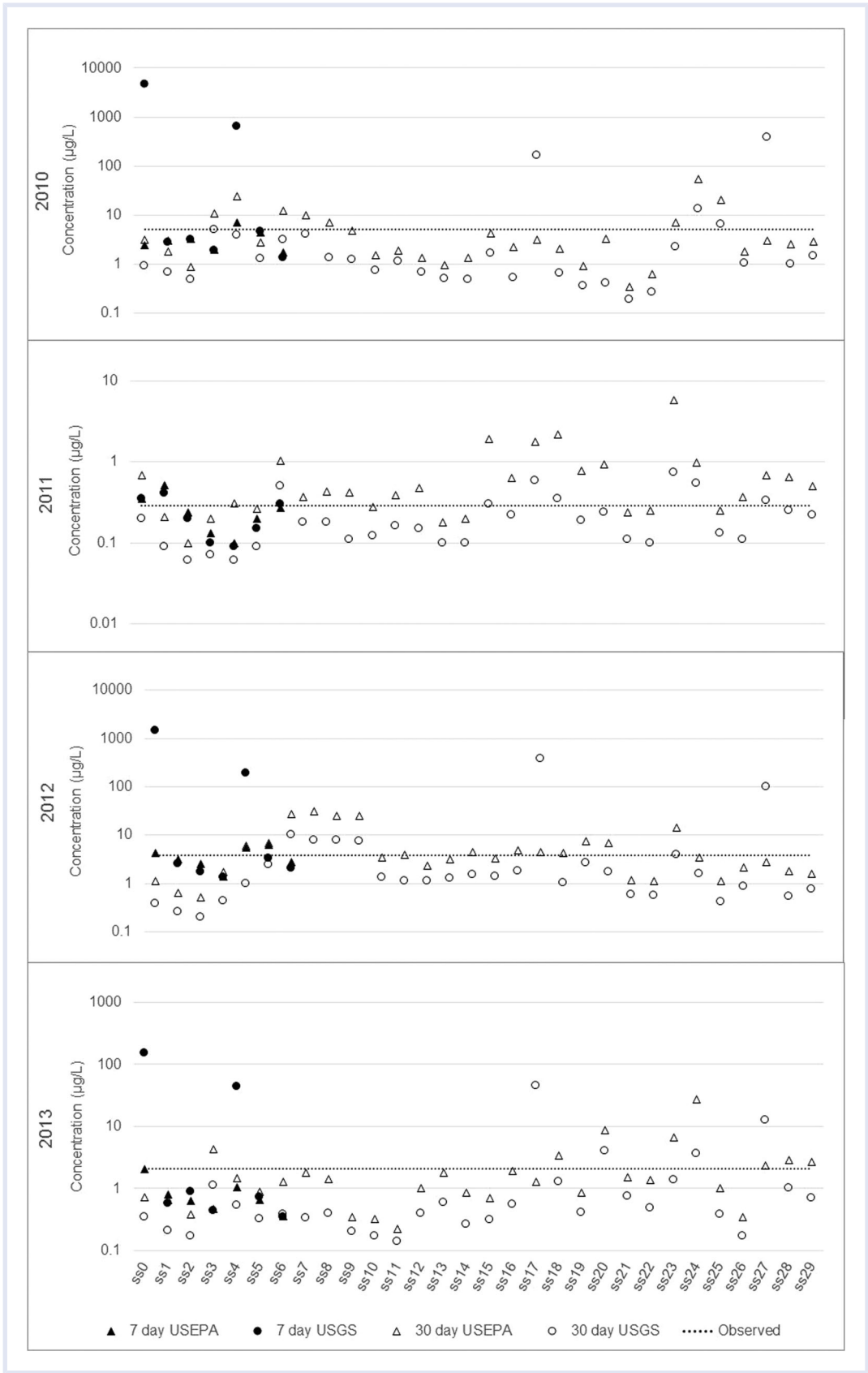

FIGURE 3 Estimates of the 1-day annual maximum concentration for flufenacet, using US Geological Survey's (USGS's) and USEPA's approaches, for the 7-day (subsamples ss0–ss6) and 30-day (subsamples ss0–ss29) intervals

 DOI: 10.1002/ieam.4688

SEAWAVE–QEX results were investigated further. The very high maximum concentrations occur during two peak flow events in November 2010 and March 2012. Those maxima occurred outside the seasonal wave (determined by SEAWAVE–QEX using the same inputs) where maximum concentrations are expected to occur. Although the USEPA's postprocessing approach effectively discards the highest unrealistic results, the USGS's approach, which takes the average of the maximum concentration from each trace, is easily skewed by an unrealistic trace.

