MSc in Environmental Contamination and Toxicology EMJMD
UPV/EHU – UBx – ULiège
UPPA – NTNU – UPO
an Erasmus+ Consortium
www.ectplus.eu

Co-funded by the Erasmus+ Programme of the European Union

MASTER THESIS PROJECT

# Evaluating the Relationship Between Pesticide Pressure and Macroinvertebrate Bioindicators in Swiss Streams


Anthony Jeanpier Fow Esteves

Swiss Water Association (VSA)

Swiss Federal Institute of Aquatic Science and Technology (EAWAG)

Zürich, Switzerland



PLENTZIA (UPV/EHU), JULY 2025


Universidad del País Vasco Euskal Herriko Unibertsitatea · université de BORDEAUX · UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR · LIÈGE université · NTNU Norwegian University of Science and Technology · U.PORTO

Dr. **Maren Ortiz-Zarragoitia** as International Coordinator of the ECT+ EMJMD of the **University of the Basque Country (UPV/EHU), Spain**

CERTIFIES:

That the research work entitled **"Evaluating the Relationship Between Pesticide Pressure and Macroinvertebrate Bioindicators in Swiss Streams"** has been carried out by **Anthony Jeanpier Fow Esteves** at **EAWAG** under the supervision of **Dr Anne Dietzel and Dr Christiane Ilg** from the **VSA and EAWAG** in order to achieve 30 ECTS as a part of the ECT+ Master program.

In Zürich, June 30th, 2025

Signed:

*Anthony Jeanpier Fow Esteves*

Supervisors(s)

*Anne Dietzel*

*Christiane Ilg*

**PLENTZIA (UPV/EHU), JULY 2025**

# TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# LIST OF ABBREVIATIONS

- **BAFU:** Bundesamt für Umwelt – German abbreviation for "Federal Office for the Environment" (Switzerland).
- **NAWA:** Nationale Beobachtung Oberflächengewässerqualität – German abbreviation for "Swiss National Surface Water Quality Monitoring Program".
- **EU WFD:** European Union Water Framework Directive.
- **SPEARpesticides:** Species at Risk index (macroinvertebrate-based).
- **IBCH:** Indice Biologique Canton Helvétique.
- **GI:** Indicator Group – a component of the IBCH index.
- **VT:** Taxonomic Diversity – a component of the IBCH index.
- **EPT:** Ephemeroptera, Plecoptera, Trichoptera.
- **RQ; *ARQ, CRQ*:** Risk Quotients – used throughout the thesis as Acute Risk Quotient (ARQ) and Chronic Risk Quotient (CRQ).
- **TU; *TU-EC, TU-NOEC*:** Toxic Units – used throughout the thesis as Toxic Unit based on Effective Concentration 50 (TU-EC) and Toxic Unit based on No Observed Effect Concentration (TU-NOEC).
- **MEC:** Measured Environmental Concentration.
- **PEC:** Predicted Environmental Concentration.
- **PNEC:** Predicted No Effect Concentration.
- **$EC_{50}$:** Effective Concentration 50.
- **$LC_{50}$:** Lethal Concentration 50.
- **NOEC:** No Observed Effect Concentration.
- **MSK:** Modular Stepwise Procedure – a set of methods for systematic analysis and assessment.
- **PCA:** Principal Component Analysis.
- **LM:** Linear Model.
- **LMM:** Linear Mixed Model.
- **GAM:** Generalized Additive Model.
- **GAMM:** Generalized Additive Mixed Model.
- **EDF:** Effective Degrees of Freedom.
- **CV:** Cross-Validation.
- **RF:** Random Forest.
- **RMSE:** Root Mean Squared Error.
- **AIC:** Akaike Information Criterion.

# ABSTRACT

Pesticide contamination in streams poses a significant threat to aquatic ecosystems, as evidenced by impacts on macroinvertebrate communities. This study quantitatively evaluated whether pesticide effects over macroinvertebrate communities, indicated by the SPEARpesticides index, relates with chemical metrics based on risk quotients (RQs) and toxic units (TUs) using data from the Swiss National Surface Water Quality Monitoring Program.

The SPEARpesticides index was evaluated alongside other bioindicators and related to chemical metrics derived from 3.5- and 14-days composite samples based on pesticide measurements. Our approach combined data aggregation, chemical metric selection and flexible statistical modeling to examine the explanatory power and stability of RQs and TUs as predictors in models across different sampling periods and temporal windows.

Results showed that both RQs and TUs were negatively related with SPEARpesticides, confirming that higher pesticide pressure corresponds to declines in sensitive macroinvertebrate taxa. However, the strength and consistency of these associations varied: RQs achieved the highest explanatory power in linear models ($R^2$=0.28) but lacked temporal stability, while TUs, particularly calculated from 3.5-days composites (acute approach), provided more consistent, though lower explanatory power ($R^2$=0.17). Incorporating environmental parameters such as proportion of agricultural land use and flow velocity further improved model performance by up to ($R^2$=0.60), underscoring the value of integrated, multi-factor approaches. Despite these advances, substantial unexplained variance persisted, especially at the extremes of the exposure gradient, reflecting the complexity of field conditions, data limitations, and the challenges of harmonizing and aligning chemical and biological monitoring.

By integrating biological and chemical metrics, this research advances understanding of pesticide impacts on macroinvertebrate communities in Swiss streams and supports the ongoing refinement of ecological assessment and monitoring strategies for freshwater ecosystems.

# RESUMEN

La contaminación por pesticidas en ríos representa una amenaza significativa para los ecosistemas acuáticos, evidenciándose en impactos sobre comunidades de macroinvertebrados. Este estudio evaluó la relación de los efectos de los pesticidas sobre comunidades de macroinvertebrados, a través del índice SPEARpesticides, con métricas químicas como los cocientes de riesgo (RQs) y unidades tóxicas (TUs) utilizando bases de datos del Programa Nacional Suizo de Monitoreo de la Calidad del Agua Superficial.

El índice SPEARpesticides se evaluó junto con otros bioindicadores y se relacionó con métricas químicas derivadas de muestras compuestas de 3.5 y 14 días basadas en concentraciones de pesticidas. La investigación combinó diferentes modelos estadísticos de regresión para explorar el poder explicativo y la estabilidad de los RQs y TUs como variables independientes en modelos a través de diferentes ventanas temporales.

Los resultados mostraron que RQs y TUs se asociaron negativamente con SPEARpesticides, confirmando que una mayor exposición a pesticidas refiere a una disminución de la respuesta en el índice. Los RQs alcanzaron el mayor poder explicativo en modelos lineales ($R^2$=0.28), pero carecieron de estabilidad a través de las ventanas temporales, mientras que los TUs calculados de muestras compuestas de 3.5 días mostraron asociaciones más consistentes pero con menor poder explicativo usando modelos generalizados ($R^2$=0.17). Además, la incorporación de parámetros ambientales como la proporción de uso de suelo para agricultura y la velocidad del flujo mejoró el desempeño de los modelos hasta en un ($R^2$=0.60). A pesar de ello, los modelos mostraron niveles de incertidumbre considerables, especialmente en los extremos del gradiente de exposición química, lo que refleja la complejidad de las condiciones reales, las limitaciones de los datos y los desafíos de armonizar el monitoreo químico y biológico.

La integración de las bases de datos biológica y química en este estudio contribuye a un mejor entendimiento del impacto de los pesticidas sobre las comunidades de macroinvertebrados en ríos suizos y respalda la mejora continua de las estrategias de monitoreo y evaluación del riesgo ecológico en ecosistemas acuáticos.

# 1. INTRODUCTION

## 1.1. General Background

### 1.1.1. Aquatic ecosystem health and micropollutant contamination

Aquatic ecosystems are critical to global biodiversity, providing essential services such as water purification, nutrient cycling, and habitat for a wide range of species (Albert et al., 2021; Haase et al., 2023). However, these systems are increasingly threatened by anthropogenic pressures. By 2000, around 80% of the world's population lived in regions where freshwater resources were at risk due to habitat destruction, eutrophication, excessive water extraction, and pollution (Vörösmarty et al., 2010). With an expected 26% increase in the world's population by 2050, the situation seems to become more complicated (FAO, 2021).

Streams, as integral components of aquatic ecosystems, are particularly important due to their ecological roles and vulnerability to stressors. They support diverse communities, especially benthic macroinvertebrates, which are central to food web dynamics and nutrient cycling, and act as natural filters that break down organic matter and contaminants, functions especially vital in agricultural landscapes (Shah et al., 2020; Yeakley et al., 2016). Streams and their riparian zones provide essential ecosystem services such as flood regulation, groundwater recharge, and water filtration, all of which contribute to environmental stability at both local and global scales. Additionally, these areas enhance ecosystem services by sequestering carbon, moderating temperatures, and buffering against extreme weather events (Vári et al., 2022; Yeakley et al., 2016).

The health and integrity of stream ecosystems are shaped by interactions among physical, chemical, and biological factors, all of which are influenced by anthropogenic activities such as agriculture, urbanization, and industrial discharges (FAO, 2021; Rasmussen et al., 2012). Streams are especially susceptible to chemical pollutants from nearby agricultural runoff and urban areas. Unlike larger water bodies, streams have limited dilution capacity, so pollutants can quickly reach ecologically harmful concentrations (Morin & Artigas, 2023).

Of particular concern are micropollutants, a diverse group of chemical compounds including pesticides, pharmaceuticals, personal care products, and heavy metals (Ginebreda et al., 2014; Munz et al., 2017; Weisner et al., 2022). These contaminants are notable for their persistence, potential for bioaccumulation, and ability to cause chronic or sub-lethal effects on aquatic organisms and ecosystem processes (SETAC, 2018). For instance, endocrine disruptors can impair fish reproduction, while heavy metals and persistent organic chemical reduce biodiversity and alter food web dynamic, undermining ecosystem stability (Abbasi et al., 2022; Ginebreda et al., 2014).

Research shows that micropollutant concentrations, such as those from wastewater treatment plants, can significantly affect vulnerable species like microorganisms and invertebrates, leading to changes in stream community structure (Burdon et al., 2019; Munz et al., 2017; Tlili et al., 2017). The combined effect of pollutants, habitat alteration, and climate change can further exacerbate impacts on aquatic organisms, emphasizing the need for integrated assessment and management strategies (Beyer et al., 2014; Ginebreda et al., 2014).

Given the centrality of micropollutants in shaping stream ecosystem health, the following section focuses on anthropogenic pesticide pressures as a critical subset of these contaminants.

#### 1.1.2. Anthropogenic (pesticide) pressure and its consequences.

Anthropogenic pressures on streams have intensified in recent decades, with pesticide contamination emerging as a particularly critical threat (FAO, 2021; Weisner et al., 2022). Pesticides reach aquatic environments through multiple pathways, including runoff from agricultural fields, leaching, and atmospheric deposition, as well as non-agricultural uses such as urban landscaping and public health applications (Beketov et al., 2013; Doppler et al., 2012, 2024). The presence of both agricultural and non-agricultural sources complicates regulatory efforts, as cumulative impacts are often underestimated when only individual inputs are considered.

Persistent pesticide residues are widely documented in small streams, posing significant risks even at low concentrations (Bai et al., 2018; Liess et al., 2021). Numerous studies have demonstrated negative correlations between pesticide concentrations and biodiversity, especially among sensitive aquatic invertebrates that serve as key bioindicators of ecosystem health (Beketov et al., 2013; Burdon et al., 2019; Ganatra et al., 2021). Both chronic and acute exposures to pesticides have been shown to impair freshwater ecosystems, resulting in the loss of sensitive species, and altered community structures (Kumar et al., 2023; Stehle & Schulz, 2015).

The ecological consequences of pesticide contamination are multifaceted, encompassing both direct and indirect effects in streams. Direct effects manifest at the individual or species level, impairing physiological functions, reproduction, or survival, and leading to marked reductions in sensitive taxa such as aquatic invertebrates, key components of freshwater food webs (Beketov et al., 2013; Liess & von der Ohe, 2005). The loss of these species disrupts ecological balance, reduces biodiversity, and diminishes the resilience of aquatic systems to further stressors. Indirect effects operate at the community and ecosystem levels, manifesting as altered predation and competition, changes in nutrient cycling, and weakened trophic interactions due to the decline or loss of key taxa (Hou et al., 2025; Shah et al., 2020).

Pesticide contamination can shift species composition and undermine ecosystem functions and integrity. In lowland streams, increased toxicity from pesticide runoff has been linked to significant reductions in species richness, with some studies reporting losses in taxa of up to 42% of the recorded taxonomic pools (Beketov et al., 2013). Notably, these adverse effects can occur at concentrations lower than those deemed "safe" by regulatory frameworks, highlighting critical gaps in current risk assessments that often focus on individual compounds rather than complex mixtures (Kienzler et al., 2016; Schriever et al., 2025).

Moreover, the presence of pesticide mixtures in streams can produce synergistic toxicological effects, amplifying risks beyond those predicted by single-chemical assessments. Addressing these challenges requires robust monitoring and mitigation strategies, as well as a nuanced understanding of ecological impacts. This often necessitates the integration of biological assessments with chemical analyses to delineate the effects of pesticides and other micropollutants on stream health (Bettinetti et al., 2020; Rico & Van den Brink, 2015). Ongoing research continues to refine models and indicators to better assess and manage the risks posed by these contaminants (Heß et al., 2024; Hunt et al., 2017).

## 1.2. Bioindicators in ecological assessment.

Biological communities are central to ecosystem health assessment because they provide an integrated measure of ecological status over time. Unlike chemical analyses, which offer a snapshot of contaminant levels, biological indicators can reflect the cumulative impacts of pollution, habitat alteration, and hydrological changes on aquatic systems (de Castro-Català et al., 2016). Among the most informative bioindicators are macroinvertebrates, fish and algae, whose varied sensitivities to environmental stressors make them reliable proxies for water quality and ecosystem resilience (Herman & Nejadhashemi, 2015; Tlili et al., 2017).

Regulatory frameworks increasingly recognize the value of biological communities for water quality assessment. The European Union Water Framework Directive (EU WFD) mandates the use of biological quality elements, including macroinvertebrates, fish, and macrophytes, as part of its approach to achieving good ecological status in water bodies (EC, 2011). This directive emphasizes the integration of hydromorphological, chemical, and biological data to support adaptive management and restoration. Similarly, Swiss water protection laws recommend the use of bioindicators alongside chemical metrics to assess ecological integrity, reflecting a commitment to biodiversity conservation and comprehensive evaluation of anthropogenic pressures (BAFU, 2013; Gewässerschutzverordnung, GSchV, 1998). Through these frameworks, the demand for robust bioindicator systems and standardized methodologies has grown, supporting more effective ecosystem management (Espinar-Herranz et al., 2025; Heß et al., 2024).

Biological indices in ecological studies generally fall into three categories: those describing basic community structure (e.g., species richness), those evaluating deviations from reference conditions, and those targeting specific stressors (Khaliq et al., 2024). Macroinvertebrate-based assessments offer practical advantages: they are relatively easy to collect and identify, and they integrate the effects of both acute and chronic pollution as well as multiple environmental stressors over time (Ganatra et al., 2021; Hunt et al., 2017). This makes them useful for long-term monitoring and for detecting subtle ecological changes that chemical monitoring might miss.

### 1.2.1. Overview of bioindicators used in this project: SPEARpesticides, IBCH and EPT.

Macroinvertebrate-based bioindicators such as SPEARpesticides, the Swiss Biological Index (IBCH) with its VT (Diversity Class Index) and GI (Indicator Group Index) components, and the EPT richness are foundational tools in stream ecological assessment, each offering distinct advantages and facing specific challenges. The SPEARpesticides (SPEcies At Risk), index developed in Germany (Liess & von der Ohe, 2005) is designed to evaluate the ecological impact of organic pollutants, particularly pesticides, on stream macroinvertebrate communities by identifying vulnerable species based on life-history traits and pesticide sensitivity. Its foundation lies in quantifying the proportion of “at risk” taxa within a community and linking observed declines to pesticide pressure, even in the presence of chemical mixtures (Beketov et al., 2013; Schäfer et al., 2013; Wogram & Liess, 2001). The index has been widely applied in European and international contexts for stream monitoring and research, detecting subtle, community-level impacts of pesticides and supporting water management strategies (Schäfer & Liess, 2013).

The IBCH (Indice Biologique Canton Helvétique or Swiss Biological Index) is Switzerland’s standardized macroinvertebrate-based index for evaluating stream health, with the VT measuring overall biodiversity and the GI focusing on the abundance of key indicator taxa. The IBCH, along with the VT and GI, is used to monitor temporal changes and compare ecological conditions across regions and stream types in Switzerland (BAFU, 2019, 2022).

The EPT richness, based on the presence and abundance of Ephemeroptera, Plecoptera, and Trichoptera, is a widely used metric for evaluating stream health and water quality worldwide (Suhaila & Che Salmah, 2017). Its sensitivity to various environmental stressors enables the detection and quantification of pollution severity, identification of contamination sources, and assessment of habitat quality (Haase et al., 2023; Khaliq et al., 2024). The EPT richness is often integrated with other indices to prioritize areas for conservation or restoration, but it should not be used alone in watercourses with naturally low EPT taxa richness, such as lowland rivers, large fine-substrate rivers, slow-flowing or stagnant water bodies, or urban streams, where pollution-tolerant groups dominate (Tubić et al., 2024); instead, a multimetric approach is recommended to assess potential confounding factors and improve the reliability and comprehensiveness of environmental evaluations (Poikane et al., 2016; Schuwirth et al., 2015).

Despite their broad application and regulatory endorsement, the use of these macroinvertebrate indicators faces challenges when attempting to elucidate their relationship with chemical pressure alone due to multistressor reality. Community structure and indicator responses can be influenced by environmental factors such as hydrology, substrate quality, and physical habitat integrity, independently of chemical stressors (Robinson et al., 2014; Villeneuve et al., 2018). To address these complexities, researchers have employed advanced statistical models and multimetric approaches that account for multiple stressors, improving the reliability and interpretive power of the assessments (Liess et al., 2021; Tampo et al., 2021). However, careful consideration of confounding factors remains key to their effective use (Schuwirth et al., 2015).

Case studies across Europe highlight the versatility and effectiveness of these indices in diverse ecological contexts, facilitating both local and international initiatives aimed at improving water quality and conserving aquatic ecosystems (Poikane et al., 2016, 2020).

### 1.3. Chemical metrics: Risk Quotients (RQs) and Toxic Units (TUs).

Risk Quotients (RQs) and Toxic Units (TUs) are foundational tools in aquatic risk assessment, enabling the evaluation of ecological risks posed by chemical substances in surface waters. RQs are typically calculated as the ratio of the measured or predicted environmental concentration (MEC or PEC) of a chemical to its predicted no-effect concentration (PNEC), typically derived from toxicity tests on aquatic organisms (Junghans et al., 2013). An RQ greater than 1 indicates a potential risk to aquatic life and warrants further investigation or regulatory action. These metrics are integrated into regulatory frameworks, such as the EU WFD, to guide restrictions, bans or mitigation strategies and are applied for both acute and chronic exposure assessments (Moe et al., 2022; Peterson, 2006). In Switzerland, RQs are also routinely used in national and cantonal water quality monitoring programs to identify priority pollutants and inform management actions (Junghans et al., 2013; Spycher et al., 2018).

RQs are not limited to single substances; they can be extended to assess the combined risk of multiple chemicals by summing individual risk, a method known as summations or mixture approach (Backhaus & Faust, 2012; Wei et al., 2022). This approach is particularly relevant for evaluating mixture effects in environments where multiple contaminants co-occur.

TUs provide a standardized metric for assessing the risk posed by individual chemicals and mixtures in aquatic systems. A TU is calculated as the ratio of a chemical's concentration in water to its effect concentration, commonly $LC_{50}$ or $EC_{50}$, with a TU value of 1 indicating that the environmental concentration matches the effect threshold for 50% of test organisms (von der Ohe & de Zwart, 2013) TUs are especially valuable for evaluating chemical mixtures, as the toxic unit summation method allows researchers to estimate the combined toxicity of complex contaminant mixtures (Backhaus & Faust, 2012; Ginebreda et al., 2014). Alternatively, the maximum toxicity approach, which uses the highest TU among mixture components, is particularly relevant in acute exposure scenarios (Hunt et al., 2017; Liess et al., 2021).

