Environment International 208 (2026) 110057

Contents lists available at ScienceDirect

Environment International

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

Full length article

Confronting pesticide exposure predictions from different models to observations from a monitoring study in small freshwater streams in Germany☆

Paula Scharlach [a], Gustaf Boström [b,c], Jörg Klasmeier [a], Amelie Leonardi [a], Andreas Focks [a,*]

[a] Institute of Mathematics and Institute of Environmental Systems Research, Osnabrück University, Barbarastraße 12, 49076 Osnabrück, Germany
[b] Department of Aquatic Sciences and Assessment, Division of Environmental Organic Chemistry and Ecotoxicology, Swedish University of Agricultural Sciences, P.O. Box 7050, SE-750-07 Uppsala, Sweden
[c] SLU Centre for Pesticides in the Environment, Swedish University of Agricultural Sciences, P.O. Box 7066, SE-750 07 Uppsala, Sweden

ARTICLE INFO

Keywords:
Pesticides
Fate models
Exposure
Environmental risk assessment
Surface water
Model evaluation

ABSTRACT

Plant protection products are integral to European agriculture but can cause unwanted environmental impacts. Before authorisation, predicted concentrations in environmental compartments are compared with effect thresholds in a regulatory risk assessment. This study evaluates the agreement between predicted and measured concentrations for the established FOCUS surface water models (Steps 1–3) and the recently published PEC-CKB model. Model results were compared with monitoring data from lowland streams in Germany, and particular attention was paid to the models' conservatism. The conservative character of FOCUS Step 1 can be confirmed, but underestimations were observed for FOCUS Step 2 and 3 models. PEC-CKB results are similar to those of the higher-tier FOCUS models, while having lower model complexity and requiring less input data. Using real application rates and landscape information generally improved model predictions by nearly halving the bias, but led to increased underestimations of measured concentrations. Linking prospective and retrospective environmental risk assessment (ERA) by incorporating real data can make prospective ERA more realistic and identify opportunities for simplification. Finally, we discuss the challenges in evaluating prediction models for pesticide concentrations in surface waters, particularly with regard to the environmental variability of measured concentrations.

1. Introduction

The use of plant protection products (PPPs) is integral to European agriculture (Alix and Capri, 2018) and contributes to food security by preventing pest-related crop losses (Carvalho, 2006; Tudi et al., 2021). However, PPP application in agricultural areas is associated with their transport into adjacent ecosystems such as surface water bodies via spray drift, drainage or runoff (Boye et al., 2012; de Souza et al., 2020; Schwarzenbach et al., 2006; Reichenberger et al., 2007). There, they may affect aquatic organisms, such as invertebrates or fish, and in turn threaten biodiversity and the ecological integrity of ecosystems (de Souza et al., 2020; Liess et al., 2021a; Schäfer et al., 2011; Stehle and Schulz, 2015a). For market authorisation, environmental risks of PPPs are subject to a prospective regulatory assessment, where pesticides are evaluated concerning their potential for environmental exposure and effects. Due to the lack of field data, exposure models with various levels of complexity that predict the occurrence and magnitude of pesticides in ecosystems have developed into an essential element within the European regulatory environmental risk assessment (ERA) (Di Guardo et al., 2018; MacLeod et al., 2010, FOCUS, 2001; Centanni et al., 2023; Fox et al., 2021). The ERA relies on the comparison of model-based estimates to effect thresholds (Forbes et al., 2009; Rico et al., 2021; Schmolke et al., 2010) and follows a tiered approach, starting from conservative and simple estimations and advancing to more realistic and complex approaches in higher tiers (FOCUS, 2001). The *FOrum for the Co-ordination of pesticide fate models and their Use* (FOCUS) agreed some

☆ This article is part of a special issue entitled: 'B-SAFE: Science for ERA' published in Environment International.

* Corresponding author.
E-mail address: andreas.focks@uni-osnabrueck.de (A. Focks).

https://doi.org/10.1016/j.envint.2026.110057
Received 12 March 2025; Received in revised form 8 January 2026; Accepted 8 January 2026
Available online 9 January 2026

decades ago upon different models and scenarios for predicting pesticide concentrations in surface waters (FOCUS, 2001). FOCUS surface water Step 1 uses basic assumptions to estimate extreme worst-case pesticide loadings. FOCUS surface water Step 2 refines transport processes, including degradation and differentiation by region, application season, and cultivated crop. FOCUS surface water Step 3 and Step 4 use a time-dynamic differential equations-based model and consider ten realistic worst-case scenarios representing European agricultural areas (FOCUS, 2001). If a PPP fails the risk assessment at lower tiers, it may be authorised if acceptable risk ratios are reached at higher tiers.

Several studies have claimed that the actual regulatory ERA does not protect the aquatic environment (Knäbel et al., 2012, 2014; Fulda et al., 2015; Stehle and Schulz, 2015b) because measured environmental concentrations (MECs) of PPPs exceed exposure predictions of the FOUCS models. Such underestimations of MECs challenge the use of exposure models and the effectiveness of a protective ERA for pesticides (Knäbel et al., 2012, 2014).

Recently, a new model for predicting environmental pesticide concentrations in surface waters, called PEC-CKB, has been introduced (Boström et al., 2019). It relies on a single equation and omits scenario assumptions, which simplifies its application compared to FOCUS models (Boström et al., 2019). PEC-CKB predictions matched observations from Swedish monitoring data similarly well as complex FOCUS Step 3 predictions (Boström et al., 2019). It remains open whether this performance applies to other European regions with different climate and topography.

The desired status of models for predicting environmental pesticide concentrations in a regulatory context is to avoid underestimation without overestimating too much. Models should account for the primary drivers of exposure, and complexity should be kept to a minimum, adhering to the principle of parsimony. In this sense, alternative methods with varying degrees of complexity for use in prospective ERA can benefit from comparisons with monitoring data, while addressing aspects such as accuracy, protectiveness, realism, complexity, and relevance of the methods (Axelman et al., 2024).

In this study, we evaluated FOCUS surface water models for tiers 1 to 3, and the PEC-CKB model for 36 PPPs against MECs from an event-driven pesticide monitoring in small German streams (Liess et al., 2021b). We focus on event-driven samples because they represent high but realistic entries (Halbach et al., 2021). Special attention was paid to the agreement between MECs and model predictions and the number of underestimations. In addition, we conducted supplementary analyses restricted to grab water samples, which reflect the approach of most governmental monitoring. Furthermore, we investigated how the simple PEC-CKB model relates to the established tiered exposure assessment methods. We also assessed the impact of landscape-specific information, in particular the quality of application data, on the comparisons with MECs using the PEC-CKB model for an extended set of 59 active substances. Finally, we discuss the relationship between model performance and complexity in regulatory risk assessment, as well as the possible consequences for an improved European ERA.

2. Materials and methods

Predicted environmental concentrations (PECs) were calculated using the FOCUS surface water Step 1, 2, and 3 models and the PEC-CKB model (Section 2.1) with realistic application rates for Germany extracted from the PAPA database (Section 2.2). The 90th percentile MECs from a German monitoring study (Section 2.4) were compared to PECs per substance for each model (Section 2.5). In addition, PEC-CKB values were calculated considering crop-specific application data and refined assumptions about the treated area in the catchments based on landscape analyses (Section 2.3).

2.1. Calculating PEC values with different models

2.1.1. Performing FOCUS calculations

European ERA follows a tiered model approach, with simplifying and conservative assumptions in the lowest tier and increasing realism and complexity for every higher tier. In the FOCUS surface water models, Step 1 results in 'worst-case' water and sediment concentrations based on summarised transport processes and crop-specific spray-drift rates. Step 2 refines the model predictions by considering a series of individual loads, discharge events occurring four days after the last application and simple degradation mechanisms (FOCUS, 2001). Both steps align with the Tier 1 and Tier 2 risk assessment guidelines outlined in Council Directive 91/414/ECC, now superseded by Regulation (EC) No. 1107/2009. We performed model calculations with the Steps-1–2 model (version 3.2) (FOCUS, 2001). Here, we selected the Northern Europe scenario and the application season from March to May for all Step 1 and 2 calculations because this period appeared most suitable for the relevant monitoring period (see SI, Section D).

The third Tier of FOCUS surface water modelling comprises a dynamic model with ordinary differential equations. It can be seen as a '[d]eterministic estimate of aquatic exposure across [the defined] scenarios' (FOCUS, 2001, p. 17, Fig. 1.3-1) and is supposed to deliver 'realistic worst-case' concentrations. The inclusion of ten scenarios for Europe aims to represent different European environmental conditions and to yield at least the 90th percentile of the highest pesticide loads in the EU (FOCUS, 2001). Additional mechanistic models are employed to facilitate more realistic predictions for pesticide input to surface waters via drainage, surface runoff and spray drift. We performed model calculations for FOCUS Step 3 with the software shell SWASH 5.3, which included TOXSWA 5.5.3, PRZM 4.3.1, and MACRO 5.5.4. Because MECs were derived from streams in Germany, we selected three of the ten FOCUS surface water scenarios as the most representative for Germany: two drainage scenarios (D3 and D4) and one run-off scenario (R1). The selection was made based on comparisons with environmental conditions in Germany and after consultation with FOCUS surface water model developers. Although the monitoring study does not provide precise information on the contribution of drainage at the different monitoring sites, we included D3 and D4 to cover drainage and run-off. Both fast transport mechanisms can be important depending on local conditions in the catchments (Leu et al., 2004). If defined for the corresponding crops, PECs were calculated for all three scenarios, but we excluded pond scenarios because monitoring data was available only for small and medium-sized streams. FOCUS Step 4 can be performed with the same software and additionally includes mitigation options, but will not be considered in the present study due to lacking data on mitigation strategies during monitoring. A more detailed comparison of the models can be found in SI, Table G.1.