The spread in estimates between the different subsamples demonstrates that results can be heavily influenced by individual events. For some sampling frequencies and site-years, the model is producing reasonable results, whereas for others, estimation of annual maximum concentrations are several orders of magnitude higher than the observed maxima (considering almost daily monitoring data). No clear pattern of subsampling sets or frequencies causing unrealistically high results could be identified, which indicates that SEAWAVE–QEX can be very sensitive to individual observed data points.

Flufenacet without point source results

Point-source-driven events provide a challenge for SEAWAVE–QEX when using streamflow as a covariate, because those events are unrelated to flow. SEAWAVE–QEX results for FFA without point sources using both USEPA's and USGS's approaches resulted in lower RMSE than results for FFA, for every year and subsampling interval, with the exception of 7-day subsamples for 2012 (Table 2). The model results for flufenacet with or without point sources had RMSE ranges from 0.12 to 1.51 for the 7-day subsamples in 2011 and 2013 (years of low and moderate flows), and RMSE ranges from 1.62 to 2.62 for the 7-day subsamples in 2010 and 2012 (years with higher streamflow).

Considering the average across all years, the RMSE of the 1-day annual maximum concentrations decreases when excluding point sources. Those results can be considered as an additional line of evidence that the point source classification conducted by Sur et al. (2018) provided realistic results. Both FFA with and without point sources performed worse at the 30-day interval (average RMSE of 1.69–6.40) than the 7-day interval sampling (average RMSE of 1.04–1.51). The USEPA's postprocessing approach yielded slightly better performance for the 7-day interval in both FFA with and without point sources, with average RMSE values 10%–20% lower than USGS average RMSE values. However, the USGS's approach performed noticeably better at the 30-day interval, because average RMSE values were two to almost three times lower than USEPA average RMSEs. The lower RMSE for FFA without point sources indicates that the point source detection was correct, and it demonstrates that SEAWAVE–QEX is sensitive to point source events in the monitoring data. However, the improvement of SEAWAVE–QEX performance was not consistent across all site-years, and there are still some cases where unreasonable annual maximum concentrations are predicted even after removing the point source events. This could be caused by not capturing all point sources in the initial analysis, drift driven concentrations, or overall SEAWAVE–QEX performance. Because most FFA scenarios without point sources resulted in reasonable results, SEAWAVE–QEX sensitivity to a particular observed concentration is likely the cause.

Flufenacet sulfonic acid results

Results from the SEAWAVE–QEX model applied to the metabolite of FFA systematically estimated 1-day annual maximum concentrations within an order of magnitude of the observed annual maximum concentration (Figure 4). Estimated values of FFA-SA were consistently higher than observed for the year 2012, but only one subsample (30-day interval) yielded unrealistic results.

The RMSE values of the 1-day annual maximum concentrations at the 7-day interval were between 0.16 and 0.77 for both postprocessing approaches (Table 3). The results

TABLE 2 Root mean squared error of 1-day annual maximum concentrations for flufenacet and flufenacet without point sources (excluding unrealistic traces)

	RMSE of 1-day annual maximum concentration							
	FFA				FFAwoPS			
	7-day interval		30-day interval		7-day interval		30-day interval	
Year	USEPA	USGS	USEPA	USGS	USEPA	USGS	USEPA	USGS
2010	2.41	2.62	10.48	4.13	1.89	2.45	6.42	3.32
2011	0.13	0.12	1.17	0.18	0.12	0.12	0.57	0.17
2012	1.62	1.78	8.91	3.11	1.96	2.32	5.63	2.93
2013	1.34	1.51	5.03	1.63	0.19	0.19	1.49	0.33
Average	1.38	1.51	6.40	2.26	1.04	1.27	3.52	1.69

Abbreviations: FFA, flufenacet; FFAwoPS, flufenacet without point sources; RMSE, root mean squared error.