Both RQs and TUs have been applied in a range of contexts, from large-scale, long-term monitoring to event-driven sampling following acute pollution episodes (Ganatra et al., 2021; Hunt et al., 2017; Kienzler et al., 2019; Rasmussen et al., 2012). Their flexibility enables researchers and regulators to tailor risk assessments to specific environmental conditions and pollution profiles (Beyer et al., 2014; Švara et al., 2021).

Despite their widespread use, RQs and TUs face some challenges. They often do not fully account for interactive, synergistic, or antagonistic effects within chemical mixtures, which can result in under- or overestimation of chemical assessment (Backhaus & Faust, 2012; Spycher et al., 2018). Additionally, uncertainties in establishing accurate PNECs and effect concentrations, often based on laboratory-derived data, can affect the reliability of these metrics under field conditions (Junghans et al., 2011; Kienzler et al., 2019). To address these challenges, there is a growing emphasis of integrating site-specific ecological factors and combining chemical metrics with biological indices to enhance the validity and applicability of risk assessments (Bettinetti et al., 2020; Burdon et al., 2019; Lee et al., 2020).

As the next section will discuss, the integration of chemical metrics, such as RQs and TUs, with biological assessment tools has been identified as a critical step for comprehensively evaluating the complexity of ecological risks in aquatic ecosystems.

## 1.4. Chemical-Biological data integration and study context.

Integrating chemical and biological data is basic for a comprehensive assessment of stream health, particularly when evaluating the impacts of pesticides and other micropollutants. While chemical indicators, such as RQs and TUs, evaluate pollutant concentrations and provide valuable information on potential ecological risks, they do not directly capture the cumulative, chronic, or sublethal effects on aquatic communities. In contrast, biological indices, especially macroinvertebrate-based metrics like SPEARpesticides, IBCH, and EPT, reflect the combined effects of pollutants and offer a more nuanced perspective on ecosystem integrity.

Due to differences in land use, hydrology, and natural gradients at the catchment level, streams are inherently variable in their chemical and biological profile. Integrating both chemical and biological data enables context-specific risk assessments that are sensitive to these local conditions, thereby improving the ecological relevance of monitoring and management strategies. Recent studies have integrated RQs and TUs with biological monitoring metrics to enhance the robustness of water quality assessments. For example, while indices like SPEARpesticides have shown negative correlations with chemical pressures assessed through Tus. However, the integration of both types of data remains important for further analysis, as biological responses reflect cumulative stressors and recovery dynamics, while chemical metrics identify specific drives of impairment (Burdon et al., 2019; Liess et al., 2021).

However, macroinvertebrate indices do not exclusively reflect chemical influences; environmental variables such as temperature, flow dynamics, and habitat complexity also play significant roles in shaping biological responses (Rico et al., 2016; Rico & Van den Brink, 2015). This complexity underscores the necessity for multi-faceted assessment frameworks that account for both chemical and non-chemical stressors (Haase et al., 2023). Researchers have addressed these challenges by employing advanced experimental designs and statistical approaches to disentangle the effects of pollutants from other environmental factors (Epele et al., 2024; Rasmussen et al., 2012). Integrating chemical metrics with biological indices in event-driven and long-term monitoring studies provides critical insights into how multiple stressors interact to impact aquatic communities across spatial and temporal contexts. This helps us better understand environmental health and practice more informed management (Schriever et al., 2025).

Within Switzerland, this integrated approach is operationalized through the Swiss National Surface Water Quality Monitoring Program (NAWA). NAWA includes periodic chemical sampling and analysis of various micropollutants, which are useful for calculating RQs. It also includes macroinvertebrate monitoring, the results of which are the main input for bioindicators such as IBCH and SPEARpesticides (BAFU, 2019). Recent Swiss studies have revealed connections between macroinvertebrate communities and ecomorphology, agricultural area in the catchment, catchment area, discharge, and insecticide application rates (Hutter et al., 2019; Ilg & Alther, 2024; Khaliq et al., 2024). Other studies have examined the importance of micropollutant pressure in streams using chemical metrics (Daouk et al., 2022; Doppler et al., 2024). Some studies have explicitly linked these biological indicators with chemical metrics (Burdon et al., 2016, 2019; Junghans et al., 2019; Munz et al., 2017). However, no study has explored the long-term national datasets of micropollutants and macroinvertebrates together.

Addressing this gap, the present research aims to use the SPEARpesticides index, alongside other bioindicators, to assess the status of Swiss streams with respect to pesticide contamination. This assessment leverages the Swiss micropollutant database and macroinvertebrate monitoring data collected from 2018 to 2023, considering the characteristics of Swiss streams and the challenges identified in recent literature (Chow et al., 2020; Doppler et al., 2024; Spycher et al., 2018). The study adapts the concept of TUs to align with 14- and 3.5-days sampling periods, guiding the selection of ecotoxicological endpoints for risk metrics and enabling a direct comparison of the performance of RQs and TUs in explaining bioindicator responses. This research applies nationally representative data to evaluate the effects of pesticide mixtures on macroinvertebrate communities in Swiss watercourses by combining bioindicator analysis with chemical metrics. The findings aim to support ecological monitoring and inform risk management in Switzerland.

# 2. OBJECTIVES AND HYPOTHESES

## 2.1. Objectives

### 2.1.1. General Objective

The general objective of this work is to quantitatively evaluate the relationship between chemical metrics (risk quotients and toxic units based on pesticide concentrations) and SPEARpesticides, along other bioindicators, to evaluate how effectively these bioindicators reflect pesticides impacts on macroinvertebrate communities in Swiss streams, considering the characteristics of the NAWA monitoring framework.

### 2.1.2. Specific Objectives

- Evaluate the effectiveness of risk quotients and toxic units under acute (3.5-days samples) and chronic (14-days samples) approaches for explaining variations in macroinvertebrate-based bioindicators.
- Investigate the influence of environmental parameters on the ability of risk quotients and toxic units to explain changes in macroinvertebrate-based bioindicators.
- Assess the influence of sampling periods and temporal integration windows in the relationship between pesticide exposure and macroinvertebrate bioindicators.

## 2.2. Hypotheses

### 2.2.1. General Hypothesis

Elevated pesticide pressure, as quantified by higher RQ and TU values, will lead to a measurable decline in sensitive macroinvertebrate taxa, resulting in lower values of bioindicators such as SPEARpesticides. This decline is expected due to the direct toxic effects of pesticides on sensitive taxa and the indirect alteration of community structure.

### 2.2.2. Specific Hypotheses

- Chemical risk metrics (RQs and TUs), particularly when temporally aligned to biological sampling, will serve as robust predictors of variation of the stressor-related bioindicators, reflecting the linkage between chemical exposure and biological effect.
- With appropriate harmonization and alignment of Swiss monitoring conditions, and accounting for key environmental variables, bioindicators more effectively capture the ecological impact of pesticide mixtures on macroinvertebrate communities than when only chemical metrics are used.
- The strength of the relationship between chemical metrics and bioindicator responses will vary with the sampling period and time window, with acute (3.5-days samples) and shorter windows capturing episodic toxicity events to which sensitive indices like SPEARpesticides are more responsive, while chronic (14-days samples) and longer windows will reflect cumulative impacts, potentially resulting in less sensitive bioindicator responses.

# 3. DATA AND METHODS

This study is based entirely on the analysis of existing datasets collected through national Swiss monitoring programs. No new fieldwork or primary data collection was conducted during this study; all results derive from the integration and statistical analysis of previously gathered chemical and biological monitoring data. All site selection and sampling were performed by the Swiss National Surface Water Quality Monitoring Program (NAWA); this study's contribution is in the downstream integration, harmonization, and analysis of these existing data sets.

Thus, the overall workflow of this study involved: (1) obtaining and harmonizing macroinvertebrate and chemical monitoring data from national databases; (2) calculating chemical metrics; (3) integrating datasets by aligning temporal and spatial sampling events; and (4) applying statistical models to evaluate relationships between pesticide exposure, bioindicator responses, and environmental parameters.

### 3.1. Data Resources and Harmonization

Two primary datasets were integrated for this study:

- **Macroinvertebrate Dataset (MI-DS)**
- Comprising data from 42 streams across 16 cantons, including 213 monitoring events conducted between 2018 and 2023. Each stream was monitored at least annually, with 70% of the samples collected during spring.
- **Micropollutant Dataset (MP-DS)**
- Encompassing data from 53 streams in 17 cantons, derived from NAWA programs (BAFU, 2013), spanning 2011–2023. The dataset includes approximately 10,600 samples and measurements for around 770 substances, with a focus on pesticides, pharmaceuticals, and industrial chemicals. Chronic and acute toxicity criteria were available for 440 substances (Swiss Centre for Applied Ecotoxicology, 2020).

To harmonize the datasets, both were prepared to be compatible in formatting, naming, and units. As a key step, MP-DS samples were categorized by composite sampling period (3.5-days for acute, 14-days for chronic exposure) and then temporally aligned to macroinvertebrate monitoring events in MI-DS for sub-sequent joint analysis.

### 3.2. Study Area and Sampling Sites

Switzerland's river network extends approximately 65,000 km, of which 43% have catchments influenced by agricultural or urban sources of micropollutants, mainly concentrated in the Swiss Plateau. In contrast, many sites with minimally expected anthropogenic pollution are found in the Alpine zones; however, these low-impact sites were not included in the NAWA monitoring framework for micropollutants (Doppler et al., 2020).

This study covers 41 streams distributed primarily across the Swiss Plateau (Figure 1), with a few additional locations in the Jura and Alpine zones, as part of the NAWA TREND

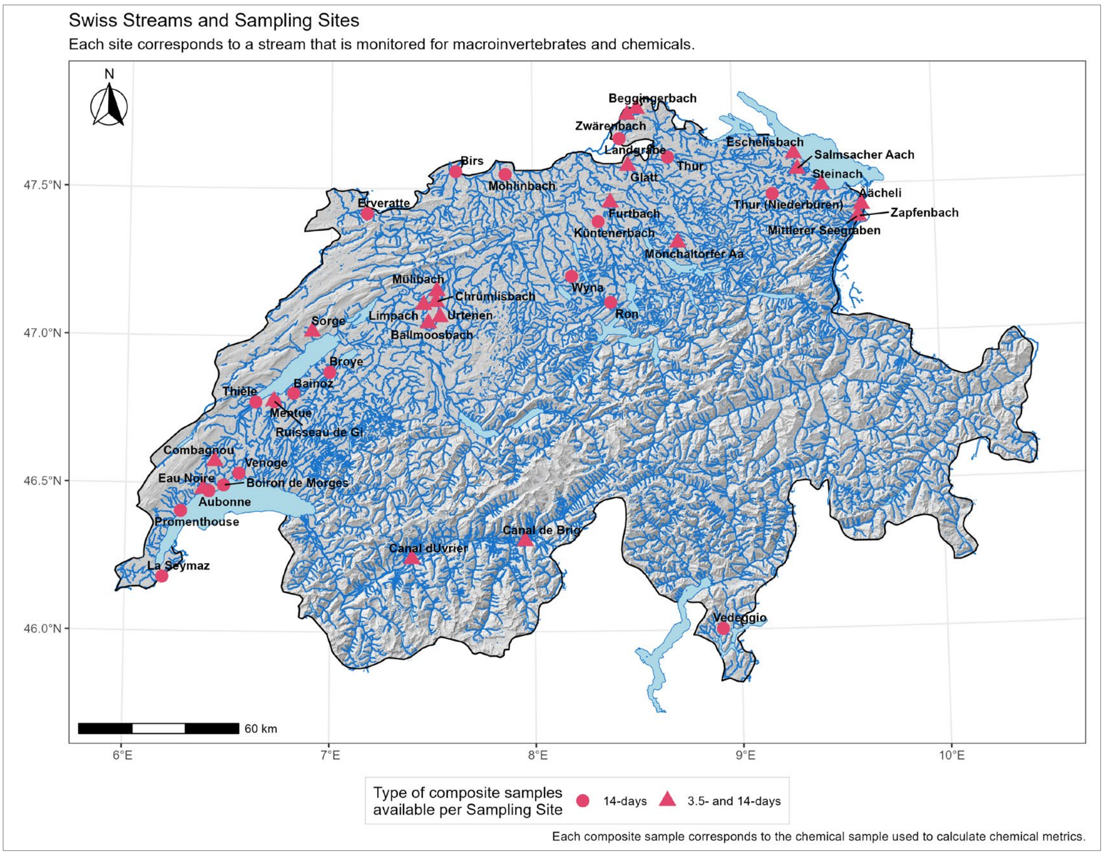


*Figure 1. Spatial distribution of the 41 streams and sampling sites considered in the study.*

(BAFU, 2013). Sampling sites span an altitudinal gradient from 250 to 660 meters above sea level.

## 3.3. Swiss Water Quality Assessment: Biological and Chemical Data Collection and Calculation

The assessment of Swiss watercourses is based on micropollutants, nutrients, heavy metals, and biological elements, with a particular focus on micropollutants and biological diversity in small and medium-sized streams. For the purposes of this study, our focus was on macroinvertebrate communities and pesticides. Monitoring is conducted under the NAWA framework (BAFU, 2013) and following the Modular Stepwise Procedure (MSK) methods (BAFU, 1998; MSK, 2025). These methods were developed collaboratively by the Federal Office for the Environment (BAFU), the Swiss Federal Institute of Aquatic

Science and Technology (EAWAG), the Swiss Water Association (VSA) and the cantons to assess different components of aquatic ecosystems.

#### 3.3.1. Biological Component

*Field sampling and data flow:*

Macroinvertebrate sampling followed standardized MSK protocols (BAFU, 2019; BDM, 2021, 2022). A section of the stream is selected, and eight points are established to span substrate and flow velocity diversity. Macroinvertebrates are collected and samples are cleaned, preserved, and labeled for traceability. In the laboratory, specimens are sorted and identified primarily to family level, and their abundance is recorded. All data is archived and entered into standardized protocols for automated bioindicator calculation.

– *Calculation of bioindicators:*

The selection of IBCH, SPEARpesticides, and EPT as primary bioindicators follows the rationale established in the introduction: these indices are widely used in Switzerland, sensitive to pesticide impacts, and supported by both regulatory and research frameworks.

*IBCH index:*

The IBCH (Swiss Biological Index) assesses ecological status using two components:

- *Diversity Class (VT value):* Number of taxa (family level), adjusted for hydrological regime and scaled from 0 to 1.
- *Indicator Group (GI value):* Sensitivity of the most water quality-sensitive taxon present, also scaled from 0 to 1.

The final IBCH is calculated as:

$$IBCH = 0.62 \times VT\ value + 0.38 \times GI\ value$$

Status classes are defined as: Very good (≥0.8), Good (0.6–0.8), Moderate (0.4–0.6), Unsatisfactory (0.2–0.4), Bad (<0.2).

*SPEARpesticides index:*

SPEARpesticides is a bioindicator designed to detect the effects of short-term, pulse pesticide pollution in agricultural streams. It classifies species based on four key traits: (i) their physiological sensitivity to insecticides and related compounds, (ii) their generation time, (iii) presence of aquatic stages and (iv) their capacity for migration and recolonization (Liess, 2023). It is calculated at the family level using the latest trait database (Liess et al., 2021). SPEARpesticides in Switzerland does not uses the normalization proposed in the last version as it is based on German reference conditions (Knillmann et al., 2018), which were considered as unsuitable for Switzerland.

The formula is:

$$SPEARpesticides = \frac{\sum_{n}^{i=1} log\ (4x_i + 1)\ \times y_i}{\sum_{n}^{i=1} log\ (4x_i + 1)} \times\ 100$$

Where $x_i$: Abundance of the taxon i; $y_i$ : 1 (sensitive) or 0 (insensitive). Status classes are defined as: Very good (>44), Good (33–44), Moderate (22–33), Unsatisfactory (11–22), Bad (≤11).

*EPT Richness:*

EPT richness is a widely used indicator of stream ecological quality (Tubić et al., 2024), reflecting the diversity of three insect orders: Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies). The number of distinct taxonomic families belonging to each EPT was counted per sample. The total EPT richness is the sum of these three counts.

### **3.3.2.** Chemical component

*Field sampling and data flow:*

The primary sampling strategies are:

- ***14-days composite samples:*** Continuous, time-proportional composites are used for chronic exposure assessment. This type of sampling is mandatory to evaluate the numerical requirements stated in the Water Protection Ordinance (Gewässerschutzverordnung, GSchV, 1998).

- ***3.5-days composite samples:*** Shorter, time-proportional, composites targeting acute pollution events, especially during peak pesticide application (Doppler et al., 2012; Spycher et al., 2018).

In the NAWA monitoring program, 14-days composite water samples are collected throughout the year using automated samplers maintained at 4°C. During the main pesticide application period (April to July), 3.5-days composite samples are also collected from some specific streams under larger agricultural pressure to capture short-term fluctuations and peaks. For consistency, four consecutive 3.5-days samples are averaged to generate a synthetic 14-days value, ensuring comparability with directly measured 14-days composites. As a result, the annual dataset comprises both measured and calculated 14-days composite samples, while 3.5-days samples are only available during periods of expected peak pesticide concentrations and only at specific streams (Daouk et al., 2022; Doppler et al., 2017).

Chemical analyses are performed primarily using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS and GC-MS/MS) for the majority of substances. For pyrethroids and some other compounds, liquid-liquid extraction (LLE) is followed by gas chromatography with tandem mass spectrometry. Specific analytical protocols and instrumentation may vary between years and laboratories, leading to differences in the number of substances analyzed and their respective limits of quantification (LOQ). These methodological details are described in Daouk et al. (2022) and Spycher et al. (2019), and further information about the methods is available in Moschet et al. (2019) and Rösch et al. (2019).

- *Calculation of chemical metrics:*

The calculation of RQs and TUs for both acute (3.5-days) and chronic (14-days) exposures directly addresses the study's secondary objective of comparing chemical metrics under different temporal aggregation or time windows, as outlined in the introduction.

*Risk Quotient (RQ) and Toxic Unit (TU) Approaches*

To assess the broader pressure of pesticides on the streams, risk quotients and toxic units were calculated for all substances categorized as insecticides, herbicides, or fungicides.

These calculations were based on environmental quality criteria (EC, 2011; Junghans et al., 2019) or critical effect concentrations (Lewis et al., 2016).

*Individual RQ*

$$RQ_i = MEC_i/QC_i$$

Where $MEC_i$ is the measured environmental concentration and $QC_i$ is the relevant acute (for 3.5-days samples) or chronic (for 14-days samples) quality criterion for calculating the Acute Risk Quotient (ARQ) and Chronic Risk Quotient (CRQ) respectively (Bai et al., 2018; Daouk et al., 2022; Spycher et al., 2018) for each substance "i" measured in each sample. It is important to remark that quality criteria were derived based on the most sensitive taxon available for each substance, which is not always based on invertebrates (EC, 2011; Junghans et al., 2011). It is also relevant to mention that only substances labeled as toxic to invertebrates and with QC derived under robustness level 1 and 2 were considered for the risk quotient calculation (Swiss Centre for Applied Ecotoxicology, 2020).

*Individual TU*

$$TU_i = MEC_i/CEC_i$$

Where $MEC_i$ is the measured environmental concentration and $CEC_i$ is the critical effect concentration for *Daphnia magna* for each substance "i" measured in each sample. For acute assessment (3.5-days samples) the "Effective Concentration 50 at 48 hours" ($EC_{50}$) values were used to calculate the TU-EC. For the chronic assessment (14-days samples) the "No Observed Effect Concentration at 21 days" (NOEC) values were used to calculate TU-NOEC. The ecotoxicological data were sourced from the Pesticide Properties Database (PPDB) (Lewis et al., 2016).

- *Aggregation of Chemical metrics per sample:*

For each sample, two aggregation approaches were applied:

*Mixture Approach:*

$$RQ_{mix} = \sum_{i=1}^{n} RQ_i \,; \; TU_{mix} = \sum_{i=1}^{n} TU_i$$

*Maximum Approach:*

$$RQ_{max} = max(RQ_1, RQ_2, RQ_3, \ldots, RQ_n)$$

$$TU_{max} = max\ (TU_1, TU_2, TU_3, \ldots, TU_n)$$

Where “n” is the number of substances detected in the sample.