2.1.2. PEC-CKB model and parameter assumptions

The PEC-CKB model consists of a single equation that estimates PECs for PPPs in surface water, briefly defined as (Boström et al., 2019):

$$PEC_{CKB} = (f/(\mathrm{N_s} \bullet \mathrm{q})) \bullet D \bullet M \bullet F_w \quad (1)$$

PECs are obtained by multiplying the annual application rate $D\,[\mathrm{mg/m^2}]$ with a dilution factor, which is built by the proportion of the catchment area that is annually sprayed $f\,[-]$ divided by the duration of the spraying period N_s $[weeks]$ and the weekly runoff factor $q\,[m/week]$. Transport processes are combined into a generic factor $M\,[-]$, representing the proportion of the applied dose that ends up in surface water via different pathways, such as drainage or runoff. To account for the specific sorption behaviour of the substance, the factor $F_w\,[-]$ calculates the fraction of the substance mass which partitions to water rather than sediment, assuming an adsorption equilibrium by

$$F_w = Z_w/(Z_w + (Z_s \bullet \gamma \bullet K_{oc} \bullet f_{oc})) \quad (2)$$

where Z_w represents the water depth $[m]$, Z_s the thickness of adsorbing sediments $[m]$, γ the bulk density $[g/cm^3]$, f_{oc} the organic carbon content $[kg/kg]$ and K_{oc} the substance-specific partitioning constant between organic carbon and water $[L/kg]$. Based on the FOCUS assumptions $Z_w = 0.3\,m$, $Z_s = 1cm$, $\gamma = 0.8\,g/cm^3$ and $f_{oc} = 0.05\,kg/kg$ and further parameter assumptions $f = 20\%$, $N_s = 4\,weeks$, $q = 0.001\,m/week$ and $M = 1\,\%$ from Boström et al., the PEC-CKB Eq. (1) can be simplified to

$$PEC = 0.5 \bullet D \bullet F_w \tag{3}$$

PECs from this simplified model equation matched well with monitoring data from freshwater samplings in four small Swedish agricultural catchments (Boström et al., 2019). We used Eq. (3) for comparison with German monitoring data and refined calculations based on land-use analysis, specifically by using more realistic *f*-values.

2.1.3. Model comparisons and input parameters

FOCUS Step 1 and PEC-CKB models share a similar structure but differ in the parametrisation of transport and dilution. PEC-CKB uses a generic loss rate of 1%, while FOCUS Step 1 includes crop-specific spray drift rates (FOCUS, 2001) and a higher loss rate of 10% through runoff and drainage. This results in PECs from Step 1 being approximately one order of magnitude higher than those of PEC-CKB (for details see SI, Section G).

FOCUS Step 2 requires degradation and crop-specific data, while FOCUS Step 3 needs additional substance properties. This information was taken from the EFSA conclusions to follow the standard assumptions in the prospective ERA (input parameters are provided in SI, Section D). If the required information was not available from these reports, default parameters were used instead. A substance database that can be added directly to SWASH is attached to this study. Aggregated German application data from the PAPA database (see Section 2.2) was used for model calculations. Other settings were kept as the default.

2.1.4. Aggregation of PEC values

From all four models, crop-specific PEC values were calculated using crop-specific application rates and, in the case of FOCUS models, crop-specific transport parameters (especially spray drift). As the concentrations measured in the monitoring study cannot be assigned to individual, crop-specific applications, we calculated average PEC values for each substance by averaging all the crop-specific PEC values. This is because we aim to cover single PPP applications for different crops in Germany in order to compare the resulting PECs to monitoring data. We also extracted maximum and minimum PEC values for each substance to account for the extremes of crop-specific application rates and risk of PPP transport to surface waters. For Step 3, the combination of the selected scenarios with varying crop-specific application rates across substances yielded a set of PECs for each substance. To calculate the quality metrics, we used the maximum PEC values across the three scenarios D3, D4 and R1 for each crop-specific application rate, as common in regulatory ERA.

2.2. Application data

Application data from the monitoring study was not available at the time of our study, so application rates were derived from average German substance- and crop-specific application data from various example farms across the country. This average German application data is provided in the PAPA database (JKI, 2023; cf. Roßberg et al., 2017). Based on the EU pesticide review reports from the EU Pesticides Database (EU, 2023) and available application data, the crops were categorised into different groups (see SI, Table B.1). We calculated average crop-specific application data (according to the defined groups) for each substance from the PAPA database for the years 2018 and 2019 and used them for the PEC calculations (see SI, Table B.2). The application seasons, e.g. BBCH codes, were taken from EFSA conclusions for the corresponding crop groups (see SI, Section D). We only used single applications, as the aggregation of data from different farmers hampers to infer the number of applications.

2.3. Land-use analysis of catchments

PPPs are approved and applied to specific crops. Because the monitoring catchments have different crop compositions, we used a geographic information system (GIS) analysis to estimate the percentages of each catchment area sprayed with pesticides. To derive such more realistic *f*-values, we determined the sizes of the different catchments based on a freely available watercourse network (DLM 250) (Bundesamt für Kartographie und Geodäsie, 2023). The catchment areas were intersected with land-use data provided by Blickensdörfer et al. (2021), and the agricultural crop cover of the defined crop groups (SI, Table B.1) was derived for each catchment (see detailed information in SI, Section H). Next, average proportions of land use across the different catchments were calculated by averaging all generated, individual *f*-values for each crop group. In a worst-case scenario, it is assumed that crops are treated with the same pesticide, but in general, the use of different PPPs is expected in one catchment, so average values provide a good approximation. The resulting *f*-values were used to calculate crop-specific PEC-CKB values, which now incorporate crop-specific application rates and *f*-values:

$$PEC_{CKB,crop} = (f_{crop}/(N_s \bullet q)) \bullet D_{crop} \bullet M \bullet F_w \tag{4}$$

We also calculated an additional overall average *f*-value by averaging all crop-specific *f*-values to adapt the simplified PEC-CKB equation (Eq. (3)) to German conditions, accounting for crop coverage within catchments.

2.4. German monitoring dataset (KgM)

We compared PECs with MECs from a lowland stream monitoring campaign (Kleingewässermonitoring, KgM) encompassing 124 different stream sections in Germany (Liess et al., 2021b). Samples were collected from April to July 2018 and 2019, through regular grab samples taken every three weeks, and event-driven samples, triggered by a water level rise of at least 5 cm (Liess et al., 2021a). We assume that substances detected in very low concentrations might not have been applied in the corresponding catchment in the monitoring period but might result from earlier applications, so we decided to exclude values below the limit of quantification (LOQ) from our analysis.

To assess the models' ability to avoid underestimating high MECs, we selected event-driven concentration measurements for the model comparisons, as increased pesticide concentrations in streams after rainfall have been reported previously (Halbach et al., 2021). Pooling of data from both sampling methods was inappropriate because the distribution of concentrations differed. We also performed the whole analysis with grab samples to check whether the sampling method or scenario definition influences the results (see SI, Table E.2). We used the 90th percentile and maximum values above the LOQ to represent higher MECs.

Of the 75 detected active substances, we selected 36 for which we had all the required input parameters for the PEC calculations (see Section 2.1; SI, Section A). These substances have different degradation rates and K_{oc} values (see SI, Fig. A.1), which ensures that our analyses include substances with different fate, accumulation and transport behaviours in the environment. 23 additional substances from the KgM study were added for further comparisons with the PEC-CKB model.

2.5. Quality measures for comparison of MECs and PECs

We compared maximum and 90th percentile MECs to average PECs derived from the different models and crop-specific application rates for each substance to evaluate model performance with respect to over- and

underestimations. Good agreement between predicted and measured values is indicated by data points smoothing around the 1:1 line in a MEC-PEC plot and is quantitatively assessed by a small bias and root mean square error (RMSE). Bias reveals systematic over- or underestimation of MECs, while RMSE measures the overall model performance and the typical error magnitude. We also compare the PEC/MEC ratios, with values closer to 1 indicating better concordance between PECs and MECs, to better understand heteroscedasticity. If predictions are correct on average, typical quality measures such as bias or RMSE tend to be low, indicating good average agreement between PECs and MECs because overestimates and underestimates equalise. Additionally, the coefficient of determination R^2, calculated from the Pearson correlation between PECs and MECs, is used to assess how well PECs explain MEC variability. Unlike RMSE, R^2 does not require a 1:1 agreement, but a low R^2 could indicate an inappropriate model approach (e.g. incomplete process descriptions) or inadequate input data.

However, from a risk assessment perspective, PECs should not be lower than MECs, particularly those exceeding effect thresholds. In consequence, relying solely on measures like bias, RMSE or similar is inappropriate for the assessment of model quality in isolation, and the number and magnitude of underestimations are also relevant. We evaluate the conservative character of the models by counting the underestimations of MECs by PECs. This evaluation is independent of any regulatory framework or protection goal, where effect values should be considered. We define four safety groups to identify models that avoid both underestimation and excessive conservatism: (1) underestimations – PEC is lower than MEC; (2) “safety margin 10” – PEC exceeds MEC by one order of magnitude; (3) “safety margin 100” – PEC exceeds MEC by one to two orders of magnitude and (4) all further overestimations. In this context, the heteroscedasticity and the model's capacity to explain variability in MECs also matter, so results from the ratio and correlation analyses can be considered here as well.