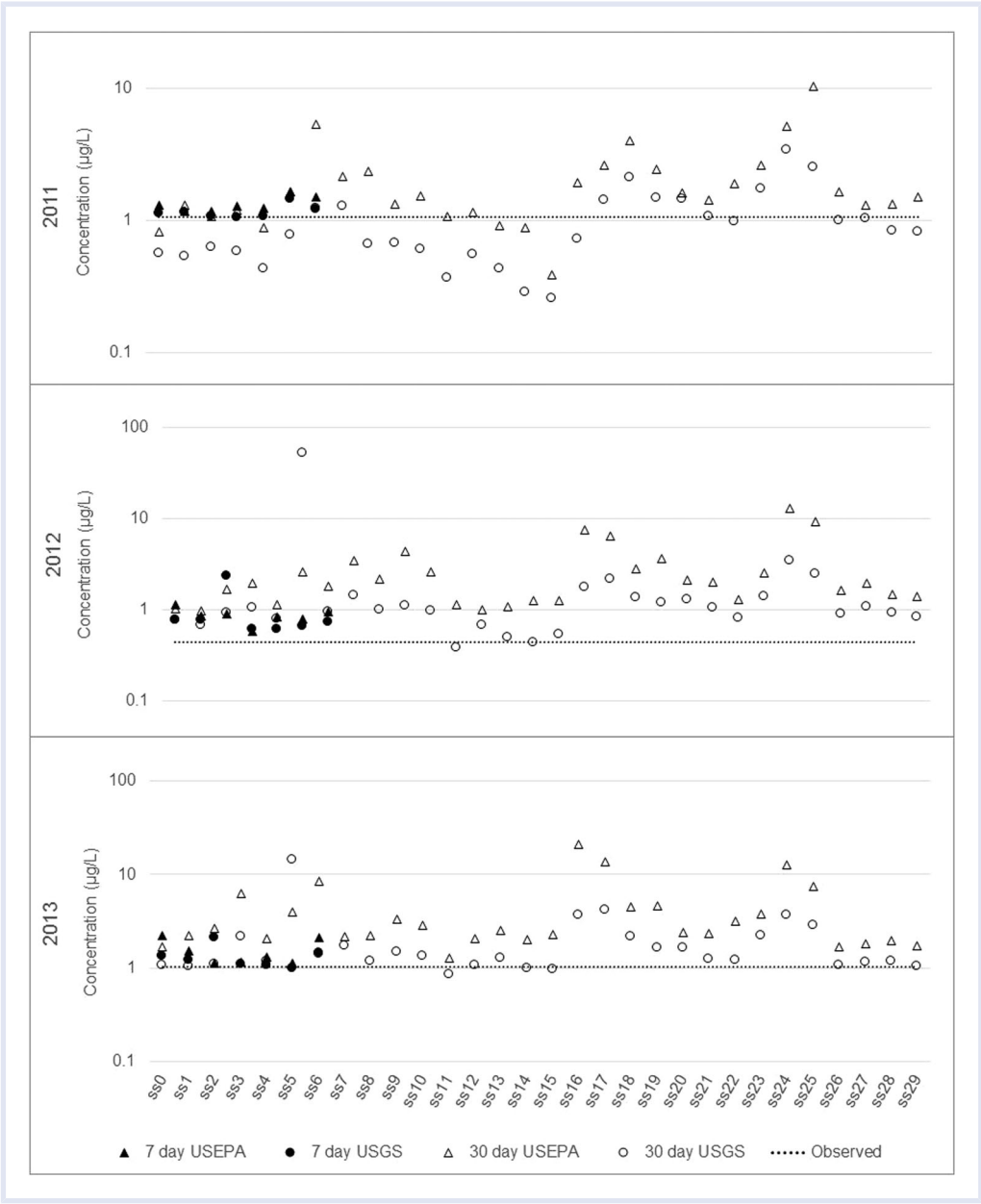

FIGURE 4 Estimates of the 1-day annual maximum concentration for flufenacet sulfonic acid using US Geological Survey's (USGS's) and USEPA's approaches, for the 7-day (subsamples ss0–ss6) and 30-day (subsamples ss0–ss29) intervals

suggest that SEAWAVE–QEX is an appropriate tool for estimating FFA-SA concentrations in this watershed using sampling frequencies of 7 days. However, the results also indicate that estimates on wet years (e.g., 2012) would probably overestimate maximum observed concentrations (Figure 4). Evaluating more years of data could verify this claim.

Regarding the difference between USGS's and USEPA's SEAWAVE–QEX postprocessing approaches, the estimates of FFA-SA's 1-day annual maximum concentration using USEPA's approach were considerably higher than those obtained with the USGS's approach (Figure 4), for all years. The RMSE results reveal that the USGS's approach performs better with the lower frequency sampling (30-day interval). Flufenacet sulfonic acid is a soil metabolite and thus its transport is driven primarily by subsurface processes; therefore, FFA-SA concentrations in the surface waterbodies are more likely to be correlated with baseflow than surface runoff, and drift or point source contributions can be ruled out (Rathjens et al., 2022). The results for FFA-SA suggest that SEAWAVE–QEX might perform well for compounds transported mainly through baseflow, but additional modeling is needed to verify this observation. For FFA-SA, SEAWAVE–QEX correctly estimated 1-day annual maximum concentrations within a 90% confidence interval, and the results

 DOI: 10.1002/ieam.4688

TABLE 3 Root mean squared error of 1-day annual maximum concentrations for flufenacet sulfonic acid, diflufenican, and isoproturon, excluding two unrealistic traces (one for flufenacet sulfonic acid and one for diflufenican)