Applying both mixture and maximum approaches allows for the assessment of cumulative (chronic) and peak (acute) exposure scenarios, which are critical for understanding the ecological risks posed by pesticide mixtures in running waters (Junghans et al., 2013; von der Ohe & de Zwart, 2013; Wei et al., 2022).

Because multiple samples were collected over time at each stream, summary statistics were calculated for each metric (mixture and maximum approaches) within the considered time windows (explained in next section) prior to the macroinvertebrate sampling:

- Mean (e.g., $RQ_{(mix;\ mean)}, RQ_{(max;\ mean)}, TU_{(mix;\ mean)}, TU_{(max;\ mean)}$).
- Median (e.g., $RQ_{(mix;\ median)}, RQ_{(max;\ median)}, TU_{(mix;\ median)}, TU_{(max;\ median)}$).
- Maximum (e.g., $RQ_{(mix;\ max)}, RQ_{(max;\ max)}, TU_{(mix;\ max)}, TU_{(max;\ max)}$).

### 3.4. Data alignment:

To relate chemical metrics to biological responses from the bioindicators values, we implemented a temporal alignment approach: for each macroinvertebrate sampling event, chemical metrics (Risk Quotients, Toxic Units) were summarized as the mean, median, and maximum over seven cumulative time windows (1 week, 2 weeks, 1 month, 2 months, 3 months, 6 months, and 1 year) preceding the macroinvertebrate sampling date. For example, the "1-week" window includes all chemical samples collected in the week prior to each macroinvertebrate sample. These samples are used to calculate and summarize chemical metrics, as previously outlined. The "2-weeks" window includes all samples from the preceding two weeks, and so forth. This method allowed us to maximize the use of available data despite the lack of perfectly time-matched chemical and biological samples. Therefore, for each stream and macroinvertebrate monitoring, these chemical metrics, calculated in both mixture and maximum forms, and aggregated as mean, median, and maximum, were systematically aligned to the corresponding bioindicator values.

## 3.5. Data and Statistical Analysis

Statistical analyses were designed to systematically test the relationships between chemical pressure (as measured by RQs and TUs across sampling periods and different lengths of the considered time window), macroinvertebrate bioindicators, and environmental parameters.

### 3.5.1. Data Preparation

All numerical variables, including environmental parameters (e.g., catchment area, discharge, land use fractions, water temperature, etc.) and chemical metrics, were standardized (z-score scaling) to ensure comparability. Chemical metrics were log(x+1)-transformed prior to scaling to mitigate skewness. No missing data was found.

### 3.5.2. Statistical and Modelling Analysis

To evaluate the relationship between chemical metrics and macroinvertebrate bioindicators, we applied a suite of statistical models, with model selection and validation tailored to data availability and structure for each sampling period. Importantly, each chemical metric (RQmix, RQmax, TUmix, and TUmax) was evaluated individually as a sole predictor in separate models for each data subset defined by time window, isolating and assessing the specific effect of each chemical metric on bioindicator responses.

### 3.5.3. Modeling approach:

We began by calculating Spearman rank correlations to assess initial associations between chemical metrics and bioindicators across all time windows and sampling periods. For regression analyses, we applied linear models (LM), generalized additive models (GAM), linear mixed-effects models (LMM), and generalized additive mixed models (GAMM), incorporating year as a random effect where appropriate. The choice of modeling approach was guided by both data characteristics and model diagnostics:

- LMMs were prioritized for datasets where repeated measurements across years could be leveraged, with year included as a random effect to account for temporal structure.
- GAMs and GAMMs were used primarily when linear models showed poor fit or violated assumptions (e.g., non-linearity, heteroscedasticity), allowing for flexible, data-driven modeling of potentially complex relationships.

Model assumptions (normality, homoscedasticity, linearity, and outliers) were checked via residual diagnostics and Cook's distance. Model selection was mainly guided by explanatory power, Akaike's Information Criterion (AIC), and ecological plausibility.

To assess model robustness and potential overfitting, we performed cross-validation (CV) for all main models. Five-fold CV was used for both sampling periods due to limited sample size. As an additional benchmark and to test for potential non-linear or interactive effects not captured by parametric models, we included Random Forest (RF) regression in the CV framework. However, the primary purpose of RF was not predictive optimization, but rather to provide a flexible, non-parametric reference for model comparison. Model performance was evaluated using cross-validated root mean squared error (RMSE) and $R^2$.

*Conceptualization for the implementation of abiotic factors in statistical models*

In this research, abiotic factors are divided into two main categories: chemical metrics (e.g., RQs and TUs) and environmental parameters (e.g., agricultural land use, water temperature, etc.). To avoid confusion, we refer to the latter group as environmental parameters throughout the document.

Environmental parameters were grouped into five major categories: land use, spatial, hydrological, morphological, and temporal. Due to their slow-changing nature and data limitations, most variables were treated as temporally static, with 2023 values used as proxies across all years (Table S1). Flow velocity and water temperature were modeled from static landscape features (Khaliq et al., 2024) and thus also treated as static. Precipitation, although dynamic, was not retained in final models due to lack of significant effect.

Candidate models were initially developed to include at least one representative variable from each environmental parameter, with year included as a random effect where feasible.

Implementation of environmental parameters and final model selection was guided by ecological plausibility, correlation and principal component analysis, and variance inflation factors, resulting in a parsimonious model structure that included the most relevant chemical metric, proportion of agricultural land use, and flow velocity, with slight variations depending on the sampling period.

#### 3.5.4. Interpretation and Reporting

Results are presented as correlation coefficients, regression coefficients, and 95% confidence intervals. Vertical forest plots visualize effect sizes across time windows, regression plots illustrate key relationships and partial effects between variables, cross-validation values assess model performance, and PCA biplots display the multivariate data structure. All analyses were conducted in R (v4.4.2) using multiple packages; the code is available in the Supplementary Information (Link 1). Only models with significant or marginally significant results ($p \leq 0.10$) and/or superior predictive performance (as indicated by cross-validated RMSE and $R^2$) were retained for primary interpretation across time windows and modelling approaches, taking ecological plausibility into account.

# 4. RESULTS

## 4.1. Overview of Data and Sampling

A total of 41 streams were included in the study, with data collected over six years (2018–2023). Macroinvertebrate data were available from 195 sampling events, unevenly distributed across streams and years. For chemical analysis, composite water samples were collected using either 14-days or 3.5-days sampling periods, depending on the site and year. The 14-days composite samples were collected systematically throughout the year at all sites, as part of the standard monitoring protocol. In contrast, the 3.5-days composite samples were targeted primarily at sites known or suspected to be impacted by pesticide use, allowing for higher temporal resolution, and targeted analysis of specific compound groups, such as pyrethroids.

The numbers in Table 1 reflect the total number of observations across all sites and years for each time window, not per site. Temporal alignment of chemical and biological data revealed that, for the 3.5-days sampling period, 22 streams had micropollutant data available within 1-week to 1-year prior to macroinvertebrate monitoring, while for the 14-days sampling, all 41 streams were represented in the 1-year window.

*Table 1. Number of observations per time window and sampling period. The numbers represent the total observations across all 41 streams and all years for each time window.*

| **3.5-days sampling period** | **1-week** | **2-weeks** | **1-month** | **2-months** | **3-months** | **6-months** | **1-year** |
|---|---|---|---|---|---|---|---|
| Macroinvertebrates monitoring | 34 | 39 | 39 | 42 | 42 | 45 | 92 |
| Streams | 13 | 13 | 13 | 15 | 15 | 16 | **22** |
| Chemical samples | 69 | 145 | 293 | 567 | 848 | 1115 | 1854 |
| Considered substances | 73 | 73 | 73 | 74 | 74 | 74 | 74 |
| Range of years | | | | 2019 – 2023 | | | |
| **14-days sampling period** | **1-week** | **2-weeks** | **1-month** | **2-months** | **3-months** | **6-months** | **1-year** |
| Macroinvertebrates monitoring | 111 | 177 | 183 | 185 | 185 | 194 | 195 |
| Streams | 38 | 40 | 40 | 40 | 40 | 41 | **41** |
| Chemical samples | 123 | 226 | 438 | 831 | 1171 | 2054 | 3281 |
| Considered substances | 109 | 110 | 110 | 110 | 111 | 111 | 111 |
| Range of years | | | | 2018 – 2023 | | | |

## 4.2. Exploratory Analysis

### 4.2.1. Correlation: Bioindicators and Chemical Metrics

As shown in Figure S1, Spearman correlation matrices revealed negative associations between chemical metrics and bioindicators, especially for SPEARpesticides and GI. In the 3.5-days sampling period dataset, SPEARpesticides correlations ranged from −0.22 to −0.61, and GI correlations ranged from −0.21 to −0.45. In the 14-days sampling period dataset, SPEARpesticides correlations ranged from −0.16 to −0.41, GI correlations ranged from −0.18 to −0.44, and IBCH correlations ranged from −0.14 to −0.17. EPT showed negative correlations ranging from −0.33 to −0.38 and from −0.15 to −0.17 in the 3.5-days and 14-days sampling datasets, but only for one time window. The VT index did not show significant or consistent correlations over time windows.

### 4.2.2. Principal Component Analysis of Abiotic Factors

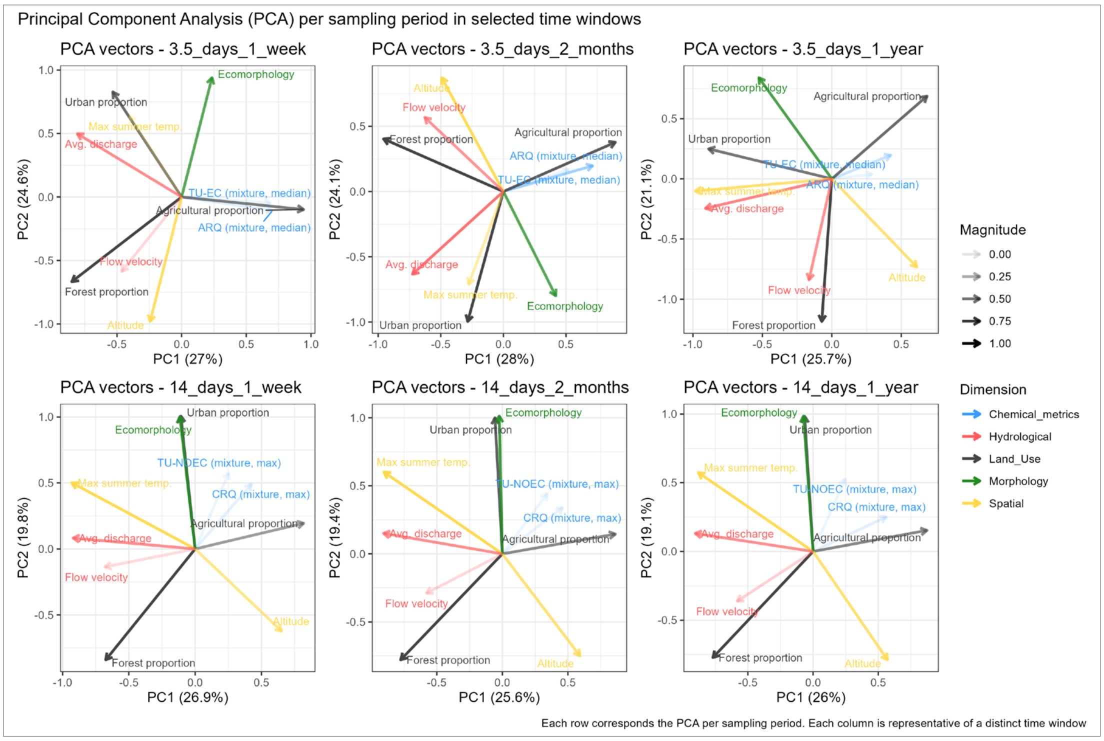


*Figure 2. PCA-Biplot for both sampling periods across 1-week, 2-months and 1-year time windows (Each abiotic factor belongs to one dimension, while magnitude reflects each variable's loading in the PCA).*

Principal component analysis (PCA) was used to reduce the dimensionality of the environmental dataset and to identify the main gradients structuring abiotic conditions across sampling periods and time windows. This approach clarifies which factors most strongly differentiate stream environments as more streams and chemical samples are included across time windows, facilitating the identification of the most pertinent environmental variables for subsequent modeling.

PCA was conducted for each time window and for both sampling periods (Figure 2). Since macroinvertebrate monitorings are related to specific streams, increasing the time window results in the inclusion of more streams and chemical samples. Except for chemical metrics, all abiotic factors were static across years, reflecting their spatial nature and data constraints.

For static variables, values of 2023 were used as proxies, so differences across time windows primarily reflect the addition of chemical data and sites. We focus on three main time windows for description, 1-week, 2-months, and 1-year, selected based on the frequency of monitoring (Table 1). The environmental variables were categorized into five dimensions (1) Land Use in the catchment: proportions of agricultural, forest, and urban land use upstream of the sampling point, (2) Spatial: modelled maximum summer water temperature (°C), altitude (m.a.s.l.), (3) Hydrological: modelled flow velocity ($m.s^{-1}$), annual average modeled discharge ($m^3.s^{-1}$), (4) Morphological: ecomorphology (classification based on water body morphology, ranging from 0 (natural) to 12 (unnatural)), and (5) Temporal: year of sampling.

In both the 3.5- and 14-days sampling periods, the first principal component (PC1) consistently reflected an "agricultural pressure" gradient, defined by high loadings for agricultural land proportion and chemical metrics, opposed by forest land proportion and flow velocity. This second association between forest and flow velocity likely represents a gradient of naturalness and hydrological regime. Conversely, urban land proportion and ecomorphology formed a distinct anthropogenic axis, particularly in the 14-days sampling period. This highlights the influence of urbanization on stream morphology and habitat quality.

The magnitude of each variable's loading in the PCA reflects its relative contribution to the principal components, with higher absolute values indicating a stronger influence on the environmental gradient (Figure 2). As the time window increased, the association between chemical metrics and proportion of agricultural land use weakened, but the influence of land use and hydrology remained consistent. Overall, these results demonstrate that agricultural activities, hydrological and forest characteristics, and urbanization are the dominant forces structuring the abiotic environment in these streams, across the evaluation of the sampling periods and time windows.

## 4.3. Bioindicator Response to Chemical Metrics

### 4.3.1. Linear Models (LM/LMM)

## *3.5-days Sampling Period*

In the 3.5-days sampling period, the relationship between pesticide pressure and macroinvertebrate bioindicators was assessed using both LMMs and LMs. LMMs initially included "year" as a random effect; however, many models exhibited singular fits (i.e., the model could not reliably estimate variance for the random effect, likely due to the limited sample size or low variation in "year"; see Figure S2-A). Consequently, LMs were used as the primary approach, providing robust estimates and confidence intervals (Figure 3).

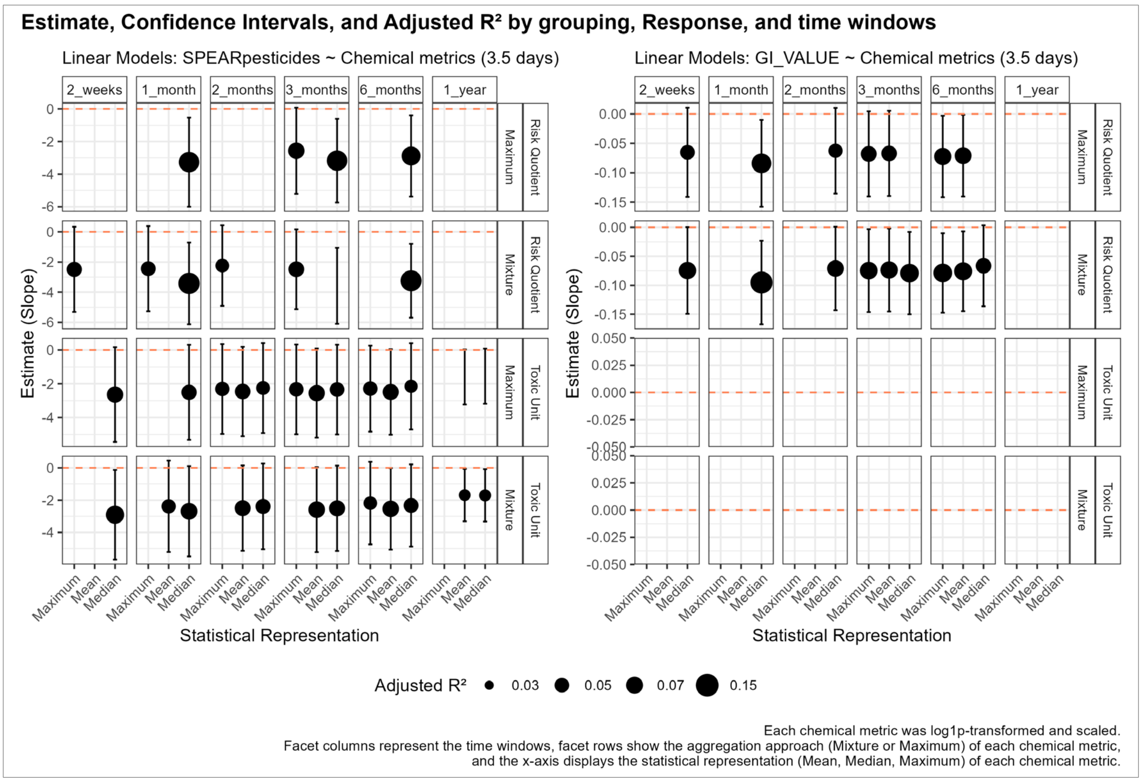


*Figure 3. Model Estimates (slopes), Confidence Intervals, and Adjusted R² for the Linear Models in the 3.5-days sampling period. Only the slopes of the models that are at least marginally significant (p < 0.1) are shown.*

All models that reached at least marginal significance ($p \leq 0.1$), including those fully significant ($p \leq 0.05$), demonstrated negative estimates. Thus, negative associations for both SPEARpesticides and GI, indicating that increased pesticide exposure corresponded to lower values of these bioindicators. Significant negative effects (confidence intervals not crossing zero in Figure 3) were detected for specific chemical metrics. Particularly for the mixture-median representation of the TU-EC and ARQ metrics at several time windows. No significant estimates were identified for IBCH, VT or EPT over the time windows in this sampling period (Figure S2-B).

The explanatory power of these models, as indicated by adjusted $R^2$, ranged from 0.03 to 0.15 for SPEARpesticides and from 0.05 to 0.09 for GI (Figure 3). Generally, ARQ-based metrics provided higher explanatory power than TU-based metrics for both bioindicators.
Nevertheless, although the LMs provided a reasonable fit overall, residuals versus fitted value plots showed that the linearity assumption was only partially met for most models. Moreover, outliers disproportionately influenced the statistical representations of chemical metrics, while the 'median' was occasionally more robust, though it did not consistently mitigate the influence of outliers across all metrics (Figure S3).
Given these limitations, including non-linearity, and modest explanatory power, we proceeded to explore non-linear relationships using generalized additive models (GAMs).

### *14-days Sampling Period*

For the 14-days sampling period, LMMs were fitted to assess the relationship between pesticide pressure and macroinvertebrate bioindicators (Figure 4). The relatively large number of observations for these models (Table 1) prevented singular fits and contributed to model stability.
Across all chemical metrics, the direction of effect (slope) was consistently negative, indicating that higher pesticide concentrations were associated with lower bioindicator values. Effect estimates and confidence intervals for each bioindicator and chemical metric are illustrated in Figure 4. Conditional $R^2$ values indicated varied explanatory power: 0.59 - 0.64 for EPT, 0.11 - 0.18 for SPEARpesticides, 0.04 - 0.10 for GI and IBCH. Marginal $R^2$ values, reflecting the variance explained by chemical metrics alone, were highest for SPEARpesticides and GI, especially in shorter time windows. The random effects increased explanatory power for SPEARpesticides by approximately 2- to 5- fold, and for EPT by approximately 22- to 60-fold.