Model quality measures were calculated using the decadic logarithm of all concentration values to facilitate comparison to the calculations reported in Boström et al. (2019). To account for varying application rates, we assessed model performance using bias and RMSE with average, minimum and maximum PEC values. We focused on 90th percentile MECs and average PECs for MEC-PEC agreement, and also considered the maximum MECs to assess conservatism. Furthermore, we benchmarked the PEC-CKB models against the FOCUS models by comparing Pearson correlation coefficients and examining the influence of input parameters on the results.

3. Results

First, we focused on the overall model performance of the three FOCUS models and the PEC-CKB model by MEC-PEC comparisons. Next, we directly compared the PEC-CKB predictions to the FOCUS models in order to group the PEC-CKB into the tiered system of exposure models in ERA. Finally, we adapted the PEC-CKB assumptions based on landscape-specific information and assessed the impact on MEC-PEC agreements.

3.1. Evaluation of the model performance

3.1.1. PEC-CKB predictions range between FOCUS Step 2 and Step 3

Predictions of environmental concentrations for 36 substances deviated considerably from 90th percentile MECs for all four tested models. The PEC-CKB performed better than the FOCUS Step 1 and Step 2, while the more complex FOCUS Step 3 showed the best MEC-PEC agreement with the lowest bias and RMSE, but also revealed one underestimation (see Table 1, Fig. 1). The positive bias across all cases indicates that the models systematically overestimate the measured concentrations (Table 1).

Overestimation was most pronounced for low MECs, whereas the MEC-PEC agreement improved as MEC values increased (Fig. 1 and SI, Fig. E.1). The distribution of ratios between average PECs and 90th percentile MECs for the four models confirms this pattern (Fig. 2). The median ratio decreases progressively from Step 1 to Step 3. For PEC-CKB, the median ratio ranges between FOCUS Step 2 and 3, while the variability of ratios is lower than the ratios of FOCUS Step 3 (Fig. 2 and SI, Fig. E.1).

Based on the bias and RMSE reductions achieved with each subsequent FOCUS step and with PEC-CKB in between Step 2 and Step 3, we conclude that MEC-PEC agreement improves in the tiered ERA. However, the correlation coefficients and corresponding R^2 values contradict this order. PEC-CKB showed the highest correlation ($r = 0.31$, $p = 0.07$) compared with the FOCUS models (Step 1: $r = 0.27$; Step 2: $r = 0.22$; Step 3: $r = 0.16$). Still, none of these correlations were statistically significant (Table 2). This changes when comparing maximum MECs with average PECs. Here, R^2 values increase, indicating a better explanation of maximum MECs by PECs (see Table 2).

3.1.2. Only FOCUS Step 3 reveals underestimations of 90th percentile MECs

For both FOCUS Step 1 and Step 2, all average PECs, as well as each individual PEC (derived from different application rates), exceed the corresponding 90th percentile MECs (Fig. 1). In FOCUS Step 1, the PECs exceed MECs by more than two orders of magnitude, so none of the results entered the ‘Safety margin 10′ group (Table 1). FOCUS Step 2 and PEC-CKB values also exceed the 90th percentile MECs for all 36 substances, and the MEC-PEC tuples are distributed almost evenly across the defined overestimation safety groups (Fig. 1, Table 1). By contrast, FOCUS Step 3 shows a single underestimation for florasulam, and crop-specific PECs underestimate MECs for two substances (dichlorprop-P and florasulam) (Fig. 1). When the three FOCUS surface-water scenarios are examined separately, up to five underestimations become apparent (Table 1 and SI, Table E.1).

To examine the conservative character for particularly high concentrations, we also compared the maximum MECs to the maximum PECs (SI, Table E.1). Taking this view, only the FOCUS Step 1 model results consistently exceed MECs, while Step 2 underestimates four of 36 maximum MEC values and Step 3 underestimates six MEC values. The PEC-CKB results underestimate the maximum MECs for only three of the 36 substances (SI, Table E.1).

Similar results were obtained when we considered only grab samples for MEC-PEC comparisons. The 90th percentile MECs decrease slightly, which modestly worsens the agreement between MECs and PECs. FOCUS Step 3 also shows a single underestimation, but the drainage scenarios show fewer underestimations compared to analyses with event-driven samples (see SI, Table E.2).

3.1.3. PEC-CKB shows best overall model performance

Taking all results into account, the PEC-CKB model shows the best overall performance among the prediction models. It aligns more closely with measured concentrations than the lower-tier FOCUS models and does not produce any underestimation of the 90th percentile MECs. Only FOCUS Step 3 shows a single underestimation of the 90th percentile MECs for 36 substances.

3.2. Comparing PEC-CKB predictions with FOCUS models

The previous results on model performance according to MECs show that the results of PEC-CKB lie between those of FOCUS Step 2 and Step 3, but PEC-CKB results correlate strongest with FOCUS Step 1 values while being at least one order of magnitude lower (Fig. 3). The high correlation coefficient indicates a strong linear association between PEC-CKB and FOCUS Step 1 (Fig. 3). This can be explained by both models' similar equation structure, while the difference in magnitude is due to the different parameterisation (see Section 2.1.3; SI, Section G). The minimal reduction of the correlation coefficient is caused by the additional crop-specific spray-drift input in FOCUS Step 1. The deviation between PEC-CKB and FOCUS Step 1 of about one order of magnitude

Table 1
Quality measures of MEC-PEC comparisons. We calculated the bias and RMSE as quality measures for different models (and adaptations) and representative MEC values (90th percentiles and maximum). Substances that over- or underestimate the MEC values by a certain amount (up to one, two or more orders of magnitude) are counted as a benchmark for MEC-PEC agreement and conservatism. Attention should be paid to different substance numbers (n) included in this analysis.

model (and scenario)	n	PEC selection	90th percentile MEC						Max. MEC					
			bias	RMSE	Under-estima-tions	Overestimations…			bias	RMSE	Under-estima-tions	Overestimations…		
						up to one order of magnitude *"Safety margin 10"*	up to two orders of magnitude *"Safety margin 100"*	by more than two orders of magnitude				up to one order of magnitude *"Safety margin 10"*	up to two orders of magnitude *"Safety margin 100"*	by more than two orders of magnitude
PEC-CKB standard f = 20%	36	average	1.6	1.78	0	10 (27.8%)	14 (38.9%)	12 (33.3%)	0.88	1.18	3 (8.3%)	20 (55.6%)	10 (27.8%)	3 (8.3%)
FOCUS Step 1	36	average	2.52	2.63	0	0	10 (27.8%)	26 (72.2%)	1.79	1.96	0	7 (19.4%)	17 (47.2%)	12 (33.3%)
FOCUS Step 2	36	average	1.78	1.96	0	9 (25.0%)	12 (33.3%)	15 (41.7%)	1.06	1.35	4 (11.1%)	16 (44.4%)	10 (27.8%)	6 (16.7%)
FOCUS Step 3	36	average	1.41	1.65	1 (1.7%)	10 (27.8%)	15 (41.7%)	10 (27.8%)	0.68	1.12	7 (19.4%)	15 (41.7%)	10 (28.6%)	4 (11.4%)
FOCUS Step 3 D3	35*	max	1.34	1.69	5 (14.3%)	9 (25.7%)	11 (31.4%)	10 (28.6%)	0.63	1.21	11 (31.4%)	10 (28.6%)	10 (28.6%)	4 (11.4%)
FOCUS Step 3 D4	35*	max	1.28	1.65	5 (14.3%)	10 (28.6%)	10 (28.6%)	10 (28.6%)	0.57	1.19	11 (31.4%)	10 (28.6%)	12 (34.3%)	2 (5.7%)
FOCUS Step 3 R1	36	max	1.54	1.78	0	12 (33.3%)	11 (30.6%)	13 (36.1%)	0.82	1.21	7 (19.4%)	12 (33.3%)	13 (36.1%)	4 (11.1%)
PEC-CKB standard f = 20%	59	average	1.47	1.66	1 (1.7%)	17 (28.8%)	26 (44.1%)	15 (25.4%)	0.67	1.04	13 (22.0%)	27 (45.8%)	16 (27.1%)	3 (5.1%)
PEC-CKB average f = 4.4%	59	average	0.81	1.12	10 (16.9%)	26 (44.1%)	17 (28.8%)	6 (10.2%)	0.01	0.80	32 (54.2%)	19 (32.2.9%)	8 (13.6%)	0
PEC-CKB crop-specific f	59	average	0.95	1.30	10 (16.9%)	25 (42.4%)	14 (23.7%)	10 (16.9%)	0.15	0.94	25 (42.4%)	25 (42.4%)	6 (10.2%)	3 (5.1%)

* Imidacloprid has only been calculated for hops, and in FOCUS Step 3, scenarios D3 and D4 are not defined for hops.