	RMSE of 1-day annual maximum concentration											
	FFA-SA				DFF				IPU			
	7-day interval		30-day interval		7-day interval		30-day interval		7-day interval		30-day interval	
Year	USEPA	USGS	USEPA	USGS	USEPA	USGS	USEPA	USGS	USEPA	USGS	USEPA	USGS
2010					0.47	0.12	0.18	0.25				
2011	0.31	0.16	2.18	0.71	0.01	0.01	0.04	0.02	7.32	7.86	27.30	8.57
2012	0.45	0.77	3.67	0.95	0.65	0.58	0.80	1.54	5.45	5.49	19.37	7.98
2013	0.65	0.46	5.49	1.07	0.01	0.04	0.11	0.14	0.73	1.49	31.33	11.04
Average	0.47	0.47	3.78	0.91	0.29	0.19	0.28	0.49	4.50	4.95	26.00	9.19

Abbreviations: DFF, diflufenican; FFA-SA, flufenacet sulfonic acid; IPU, isoproturon; RMSE, root mean squared error.

indicate that a sampling monitoring program could be set up with a decreased sampling frequency (e.g., weekly).

Diflufenican results

Comparing 1-day annual maximum concentrations of DFF obtained from the two postprocessing methods revealed inconsistent over- and underestimation of the observed 1-day annual maximum concentrations (Figure 5). No trends are discernible between the two postprocessing approaches across sampling frequencies and years. Only one DFF subsample, for the 30-day interval, had unrealistically high results, six orders of magnitude higher than the maximum observed concentration in 2010 and four orders of magnitude higher than the maximum observed concentration in 2012.

Average RMSE values for DFF were less than 1, even with decreased sampling frequency and with both postprocessing approaches (Table 3). In contrast, the model performance varied across the years evaluated. All scenarios overestimated the maximum concentration in 2010, although performance was good in 2011, and consistent underestimation in 2012 and 2013. Monitoring data for 2010 were only available from May, and the actual maximum concentration might not have been captured by the monitoring data. SEAWAVE–QEX assumed a double seasonal wave for this chemical based on years 2011–2013 of observations. Therefore, it is reasonable that, in 2010, with missing data for the first part of the year, there is greater spread of potential maximum concentrations.

The RMSE of the 1-day annual maximum DFF concentration revealed that the over- and underestimation, if existent, in 2010 and 2013, respectively, were generally low, and within the range of observed values at the site (estimated 1-day annual maximum concentrations were within 0.5–2 times observed annual maximum concentration). However, the underestimation in 2012 was more significant and was noticeably higher for the lower frequency subsamples using USGS's approach. Diflufenican has no clear seasonality in its use in the GKb watershed, and there is consequently a less than ideal model fit because SEAWAVE–QEX must assign up to two seasonal waves to the observed data. Nonetheless, these results suggest that the SEAWAVE–QEX model might be a suitable tool to estimate 1-day annual maximum concentrations of DFF in this watershed if aggregating data over many years.

Isoproturon results

SEAWAVE–QEX annual 1-day maximum estimates for IPU using USGS's approach ranged from 1.71 to 14 µg/L (7-day subsamples) and 0.5 to 55 ug/L (30-day subsamples) so were generally within an order of magnitude of the observed annual maximum concentrations of 3.99–13.55 µg/L (Figure 6). The USEPA's approach to estimating the annual maximum concentration resulted in similar estimates for the 7-day subsamples (2.81–14.17 µg/L) and higher overall estimates for the 30-day subsamples (1.11–123.19 µg/L). There were no traces with unrealistic results. Application of SEAWAVE–QEX to IPU subsampled datasets resulted in consistent underestimation of 1-day annual maximum concentrations in 2011 and 2012, regardless of sampling frequency or approach.

Evaluation of RMSE for the three years revealed that only the 7-day subsamples for the year 2013 were close to the observed IPU concentration (Table 3). The 30-day interval RMSEs were worse for all years and both processing approaches. In fact, IPU had the highest RMSE calculated across the compounds studied. A decrease in sampling frequency did not affect the RMSE significantly for the years 2011 and 2012. The large spread in IPU results with 30-day subsampling suggests sensitivity of the model to specific observations, as is the case with FFA.

Implications for environmental management and pesticide risk assessments

SEAWAVE–QEX was developed as a tool that environmental managers and risk assessors could use to estimate concentration peaks for flowing water bodies where sufficient flow, but only sparse pesticide monitoring data, are available. In this study, SEAWAVE–QEX demonstrated high sensitivity to individual observed FFA data points for certain sampling frequency combinations. Those combinations