Across all statistical representations (mean, median, maximum), effect sizes and significance levels were similar. The “mixture” approach was found to be sufficient, and CRQs were the most prominent chemical metrics. Despite the consistent negative relationships, the overall explanatory power of these models was limited. It is important to note that model assumptions were not fully met for GI or EPT, and only marginally for IBCH, exhibiting non-normal residuals and heteroscedasticity (Figure S4). Therefore, results for these indices should be interpreted with caution, and primary emphasis should be placed on the more robust findings for SPEARpesticides.

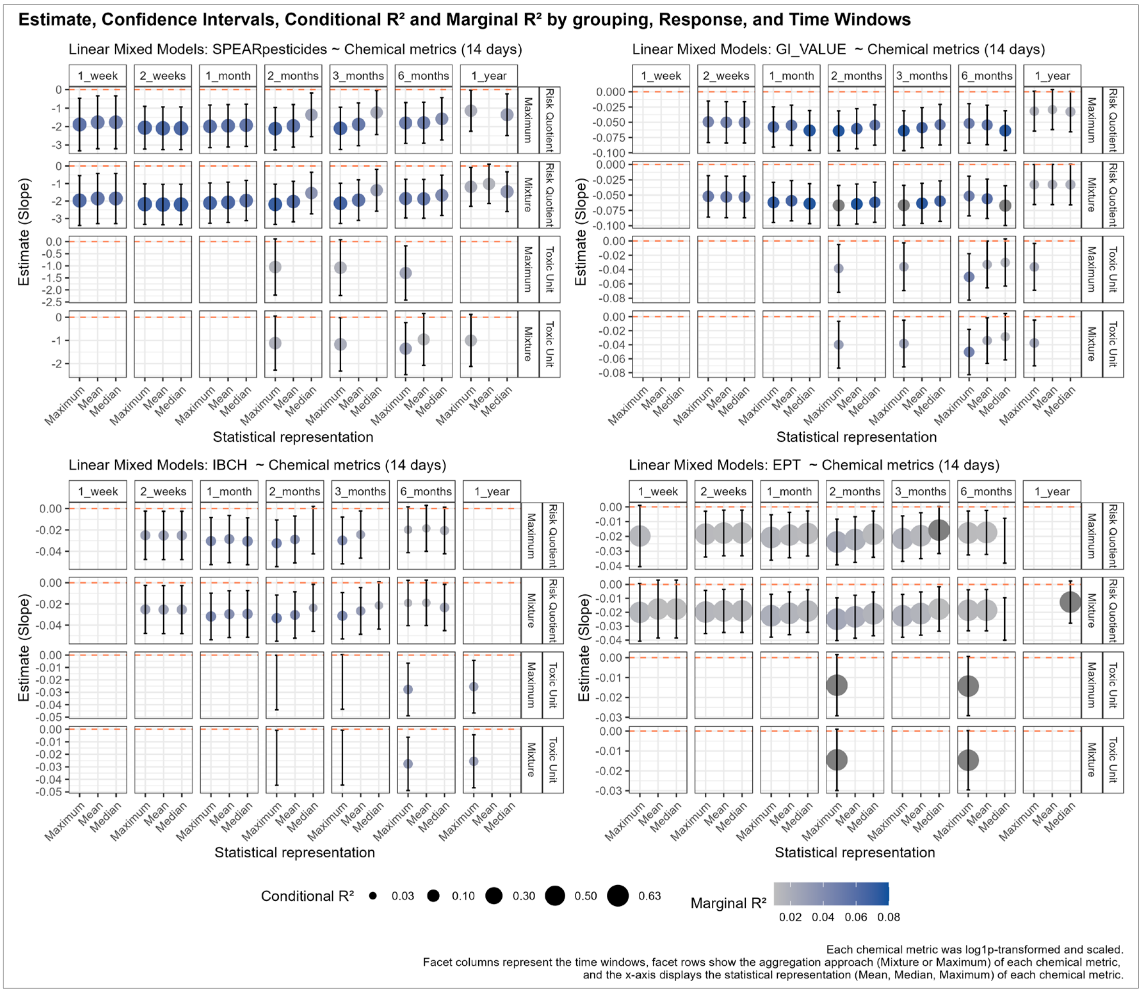


*Figure 4. Model Estimates, Confidence Intervals, Conditional and Marginal $R^2$ for the Linear Mixed Models in the 14-days sampling. Only the slopes of the models that are at least marginally significant ($p < 0.1$) are shown.*

### 4.3.2. Non-linear Modeling of Pesticide-Bioindicator Relationships

*3.5-days sampling period: Generalized Additive Models (GAMs)*

We fitted GAMs to capture potential nonlinear relationships between pesticides pressure and bioindicator response. GAMs yielded a wider range of significant models for ARQ and TU-EC metrics across all time windows for SPEARpesticides and GI (Figures S5 and S6). However, in the 1-week window, although some models showed reasonable fit ($R^2 = 0.14 - 0.20$), the estimated values for SPEARpesticides or GI were inconsistent and often ecologically implausible based on the response of the bioindicator. This shows the importance of not relying solely on fit metrics, particularly at the extremes of the predictor range, and underscores the need for careful graphical interpretation.

In the 2-weeks window, graphical inspection revealed that GI's fit was largely driven by the smoother's adaptation to a few extreme values at high pesticide pressure, with a clearer negative trend only apparent at lower exposure levels (Figure S5). For SPEARpesticides, the negative trend was clearer and more consistent, particularly for ARQ(mix, median) and TU-EC(mix, median), even in the presence of outliers (Figure 5).

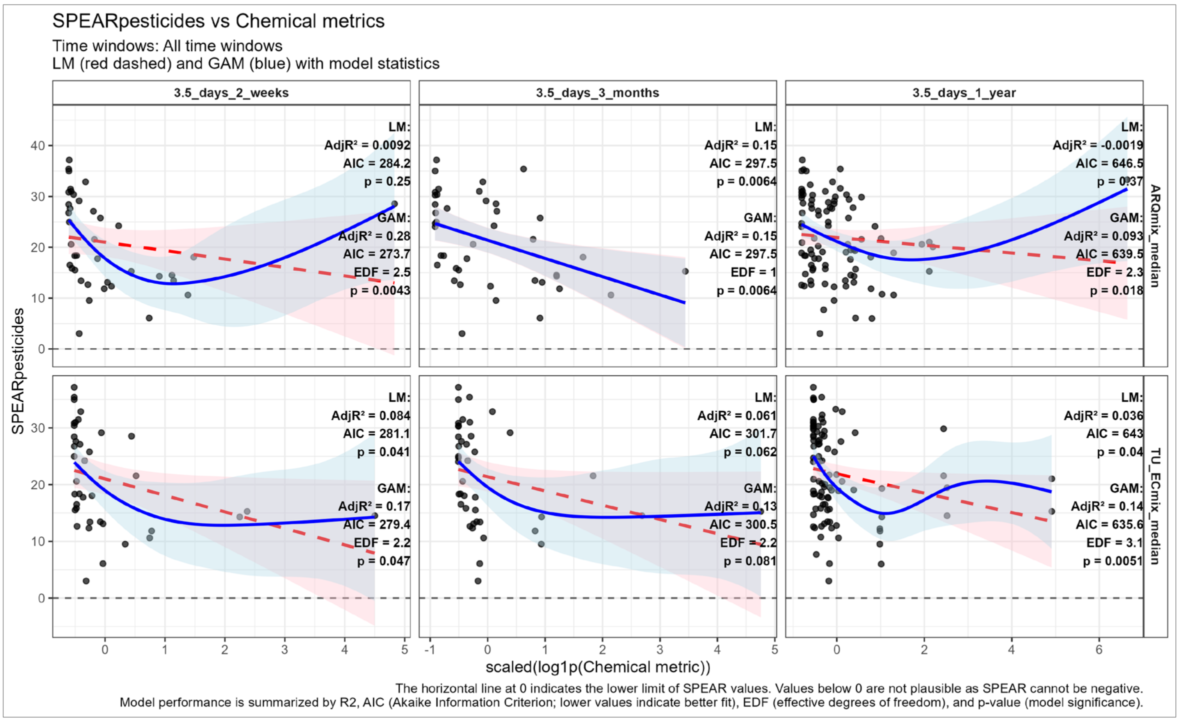


*Figure 5. Comparison of Linear (red line) and Generalized Additive Models (blue line) for SPEARpesticides vs. TU(mix, median) and ARQ(mix, median) in 2-weeks, 3-months, and 1-year time windows.*

From the 1-month to the 6-months windows, similar patterns persisted (Figures S5 and S6). For example, in the 3-months window, both "max" and "median" statistical representations were significant or marginally significant for SPEARpesticides. ARQ-based metrics yielded a marginally significant and negative relationship, with values matching those from their LM counterparts ($R^2$ = 0.15, $p$ = 0.06) (Figure 5). In these cases, the effective degrees of freedom (EDF) were equal to one, indicating that the GAMs recognized and retained the linearity of the relationship when appropriate. TU-EC(mix, median) showed improved explanatory power from LM ($R^2$ = 0.06) to GAM ($R^2$ = 0.13) in the same window (Figure 5). In contrast, GI models became less plausible, highlighting their limited sensitivity to chemical metrics (Figure S5).

In the 1-year window, no significant GAM was obtained for GI (Figure S5). For SPEARpesticides, TU-EC(mix; median) demonstrate an increased explanatory power from LM ($R^2$ = 0.04) to GAM ($R^2$ = 0.14). Other significant chemical metrics, such as ARQ(mix; median), had $R^2$ below 0.10. Nevertheless, a negative trend between SPEARpesticides and chemical metrics remained evident (Figure 5).

Overall, the explanatory power for SPEARpesticides was improved by GAMs ($R^2$ = 0.03–0.15 in LMs and 0.10–0.28 in GAMs). However, chemical metrics still explain only a moderate proportion of the variability in the bioindicator for most GAMs, the presence of an isolated observation above 4 units (in log1p and scaled units) caused the curve to rise and uncertainty to increase at higher pesticide concentrations. TU-EC metrics, however, appeared less affected by this outlier, with the negative relationship most apparent where observations were densely aggregated and prediction intervals were narrower. Across all time windows, SPEARpesticides was more consistently and robustly explained by TU-EC and ARQ metrics in their (mix, median) representations compared to GI (Figures S5 and S6). Even when models were statistically significant and exhibited high explanatory power, graphical evaluation was essential to check direction and ecological plausibility of the relationships.

### *14-days sampling period: Generalized Additive Mixed Models (GAMMs)*

Following the LMM analysis of the 14-days sampling period, which revealed that EPT was primarily explained by the random effect (year) and that SPEARpesticides was best explained by the fixed effects (chemical metrics), we further explored potential nonlinear relationships using GAMMs, motivated by the observed deviations of the model assumptions, as well as the pronounced influence of the random effect.

GAMMs were fitted with year as a random effect. For SPEARpesticides, the difference between marginal and conditional $R^2$ remained similar to the LMMs, and the overall explanatory power did not increase. Model selection criteria (AIC) was higher for the GAMMs than for the LMMs, with differences of 4-6 units. This indicates no meaningful advantage in model complexity (Figure S7).

For IBCH, GI, and EPT, GAMMs failed to resolve the diagnostic issues observed in LMMs. QQ plots continued to display step-like patterns, and residual plots showed persistent arcs or systematic deviations, particularly for GI and IBCH (Figures S4). EPT exhibited a slight improvement in residual distribution, but not in AIC or model interpretability. Fitted smooth terms for GI, IBCH and EPT were flat or lacked a discernible trend across chemical metrics. GAMMs did not improve model fit, explanatory power or performance for any bioindicators compared to LMMs. Marginal and conditional $R^2$ values and fitted relationship shapes remained largely unchanged. For SPEARpesticides, results were consistent enough with LMMs; for GI, IBCH and EPT, models were limited by assumption violations and lack of clear association with chemical metrics.

### 4.3.3. Synthesis and further modeling

Given the contrasting performance of statistical models across sampling periods, we addressed our analytical approach to best capture the relationships between pesticide exposure and bioindicator response. For the 3.5-days composite samples, GAMs were employed due to evidence of non-linearity, and outlier influence. The improved explanatory power and interpretability observed with this flexible modeling framework also led us to select it over LMs. In contrast, for the 14-days sampling period, LMMs proved most appropriate, as the relationship between chemical metrics and bioindicators, particularly SPEARpesticides, was adequately described by linear associations, and the inclusion of random effects accounted for substantial variance.

These differences in model fit likely reflect both the temporal resolution of the chemical data and the ecological processes captured at each sampling scale even when being composite samples.

## 4.4. Integration of Environmental Parameters

### 4.4.1. Selection of Environmental Parameters for Integrated Regression Modeling

The integration of environmental parameters into our regression models was motivated by two main considerations. First, initial analyses using only chemical metrics as predictors for SPEARpesticides revealed limited explanatory power, suggesting that other sources of variability were influencing macroinvertebrate responses. Second, our PCA results highlighted key environmental gradients, especially those related to proportion of agricultural land use (associated with pesticide pressure), hydrology, and urbanization. This underscored the need to move beyond a purely chemical perspective and to account for the broader environmental context in which pesticide pressure and biological responses occur.

The potential relationships between each parameter, and its influence on pesticide pressure and macroinvertebrate response, were conceptually evaluated. For instance, water temperature is modeled by altitude and catchment area; flow velocity is determined by slope, discharge, and width (Khaliq et al., 2024); and discharge is a function of catchment area (BAFU, 2020). These inherent relationships required caution to avoid collinearity in our model construction. Notably, a conceptual link exists between proportion of agricultural land use and chemical metrics, as greater agricultural land use increases the probability of pesticide application and runoff, a relationship acknowledged in the definition of SPEARpesticides (Liess et al., 2008, 2021). Similarly, hydrological variables such as discharge and flow velocity are interrelated, and their influence on pesticide concentrations is mediated by both dilution and dispersal processes, as well as by the timing and nature of runoff events (Doppler et al., 2012).

To empirically assess these relationships, we conducted correlation analyses between SPEARpesticides and abiotic factors for each time window on the 3.5-days sampling period. Consistent negative correlations were observed between SPEARpesticides and both proportion of agricultural land use and chemical metrics, while moderate positive correlations were found with forest proportion, altitude, flow velocity, and discharge (Figure S1-A). The same correlation structure was observed for the 14-days sampling period, with the addition of urban proportion and ecomorphology showing negative associations with SPEARpesticides, consistent with PCA results (Figure S1-B).

Correlation analyses among abiotic factors revealed strong associations within each dimension (e.g., among land use or hydrological variables) and modest positive correlations between proportion of agricultural land use and chemical metrics, as well as negative associations between agricultural and forest proportions, and between urban and forest land uses (Figure S8). These findings suggest that the positive relationship between SPEARpesticides and forest proportion may reflect the spatial opposition to proportion of agricultural land use, rather than a direct causal effect.

Based on these insights, we selected proportion of agricultural land use, chemical metric, and flow velocity as variables for regression modelling in the 3.5-days period, as they loaded strongly on the main axes in the PCA and defined clear environmental gradients. For the 14-days period, we chose proportion of agricultural and urban land use, chemical metric, and flow velocity, as these variables were consistently important across time windows and could capture chemical and spatial drivers of macroinvertebrate community structure.

Finally, this approach aligns with our second specific objective: to evaluate the influence of environmental parameters on the relationship between chemical metrics and biological response. At this stage, SPEARpesticides stood as the only bioindicator that exhibited consistent responses to chemical metrics in prior analyses. Consequently, it was the exclusive bioindicator utilized in the subsequent analyses.

### 4.4.2. Integration of Environmental Parameters in the Regression Models

*3.5-days Sampling Period*

We assessed the impact of pesticides and abiotic factors on macroinvertebrates, using GAMs and SPEARpesticides as the bioindicator. The predictors included TU-EC(mix, median) and ARQ(mix, median), alongside proportion of agricultural land use and flow velocity.

*Model Selection and Diagnostics*

Initial models incorporated TU-EC(mix, median), proportion of agricultural land use, and flow velocity. In the Table S2, which presents the results of the GAMs across all the time windows, the models demonstrated robust explanatory power, with adjusted $R^2$ values above 40% and reaching up to 70%. Model selection criteria, as indicated by AIC values, remained relatively stable across time windows (ranging from 222 to 278), but increased substantially in the 1-year window (AIC = 600), suggesting a decrease in model parsimony for longer integration periods.

Statistical significance was consistently observed for proportion of agricultural land use and flow velocity ($p < 0.05$), while TU-EC(mix, median) was significant or marginally significant in most windows (except at 1 week, where $p = 0.18$, still indicating a potential effect). The effective degrees of freedom (EDF) for TU-EC ranged between 1 and 2.5, indicating a predominantly linear or gently curved relationship. For proportion of agricultural land use, EDF values ranged from 1 to 4.5, with higher values (and more complex, “wiggly” fits) observed when TU-EC’s EDF was 1, and vice versa. Flow velocity consistently exhibited EDF values around 4. F-statistics across all models ranked the predictors in the following order of explanatory strength: proportion of agricultural land use > flow velocity > TU-EC(mix, median).

*Visual and Ecological Interpretation*

Visual inspection of the fitted GAMs (Figure 6) revealed that proportion of agricultural land use exerted a linear and consistently negative effect on SPEARpesticides, underscoring their role as a dominant static environmental parameter. TU-EC(mix, median) also showed a negative association, with a smoother fit and reduced uncertainty intervals compared to models including only the chemical metric (Figure S9). Flow velocity, while a significant predictor, did not display a strongly positive trend; its inclusion nonetheless improved

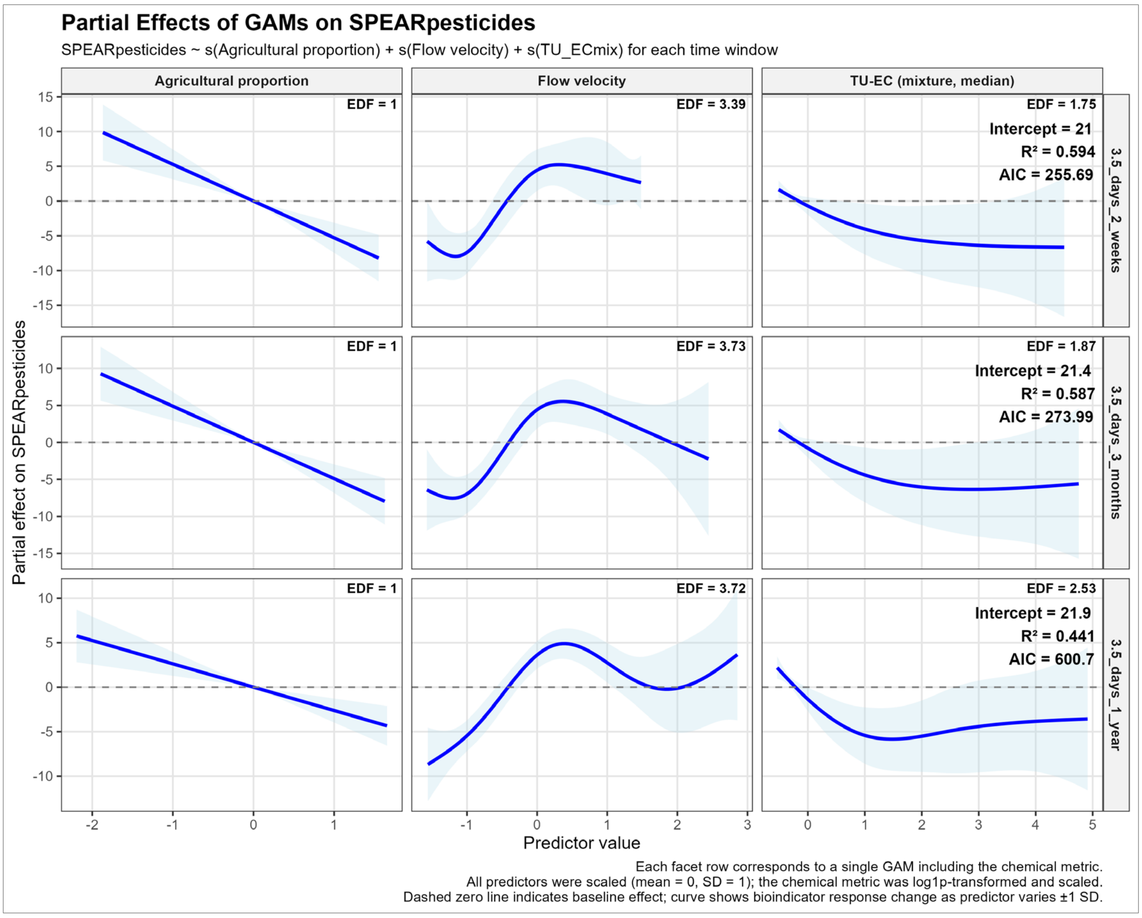


*Figure 6. Partial effects of TU-EC(mix, median), proportion of agricultural land use, and flow velocity on SPEARpesticides across selected scenarios (see row labels) using multivariate GAMs. The values ($R^2$, AIC, and intercept) for each model (facet row) are displayed in the TU box because they are identical for the other two predictors, as they belong to the same model.*

model fit and reduced uncertainty, reflecting its relevance as a hydrological driver. This suggests that the inclusion of both proportion of agricultural land use and flow velocity helps absorb unexplained variance and provides a more stable estimate of pesticide pressure effects.