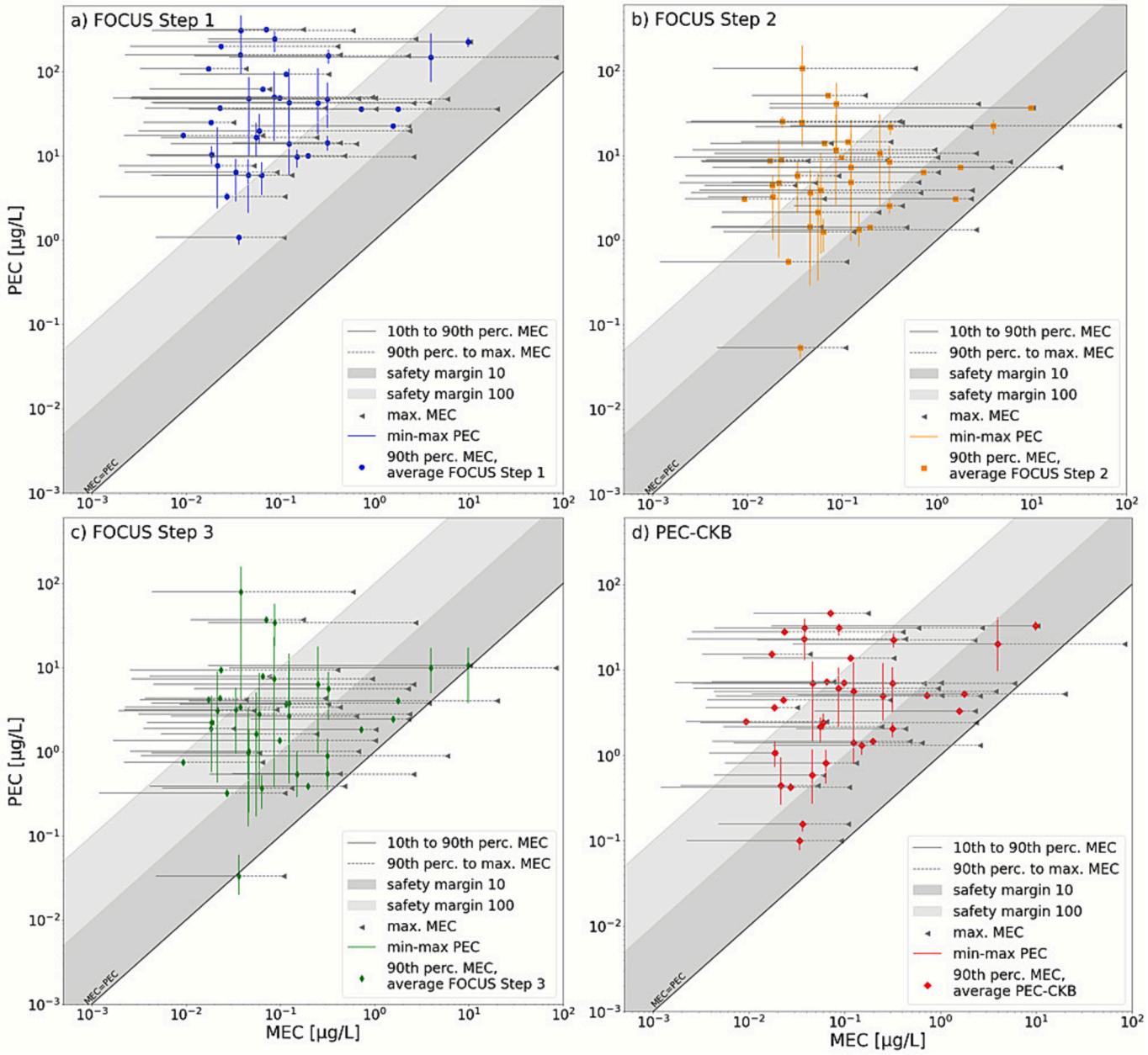

Fig. 1. Comparison of measured environmental concentrations (MECs) and predicted environmental concentrations (PECs) for the FOCUS Step 1, 2 and 3 models and PEC-CKB for 36 pesticides. We compared average PEC values based on different crop-specific application rates and scenarios (see chapter 2.1 and 2.2) with the 90th percentiles of the monitoring data. The FOCUS model calculations were adapted to German environmental conditions and Step 3 was only calculated for the R1, D3 and D4 stream scenarios (more information on the scenarios can be found in chapter 2.1.1). The range of PEC values is shown by the coloured vertical lines (minimum to maximum PEC), and the MEC values are shown by a solid, horizontal line for the 10th to 90th percentile and by a dashed line for the 90th percentile to the maximum measured concentration. The safety margins represent overestimation of a factor of 10 and 100, respectively.

can be mainly explained by the factor that summarises all transport processes (M in Eq. (1); see also Section 2.1.3).

3.3. Refinement of PEC-CKB parameters and impact on prediction quality

A comparison of the original PEC-CKB model predictions based on Eq. (3) with German monitoring data (Liess et al., 2021b) shows a wide range of PEC values (Fig. 4) driven by varying crop-specific application rates. Considering 59 substances, only isopyrazam is underestimated, but this is one of the substances with less than ten concentration measurements above LOQ, implying that the 90th percentile value may be less reliable (Fig. 4a; SI, Table A.1). The majority of substances exceed MECs by factors between one and two orders of magnitude (Table 1).

As model parameterisation mainly relies on former analyses based on Swedish monitoring data, we aim to account for the differences between the German and Swedish monitoring catchments (see also SI, Section A) and possibly improve PEC-CKB predictions. For this, we refined the model parameters with a GIS analysis of landscape data (see Section 2.3). The results indicate a lower degree of agricultural land use than in the Swedish study, supporting a lower f-value than 20% (see SI, Table H.1). Averaging all crop-specific f-values yields an average crop cover of 4.4%, approximately one-fourth of the treatment proportion of 20% proposed by the FOCUS surface water working group and set in the original PEC-CKB definition. Using the refined f-value in the PEC-CKB model improved MEC-PEC agreement (bias: 0.81, RMSE: 1.12) (Table 1), while maintaining similar variability in PECs, visible as a

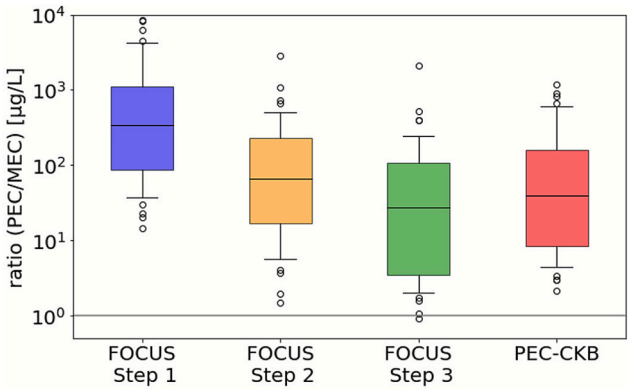

Fig. 2. Distribution of residuals from the three FOCUS models Step 1–3 and PEC-CKB, illustrating the range and central tendency of prediction errors (differences between PEC and MEC values). Boxplots show the median (line within the box), interquartile range (box boundaries) and 10th and 90th percentiles (whiskers).

Table 2
Linear correlation analysis results and coefficients of determination R^2. R^2 is determined by squaring the correlation coefficient r. We used Pearson correlation coefficients, accounting for significance. The null hypothesis of no linear correlation is rejected if $p < 0.05$. We differentiated between comparing 90th percentile MECs and maximum MECs to average PECs.

model	90th percentile MEC			maximum MEC		
	Pearson's r	p-value	R^2	Pearson's r	p-value	R^2
PEC-CKB	0.31	0.07*	0.096	0.44	0.01	0.194
FOCUS Step 1	0.27	0.11*	0.073	0.42	0.01	0.176
FOCUS Step 2	0.22	0.20*	0.048	0.34	0.04	0.116
FOCUS Step 3	0.16	0.36*	0.026	0.27	0.11*	0.073

* No statistically significant correlation.

"shift down" of PEC-CKB predictions (Fig. 4b). Of course, using such an overall average *f*-value is a simplification of the reality of land use in Germany, and the high standard deviation and coefficient of variation (200%) of this *f*-value relative to the mean value represent the wide range of different crop-specific *f*-values. In the refined PEC-CKB calculations, we have ten underestimations when compared with 90th percentile MECs and 32 underestimations when compared with the maximum MECs. In the latter, maximum MECs were underestimated by up to two orders of magnitude.

To fully exploit the GIS analyses, we also investigated the influence of using crop-specific *f*-values (SI, Table H.1) in combination with crop-specific application rates on PEC-CKB model predictions (Eq. (4)). Using such crop-specific *f*-values resulted in increased variability of PECs, which spanned one to two orders of magnitude (Fig. 4c). The agreement between PECs and MECs improved in comparison to the standard assumption of $f = 0.2$ for all substances and crops, but got worse compared with the use of the average *f*-value of 4.4% (Table 1). The 90th percentile MECs for ten substances and the maximum MECs for 25 of the 59 substances were underestimated by the average PEC-CKB predictions with crop-specific *f*-values, yielding similar results to those obtained with the average *f*-value of 4.4%.

We also tested whether basic physico-chemical substance properties could explain the agreement of PEC-CKB predictions with monitoring data. Former analyses of the MACRO model showed that the distribution coefficient between water and organic carbon (K_{oc}) and the degradation rate (DT_{50}) of the substance in soil can strongly impact the model results (Dubus and Brown, 2002). The PEC-CKB model directly incorporates K_{oc}, but does not include the degradation rate. Generally, larger deviations between MECs and PECs occur with smaller K_{oc} and DT_{50}. The model performs better for persistent substances while substances with fast degradation rates exhibit larger discrepancies (see SI, Figs. C.1–C.4), highlighting the importance of temporal resolution in the models. However, detecting substances with small DT_{50} values in water samples is more difficult, because the application periods are often unknown.