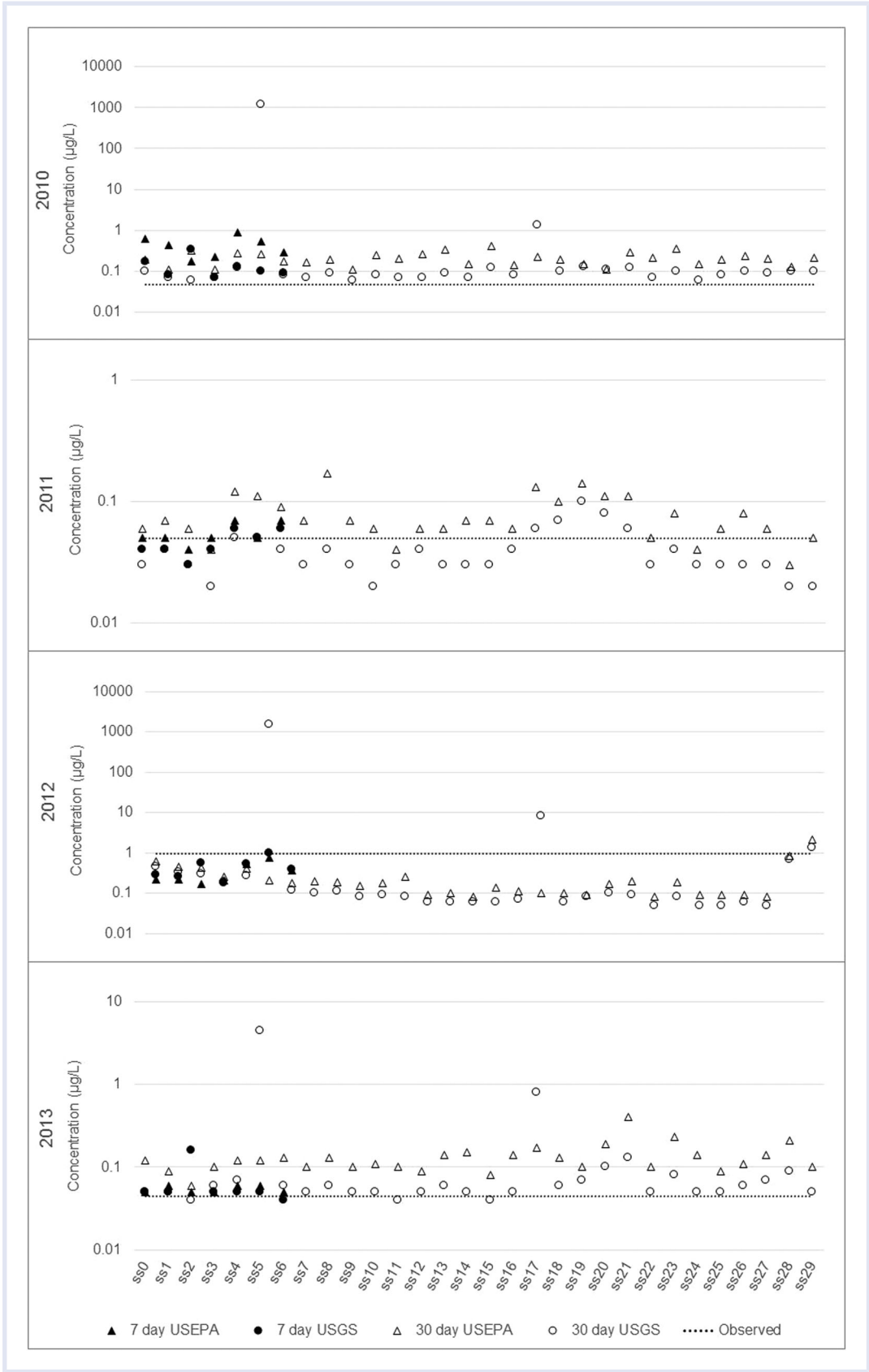

FIGURE 5 Estimates of the 1-day annual maximum concentration for diflufenican using US Geological Survey's (USGS's) and USEPA's approaches, for the 7-day (subsamples ss0–ss6) and 30-day (subsamples ss0–ss29) intervals