Alternative models, including only flow velocity or proportion of agricultural land use, yielded lower explanatory power, and produced unstable fits. The closest alternative in terms of AIC and $R^2$ was the model with proportion of agricultural land use and flow velocity, but without the chemical metric. However, the EDF for proportion of agricultural exceeded 4.3, suggesting overfitting (Table S2, Figure S10).

Additional predictors (e.g., urban proportion and ecomorphology) were tested but did not meaningfully improve the model or provide interpretable relationships with SPEARpesticides. Similarly, substituting ARQ(mix, median) for TU-EC(mix, median) resulted in non-significant effects and only a weak negative trend, suggesting that TU-EC is a more sensitive and ecologically relevant indicator of pesticide pressure in this context (Table S2). Other bioindicators did not yield plausible or interpretable results and were therefore not considered further.

### *14-days Sampling Period*

We constructed a series of LMMs using proportion of agricultural and urban land use, and flow velocity, alongside the chemical metric CRQ(mix, max). The choice of this chemical metric was supported by our observation that a more specific chemical signal was necessary to capture the impact of pesticide pressure in composite samples where peak pesticide events are likely smoothed. This concept is further explained below.

### *Model Selection and Diagnostics*

Nine candidate models were evaluated, each representing different combinations of the main abiotic gradients and the chemical metric (Figure S11). Models containing only one abiotic variable plus the chemical metric consistently exhibited low explanatory power, with adjusted $R^2$ values below 20%. The best-performing model included proportion of agricultural and urban land use, flow velocity, and CRQ(mix, max) across all time windows (Model 2 in Figure S11). This model demonstrated superior parsimony (lowest AIC by 5–6 units compared to the next best model) and higher explanatory power (marginal $R^2$ always 2–3% above the second-best model).

The explanatory power of the selected model was consistent across time windows, with marginal and conditional $R^2$ values nearly identical, ranging from 0.29 and 0.30 in the 1-year window to 0.40 and 0.42 in the 1-week window, indicating that the inclusion of environmental parameters effectively reduced the influence of the random effect (year). This contrasts with models considering only chemical metrics, where the random effect accounted for a much larger proportion of the explained variance.

The statistical and absolute importance of the predictors, based on their estimates (slopes), followed a consistent order: agricultural proportion > flow velocity > urban proportion > chemical metric. CRQ(mix, max) always exhibited a negative estimate, with effect sizes ranging from -0.3 (1-year window) to -1.48 (3-months window). The effect of the chemical metric increased from the 1-week to the 2-month window (Figure S11), then gradually decreased as the time window lengthened, mirroring the pattern observed in the PCA, where the chemical metric's relevance diminished as more samples and spatial variability were included.

Removing the chemical metric resulted in a 2–3% reduction in $R^2$, with minimal change in AIC (Figure S13), while attempts to add random slopes or interactions did not improve model fit due to limited observations per year.

All predictor estimates were ecologically plausible based on the effect over macroinvertebrate communities: negative for the chemical metric, proportion of agricultural and urban land use, while positive for flow velocity. This aligns with the expectation that increased pesticide pressure and anthropogenic land use reduce macroinvertebrate community integrity, while greater flow velocity may reflect more natural or less impacted stream conditions as shown in the PCA analysis.

## *Visual and Ecological Interpretation*

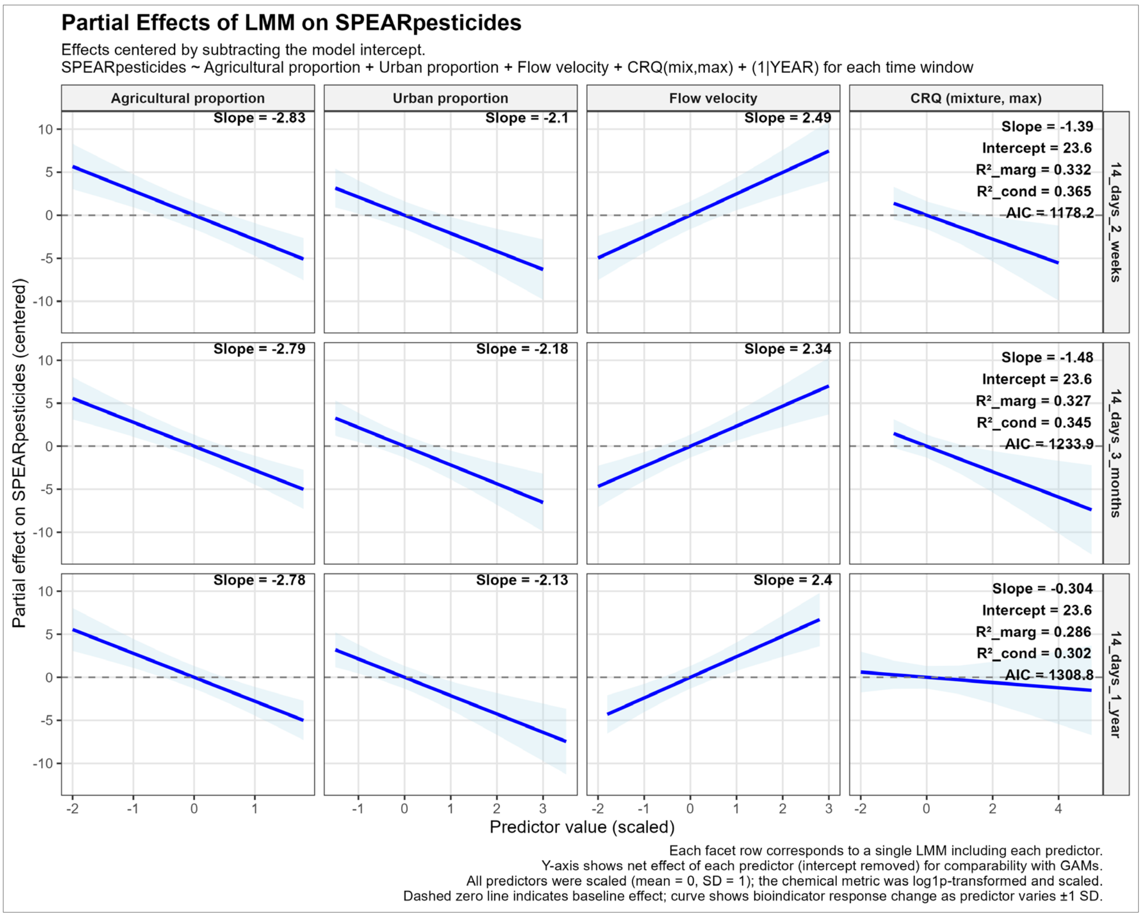


*Figure 7. Partial effects of CRQ(mix, max), proportion of agricultural and urban land use, and flow velocity on SPEARpesticides across selected scenarios (see row labels) using multivariate LMMs. The values ($R^2$, AIC, and intercept) for each model (facet row) are displayed in the CRQ box because they are identical for the other two predictors, as they belong to the same model.*

Visual inspection revealed (Figure 7) that the inclusion of environmental parameters led to a reduction in the uncertainty around the estimated effect of the chemical metric CRQ(mix, max) in all the time windows (Figure S12). In contrast, the uncertainty associated with the environmental parameters themselves did not show a comparable reduction, and their effect estimates remained relatively stable and robust across time windows (Figure S13). This suggests that, while environmental parameters are essential for capturing the broader environmental context and improving model fit, their predictive precision does not benefit as markedly from the integrated model as does the chemical metric in this context.

## 4.5. Cross-Validation and Predictive Performance

To further assess the robustness and generalizability of our selected models, we conducted 5-fold CV for both the 3.5-days and 14-days sampling periods. In addition to our best models (GAM, LMM), we implemented Random Forest (RF) as a non-linear, flexible benchmark to test whether more complex relationships or interactions might exist between SPEARpesticides and the predictor variables. It is important to note that the primary aim of this thesis is to infer relationships between pesticide pressure, environmental parameters, and bioindicators, rather than to maximize predictive accuracy. Nevertheless, CV offers a transparent assessment of model stability and helps identify potential limitations in model structure.

### 4.5.1. 3.5-days Sampling Period

For the 3.5-days period, the CV was performed as follows: SPEARpesticides ~ TU-ECmix,median + proportion of agricultural land use + flow velocity, using GAM, LM, and RF approaches (Figure 8). In the 1-week window, the GAM produced a negative $R^2$, likely due to the small sample size ($n \approx 30$), but from 2 weeks onward, both GAM and RF yielded plausible results. Across most time windows, RF and GAM outperformed LM in both RMSE and $R^2$, particularly in larger datasets (e.g., 3–6 months, 1 year). This suggests that flexible, non-linear approaches can better capture the relationships in the data when sufficient observations are available.

When the chemical metric was removed, the stability and explanatory power of the GAM decreased, especially in intermediate time windows. Additionally, when using only chemical metrics, both GAM and RF produced implausible results (negative $R^2$ across all time windows), highlighting the relevance of integrating both chemical and environmental parameters.

### 4.5.2. 14-days Sampling Period

For the 14-days period, the model: SPEARpesticides ~ CRQmix,max + proportion of agricultural land use + urban proportion + flow velocity, was evaluated using LMM (with year as a random effect) and RF (with year as a predictor) (Figure 8). However, the LMM was singular in 6 out of 7 time-windows, likely due to the reduced sample size in each CV fold and the complexity of the model. Simplifying the model by using LM (with year as a fixed effect) or including interaction terms (e.g., Proportion of agricultural land use × Chemical Metric) did not improve performance. In all cases, RF outperformed the parametric models, with $R^2$ values around 0.6 across time windows, compared to 0.3–0.4 for LMM/LM (Figure 8). This suggests the presence of complex, non-linear, or interactive effects that cannot be robustly identified with the current dataset and parametric approaches.

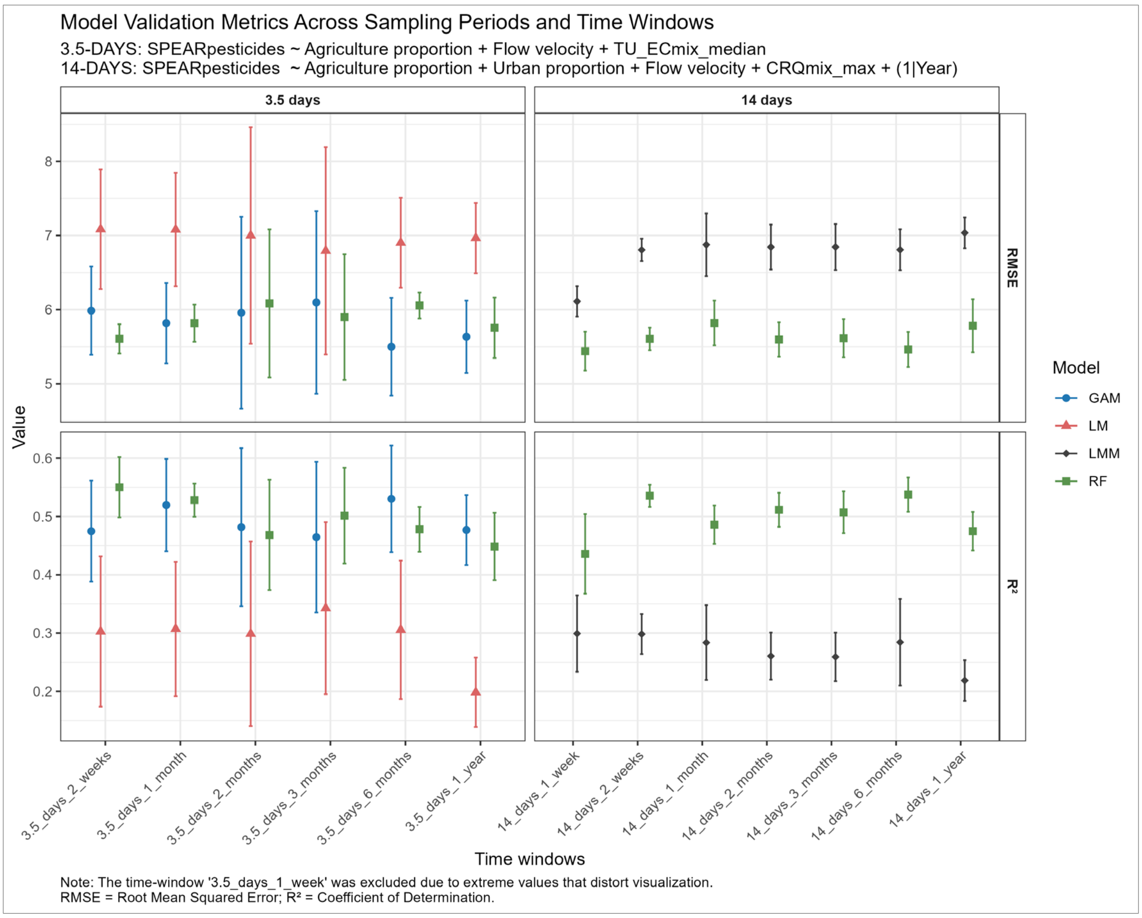


*Figure 8. Cross-Validation Performance of the Linear Model, Linear Mixed Model, Generalized Additive Model and Random Forest in the 3.5- and 14-days Sampling Period.*

# 5. DISCUSSION

## 5.1. Main Findings in Relation to Objectives and Hypotheses

This study set out to evaluate the applicability of SPEARpesticides, alongside other bioindicators, for reflecting pesticide impacts on macroinvertebrate communities in Swiss streams, within the context of the NAWA framework. The results support the primary hypothesis: *elevated pesticide pressure, as quantified by risk quotients (RQs) and toxic units (TUs), is associated with a measurable decline in sensitive macroinvertebrate taxa*. This effect, most clearly captured by the SPEARpesticides index, demonstrates a negative relationship between pesticide pressure and the bioindicator variability. These relationships yield a range of $R^2$ from 0.1 to 0.3. The strength depends on the chemical metric, sampling period, time window, and statistical modeling approach selected. These results align with those of previous research that used SPEARpesticides to reflect pesticide impacts across Europe (Beketov et al., 2013; Kuzmanović et al., 2016; Liess et al., 2021; Rasmussen et al., 2012), South America (Hunt et al., 2017) and East Africa (Ganatra et al., 2021).

RQs, being based on Environmental Quality Criteria designed to protect a broad range of organisms (Peterson, 2006; von der Ohe et al., 2008) may be preferable when regulatory thresholds and linear models are the focus. In contrast, TUs are calculated using effect concentrations for the aquatic species under analysis, often invertebrates such as *D. magna*, making them more ecologically targeted, providing insights under more flexible, data-driven modeling (Ginebreda et al., 2014; von der Ohe & de Zwart, 2013). This specificity allowed TU-EC to capture more nuanced, non-linear relationships in flexible models such as GAMs in our 3.5-days sampling period. However, we cannot neglect that in many cases the non-linear behavior captured by the GAMs was due to an isolated higher pesticide pressure value, while in the area where the most of our observations are aggregated, a clear negative relationship, characterized by a nearly linear trend, is evident.

A significant aspect of our findings is the considerable uncertainty observed in the regression analyses, especially at the extremes of the exposure gradient. Isolated, rare high-exposure events increased variability in model predictions and reduced confidence in these regions. This pattern is likely the result of factors such as data aggregation or the use of composite samples. The necessity for broader data coverage across the exposure gradient, sensitivity analyses (e.g., excluding rare events), and cautious interpretation is underscored by this pattern. These uncertainty levels may also be caused by unseen factors in our approach. Therefore, continued methodological refinement is crucial to better quantifying and communicating confidence in ecological risk assessments (Moe et al., 2022).

With respect to the outcomes based on sampling periods, we obtained more information from the 3.5-days samples, reflecting our acute approach, with different strengths of the negative relationship that spanned an $R^2$ from 0.1 to 0.3. Conversely, within the 14-day sampling period, consistent with our chronic approach, the slopes and $R^2$ ranged from 0.1 to 0.2, indicative of a more stable negative relationship. This is presumably due to the prolonged composite samples, which tend to smooth out exposure events, resulting in a conservative value of the chemical metric (Peterson, 2006; Spycher et al., 2018). Ultimately, this proves insufficient for the SPEARpesticides index, underscoring a limitation in the bioindicator sensitivity. Nevertheless, we cannot neglect the distinction between the 3.5-day and 14-day sampling periods. The primary objective of these samples differs, as the 3.5-day samples are collected in locations where the exposure to pesticides is presumed to occur, while the 14-day samples are part of a more comprehensive framework. This distinction could be the reason why the chemical metrics derived from the 3.5-day samples are more prone to predict the SPEARpesticides results.

In accordance with our second specific objective, we hypothesized that the smoother value of the 14-days sample would be comparable to the other environmental parameters and that including them would improve the model's power. This was confirmed by the inclusion of land use and hydrological variables, which were observed to be relevant predictors. This observation aligns with the findings of Rico et al. (2016), who reported a similar relationship between a bioindicator and a chemical metric in a multistressor scenario. Indeed, the PCA analysis showed that the "agricultural pressure gradient" was an indicator of the effectiveness of the chemical metrics derived from the chronic approach. The link between chemical metrics and the proportion of agricultural land use became clearer during the 3.5-day sampling period than during the 14-day sampling period.

Finally, both the regression analysis and the PCA exhibited a consensus, albeit at varying levels, in reflecting the temporal variations across designated time-windows. The higher explanatory power of the relationship and the robust association of the agricultural pressure gradient were consistently observed to be dependent on the aggregation of samples that were more proximate to the macroinvertebrate monitoring. This finding served to conclude the third objective of our research approach.

## 5.2. Interpretation of Bioindicator Responses

A key characteristic of this study is its approach to calculating chemical metrics. Only substances classified as insecticides, herbicides, and fungicides were considered, and for risk quotients, quality criteria were filtered to include only those labeled as acutely or chronically toxic for invertebrates. For toxic units, the ecotoxicological endpoints were selected to be as close as possible to macroinvertebrate sensitivity, using *D. magna* as a reference organism (von der Ohe & de Zwart, 2013; Wei et al., 2022). Although this filtering may potentially penalize the possible relationship with more bioindicators, it was necessary given our main goal. Studies have shown that filtering specific groups of pesticides enhances the relationship between chemical metrics and bioindicators (Ganatra et al., 2021; Hunt et al., 2017; Rasmussen et al., 2012).

SPEARpesticides' performance compared to the other bioindicators can be attributed to its trait-based design, which explicitly links community composition to pesticide sensitivity, generation time, dispersal ability, and exposure probability as noted in (Knillmann et al., 2018; Liess et al., 2021). This specificity allowed SPEARpesticides to detect changes in community structure even when other environmental stressors are present, a property that has been observed on multiple occasions under event-driven chemical sampling (Liess et al., 2021; Rasmussen et al., 2013; Schäfer et al., 2011). This distinction serves to highlight the importance of our methodical temporal alignment in detecting relationships. It is also important to acknowledge that the capacity of SPEARpesticides to reflect pesticide impacts is contingent not solely on its sensitivity, but also on the resolution and specificity of the chemical metrics employed (Leyva-Morales et al., 2024).