4. Discussion

The preceding analyses depend on model assumptions and decisions made during the evaluation process. Nevertheless, some general conclusions can be drawn from the results for a prospective ERA. In the following, we discuss several methodological uncertainties and challenges in model evaluation and outline potential implications for a future ERA. In doing so, we will also take a closer look at the position of the PEC-CKB model in comparison to the established FOCUS models.

4.1. Challenges in model evaluation

Assessing model performance is particularly challenging under highly variable environmental conditions. In addition to selecting suitable monitoring data, information on the emissions and model assumptions plays a decisive role in the final assessment. Furthermore, we discuss model extensions and national requirements that were not considered for this study.

4.1.1. Handling of monitoring data

4.1.1.1. Monitoring can only deliver a snapshot of reality. Monitoring campaigns provide only a snapshot of the actual environmental situation (la Cecilia et al., 2021), and MECs can vary strongly (Fig. 1, range between 10th percentile and maximum MEC). Model evaluation requires the aggregation of MEC values into single statistical descriptors, which simplify the variability and describe only one point in time and space across all measurements.

In this study, we focus on using 90th percentile MECs for model evaluation because they represent high measured concentrations without focusing on extreme values. Using the 90th percentile provides greater robustness against statistical outliers, but it also ignores up to 10% of measured values during evaluation. While higher percentiles would generally be preferable, they were not suitable in this context due to, in parts, too small sample sizes (SI, Table A.1).

4.1.1.2. Sampling type does not strongly influence the results. The decision to consider only MECs from event-driven samples to capture potentially higher concentrations could have influenced the analyses, since rain events are known to transport PPPs to adjacent surface waters via runoff. Depending on factors such as the amount and intensity of rainfall, the size of the water body and the sorption behaviour (K_{oc}), the PPP concentration in samples can be correspondingly higher or lower due to dilution. Nevertheless, similar analyses on grab samples revealed only minor differences and confirmed the conservative character of all models, with a few exceptions in FOCUS Step 3 (see SI, Section E-b).

4.1.1.3. Limited measurability of concentrations influences results. We only considered measured concentrations above the LOQ. Due to the lack of application data, we assume that very low concentrations below LOQ are not attributable to an application during the monitoring period. Including non-detections and values below the LOQ resulted in lower 90th percentiles MECs, increased overestimations and thus worse agreement with monitoring data, but without underestimations (see additional analyses in SI, Section F). This latter approach can also be criticised, as 1) non-detections may simply originate from non-application of a PPP and 2) values < LOQ are highly uncertain.

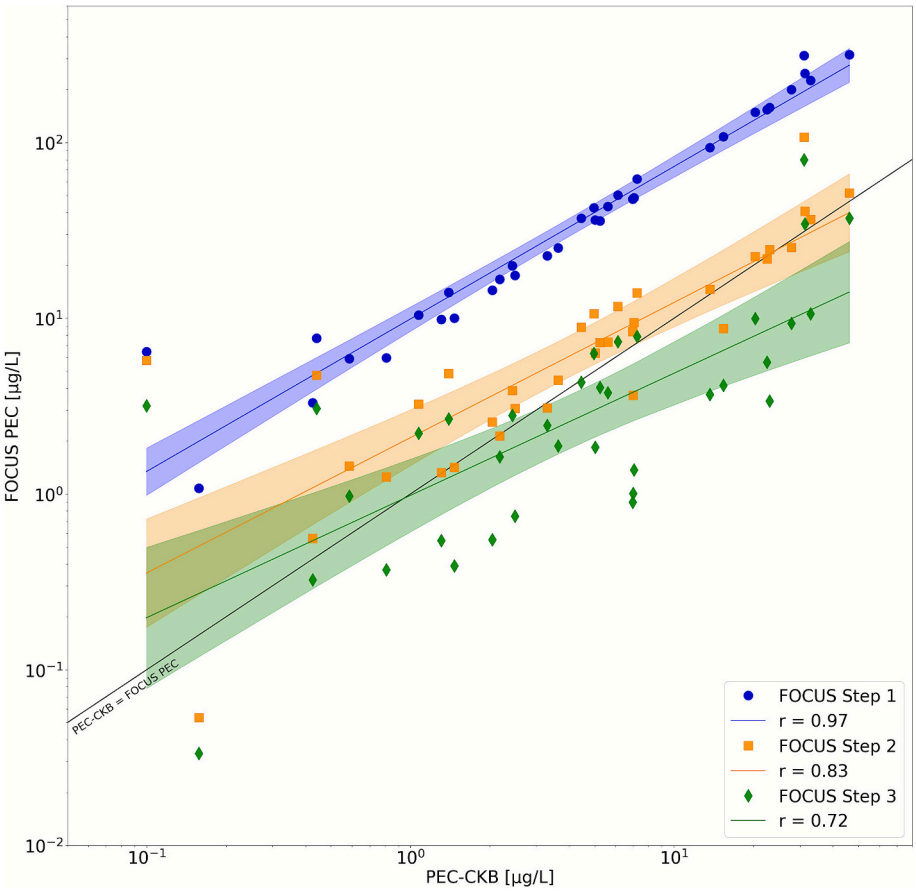

Fig. 3. Comparison of PEC-CKB to FOCUS model calculations based on the model results from previous analyses. We compared average PEC values based on different application doses and scenario calculations (FOCUS). Correlation analysis (Pearson correlation coefficient r) shows the highest value for FOCUS Step 1, but visual representation indicates that Step 2 matches PEC-CKB calculations best, as it is closer to the 1:1 line.

4.1.1.4. Extreme conditions could influence maximum MECs. Using the maximum MECs helps avoid underestimation, but they can be skewed by extreme climatic and geographical conditions or PPP misuse (Kreuger, 1998; Schriever et al., 2025; Wittmer et al., 2010). Situations that are not accounted for in the authorisation process (e.g. cleaning of PPP application machinery, incorrect PPP application, extreme hydrological conditions) can strongly affect MEC-PEC comparisons, calling into question the use of maximum MECs as a benchmark for model evaluation (Schriever et al., 2025). Nevertheless, it is unlikely that high MECs are mainly driven by malpractices or urban sources (Schriever et al., 2025; Wittmer et al., 2010), because that would require widespread farmer misuse. Multiple fields in the catchment, the repeated use of the same active ingredient across several products and crops, as well as spray series, can also raise MEC values. For this reason, we further assessed conservatism by examining maximum MECs and found that in this case, only FOCUS Step 1 showed no underestimations (see Section 3.1; SI, Table E.1).

4.1.2. Model assumptions and their impact on the results

Models are based on assumptions that simplify reality. The selection of assumptions significantly influences the results and needs to be reported transparently in the context of ERA.

4.1.2.1. Generic application data. The lack of specific application data led us to use generic German application rates and assume single applications in our study. We have no information about the representativeness of this data for our catchments from the monitoring study. By averaging the application data in the PAPA database, we could generate application rates for crop-specific single applications. From Tier 2 onwards, multiple applications play an important role, and degradation becomes more relevant. Especially for substances with high DT_{50} values relative to the period between two applications, multiple applications may result in higher PECs, because concentrations from previous applications persist.

4.1.2.2. Scenario selection. Selecting scenarios from a retrospective perspective for a prospective ERA introduces limitations. By selecting some and omitting other scenarios in the PEC calculation, we accept that we only incompletely represent the distribution of real conditions. The quality of the FOCUS model predictions could have been improved by using more specific and appropriate scenarios.

However, this study aimed to link the retrospective and prospective ERA perspectives, so that we selected application regions, time periods, and pre-defined drainage and runoff scenarios that match the monitoring data best from a theoretical point of view. Although the event-driven measurements in the KgM study have been attributed to runoff by the authors, input via drainage may also have influenced the measurements (see 2.4 and 3.1). The results confirm that models cannot generate accurate predictions if the best-matching environmental scenarios at the corresponding sites are not known.