 DOI: 10.1002/ieam.4688

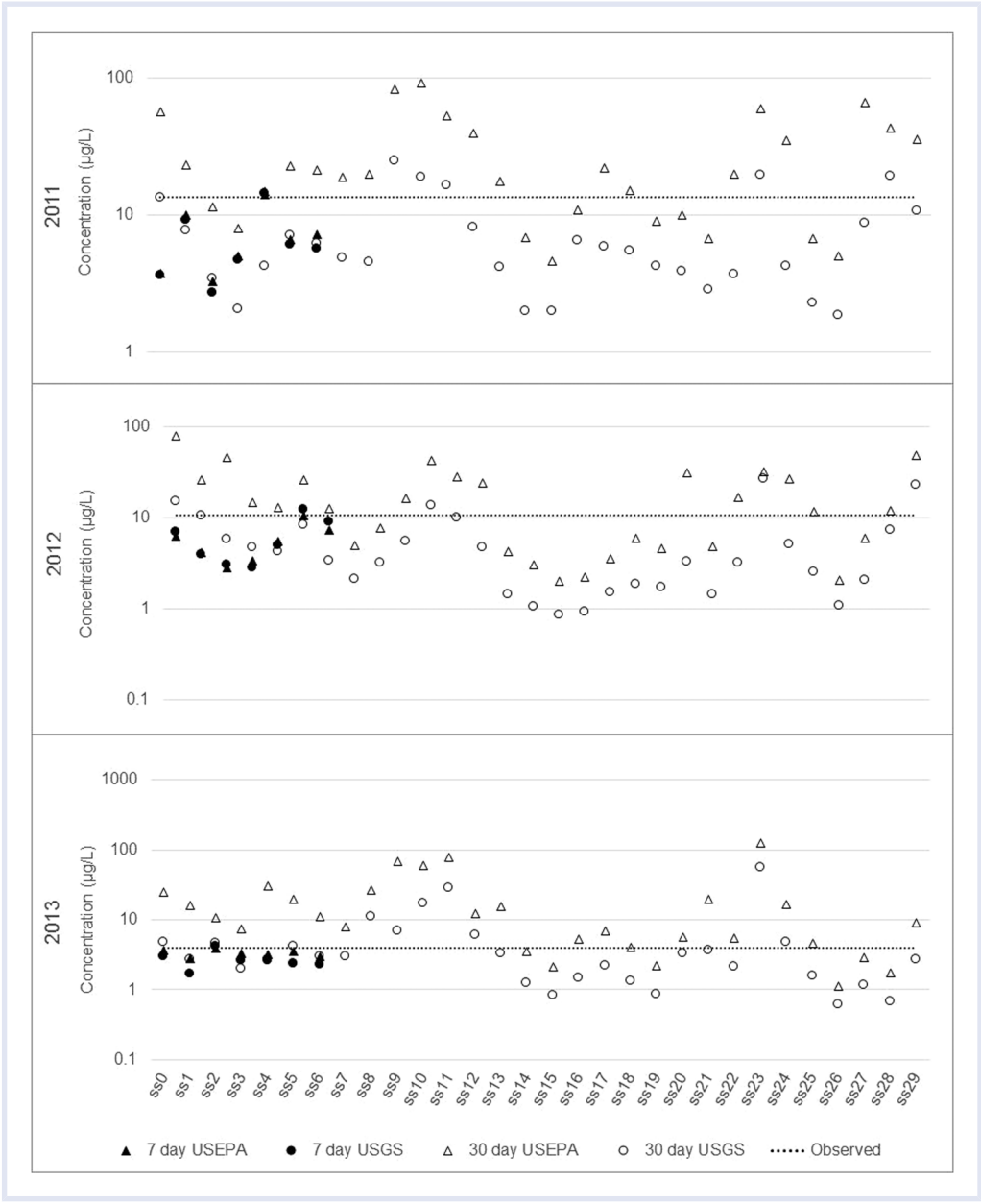

FIGURE 6 Estimates of the 1-day annual maximum concentration for isoproturon using US Geological Survey's (USGS's) and USEPA's approaches, for the 7-day (subsamples ss0–ss6) and 30-day (subsamples ss0–ss6) intervals

yielded results orders of magnitude above observed concentrations. For DFF, SEAWAVE–QEX results had interannual variations of under- or overestimations of 1-day annual maximum concentrations, with results within an order of magnitude of observed results. Application of SEAWAVE–QEX to IPU subsampled datasets resulted in consistent underestimation regardless of method or sampling frequency for two out of three years evaluated. Thus, the results presented and discussed above revealed that SEAWAVE–QEX cannot reliably predict peak concentrations in the GKb watershed. This is especially true for compounds whose concentrations are surface-runoff-driven and where occasional nonrunoff-driven events are present (e.g., events driven by drift or point sources). The evaluation of the SEAWAVE–QEX results could not determine a systematic pattern of under- or overestimation of peak concentrations. For FFA, SEAWAVE–QEX tended to overestimate peak concentrations whereas the opposite was found for IPU.

However, for metabolites, whose concentrations are usually driven by subsurface processes, the results from the FFA-SA evaluation suggest that SEAWAVE–QEX would yield reasonable estimates of daily concentrations and annual maximum concentrations. For similar compounds, SEAWAVE–QEX might be a valuable tool for environmental managers and could replace high-frequency monitoring programs.

Estimates of the target quantity vary significantly between USEPA's and USGS's postprocessing approaches to the pesticides studied, particularly at a lower sampling frequency, that is, monthly. In this study, a percentile-based approach (USEPA's approach) was more robust than the USGS's approach to estimating 1-day annual maximum concentrations when there is a possibility for unrealistically high results (e.g., FFA and DFF). However, the 99th percentile used in USEPA's approach did not yield better estimates for IPU and FFA-SA. Further, USGS's approach performed significantly better with the 30-day sampling frequency. A robust SEAWAVE–QEX postprocessing method for extracting annual maximum exposure concentrations is critical to establishing SEAWAVE–QEX as a tool in pesticide risk assessments.