### 5.3. Integration of Environmental Parameters in Regression Models

Integrating environmental parameters alongside chemical metrics, supported by their conceptualization and previous analyses, was important to explaining the variability in the responses of macroinvertebrate communities to pesticide pressure, without making the model proposal overly complex and acknowledging the potential impact of external factors (Chollet Ramampiandra et al., 2023; Schuwirth et al., 2015).

In the 3.5-days sampling period, the combination of TU-EC(mix, median), proportion of agricultural land use, and flow velocity consistently yielded the most robust and interpretable models for SPEARpesticides. This outcome highlights the ecological reality that both static landscape attributes and dynamic stressors jointly shape biological communities in agricultural streams. The negative, nearly linear effect of proportion of agricultural land use on SPEARpesticides aligns with trait-based theory and empirical evidence, reinforcing the persistent role of land use as a stressor (Liess, 2023). The inclusion of TU-EC(mix, median) was further justified by its specificity, directly quantifying invertebrate-toxic pressure from pesticides, and capturing dynamic exposure events that static variables cannot (de Castro-Català et al., 2016; Leyva-Morales et al., 2024; Švara et al., 2021). Flow velocity, although not always exhibiting a strong directional effect, contributed to model performance by accounting for hydrological variability that can influence both pesticide transport and habitat suitability for macroinvertebrates (Burdon et al., 2019; Rico et al., 2016).

In the 14-days sampling period, the prominent chemical metric shifted from TUs to RQs. CRQ(mix, max) emerged as the most informative chemical predictor, and the relevance of environmental parameters, particularly flow velocity, and proportion of agricultural and urban land use, became more prominent than the chemical metric, as it was observed in other studies (Rasmussen et al., 2013; Rico et al., 2016). The 14-days sampling period, apparently smoothing short-term peaks in pesticide concentrations (Ashauer et al., 2020; Backhaus & Faust, 2012), reduced the detectability of acute exposure events and make this effect comparable to the influence of spatial and land-use gradients. In this context, the use of a broader, more conservative chemical metric (CRQ) and its maximum value within the sampling period proved more effective for capturing the effect of pesticides on SPEARpesticides.

The integration of environmental parameters into regression models reduced unexplained variance and highlighted the multi-stressor reality of stream ecosystems, where land use, hydrology, and chemical exposure interact to determine biological outcomes (Erasmus et al., 2021; Liess et al., 2016; Villeneuve et al., 2018). As alternative models we evaluated our best model structures without the chemical metric as a predictor leading to reduction in the explanatory power and, in the 3.5 days sampling period, potential overfitting of the environmental parameters.

These findings underscore a methodological insight for our datasets: neither chemical metrics nor environmental parameters alone can fully explain variability in bioindicator responses. Instead, integrated models that account for both types of predictors are necessary to capture the complexity of ecological responses in real-world, multi-stressor environments (Burdon et al., 2019; Heß et al., 2024; Rasmussen et al., 2012; Rico et al., 2016; Villeneuve et al., 2018). The approach taken here, selecting, and temporal aligning chemical metrics, integrating key environmental parameters, and adapting model structure to the characteristics of each sampling period, offers a reasonable framework for ecological assessment and improving the diagnostic power of both monitoring programs performed under the NAWA framework.

### 5.4. Modeling Approaches and Interpretation

The modeling strategy adopted in this study was intentionally flexible, reflecting the complexity and heterogeneity of the data as well as the need to maximize ecological interpretability. The assessment of transition from linear models (LM/LM) to generalized additive models (GAM/GAMM) was driven by diagnostic evidence, outlier sensitivity and modest explanatory power. While using median-based chemical metrics as predictors provided some robustness against outliers, linear models frequently failed to capture the nuanced relationships between pesticide exposure and bioindicator responses, as indicated by residual diagnostics and limited explanatory power (Gotelli & Ellison, 2013; QCBS, 2024).

The use of GAMs, specifically for the 3.5-days sampling period, allowed for a more data-driven and adaptive exploration of these relationships, revealing that significant and ecologically interpretable associations for SPEARpesticides were more consistently detected with this approach (Moe et al., 2025). In the case of the 14-days sampling period, it was proven that implementing GAMMs did not improve the overall model performance. This is likely because the chemical metric in those datasets was not strong enough to express the variability of the bioindicator. This shows that SPEARpesticides is not sensitive enough for this type of sample. Nevertheless, all of the selected statistical modeling approaches consistently reflected the negative trend between pesticide pressure and sensitive macroinvertebrate taxa. However, the emergence of different metrics as primary predictors highlights the importance of aligning model structure and metric selection with the ecological question and data characteristics.

Visual inspection of model fits and residuals was crucial for ensuring ecological plausibility, particularly in avoiding over-interpretation of statistically significant but biologically spurious patterns, an issue occasionally observed for broader indices such as GI. This emphasis on visual and ecological validation, alongside statistical criteria, aligns with best practices in ecological modeling and risk assessment (Moe et al., 2025; QCBS, 2024).

Despite the advances provided by GAMs, the overall explanatory power of chemical metrics remained moderate, with a maximum of approximately 30% of SPEARpesticides' variability explained in the best 3.5-days sampling period model. In this case, the implementation of environmental parameters proved to be of critical importance. This implementation led to a significant enhancement in the explanatory power of the model, with the proportion of variability explained rising to 60%. Additionally, it contributed to a reduction in the uncertainty surrounding the partial effect of the chemical metric. This reduction in uncertainty enabled us to articulate with an improved degree of confidence the number of units of the bioindicator that are diminished by an increase of one standard deviation in the chemical metric. This was permitted in view of the additive character of the model structure that was defined, wherein the baseline levels of the other predictors and the intercept of each model are to be acknowledged to evaluate the predictions (QCBS, 2024). This interpretation demonstrates the limitations of sample size, but also the opportunities for improving the model to take potential actions regarding the overall problem addressed by the thesis. A simple exercise to move one standard deviation in the chemical metric to determine the decrease of SPEARpesticide response demonstrates that "pesticide pressure" is indeed the variable to take further immediate actions, followed by land use, to improve the ecological status of streams, as expressed by SPEARpesticides through sensitive taxa.

The levels of explained variability before and after the implementation of environmental variables are comparable to those reported in similar studies (Hunt et al., 2017; Liess et al., 2021), but it is important to recognize that achieving these results required substantial data harmonization, metric refinement, and flexible modeling, reflecting the challenges of integrating datasets with differing temporal, spatial, and methodological characteristics. The need for such analytical effort highlights the advantage of monitoring programs that coordinate chemical and biological sampling in a temporally consistent manner, as this alignment can enhance the detectability of stressor–response relationships and reduce the need for post hoc data manipulation (Spycher et al., 2018).

In summary, the modeling approaches employed here demonstrate the value of flexibility and ecological reasoning in uncovering relationships between pesticide exposure and macroinvertebrate communities, while also illustrating the inherent limitations imposed by data structure and sampling design.

## 5.5. Cross-Validation and Model Robustness

The application of cross-validation (CV) and Random Forest (RF) benchmarking provided a transparent assessment of model robustness and generalizability, complementing the inferential focus of this study. While the primary goal was not to maximize predictive accuracy, CV results highlighted the strengths and limitations of the selected modeling approaches under the constraints of the datasets.

For the 3.5-days sampling period, both GAM and RF models demonstrated reasonable performance, particularly as the number of observations increased. The inclusion of the chemical metric TU-EC(mix, median) was vital for maintaining model stability and explanatory power; its omission led to reduced performance and instability. Notably, RF models often outperformed parametric models in predictive metrics, suggesting the presence of non-linear or interactive relationships that flexible, non-parametric approaches can better accommodate (Leigh & Datry, 2017). However, the overall consistency between GAM and RF results supports the adequacy of the selected GAM structure for inferential purposes, especially when both chemical and environmental parameters were included.

In the 14-days sampling period, RF models outperformed linear mixed models (LMM) and linear models (LM), with RF achieving $R^2$ values nearly double those of parametric approaches. The frequent singularity of LMMs during CV underscored the challenges posed by limited sample size, data structure, and the inclusion of random effects. Attempts to model interactions or simplify the structure did not yield improvements, indicating that more complex relationships may exist but cannot be robustly captured with the current dataset (Chollet Ramampiandra et al., 2023; Leigh & Datry, 2017).

Overall, the CV and RF benchmarking not only test the inferential findings of this thesis but also revealed the limitations inherent in the current dataset, such as small sample sizes, lack of pristine reference sites, unbalanced monitoring across streams and years, and the use of static abiotic variables. These insights point to the need for future studies to employ larger, more diverse datasets and to consider different alternatives of regression models to better capture the complexity of ecological relationships in stream ecosystems.

### 5.6. Time Windows – Temporal Integration Effects

The time windows, initially conceptualized to optimize the aggregation of chemical samples for explaining bioindicator variability, emerged as a critical factor influencing the strength and nature of the relationships between pesticide exposure and macroinvertebrate bioindicators.

In the 3.5-days sampling period the stronger and more robust negative association between chemical metrics and SPEARpesticides were found in the first time-windows (from 1-week to 2-months). This effect was further enhanced by the inclusion of environmental parameters, which allowed for a more consistent explanation of the variability in bioindicator responses across the exposure gradient, as seen in the partial effect plots. Conversely, in the 14-days sampling period, chemical metrics alone were insufficient to consistently explain declines in SPEARpesticides beyond one standard deviation of exposure. The addition of more samples and abiotic variables only modestly reduced model uncertainty and did not substantially improve predictive power within the observed data range. This pattern aligns with previous studies (Ashauer et al., 2020; Backhaus & Faust, 2012; Spycher et al., 2018), which have highlighted the limitations of chronic, composite sampling in capturing the ecological relevance of short-term exposure peaks.

The temporal aggregation in the 14-day composites evens out the concentration over episodic toxicity events. This affects the importance of spatial and land-use variables as environmental gradients and shifts the explanatory power toward broader, more conservative chemical metrics, such as CRQ(mix, max). While this approach can improve model stability with a similar $R^2$ across time windows, it may also dilute the direct attribution of macroinvertebrate responses to specific pesticide events. This demonstrates the reduced sensitivity of indices like SPEARpesticides for these types of samples.

In the CV assessment, both GAM and RF provided acceptable explanatory power and consistency in corroborating the findings, while LMM proved less effective. These results suggest that the observed improvements may be driven by the chemical exposure signal inherent to the sampling period to which SPEARpesticides is sensitive, rather than by the statistical model chosen, the amount of data, or how the data are integrated. It is important to note that the 3.5- and 14-days sampling periods differ in more than just temporal resolution. They also differ in the substances analyzed and their main sampling objective, as discussed in previous sections.

These findings highlight a fundamental trade-off in ecological monitoring design: acute, higher-resolution sampling (3.5-days in comparison to the 14-days sample) is more sensitive to episodic events but may be more susceptible to noise and temporal mismatches, while chronic, lower-resolution sampling offers stability at the cost of reduced sensitivity to acute impacts.

### 5.7. Implications for Monitoring, Management, and Future Research

First, the demonstrated specificity of the SPEARpesticides, especially when paired with carefully selected chemical metrics and key environmental parameters, support its continued use and further development as a diagnostic tool for pesticide impact assessment in Swiss streams. The “GI index” and “EPT richness” also showed potential as a supplementary indicator for pesticide-related pressure, though its broader ecological scope makes it less specific to this stress when using the chemical measurement from the composite samples.

These results emphasize the necessity of integrating chemical and biological data streams into monitoring programs. Temporal alignment between chemical and biological sampling is particularly important to maximize the detectability of pesticide effects, especially for acute exposure events.

The study also highlights the value of including key environmental parameters, such as proportion of agricultural and urban land use and flow velocity, in both routine monitoring and ecological assessment frameworks. These variables not only improved model performance but also provided crucial ecological context for interpreting bioindicator responses. The consistent influence of proportion of agricultural land use across scales suggests that land management practices remain a central approach for mitigating pesticide impacts in agricultural catchments (Ilg & Alther, 2024; Khaliq et al., 2024; Nguyen et al., 2023).

Despite these advances, several limitations must be acknowledged. The reliance on composite chemical samples, the absence of pristine reference sites, static abiotic variables, and unbalanced monitoring designs all constrained the explanatory power and generalizability of the models. Once a sensitivity analysis of the “rare events” that generated that rise in the curves is performed we might be able to argue better where the main monitoring efforts should be addressed to clarify such specific relationship “SPEARpesticides and chemical metrics”, since the present analysis showed that the closer the chemical monitoring to the biological monitoring is better and the aggregation of more data is not necessarily beneficial for the relationship. For example, event-driven, continuous high-resolution measurements, increasing sampling size and spatial coverage could provide a better alternative capturing short-term fluctuation episodes pollution events that current information from our databases may miss. However, the cost could be much higher for only one purpose  (Spycher et al., 2018). Therefore, long-term monitoring programs that are well-designed and capable of assessing multiple stressors, including those not initially targeted, are crucial for a comprehensive understanding of ecosystem health. Purpose-designed monitoring campaigns that harmonize chemical and biological sampling in time and space, and that are flexible enough to address both known and unknown stressors, will be especially valuable for disentangling the complex drivers of macroinvertebrate community change.

On the other hand, as research on diverse modes of action expands, these may not be fully captured in the analysis for the Species At Risk used to determine the SPEARpesticides. Greater attention to taxonomic resolution is also warranted, as this underpins index calculations and ecological interpretation. While this study took a broad approach, future work would benefit from more detailed taxonomic and contaminant-specific analyses to better understand stressor impacts on aquatic communities.

# 6. CONCLUSION

This study quantitatively evaluated the relationship between chemical metrics and the SPEARpesticides index, alongside other bioindicators, to evaluate how effectively these tools reflect pesticide impacts on macroinvertebrate communities in Swiss streams within the NAWA framework.

First, the results confirm that pesticide exposure, as measured by both RQs and TUs, is associated with declines in sensitive macroinvertebrate taxa, most clearly captured by the SPEARpesticides index. Both RQs and TUs demonstrated negative associations with bioindicator values, but their relative performance depended on the modeling approach, reaching 28% and 18% of the variability explained in the 3.5- and 14-days sampling periods. RQs, being based on broad environmental quality criteria, performed well in linear models and provided a conservative, regulatory-oriented signal. In contrast, TUs, derived from effect concentrations for sensitive aquatic species, proved more nuanced and ecologically targeted, especially when flexible, non-linear models were employed.

Second, the integration of environmental parameters, particularly agricultural land use and flow velocity, improved the explanatory power of models, supporting the hypothesis that harmonized, multi-factor approaches more effectively capture the ecological impacts of pesticide mixtures than chemical metrics alone. The inclusion of other abiotic factors allowed for a broader and more consistent explanation of variance across the exposure gradient reaching around 60% and 36% of the variability explained in the 3.5- and 14-days sampling periods respectively, underscoring the value of integrated monitoring strategies.

Third, the comparative analysis of acute (3.5-days) and chronic (14-days) sampling periods and the time windows used for chemical samples aggregation revealed their critical influence in the sensitivity and stability of bioindicator responses. The 3.5-days samples and the shorter time windows were more effective at detecting episodic pesticide events and yielded stronger explanatory relationships, particularly for SPEARpesticides. In contrast, longer, chronic sampling periods provided greater model stability but diluted acute exposure signals, resulting in less sensitive bioindicator responses.

Despite these advances, a considerable amount of unexplained variance remained in all models. This is attributable to inherent complexities in field conditions, measurement limitations, the structure of composite samples, and the lack of pristine reference sites, factors well recognized in ecological risk assessment. The presence of rare or extreme events, particularly at the margins of the exposure gradient, further increased model uncertainty and highlighted the need for expanded data coverage and targeted sensitivity analyses.

In summary, the present study demonstrates that meticulous data processing, temporal alignment, data integration, and careful selection of chemical metrics with bioindicators enhance the diagnostic power of ecological risk assessments while using the Swiss databases. The extant data demonstrate a robust and consistent negative relationship between TU-EC(mix, median) and SPEARpesticides. This relationship is indicative of pesticide pressure on macroinvertebrate communities in Swiss streams. This is particularly relevant when employing chemical samples that are closely aligned with the biological monitoring, with the objective of reflecting specific episodes of pesticide exposure and acknowledging the multistressor reality. The findings advocate for continued refinement of monitoring programs, emphasizing methodological transparency, harmonization, and the adoption of integrated, multi-factor approaches to manage pesticide risks in freshwater ecosystems.

## 7. ACKNOWLEDGEMENTS

I sincerely thank the VSA and SIAM departments for hosting me over the past 6 months and for their comprehensive support and facilities, including both human and IT resources. This project was made possible by the fund and infrastructure provided by Eawag and the Federal Office for the Environment.

I am especially grateful to my supervisors, Anne Dietzel, Christiane Ilg, and Nele Schuwirth, for their ongoing support, openness, and for connecting me with key colleagues who helped shape this work. I also thank Matthias Liess for his readiness to review my progress and offer valuable feedback.

Special thanks to Tobias Doppler and Sophia Barth (VSA), and Andreas Scheidegger and Rosi Siber (SIAM), for their willingness to discuss ideas and share their expertise throughout this project.