4.1.2.3. Incorporation of landscape-specific information. One step towards better-matching environmental scenarios is the consideration of the crop distribution in the different monitoring catchments. Using landscape information in our analyses has improved model performance. The model predictions were more consistent with the measured concentrations when crop-specific *f*-values were utilised, particularly the reduced, aggregated *f*-value of 4.4%. But averaging *f*-values from the

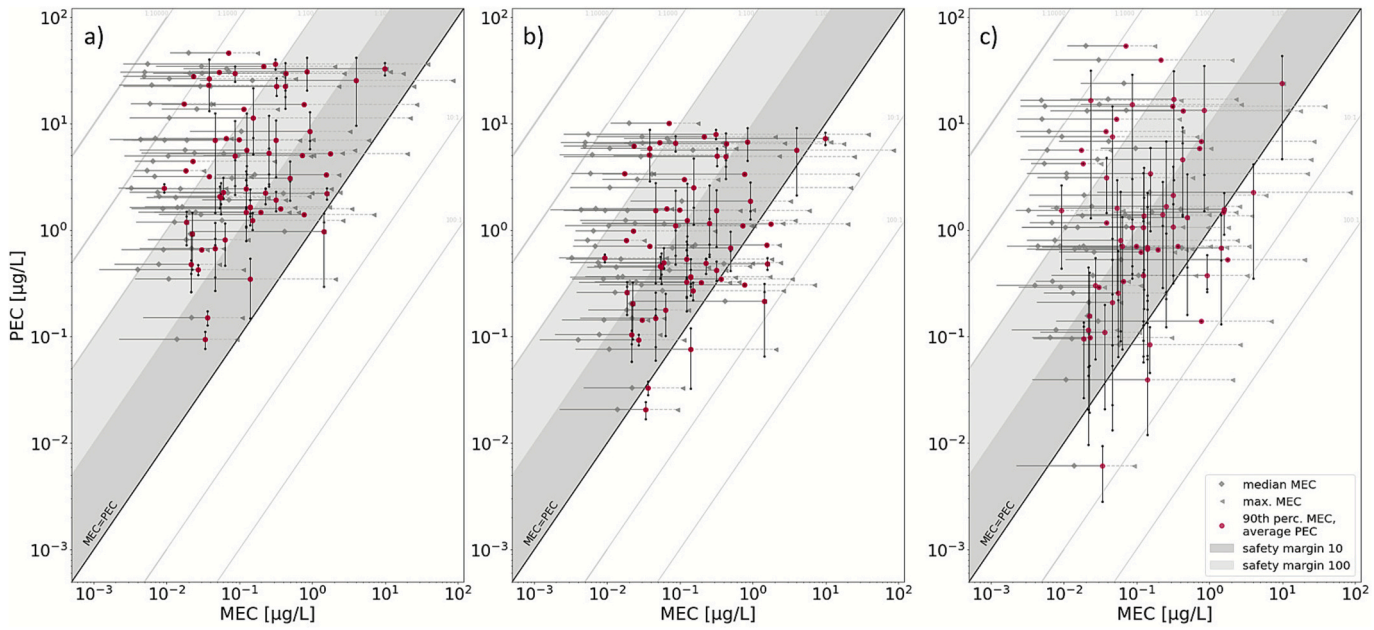

Fig. 4. Comparison of MECs from a comprehensive pesticide monitoring in small German streams (Liess et al., 2021b) (represented by lines from 10th to 90th percentiles as well as median and maximum) with PECs resulting from the PEC-CKB. The 90th percentile of MECs are marked with red dots. Figure a) represents the PEC calculations using the crop-specific average application doses in combination with the standard f-value of 0.2. Figure b) and c) represent PEC calculation using crop-specific average f-values based on a GIS analysis. Figure b) shows PEC values calculated with an average crop-specific f-value of 0.044 and Figure c) shows the use of individual crop-specific f-values. MEC data is plotted against the average PEC value, whereas the range of different PEC values due to different input parameters (D and f) is represented by vertical lines. 59 substances are included in this analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

GIS analysis introduces bias, particularly for crops with uneven distribution. While some crops, such as grassland, winter cereals or oilseed rape, are consistently present, others, such as vines, sunflowers or orchards, vary significantly between catchments. This variability, indicated by high coefficients of variation (SI, Table H.1), shows that an average *f*-value does not accurately reflect conditions in regions where these crops dominate, nor regions where they are absent (e.g. wine regions in Germany). Also, monocultures cannot be meaningfully described using averaged treatment proportions. Nevertheless, the results of the analyses are helpful for testing the standard assumption of a treated area of 20% (FOCUS, 2001). This value is indeed nearly a worst-case assumption; only winter cereals resulted in an average proportion of more than 20% of the catchments in our analyses, while all other crops ranged below 10%.

4.1.3. Dismissed model extensions

4.1.3.1. FOCUS repair scenarios. In the last few years, efforts have been made to revise the FOCUS scenarios, resulting in the so-called 'FOCUS repair' scenarios, with prolonged time periods and adapted application dates based on corresponding weather data (EFSA, 2020). Here, a moving 3-day average air temperature represents the daily water temperature. Crop interception values based on the actual developmental state allow a better comparison of model predictions to other exposure models. The FOCUS repair actions will most likely result in similar PECs with reduced variability. In particular, the annual maximum PECs are expected to fall within the range of previous results (EFSA, 2020). We do not expect major changes when using the 'repaired' scenarios, but their use might positively influence overall model performance. For future analyses, the revised FOCUS version should be utilised.

4.1.3.2. Potential mitigation strategies. In addition, some of the analysed substances will require mitigation strategies such as drift-reducing technologies, e.g. nozzles and runoff buffers, which are not relevant in FOCUS Steps 1 to 3 simulations or the PEC-CKB model. In general, the consideration of mitigation strategies for PEC calculations, for example by using FOCUS Step 4, would decrease PECs and consequently most likely lead to worse MEC-PEC-agreement.

4.1.3.3. National registration requirements. Further on, we ignored in our analyses that national registration procedures might be based on other models, for example, the models EXPOSIT 3.02 and EVA3 are used in Germany. Nevertheless, models used in national registration processes could be evaluated in a similar manner. For the German surface water PEC model EXPOSIT, predictions showed a tendency to underestimate MEC values (Liess et al., 2021a; Weisner, 2022, p. 215).

4.2. The tiered model approaches

Using a tiered approach in ERA intends to stay with simple, but safe assessments for uncritical chemicals, and to increase complexity and computational effort step-by-step in higher tiers. However, previous studies have indicated that models at higher tiers are not necessarily more precise and can lead to underestimations. The latter is particularly concerning in ERA and requires a closer examination.

4.2.1. Comparing results to previous analyses

In general, comparing PECs with the 90th percentile MECs is consistent with the FOCUS philosophy. Our study confirms that PECs are generally higher than the 90th percentile MECs that the FOCUS models should represent. Comparing these findings with those from the studies by Knäbel et al. (2012, 2014), we can confirm, based on our MEC-PEC comparisons, that Steps 1 and 2 produce conservative model results. Our analyses also show two underestimations of 90th percentile MECs in Step 3 when considering different crop-specific PECs. However, the FOCUS Step 3 model performs clearly better in our comparisons concerning the overall model performance (based on MEC/PEC ratios) and the number of underestimations. One possible reason is that Knäbel and colleagues included fewer substances in their comparisons but considered more studies and, therefore, more individual MEC values (2012: 17

substances, 77 MECs and 2014: 38 substances, 417 MECs). The MECs from the German KgM study (Liess et al., 2021b) suggest a high degree of applicability for FOCUS models and scenarios designed for EU countries. In contrast, Knäbel and colleagues included data from multiple international studies, which increases variability due to differing measurement methods, different PPP usages or agricultural practices. While FOCUS scenarios partially accounted for environmental variability, another main difference is that Knäbel and colleagues compared MECs directly to PECs, whereas we compared statistical MEC values with corresponding PEC values.

4.2.2. Targets of tiered ERA only partially achieved

The estimation of PECs for regulatory ERA in the EU is performed in a tiered system, with increasing complexity and the ambition to deliver more accurate predictions at higher tiers. Our findings align with the original conservative design of FOCUS, as Step 1 yields worst-case estimates (FOCUS, 2001), while FOCUS Step 2 produces less conservative, more realistic PECs and shows no underestimation of 90th percentile MECs. FOCUS Step 3 is considered a realistic worst-case model, as it was developed to account for local conditions more specifically and includes a number of scenarios that should represent environmental conditions for one-third of the total area of the EU (FOCUS, 2001). However, our analysis of the KgM dataset does not consistently show worst-case estimates.

Generally, regulatory exposure models aim to predict exposure in generic scenarios rather than specific site conditions. Also, FOCUS Step 3 was not designed for MEC-PEC comparisons on a regional scale, as it was done in this study (Bach et al., 2016), and part of the deviation of model predictions and MECs may arise from a mismatch between environmental site conditions and their representation in the scenarios.

Model refinement in higher tiers follows the scientific ambition to be more accurate on average, meaning that over- and underestimations equalise. This can be in direct conflict with the regulatory aim of always producing conservative predictions, so that PPPs can be released to the market in a safe and cost-effective way, while allowing for understanding the model results (Enquist et al., 2024). Ideally, from a regulatory perspective, PEC models should generate values higher than MECs with only small deviations from them. This could be reached, for example, by applying safety factors to scientifically accurate models.

4.3. Complexity of models: PEC-CKB instead of FOCUS models?

One of the principal challenges of regulatory ERA for pesticides and other chemicals is the development of exposure prediction models that deliver realistic but conservative predictions of pesticide concentrations at the same time, ideally at low to moderate model and scenario complexity. The complexity of the model calculations and the related scenarios is also relevant because the capacity for model evaluation is often constrained for practical work in environmental authorities. Complex models tend to cause high efforts in maintenance and evaluation and are not useful from a regulatory perspective because they entail the risk of overfitting singular environmental situations but lack generality. The complexity of models need to be assessed conserving the principle of parsimony, a basic concept that is also often used in ecological studies (Coelho et al., 2019), and that considers simpler models being preferable when having an incomplete understanding of complex processes, to avoid a false impression of risk understanding and growing modelling effort (Etterson, 2022; Hill et al., 2000).

The PEC-CKB is a simple PEC modelling approach that performs better than FOCUS Step 2 while being less complex than FOCUS Step 1, as transport processes are summarised in a single generic factor (Sections 3.1 and 3.2). FOCUS Step 3 shows slightly better agreement between PEC and MEC values, but more underestimations. Compared to FOCUS Step 3, the PEC-CKB is much simpler, consisting of a single formula rather than a set of differential equations, and the difference in the required input is even more evident. Here, the FOCUS Step 3 model requires input from several complex models, such as PRZM and MACRO, and considers 10 scenarios. In contrast, PEC-CKB has a smaller parameter set, and we improved its performance in our study by incorporating landscape information.