CONCLUSIONS

This study evaluated the use of SEAWAVE–QEX to estimate annual maximum concentrations of three commonly used herbicides and one degradation product in a Belgian catchment with high agricultural land use. Results from this study demonstrate that using SEAWAVE–QEX to estimate annual maximum concentrations or to fill gaps in daily pesticide concentration time series should be approached with caution. The appropriateness of using SEAWAVE–QEX to estimate daily pesticide concentration series depends on pesticide characteristics, application, and probably other factors that were not analyzed in this study. Drift or point-source-driven events (e.g., spillage on hard surfaces during filling or wash-off from rinsing spraying equipment, unreported or misuse of the product) can be a major problem for SEAWAVE–QEX performance. In this context, an uncalibrated mechanistic model such as SWAT can provide more confident results, which has the additional advantage of obtaining a rigorous understanding of the chemical entry paths and can be set up to account for drift (Winchell, Pai, et al., 2018). A model framework combining a statistical approach (e.g., SEAWAVE–QEX) with an uncalibrated mechanistic model (e.g., SWAT as parameterized in Winchell et al., 2017) could leverage the advantages of both approaches.

Results from this study suggest that the appropriate SEAWAVE–QEX postprocessing approach might be compound- and/or use-pattern-specific. Further research is needed to investigate if use-pattern- and compound-specific appropriate percentiles exist that result in reasonable estimates of annual maximum concentrations across compounds, watersheds, and geographical regions.

Based on the evaluation of the sampling frequency, and the postprocessing approaches, it was concluded that SEAWAVE–QEX should not be used in the GKb watershed to estimate 1-day annual maximum concentrations from a weekly or sparser monitoring dataset, except for the mobile metabolite FFA-SA. A possible approach to assessing whether a SEAWAVE–QEX application is appropriate could be based on running the model on multiple subsamples of the available monitoring data and evaluating how stable the SEAWAVE–QEX results are, both within the different modeled results and compared with the observed data. For stable results (as seen for FFA-SA in this study), a SEAWAVE–QEX application is likely appropriate. Further research to develop universally applied criteria when SEAWAVE–QEX is likely to provide realistic estimates of annual maximum concentrations is needed before the tool is used by environmental managers and risk assessors.

The evaluation of SEAWAVE–QEX on FFA monitoring data with and without including FFA point source events revealed that removing point source events from the monitoring data improved the FFA SEAWAVE–QEX estimates. However, SEAWAVE–QEX cannot be used to identify point sources because it is not clear whether pesticide detections that do not co-occur with a runoff or precipitation event are caused by a point source or drift event.

More years of monitoring data would help increase confidence in the results obtained in this evaluation. Further research is needed to determine if the findings of this study apply to other scenarios (different size and characteristic watersheds and availability of monitoring data) or are limited to small-scale, high-intensity agricultural watersheds for which a relatively short period of monitoring data is available. Further research is also needed to identify whether specific catchment characteristics (e.g., size, landscape, climate, agronomic practices) can be used to predict the relative performance of the SEAWAVE–QEX model. All of these results are tied to the GKb catchment, and additional research is needed to assess whether the results are representative of other agriculturally intensive catchments in central Europe.

ACKNOWLEDGMENT

No additional assistance from third parties was received by the authors for the preparation of this manuscript. This work was supported by Bayer AG, Crop Science division.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Hendrik Rathjens: Conceptualization; Methodology; Project administration; Writing—review & editing. **Maria Bettina Miguez**: Data curation; Formal analysis; Methodology; Visualization; Writing—original draft. **Michael Winchell**: Conceptualization; Supervision. **Robin Sur**: Conceptualization; Data curation; Funding acquisition; Supervision.

DATA AVAILABILITY STATEMENT

SEAWAVE–QEX outputs are available upon request from corresponding author Maria Bettina Miguez (bmiguez@stone-env.com). Monitoring data in the GKb watershed is the property of Bayer AG, Crop Science Division and dissemination is contingent on the approval of Bayer AG's legal team. Thus, monitoring data will be made available upon request on a case-by-case basis to be decided by Bayer AG Crop Science Division.

 DOI: 10.1002/ieam.4688

SUPPORTING INFORMATION

Figure A.1. Hydrograph at GKb outlet.

Figure A.2. Chemograph for flufenacet at GKb outlet, with point sources identified.

Figure A.3. Chemograph for flufenacet sulfonic acid at GKb outlet.

Figure A.4. Chemograph for diflufenican at GKb outlet.

Figure A.5. Chemograph for isoproturon at GKb outlet.