# REFERENCES


Abbasi, N. A., Shahid, S. U., Majid, M., & Tahir, A. (2022). Ecotoxicological risk assessment of environmental micropollutants. In *Environmental Micropollutants: A Volume in Advances in Pollution Research* (pp. 331–337). Elsevier. https://doi.org/10.1016/B978-0-323-90555-8.00004-0

Albert, J. S., Destouni, G., Duke-Sylvester, S. M., Magurran, A. E., Oberdorff, T., Reis, R. E., Winemiller, K. O., & Ripple, W. J. (2021). Scientists' warning to humanity on the freshwater biodiversity crisis. *Ambio*, *50*(1), 85–94. https://doi.org/10.1007/S13280-020-01318-8

Ashauer, R., Kuhl, R., Zimmer, E., & Junghans, M. (2020). Effect Modeling Quantifies the Difference Between the Toxicity of Average Pesticide Concentrations and Time-Variable Exposures from Water Quality Monitoring. *Environmental Toxicology and Chemistry*, *39*(11), 2158–2168. https://doi.org/10.1002/etc.4838

Backhaus, T., & Faust, M. (2012). Predictive environmental risk assessment of chemical mixtures: A conceptual framework. *Environmental Science and Technology*, *46*(5), 2564–2573. https://doi.org/10.1021/ES2034125

BAFU. (1998). *Methoden zur Untersuchung und Beurteilung der Fliessgewässer in der Schweiz - Modul-Stufen- Konzept*. http://www.admin.ch/buwal/publikat/d/

BAFU. (2013). *NAWA – Nationale Beobachtung Oberflächengewässer- qualität. Konzept Fliessgewässer*. www.bafu.admin.ch/uw-1327-d

BAFU. (2019). *Methoden zur Untersuchung und Beurteilung von Fliessgewässern (IBCH_2019). Makrozoobenthos - Stufe F. 1. aktualisierte Ausgabe*. www.bafu.admin.ch/uv-1026-d

BAFU. (2020). *Mean discharge values and discharge regime type for the river network in Switzerland*. Mean Monthly and Annual Discharge. https://www.bafu.admin.ch/bafu/en/home/topics/water/state-of-watercourses/water-flow-and-flow-regime-in-watercourses/mean-monthly-and-annual-discharge.html#906411347

BAFU. (2022). *Gewässer in der Schweiz. Zustand und Massnahmen*. http://www.bafu.admin.ch/uz-2207-d

Bai, Y., Ruan, X., & van der Hoek, J. P. (2018). Residues of organochlorine pesticides (OCPs) in aquatic environment and risk assessment along Shaying River, China. *Environmental Geochemistry and Health*, *40*(6), 2525–2538. https://doi.org/10.1007/s10653-018-0117-9

BDM. (2021). *Anleitung für die Feldmitarbeiter:innen zum Indikator «Z9-Gewässerinsekten»*. http://www.biodiversitymonitoring.ch

BDM. (2022). *Anleitung für die Laborarbeit zum Indikator «Z9-Gewässerinsekten»*. https://www.biodiversitymonitoring.ch/

Beketov, M. A., Kefford, B. J., Schäfer, R. B., & Liess, M. (2013). Pesticides reduce regional biodiversity of stream invertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11039–11043. https://doi.org/10.1073/pnas.1305618110

Bettinetti, R., Zaupa, S., Fontaneto, D., & Boggero, A. (2020). Biological, Chemical, and Ecotoxicological Assessments Using Benthos Provide Different and Complementary Measures of Lake Ecological Status. *Water 2020, Vol. 12, Page 1140*, *12*(4), 1140. https://doi.org/10.3390/W12041140

Beyer, J., Petersen, K., Song, Y., Ruus, A., Grung, M., Bakke, T., & Tollefsen, K. E. (2014). Environmental risk assessment of combined effects in aquatic ecotoxicology: A discussion paper. *Marine Environmental Research*, *96*, 81–91. https://doi.org/10.1016/J.MARENVRES.2013.10.008

Burdon, F. J., Munz, N. A., Reyes, M., Focks, A., Joss, A., Räsänen, K., Altermatt, F., Eggen, R. I. L., & Stamm, C. (2019). Agriculture versus wastewater pollution as drivers of macroinvertebrate community structure in streams. *Science of the Total Environment*, *659*, 1256–1265. https://doi.org/10.1016/j.scitotenv.2018.12.372

Burdon, F. J., Reyes, M., Alder, A. C., Joss, A., Ort, C., Räsänen, K., Jokela, J., Eggen, R. I. L., & Stamm, C. (2016). Environmental context and magnitude of disturbance influence trait-mediated community responses to wastewater in streams. *Ecology and Evolution*, *6*(12), 3923–3939. https://doi.org/10.1002/ece3.2165

Chollet Ramampiandra, E., Scheidegger, A., Wydler, J., & Schuwirth, N. (2023). A comparison of machine learning and statistical species distribution models: Quantifying overfitting supports model interpretation. *Ecological Modelling*, *481*, 110353. https://doi.org/10.1016/j.ecolmodel.2023.110353

Chow, R., Scheidegger, R., Doppler, T., Dietzel, A., Fenicia, F., & Stamm, C. (2020). A review of long-term pesticide monitoring studies to assess surface water quality trends. *Water Research X*, *9*. https://doi.org/10.1016/j.wroa.2020.100064

Daouk, S., Doppler, T., Scheidegger, R., Kroll, A., Junghans, M., Moschet, C., & Singer, H. (2022). Insektizide In Schweizer Fliessgewässern. Welche Risiken Gehen Von Pyrethroiden Und Organophosphaten Aus? *Aqua & Gas*, *4*.

de Castro-Català, N., Kuzmanovic, M., Roig, N., Sierra, J., Ginebreda, A., Barceló, D., Pérez, S., Petrovic, M., Picó, Y., Schuhmacher, M., & Muñoz, I. (2016). Ecotoxicity of sediments in rivers: Invertebrate community, toxicity bioassays and the toxic unit approach as complementary assessment tools. *Science of The Total Environment*, *540*, 297–306. https://doi.org/10.1016/J.SCITOTENV.2015.06.071

Doppler, T., Camenzuli, L., Hirzel, G., Krauss, M., Lück, A., & Stamm, C. (2012). Spatial variability of herbicide mobilisation and transport at catchment scale: insights from a field experiment. *Hydrology and Earth System Sciences*, *16*(7), 1947–1967. https://doi.org/10.5194/hess-16-1947-2012

Doppler, T., Dietzel, A., & Stamm, C. (2024). Pestizide in Bächen und Flüssen. Wirkung des Aktionsplans Pflanzenschutzmittel. *Aqua & Gas*, *104*(7+8), 63–69.

Doppler, T., Dietzel, A., Wittmer, I., Grelot, J., Päivi Rinta, & Kunz, M. (2020). Mikroverunreinigungen im Gewässermonitoring. *Aqua & Gas*, *7/8*, 4453.

Doppler, T., Mangold, S., Wittmer, I., Spycher, S., Comte, R., Stamm, C., Singer, H., Junghans, M., & Kunz, M. (2017). Hohe PSM-Belastung In Schweizer Bächen NAWA-SPEZ-Kampagne Untersucht Bäche In Gebieten Intensiver Landwirtschaftlicher Nutzung. *Aqua & Gas*, *4*, 46–56.

EC. (2011). *Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Technical guidance for deriving environmental quality standards*. Publications Office. https://commission.europa.eu/index_en

Epele, L. B., Williams-Subiza, E. A., Bird, M. S., Boissezon, A., Boix, D., Demierre, E., Fair, C. G., García, P. E., Gascón, S., Grech, M. G., Greig, H. S., Jeffries, M., Kneitel, J. M., Loskutova, O., Maltchik, L., Manzo, L. M., Mataloni, G., McLean, K., Mlambo, M. C., … Batzer, D. P. (2024). A global assessment of environmental and climate influences on wetland macroinvertebrate community structure and function. *Global Change Biology*, *30*(2), e17173. https://doi.org/10.1111/GCB.17173

Erasmus, J. H., Lorenz, A. W., Zimmermann, S., Wepener, V., Sures, B., Smit, N. J., & Malherbe, W. (2021). A diversity and functional approach to evaluate the macroinvertebrate responses to multiple stressors in a small subtropical austral river. *Ecological Indicators*, *131*, 108206. https://doi.org/10.1016/J.ECOLIND.2021.108206

Espinar-Herranz, S., Velázquez, J., Pérez-Andueza, G., Çiçek, K., & Rincón, V. (2025). Aquatic Macroinvertebrates as Indicators of Climate Change in the Riparian Environments of the Mediterranean Region. *Land 2025, Vol. 14, Page 521*, *14*(3), 521. https://doi.org/10.3390/LAND14030521

FAO. (2021). *The state of the world's land and water resources for food and agriculture – Systems at breaking point. Synthesis report 2021.* . https://doi.org/10.4060/cb7654en

Ganatra, A. A., Kandie, F. J., Fillinger, U., McOdimba, F., Torto, B., Brack, W., Liess, M., Hollert, H., & Becker, J. M. (2021). Calibration of the SPEARpesticides bioindicator for cost-effective pesticide monitoring in East African streams. *Environmental Sciences Europe*, *33*(1). https://doi.org/10.1186/s12302-021-00497-9

Gewässerschutzverordnung, GSchV, Pub. L. No. SR 814.201 (1998). https://www.fedlex.admin.ch/eli/cc/1998/2863_2863_2863/en

Ginebreda, A., Kuzmanovic, M., Guasch, H., de Alda, M. L., López-Doval, J. C., Muñoz, I., Ricart, M., Romaní, A. M., Sabater, S., & Barceló, D. (2014). Assessment of multi-chemical pollution in aquatic ecosystems using toxic units: Compound prioritization, mixture characterization and relationships with biological descriptors. *Science of The Total Environment*, *468–469*, 715–723. https://doi.org/10.1016/J.SCITOTENV.2013.08.086

Gotelli, N. J., & Ellison, A. M. (2013). *A Primer of Ecological Statistics - Second Edition*. Sinauer Associates, Inc. https://learninglink.oup.com/access/gotelli-a-primer-of-ecological-statistics-2e

Haase, P., Bowler, D. E., Baker, N. J., Bonada, N., Domisch, S., Garcia Marquez, J. R., Heino, J., Hering, D., Jähnig, S. C., Schmidt-Kloiber, A., Stubbington, R., Altermatt, F., Álvarez-Cabria, M., Amatulli, G., Angeler, D. G., Archambaud-Suard, G., Jorrín, I. A., Aspin, T., Azpiroz, I., … Welti, E. A. R. (2023). The recovery of European freshwater biodiversity has come to a halt. *Nature*, *620*(7974), 582–588. https://doi.org/10.1038/S41586-023-06400-1

Herman, M. R., & Nejadhashemi, A. P. (2015). A review of macroinvertebrate- and fish-based stream health indices. *Ecohydrology & Hydrobiology*, *15*(2), 53–67. https://doi.org/10.1016/J.ECOHYD.2015.04.001

Heß, S., Hof, D., Oetken, M., & Sundermann, A. (2024). Macroinvertebrate communities respond strongly but non-specifically to a toxicity gradient derived by effect-based methods. *Environmental Pollution*, *356*, 124330. https://doi.org/10.1016/J.ENVPOL.2024.124330

Hou, L., Xiong, W., Chen, M., Xu, J., Johnson, A. C., Zhan, A., & Jin, X. (2025). Pesticide Pollution Reduces the Functional Diversity of Macroinvertebrates in Urban Aquatic Ecosystems. *Environmental Science and Technology*, *59*, 8577. https://doi.org/10.1021/ACS.EST.5C01093

Hunt, L., Bonetto, C., Marrochi, N., Scalise, A., Fanelli, S., Liess, M., Lydy, M. J., Chiu, M. C., & Resh, V. H. (2017). Species at Risk (SPEAR) index indicates effects of insecticides on stream invertebrate communities in soy production regions of the Argentine Pampas. *Science of the Total Environment*, *580*, 699–709. https://doi.org/10.1016/j.scitotenv.2016.12.016

Hutter, P., Roth, T., Martinez, N., & Stucki, P. (2019). Fliessgewässer - Fauna unter druck. Erste trends aus dem biodiversitätsmonitoring Schweiz (BDM). *Aqua & Gas*, *7/8*.

Ilg, C., & Alther, R. (2024). Ökologischer Zustand von Schweizer Bächen. Die meisten der untersuchten Bäche erfüllen ihre Rolle als Lebensraum für Tiere nur eingeschränkt. *Aqua & Gas*, *4*. https://www.dora.lib4ri.ch/eawag/islandora/object/eawag%3A32727

Junghans, M., Kunz, P., & Werner, I. (2013). Toxizität Von Mischungen. Aktuelle, Praxisorientierte Ansätze für die Beurteilung von Gewässerproben. *Aqua & Gas*, *5*. https://www.ecotoxcentre.ch

Junghans, M., Langer, M., Baumgartnet, C., Vermeirseen, E., & Werner, I. (2019). Ökotoxikologische Untersuchungen: Risiko Von PSM Bestätigt. NAWA-SPEZ-Studie 2017 zeigt Beeinträchtigung von Gewässerorganismen. *Aqua & Gas*, *99*(4), 26–34. https://www.dora.lib4ri.ch/eawag/islandora/object/eawag%3A19545

Junghans, M., N. Chèvre, C. Di Paolo, R.I.L. Eggen, R. Gälli, V. Gregorio, A. Häner, N. Homazava, C. Perazzolo, & R. Kase. (2011). *Aquatic Risks of Plant Protection Products: A Comparison of Different Hazard Assessment Strategies for Surface Waters in Switzerland*. www.oekotoxzentrum.ch

Khaliq, I., Chollet Ramampiandra, E., Vorburger, C., Narwani, A., & Schuwirth, N. (2024). The effect of water temperature changes on biological water quality assessment. *Ecological Indicators*, *159*. https://doi.org/10.1016/j.ecolind.2024.111652

Kienzler, A., Bopp, S., Halder, M., Embry, M., & Worth, A. (2019). Application of new statistical distribution approaches for environmental mixture risk assessment: A case study. *Science of the Total Environment*, *693*. https://doi.org/10.1016/j.scitotenv.2019.07.316

Kienzler, A., Bopp, S. K., van der Linden, S., Berggren, E., & Worth, A. (2016). Regulatory assessment of chemical mixtures: Requirements, current approaches and future perspectives. *Regulatory Toxicology and Pharmacology*, *80*, 321–334. https://doi.org/10.1016/J.YRTPH.2016.05.020

Knillmann, S., Orlinskiy, P., Kaske, O., Foit, K., & Liess, M. (2018). Indication of pesticide effects and recolonization in streams. *Science of the Total Environment*, *630*, 1619–1627. https://doi.org/10.1016/j.scitotenv.2018.02.056

Kumar, A., Jasrotia, S., Dutta, J., & Kyzas, G. Z. (2023). Pyrethroids toxicity in vertebrates and invertebrates and amelioration by bioactive compounds: A review. *Pesticide Biochemistry and Physiology*, *196*, 105615. https://doi.org/10.1016/J.PESTBP.2023.105615

Kuzmanović, M., López-Doval, J. C., De Castro-Català, N., Guasch, H., Petrović, M., Muñoz, I., Ginebreda, A., & Barceló, D. (2016). Ecotoxicological risk assessment of chemical pollution in four Iberian river basins and its relationship with the aquatic macroinvertebrate community status. *Science of The Total Environment*, *540*, 324–333. https://doi.org/10.1016/J.SCITOTENV.2015.06.112

Lee, J. W., Lee, S. W., An, K. J., Hwang, S. J., & Kim, N. Y. (2020). An Estimated Structural Equation Model to Assess the Effects of Land Use on Water Quality and Benthic Macroinvertebrates in Streams of the Nam-Han River System, South Korea. *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 2116*, *17*(6), 2116. https://doi.org/10.3390/IJERPH17062116

Leigh, C., & Datry, T. (2017). Drying as a primary hydrological determinant of biodiversity in river systems: a broad-scale analysis. *Ecography*, *40*(4), 487–499. https://doi.org/10.1111/ECOG.02230

Lewis, K. A., Tzilivakis, J., Warner, D. J., & Green, A. (2016). An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment: An International Journal*, *22*(4), 1050–1064. https://doi.org/10.1080/10807039.2015.1133242

Leyva-Morales, J. B., Bastidas-Bastidas, P. de J., Salvatierra-Stamp, V. del C., González-Márquez, L. C., Davizon, Y. A., Márquez-Pacheco, H., Amillano-Cisneros, J. M., Cruz-Acevedo, E., Amézquita-López, B. A., Romo Gómez, C., De la Torre Espinosa, Z. Y., & Rodríguez-Aguilar, B. A. (2024). Temporal distribution and ecological risk assessment for pesticides in water from the north-central coastal zone of Sinaloa, Mexico. *Heliyon*, *10*(15), e35207. https://doi.org/10.1016/J.HELIYON.2024.E35207

Liess, M. (2023, January 27). *Effect monitoring (SPEAR)*. Chemicals in the Environment. https://www.ufz.de/index.php?de=38122

Liess, M., Foit, K., Knillmann, S., Schäfer, R. B., & Liess, H. D. (2016). Predicting the synergy of multiple stress effects. *Scientific Reports*, *6*. https://doi.org/10.1038/srep32965

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. (2021). Pesticides are the dominant stressors for vulnerable insects in lowland streams. *Water Research*, *201*. https://doi.org/10.1016/j.watres.2021.117262

Liess, M., Schäfer, R. B., & Schriever, C. A. (2008). The footprint of pesticide stress in communities-Species traits reveal community effects of toxicants. *Science of the Total Environment*, *406*(3), 484–490. https://doi.org/10.1016/j.scitotenv.2008.05.054

Liess, M., & von der Ohe. (2005). Analyzing effects of pesticides on invertebrate communities in streams. *Environmental Toxicology and Chemistry*, *24*(4), 954–965. https://academic.oup.com/etc/article/24/4/954/7770594

Moe, S. J., Benestad, R. E., & Landis, W. G. (2022). Robust risk assessments require probabilistic approaches. *Integrated Environmental Assessment and Management*, *18*(5), 1133–1134. https://doi.org/10.1002/IEAM.4660

Moe, S. J., Fox, D. R., & Wolf, R. (2025). Updating statistical practice in ecotoxicology: reflections and recommendations Editorial. *Integrated Environmental Assessment and Management*, *2025*(3), 475–476. https://doi.org/10.1093/inteam/vjaf029

Morin, S., & Artigas, J. (2023). Twenty Years of Research in Ecosystem Functions in Aquatic Microbial Ecotoxicology. *Environmental Toxicology and Chemistry*, *42*(9), 1867–1888. https://doi.org/10.1002/ETC.5708

Moschet, C., Deola, L., Ruff, M., Söser, R., Mühlemann, S., Altherr, T., Götz, C., Wyss, A., Singer, H., Rösh, A., & Daouk, S. (2019). Task-force «Pyrethroide»: Erfolgreiche Zusammenarbeit der Kantone auf der Suche nach toxischen Insektiziden im Gewässer. *Aqua & Gas*, *11*, 68–73.

MSK. (2025). *Methods for the analysis and assessment of surface waters in Switzerland*. Methods. https://modul-stufen-konzept.ch/en/modular-stepwise-procedure/

Munz, N. A., Burdon, F. J., de Zwart, D., Junghans, M., Melo, L., Reyes, M., Schönenberger, U., Singer, H. P., Spycher, B., Hollender, J., & Stamm, C. (2017). Pesticides drive risk of micropollutants in wastewater-impacted streams during low flow conditions. *Water Research*, *110*, 366–377. https://doi.org/10.1016/J.WATRES.2016.11.001

Nguyen, H. H., Welti, E. A. R., Haubrock, P. J., & Haase, P. (2023). Long-term trends in stream benthic macroinvertebrate communities are driven by chemicals. *Environmental Sciences Europe*, *35*(1), 1–15. https://doi.org/10.1186/S12302-023-00820-6/METRICS

Peterson, R. K. D. (2006). Comparing ecological risks of pesticides: the utility of a Risk Quotient ranking approach across refinements of exposure. *Pest Management Science*, *62*(1), 46–56. https://doi.org/10.1002/PS.1126

Poikane, S., Johnson, R. K., Sandin, L., Schartau, A. K., Solimini, A. G., Urbanič, G., Arbačiauskas, K. stutis, Aroviita, J., Gabriels, W., Miler, O., Pusch, M. T., Tim, H., & Böhmer, J. (2016). Benthic macroinvertebrates in lake ecological assessment: A review of methods, intercalibration and practical recommendations. *Science of The Total Environment*, *543*, 123–134. https://doi.org/10.1016/J.SCITOTENV.2015.11.021

Poikane, S., Salas Herrero, F., Kelly, M. G., Borja, A., Birk, S., & van de Bund, W. (2020). European aquatic ecological assessment methods: A critical review of their sensitivity to key pressures. *The Science of the Total Environment*, *740*, 140075. https://doi.org/10.1016/J.SCITOTENV.2020.140075

QCBS. (2024). *Open-access peer-reviewed R workshops for statistical analyses in biology and ecology*. GitHub. https://github.com/QCBSRworkshops

Rasmussen, J. J., McKnight, U. S., Loinaz, M. C., Thomsen, N. I., Olsson, M. E., Bjerg, P. L., Binning, P. J., & Kronvang, B. (2013). A catchment scale evaluation of multiple stressor effects in headwater streams. *Science of the Total Environment*, *442*, 420–431. https://doi.org/10.1016/j.scitotenv.2012.10.076

Rasmussen, J. J., Wiberg-Larsen, P., Baattrup-Pedersen, A., Friberg, N., & Kronvang, B. (2012). Stream habitat structure influences macroinvertebrate response to pesticides. *Environmental Pollution*, *164*, 142–149. https://doi.org/10.1016/j.envpol.2012.01.007

Rico, A., & Van den Brink, P. J. (2015). Evaluating aquatic invertebrate vulnerability to insecticides based on intrinsic sensitivity, biological traits, and toxic mode of action. *Environmental Toxicology and Chemistry*, *34*(8), 1907–1917. https://doi.org/10.1002/ETC.3008

Rico, A., Van den Brink, P. J., Leitner, P., Graf, W., & Focks, A. (2016). Relative influence of chemical and non-chemical stressors on invertebrate communities: a case study in the Danube River. *Science of The Total Environment*, *571*, 1370–1382. https://doi.org/10.1016/J.SCITOTENV.2016.07.087

Robinson, C. T., Schuwirth, N., Baumgartner, S., & Stamm, C. (2014). Spatial relationships between land-use, habitat, water quality and lotic macroinvertebrates in two Swiss catchments. *Aquatic Sciences*, *76*(3), 375–392. https://doi.org/10.1007/s00027-014-0341-z

Rösch, A., Beck, B., Hollender, J., Stamm, C., Singer, H., Doppler, T., & Junghans, M. (2019). Geringe Konzentrationen Mit Grosser Wirkung. Nachweis Von Pyrethroid- Und Organophosphat-insektiziden In Schweizer Bächen Im pg l–1-Bereich. *Aqua & Gas*, *11*, 54–66.