However, we cannot exclude that the agreement between PEC-CKB predictions and MECs is good for the wrong reason. Parameter assumptions for PEC-CKB are based on Swedish monitoring data (Boström et al., 2019) but do not represent environmental conditions in the KgM study. The runoff is defined as a dilution in a water column with a height of 4 mm (see Eq. (1)). This would mean that almost no rain falls during the monitoring period. The KgM event-driven data was collected at a water-level rise of 5 cm, corresponding to a deviating dilution factor of 12.5.

Nevertheless, it is difficult to sort PEC-CKB into the tiers of the FOCUS models: The similarity of the model with FOCUS Step 1 and the high correlation would support considering PEC-CKB as a Tier 1 method, but concerning the absolute PEC values, PEC-CKB would range between Tier 2 and 3 (see Sections 3.1 and 3.2).

In the context of the regulatory risk assessment system, the results of the PEC-CKB model lack information that is provided by the FOCUS surface water models, most importantly, predictions of sediment concentrations and temporal profiles. The PEC-CKB model produces only point estimates after one week, whereas it should be easy to adapt PEC-CKB by incorporating time-dependency, since the equation is very similar to the time-explicit FOCUS Step 1 (FOCUS, 2001; SI, Section G). Having time-explicit PECs is very relevant, for example, in the context of mixture toxicity assessment, which is not yet part of regulatory ERA, but that is often discussed to improve ERA realism. Furthermore, the dose addition used in PEC-CKB appears similar to FOCUS Step 1. However, it does not reflect the reality of multiple exposures to the same substance over time, where transport and loss processes may occur between applications (Reinert et al., 2002). The model also does not capture spatial and temporal exposure patterns or the complexity of agricultural landscapes, which could offer options to combine risk assessment with risk management in a future ERA (Focks et al., 2014). Taking all this into account when considering PEC-CKB as an alternative to the currently used FOCUS models, it appears to be a viable option only for isolated use in predicting point estimates of surface water concentrations.

5. Conclusions

The use of predictive models in connection with regulatory risk assessment requires that the models meet three criteria, which are very difficult to fulfil simultaneously: realism, conservatism, and low effort. In our analyses of the FOCUS models, we found a general trade-off between realism and conservatism that is closely linked to the modelling effort. At low modelling efforts, low accuracy is tolerated as long as the predictions remain conservative. The principle in the tiered regulatory approach that lower tiers show higher conservatism was confirmed by our analyses. While increasing model complexity generally enhanced model accuracy, it also increased the risk of underestimating MECs. In fact, the agreement between model predictions and measured values improved at higher MEC values, but this was accompanied by an increased risk of underestimation, which is particularly important for risk assessment purposes.

Given our findings of trade-offs between realism and conservatism, risk assessors should be clear about their main priorities: Is it most important to estimate realistic or conservative estimates of MECs? Are low modelling efforts more important than accuracy and realism?

In that context, the PEC-CKB model achieved good performance through appropriate assumptions despite its simple structure and scenario. We could show that incorporating more precise land-use information improved the agreement between PEC-CKB results and MECs, but at the same time led to more underestimations, again indicating a trade-off. Overall, the study highlights several strengths of the PEC-CKB model that could be refined further and made country- or even site-

specific, and appeared as the best-balanced considering all three categories. Nevertheless, the PEC-CKB would not allow for the analysis of co-occurrences of concentration peaks of different substances, and therefore would not be suitable for risk assessments that consider chemical mixtures (Brock, 2013; Posthuma et al., 2019).

In general, knowledge of crop-specific application rates and information on catchment crop composition appears more beneficial for improving model predictions than increasing the complexity of process descriptions. We conclude that better knowledge of actual PPP applications and improved access to high-resolution data on agricultural practices could help to close the gap between prospective and retrospective exposure assessment in ERA for PPPs and improve the understanding of variabilities in MEC values.

Future work should focus on essential process descriptions without getting lost in details, emphasising robust parameters such as agronomic PPP application, physico-chemical properties and spatio-temporal exposure patterns.

Funding sources

This work was carried out in the framework of the European Partnership for the Assessment of Risks from Chemicals (PARC) and has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101057014. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

CRediT authorship contribution statement

Paula Scharlach: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gustaf Boström:** Writing – review & editing, Visualization, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Jörg Klasmeier:** Writing – review & editing, Visualization, Methodology, Conceptualization. **Amelie Leonardi:** Software, Investigation, Formal analysis. **Andreas Focks:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could appear to have influenced the work reported in this publication.

Acknowledgements

We would like to thank the whole Exposure subproject group of PARC 6.4.4 (Grand Agreement No 101057014) for many internal discussions that have contributed to these results. In particular, we would like to mention here: Mikaela Gönczi, Barbara Gworek, Larissa Hermann, Radosław Kalinowski, Marta Kijeńska-Markuszewska, Judith Klein, Beata Kubiak, Ralf Schulz and Dimitrios Skodras. We additionally thank the regulatory core group of PARC 6.4.4 (especially Annette Aldrich, Johan Axelman and Sabine Duquesne), whose contributions have facilitated the advancement of our work through valuable suggestions and critical inquiries. We also thank Jörn Strassemeyer and the Julius Kühn Institute for providing substance- and crop-specific application data for Germany from the PAPA database, and Matthias Liess and his working group from the Helmholtz Centre for Environmental Research for providing additional information and data on the KgM campaign.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2026.110057.

Data availability

Data will be made available on request.

All data needed to evaluate the conclusions of this publication can be found in the paper and/ or in SI and/or in the listed references.

References

Alix, A., Capri, E., 2018. Modern Agriculture in Europe and the Role of Pesticides. In: Advances in Chemical Pollution, Environmental Management and Protection. Elsevier, pp. 1–22. https://doi.org/10.1016/bs.apmp.2018.04.001.

Axelman, J., Aldrich, A., Duquesne, S., Backhaus, T., Brendel, S., Focks, A., Holz, S., Knillmann, S., Pieper, S., Silva, E., Schmied-Tobies, M., Topping, C.J., Wipfler, L., Williams, J., Sousa, J.P., 2024. A systems-based analysis to rethink the European environmental risk assessment of regulated chemicals using pesticides as a pilot case. Sci. Total Environ. 174526. https://doi.org/10.1016/j.scitotenv.2024.174526.

Bach, M., Diesner, M., Großmann, D., Guerniche, D., Hommen, U., Klein, M., Kubiak, R., Müller, A., Priegnitz, J., Reichenberger, S., Thomas, K., Trapp, M., 2016. Pesticide exposure assessment for surface waters in the EU. Part 1: some comments on the current procedure: pesticide exposure for surface waters in the EU. Pest Manag. Sci. 72 (7), 1279–1284. https://doi.org/10.1002/ps.4281.

Boström, G., Jarvis, N., Gönczi, M., & Kreuger, J. (2019). A proposed new method for calculating predicted environmental concentrations (PEC) for plant protection products in surface water. https://www.slu.se/globalassets/ew/org/centrb/ckb/publikationer/mo-rapporter/a-proposed-new-method-for-pec-2019.pdf.

Boye, K., Jarvis, N., Moeys, J., Gönczi, M., Kreuger, J., 2012. *Pesticide run-off to Swedish surface waters and appropriate mitigation strategies: a review of the knowledge focusing on vegetated buffer strips* [Elektronisk resurs]. Swedish University of Agricultural Sciences, Centre for Chemical Pesticides.

Brock, T.C., 2013. Priorities to improve the ecological risk assessment and management for pesticides in surface water. Integr. Environ. Assess. Manag. 9 (3). https://doi.org/10.1002/ieam.1429.

Bundesamt für Kartographie und Geodäsie. (2023, June 20). Digitales Landschaftsmodell 1:250000 (Ebenen) (DLM250). https://gdz.bkg.bund.de/index.php/default/digitales-landschaftsmodell-1-250-000-ebenen-dlm250-ebenen.html.

Carvalho, F.P., 2006. Agriculture, pesticides, food security and food safety. Environ. Sci. Policy 9 (7–8), 685–692. https://doi.org/10.1016/j.envsci.2006.08.002.

Centanni, M., Ricci, G.F., De Girolamo, A.M., Romano, G., Gentile, F., 2023. A review of modeling pesticides in freshwaters: current status, progress achieved and desirable improvements. Environ. Poll. 316, 120553. https://doi.org/10.1016/j.envpol.2022.120553.

Coelho, M.T.P., Diniz-Filho, J.A., Rangel, T.F., 2019. A parsimonious view of the parsimony principle in ecology and evolution. Ecography 42 (5), 968–976. https://doi.org/10.1111/ecog.04228.

de Souza, R.M., Seibert, D., Quesada, H.B., de Jesus Bassetti, F., Fagundes-Klen, M.R., Bergamasco, R., 2020. Occurrence, impacts and general aspects of pesticides in surface water: a review. Process Safety Environ. Protect. 135, 22–37. https://doi.org/10.1016/j.psep.2019.12.035.

Di Guardo, A., Gouin, T., MacLeod, M., Scheringer, M., 2018. Environmental fate and exposure models: advances and challenges in 21st century chemical risk assessment. Environ. Sci.: Process. Impacts 20 (1), 58–71. https://doi.org/10.1039/C7EM00568G.