ORCID

Hendrik Rathjens https://orcid.org/0000-0002-7152-9767
Maria Bettina Miguez http://orcid.org/0000-0002-0140-0247
Robin Sur http://orcid.org/0000-0002-3403-8168

REFERENCES

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large-area hydrologic modeling and assessment: Part I. Model development. *Journal of the American Water Resources Association*, *34*(1), 73–89.

Baets, D., Sur, R., Krebber, R., & Lembrich, D. (2019). High-resolution water monitoring program gives further insights on sources of residues from herbicides in surface water. *Communications in Agricultural and Applied Biological Sciences*, *83*, 326–335.

European Food Safety Authority. (2007). *Conclusion regarding the peer review of the pesticide risk assessment of the active substance diflufenican (EFSA Scientific Report 122)*. Finalized: 17 December 2007, version: 11 February 2008.

European Food Safety Authority. (2022). *Ongoing. EFSA peer review (EFSA Conclusion) on the active substance flufenacet according to Article 13 of Regulation (EU) No. 844/2012*.

European Union (EU). (2002). *Review report for the active substance isoproturon* (Document No.: EU Doc. Ref. SANCO/3045/99-final).

Larson, S. J., & Gilliom, R. J. (2001). Regression models for estimating herbicide concentrations in US streams from watershed characteristics. *Journal of the American Water Resources Association*, *37*, 1349–1367. https://doi.org/10.1111/j.1752-1688.2001.tb03644.x

Moriasi, D. N., Arnold, J. G., van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Mosquin, P. L., Aldworth, J., & Chen, W. (2018). Evaluation of the use of bias factors with water monitoring data. *Environmental Toxicology and Chemistry*, *37*, 1864–1876. https://doi.org/10.1002/etc.4154

Rathjens, H., Kiesel, J., Miguez, M. B., Winchell, M., Arnold, J. G., & Sur, R. (2022). Simulation of pesticide and metabolite concentrations using SWAT+ landscape routing and conditional management applications. *Water*, *14*(9), 1332. https://doi.org/10.3390/w14091332

Runkel, R. L., Crawford, C. G., & Cohn, T. A. (2004). *Chapter A5: Load estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers* (pp. 1–69). US Geological Survey Techniques and Methods, Book 4.

Sharifi, A., Yen, H., Wallace, C. W., McCarty, G., Crow, W., Momen, B., Lang, M. W., Sadeghi, A., Lee, S., Denver, J., & Rabenhorst, M. C. (2017). Effect of water quality sampling approaches on nitrate load predictions of a prominent regression-based model. *Water*, *9*(11), 895. https://doi.org/10.3390/w9110895

Stone, W. W., Crawford, C. G., & Gilliom, R. J. (2013). Watershed regressions for pesticides (WARP) models for predicting stream concentrations of multiple pesticides. *Journal of Environmental Quality*, *42*, 1838–1851. https://doi.org/10.2134/jeq.2013.05.0179

Sur, R., Winchell, M., Rathjens, H., Baets, D., Krebs, F., & Lembrich, D. (2018). Identification of herbicide source areas and exposure pathways in a watershed based on landscape modeling and high-resolution monitoring data. *Communications in Agricultural and Applied Biological Sciences*, *83*, 336–340.

US Environmental Protection Agency (USEPA). (2019a). Approaches for quantitative use of surface water monitoring data in pesticide drinking water assessments (EPA-HQ-OPP-2019-0417-0003).

US Environmental Protection Agency (USEPA). (2019b). Supplemental documents for November 19–22, 2019 SAP meeting. Retrieved July 25, 2022, from: https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/supplemental-documents-november-19-22-2019-sap

Vecchia, A. V. (2018). *Model methodology for estimating pesticide concentration extremes based on sparse monitoring data* (US Geological Survey. Scientific Investigations Report 2017–5159). Reston, VA. https://doi.org/10.3133/sir20175159

Willkommen, S., Pfannerstill, M., Ulrich, U., Guse, B., & Fohrer, N. (2019). How weather conditions and physico-chemical properties control the leaching of flufenacet, diflufenican, and pendimethalin in a tile-drained landscape. *Agriculture, Ecosystems & Environment*, *278*, 107–116. https://doi.org/10.1016/j.agee.2019.03.017

Winchell, M. F., Pai, N., Brayden, B. H., Stone, C., Whatling, P., Hanzas, J. P., & Stryker, J. J. (2018). Evaluation of watershed-scale simulations of in-stream pesticide concentrations from off-target spray drift. *Journal of Environmental Quality*, *47*(1), 79–87. https://doi.org/10.2134/jeq.2017.06.0238

Winchell, M. F., Peranginangin, N., Srinivasan, R., & Chen, W. (2018). Soil and water assessment tool model predictions of annual maximum pesticide concentrations in high vulnerability watersheds: SWAT predictions of annual maximum pesticide concentrations. *Integrated Environmental Assessment and Management*, *14*, 358–368. https://doi.org/10.1002/ieam.2014