Dubus, I.G., Brown, C.D., 2002. Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. J. Environ. Quality 31 (1), 227–240. https://doi.org/10.2134/jeq2002.2270.

Enquist, B.J., Kempes, C.P., West, G.B., 2024. Developing a predictive science of the biosphere requires the integration of scientific cultures. Proceed. Natl. Acad. Sci. 121 (19), e2209196121. https://doi.org/10.1073/pnas.2209196121.

Etterson, M.A., 2022. Realism, conservatism, and tiered ecological risk assessment. Ecologies 3 (2), 131–144. https://doi.org/10.3390/ecologies3020011.

EU. (2023). *EU Pesticide Database*. https://food.ec.europa.eu/plants/pesticides/eu-pesticides-database_en.

European Food Safety Authority (EFSA), Adriaanse, P., Boivin, A., Klein, M., Jarvis, N., Stemmer, M., Fait, G., & Egsmose, M. (2020). Scientific report of EFSA on the 'repair action' of the FOCUS surface water scenarios. EFSA Journal, 18(6). doi: 10.2903/j.efsa.2020.6119.

Focks, A., ter Horst, M., van den Berg, E., Baveco, H., van den Brink, P.J., 2014. Integrating chemical fate and population-level effect models for pesticides at landscape scale: new options for risk assessment. Ecol. Model. 280, 102–116. https://doi.org/10.1016/j.ecolmodel.2013.09.023.

FOCUS. (2001). FOCUS surface water scenarios in the EU evaluation process under 91/414/EEC. Report of the FOCUS Working Group on Surface Water Scenarios, EC Document Reference SANCO/4802/2001-rev.2. 245 pp.

Forbes, V.E., Hommen, U., Thorbek, P., Heimbach, F., Van den Brink, P.J., Wogram, J., Thulke, H., Grimm, V., 2009. Ecological models in support of regulatory risk assessments of pesticides: developing a strategy for the future. Integr. Environ. Assess. Manag. 5 (1), 167–172. https://doi.org/10.1897/IEAM_2008-029.1.

Fox, G.A., Muñoz-Carpena, R., Brooks, B., Hall, T., 2021. Advancing surface water pesticide exposure assessments for ecosystem protection. Transactions of the ASABE 64 (2), 377–387. https://doi.org/10.13031/trans.14225.

Fulda, B., Aldrich, A., Kastell, R., Balmer, M., & Poiger, T. (2015). Pflanzenschutzmittel in Fliessgewässern—Vergleich zwischen Expositionsmodell und Monitoringdaten. Aqua&Gas No. 9, 91ff., https://www.agroscope.admin.ch/agroscope/de/home/themen/pflanzenbau/pflanzenschutz/herbologie/allelopathie/_jcr_content/par/columncontrols/items/1/column/externalcontent_825221443.bitexternalcontent.exturl.html/aHR0cHM6Ly9pcmEuYWdyb3Njb3BlLmNoL2RlLUNIL3B1YmxpY2/F0aW9uLzM1MTg3.html.

Halbach, K., Möder, M., Schrader, S., Liebmann, L., Schäfer, R.B., Schneeweiss, A., Schreiner, V.C., Vormeier, P., Weisner, O., Liess, M., Reemtsma, T., 2021. Small streams–large concentrations? Pesticide monitoring in small agricultural streams in Germany during dry weather and rainfall. Water Res. 203, 117535. https://doi.org/10.1016/j.watres.2021.117535.

Hill, R.A., Chapman, P.M., Mann, G.S., Lawrence, G.S., 2000. Level of detail in ecological risk assessments. Marine Poll. Bull. 40 (6), 471–477. https://doi.org/10.1016/S0025-326X(00)00036-9.

JKI. (2023). Statistische Erhebungen zur Anwendung von Pflanzenschutzmitteln in der Praxis. https://papa.julius-kuehn.de/.

Knäbel, A., Meyer, K., Rapp, J., Schulz, R., 2014. Fungicide field concentrations exceed FOCUS surface water predictions: urgent need of model improvement. Environ. Sci. Technol. 48 (1), 455–463. https://doi.org/10.1021/es4048329.

Knäbel, A., Stehle, S., Schäfer, R.B., Schulz, R., 2012. Regulatory FOCUS Surface Water Models Fail to Predict Insecticide Concentrations in the Field. Environ. Sci. Technol. 46 (15), 8397–8404. https://doi.org/10.1021/es301649w.

Kreuger, J., 1998. Pesticides in stream water within an agricultural catchment in southern Sweden, 1990–1996. Sci. Total Environ. 216 (3), 227–251. https://doi.org/10.1016/s0048-9697(98)00155-7.

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 Research X 13, 100125. https://doi.org/10.1016/j.wroa.2021.100125.

Leu, C., Singer, H., Stamm, C., Müller, S.R., Schwarzenbach, R.P., 2004. Simultaneous assessment of sources, processes, and factors influencing herbicide losses to surface waters in a small agricultural catchment. Environmental Science & Technology 38 (14), 3827–3834. https://doi.org/10.1021/es0499602.

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., Reemtsma, T., 2021a. Pesticides are the dominant stressors for vulnerable insects in lowland streams. Water Res. 201, 117262. https://doi.org/10.1016/j.watres.2021.117262.

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., … Reemtsma, T. (2021b). The lowland stream monitoring dataset (KgM, Kleingewässer-Monitoring) 2018, 2019 (p. 52.3 MBytes) [Application/zip]. PANGAEA. doi: 10.1594/PANGAEA.931673.

MacLeod, M., Scheringer, M., McKone, T.E., Hungerbuhler, K., 2010. The state of multimedia mass-balance modeling in environmental science and decision-making. Environ. Sci. Technol. 44 (22), 8360–8364. https://doi.org/10.1021/es100968w.

Posthuma, L., Brack, W., van Gils, J., Focks, A., Müller, C., de Zwart, D., Birk, S., 2019. Mixtures of chemicals are important drivers of impacts on ecological status in European surface waters. Environ. Sci. Europe 31 (1), 71. https://doi.org/10.1186/s12302-019-0247-4.

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. https://doi.org/10.1016/j.scitotenv.2007.04.046.

Reinert, K.H., Giddings, J.M., Judd, L., 2002. Effects analysis of time-varying or repeated exposures in aquatic ecological risk assessment of agrochemicals. Environ. Toxicol. Chem. 21 (9), 1977–1992. https://doi.org/10.1002/etc.5620210928.

Rico, A., Dafouz, R., Vighi, M., Rodríguez-Gil, J.L., Daam, M.A., 2021. Use of postregistration monitoring data to evaluate the ecotoxicological risks of pesticides to surface waters: a case study with chlorpyrifos in the Iberian Peninsula. Environ. Toxicol. Chem. 40 (2), 500–512. https://doi.org/10.1002/etc.4927.

Roßberg, D., Aeckerle, N., Stockfisch, N., 2017. Erhebungen zur Anwendung von chemischen Pflanzenschutzmitteln in Zuckerrüben. Gesunde Pflanzen 69 (2), 59–66. https://doi.org/10.1007/s10343-017-0389-5.

Schäfer, R.B., von der Ohe, P.C., Kühne, R., Schüürmann, G., Liess, M., 2011. Occurrence and toxicity of 331 organic pollutants in large rivers of north Germany over a decade (1994 to 2004). Environ. Sci. Technol. 45 (14), 6167–6174. https://doi.org/10.1021/es2013006.

Schmolke, A., Thorbek, P., Chapman, P., Grimm, V., 2010. Ecological models and pesticide risk assessment: current modeling practice. Environ. Toxicol. Chem. 29 (4), 1006–1012. https://doi.org/10.1002/etc.120.

Schriever, C., Jene, B., Resseler, H., Spatz, R., Sur, R., Weyers, A., Winter, M., 2025. The European regulatory system for plant protection products—cause of a "Silent Spring" or highly advanced and protective? Integr. Environ. Assess. Manag., vjae007. https://doi.org/10.1093/inteam/vjae007.

Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313 (5790), 1072–1077. https://doi.org/10.1126/science.1127291.

Stehle, S., Schulz, R., 2015a. Agricultural insecticides threaten surface waters at the global scale. Proc. Natl. Acad. Scie. 112 (18), 5750–5755. https://doi.org/10.1073/pnas.1500232112.

Stehle, S., Schulz, R., 2015b. Pesticide authorization in the EU—environment unprotected? Environ. Sci. Poll. Res. 22 (24), 19632–19647. https://doi.org/10.1007/s11356-015-5148-5.

Tudi, M., Daniel Ruan, H., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., Phung, D.T., 2021. Agriculture development, pesticide application and its impact on the environment. Int. J. Environ. Res. Public Health 18 (3), 1112. https://doi.org/10.3390/ijerph18031112.

Weisner, O. (2022). The 'Kleingewässer-Monitoring" (KgM)—A monitoring of German small streams and its implications for the environmental risk assessment of pesticides [Dissertation, University Koblenz-Landau]. https://uko.opus.hbz-nrw.de/frontdoor/deliver/index/docId/2284/file/Dissertation_Weisner.pdf.

Wittmer, I.K., Bader, H.-P., Scheidegger, R., Singer, H., Lück, A., Hanke, I., Carlsson, C., Stamm, C., 2010. Significance of urban and agricultural land use for biocide and pesticide dynamics in surface waters. Water Res. 44 (9), 2850–2862. https://doi.org/10.1016/j.watres.2010.01.030.