Schäfer, R. B., Gerner, N., Kefford, B. J., Rasmussen, J. J., Beketov, M. A., De Zwart, D., Liess, M., & Von Der Ohe, P. C. (2013). How to characterize chemical exposure to predict ecologic effects on aquatic communities? *Environmental Science and Technology*, *47*(14), 7996–8004. https://doi.org/10.1021/ES4014954

Schäfer, R. B., Kefford, B. J., Metzeling, L., Liess, M., Burgert, S., Marchant, R., Pettigrove, V., Goonan, P., & Nugegoda, D. (2011). A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in South-East Australia. *Science of the Total Environment*, *409*(11), 2055–2063. https://doi.org/10.1016/j.scitotenv.2011.01.053

Schäfer, R. B., & Liess, M. (2013). Species at Risk (SPEAR) Biomonitoring Indicators. In JF. Férard & C. Blaise (Eds.), *Encyclopedia of Aquatic Ecotoxicology* (pp. 1063–1072). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-5704-2_96

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? *Integrated Environmental Assessment and Management*, *21*(1), 3–19. https://doi.org/10.1093/inteam/vjae007

Schuwirth, N., Kattwinkel, M., & Stamm, C. (2015). How stressor specific are trait-based ecological indices for ecosystem management? *Science of the Total Environment*, *505*, 565–572. https://doi.org/10.1016/j.scitotenv.2014.10.029

SETAC. (2018). *Technical Issue Paper: Environmental Risk Assessment of Chemicals.* https://www.setac.org/asset/8C20F72C-5BBA-46DE-AB3828C43810D98A/

Shah, R. D. T., Sharma, S., Shah, D. N., & Rijal, D. (2020). Structure of Benthic Macroinvertebrate Communities in the Rivers of Western Himalaya, Nepal. *Geosciences 2020, Vol. 10, Page 150*, *10*(4), 150. https://doi.org/10.3390/GEOSCIENCES10040150

Spycher, S., Mangold, S., Doppler, T., Junghans, M., Wittmer, I., Stamm, C., & Singer, H. (2018). Pesticide Risks in Small Streams - How to Get as Close as Possible to the Stress Imposed on Aquatic Organisms. *Environmental Science and Technology*, *52*(8), 4526–4535. https://doi.org/10.1021/acs.est.8b00077

Spycher, S., Teichler, R., Vonwyl, E., Longrée, P., Stamm, C., Singer, H., Daouk, S., Doppler, T., & Junghans, M. (2019). Anhaltend hohe PSM-Belastung in Bächen NAWA SPEZ 2017: Kleine Gewässer In Gebieten mit Intensiver Landwirtschaft Verbreitet Metroffen. *Aqua & Gas*, *4*, 14–25.

Stehle, S., & Schulz, R. (2015). Agricultural insecticides threaten surface waters at the global scale. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(18), 5750–5755. https://doi.org/10.1073/PNAS.1500232112

Suhaila, A. H., & Che Salmah, M. R. (2017). Application of aquatic insects (Ephemeroptera, plecoptera and trichoptera) in water quality assessment of Malaysian headwater. *Tropical Life Sciences Research*, *28*(2), 143–162. https://doi.org/10.21315/TLSR2017.28.2.11,

Švara, V., Krauss, M., Michalski, S. G., Altenburger, R., Brack, W., & Luckenbach, T. (2021). Chemical Pollution Levels in a River Explain Site-Specific Sensitivities to Micropollutants within a Genetically Homogeneous Population of Freshwater Amphipods. *Environmental Science & Technology*, *55*(9), 6087–6096. https://doi.org/10.1021/acs.est.0c07839

Swiss Centre for Applied Ecotoxicology. (2020). *Proposals for Quality Criteria for Surface Waters*. Quality Criteria for Surface Waters. https://www.ecotoxcentre.ch/expert-service/quality-criteria/quality-criteria-for-surface-waters

Tampo, L., Kaboré, I., Alhassan, E. H., Ouéda, A., Bawa, L. M., & Djaneye-Boundjou, G. (2021). Benthic Macroinvertebrates as Ecological Indicators: Their Sensitivity to the Water Quality and Human Disturbances in a Tropical River. *Frontiers in Water*, *3*, 662765. https://doi.org/10.3389/FRWA.2021.662765/BIBTEX

Tlili, A., Hollender, J., Kienle, C., & Behra, R. (2017). Micropollutant-induced tolerance of in situ periphyton: Establishing causality in wastewater-impacted streams. *Water Research*, *111*, 185–194. https://doi.org/10.1016/J.WATRES.2017.01.016

Tubić, B., Andjus, S., Zorić, K., Vasiljević, B., Jovičić, K., Čanak Atlagić, J., & Paunović, M. (2024). Aquatic Insects (Ephemeroptera, Plecoptera and Trichoptera) Metric as an Important Tool in Water Quality Assessment in Hilly and Mountain Streams. *Water*, *16*(6), 849. https://doi.org/10.3390/w16060849

Vári, Á., Podschun, S. A., Erős, T., Hein, T., Pataki, B., Iojă, I. C., Adamescu, C. M., Gerhardt, A., Gruber, T., Dedić, A., Ćirić, M., Gavrilović, B., & Báldi, A. (2022). Freshwater systems and ecosystem services: Challenges and chances for cross-fertilization of disciplines. *Ambio*, *51*(1), 135–151. https://doi.org/10.1007/S13280-021-01556-4/METRICS

Villeneuve, B., Piffady, J., Valette, L., Souchon, Y., & Usseglio-Polatera, P. (2018). Direct and indirect effects of multiple stressors on stream invertebrates across watershed, reach and site scales: A structural equation modelling better informing on hydromorphological impacts. *Science of The Total Environment*, *612*, 660–671. https://doi.org/10.1016/J.SCITOTENV.2017.08.197

von der Ohe, De Deckere, E., Prüß, A., Muñoz, I., Wolfram, G., Villagrasa, M., Ginebreda, A., Hein, M., & Brack, W. (2008). Toward an Integrated Assessment of the Ecological and Chemical Status of European River Basins. *Integrated Environmental Assessment and Management*, *5*(1), 50–61. https://academic.oup.com/ieam/article/5/1/50/7746398

von der Ohe, P., & de Zwart, D. (2013). Toxic Units (TU) Indicators. In J.-F. Férard & C. Blaise (Eds.), *Encyclopedia of Aquatic Ecotoxicology* (pp. 1162–1170). Springer Netherlands. https://doi.org/10.1007/978-94-007-5704-2

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature 2010 467:7315*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440

Wei, R., Escher, B. I., Glaser, C., König, M., Schlichting, R., Schmitt, M., Störiko, A., Viswanathan, M., & Zarfl, C. (2022). Modeling the Dynamics of Mixture Toxicity and Effects of Organic Micropollutants in a Small River under Unsteady Flow Conditions. *Environmental Science & Technology*, *56*(20), 14397–14408. https://doi.org/10.1021/acs.est.2c02824

Weisner, O., Arle, J., Liebmann, L., Link, M., Schäfer, R. B., Schneeweiss, A., Schreiner, V. C., Vormeier, P., & Liess, M. (2022). Three reasons why the Water Framework Directive (WFD) fails to identify pesticide risks. *Water Research*, *208*. https://doi.org/10.1016/j.watres.2021.117848

Wogram, J., & Liess, M. (2001). Rank Ordering of Macroinvertebrate Species Sensitivity to Toxic Compounds by Comparison with That of Daphnia magna. *Bulletin of Environmental Contamination and Toxicology*, *67*(3), 0360–0367. https://doi.org/10.1007/s00128-001-0133-8

Yeakley, J. A., Ervin, D., Chang, H., Granek, E. F., Dujon, V., Shandas, V., & Brown, D. (2016). Ecosystem Services of Streams and Rivers. In *River Science: Research and Management for the 21st Century* (pp. 335–352). Wiley Blackwell. https://doi.org/10.1002/9781118643525.CH17

# APPENDIX

## SUPPLEMENTARY FIGURES

## SUPPLEMENTARY TABLES



## SUPPLEMENTARY RESOURCE

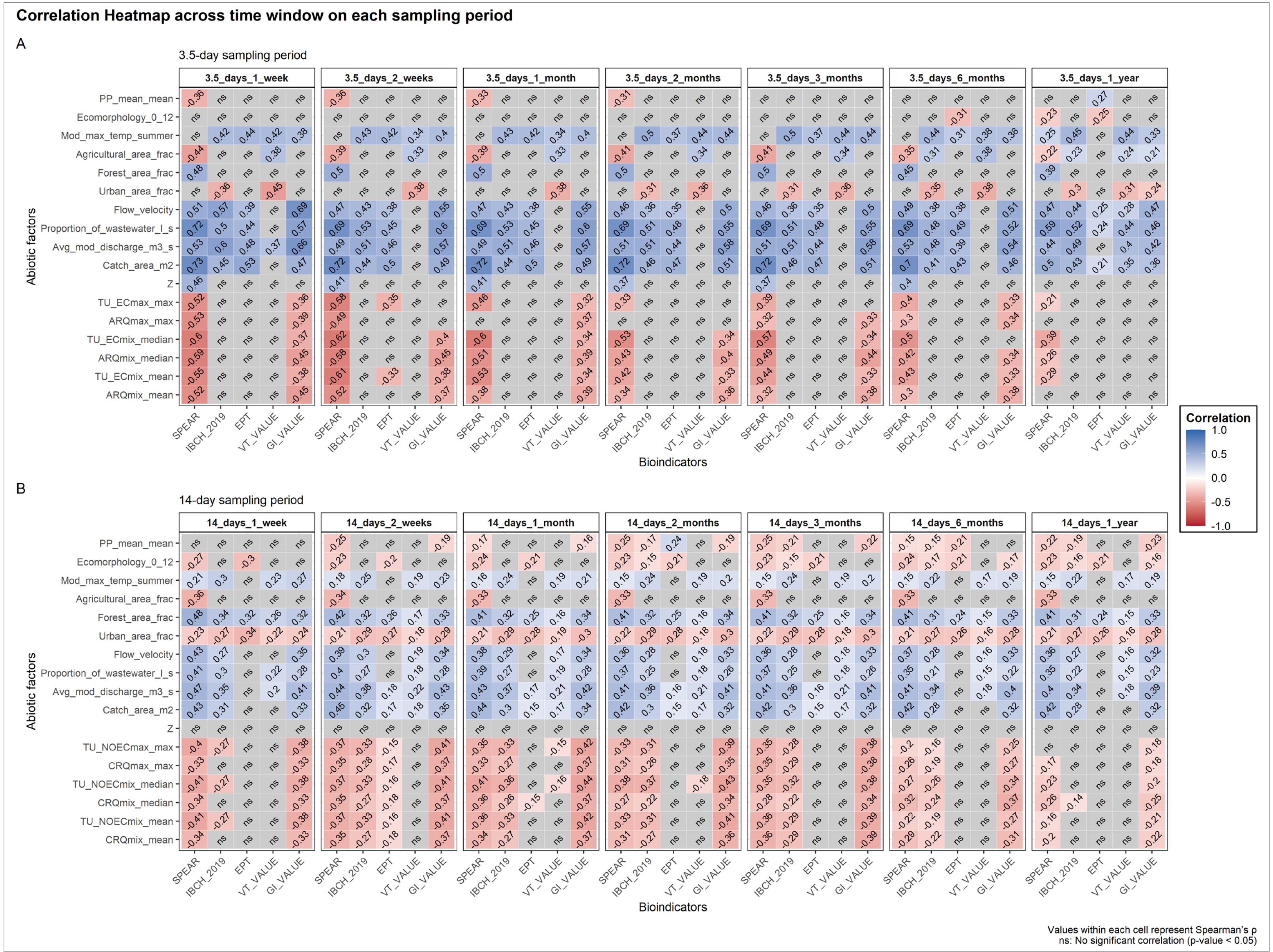


*Figure S 1. Correlogram Abiotic factors and Bioindicators (Chemical metrics represented by TUs and RQs (max, max) and (mix, median)) in both 3.5-days (A) and 14-days (B) sampling period across time windows.*

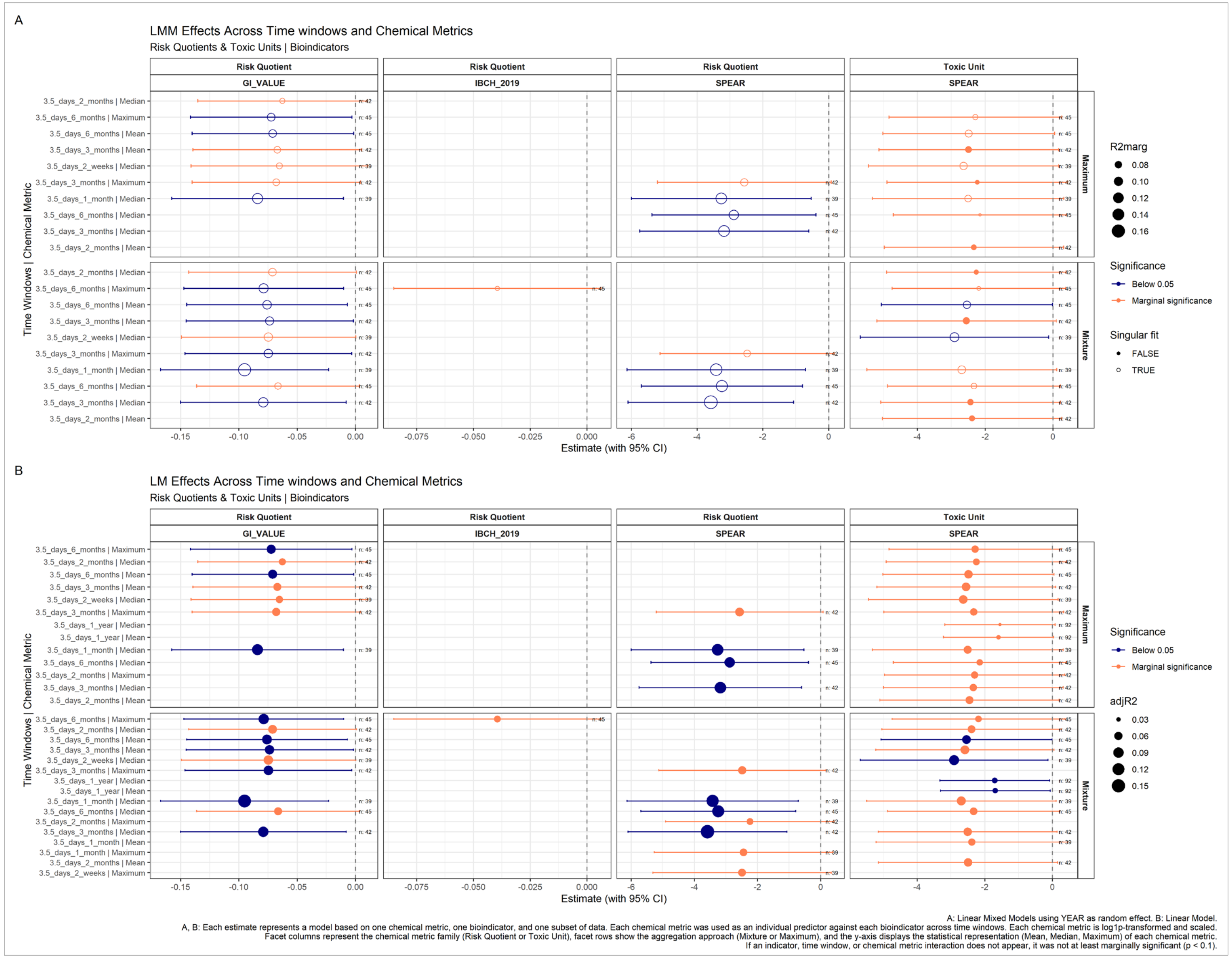


*Figure S 2. Estimates (95% CIs) of chemical metrics used as predictors in (A) Linear Mixed Models (including "year" as random effect) and (B) Linear Models against bioindicators.*

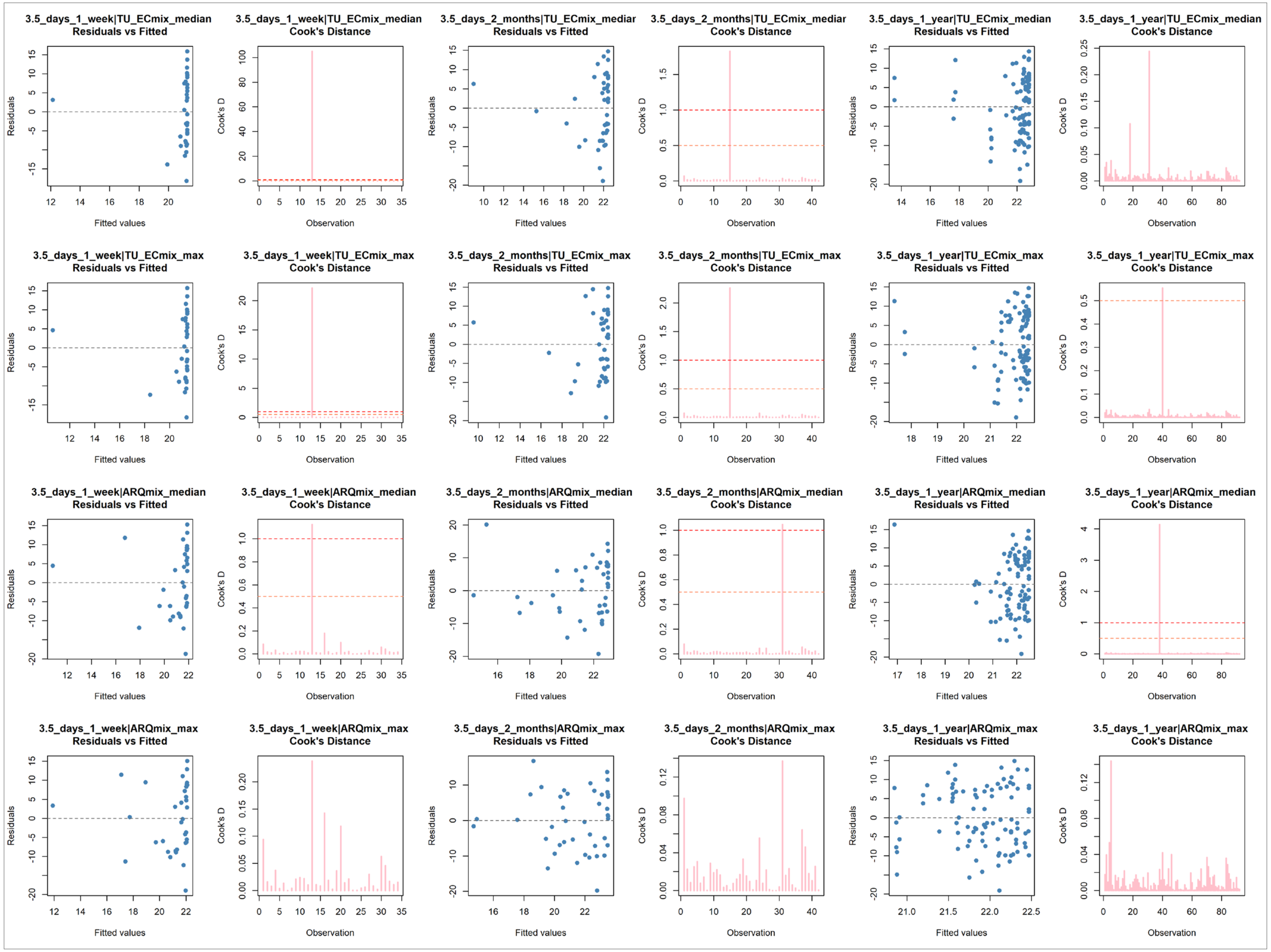


*Figure S 3. Residual vs. Fitted values and Cook's distance plots for SPEARpesticides (response variable) and chemical metrics (predictors) in 1-week, 2-week, and 1-year time windows for linear models in 3.5 days sampling period.*

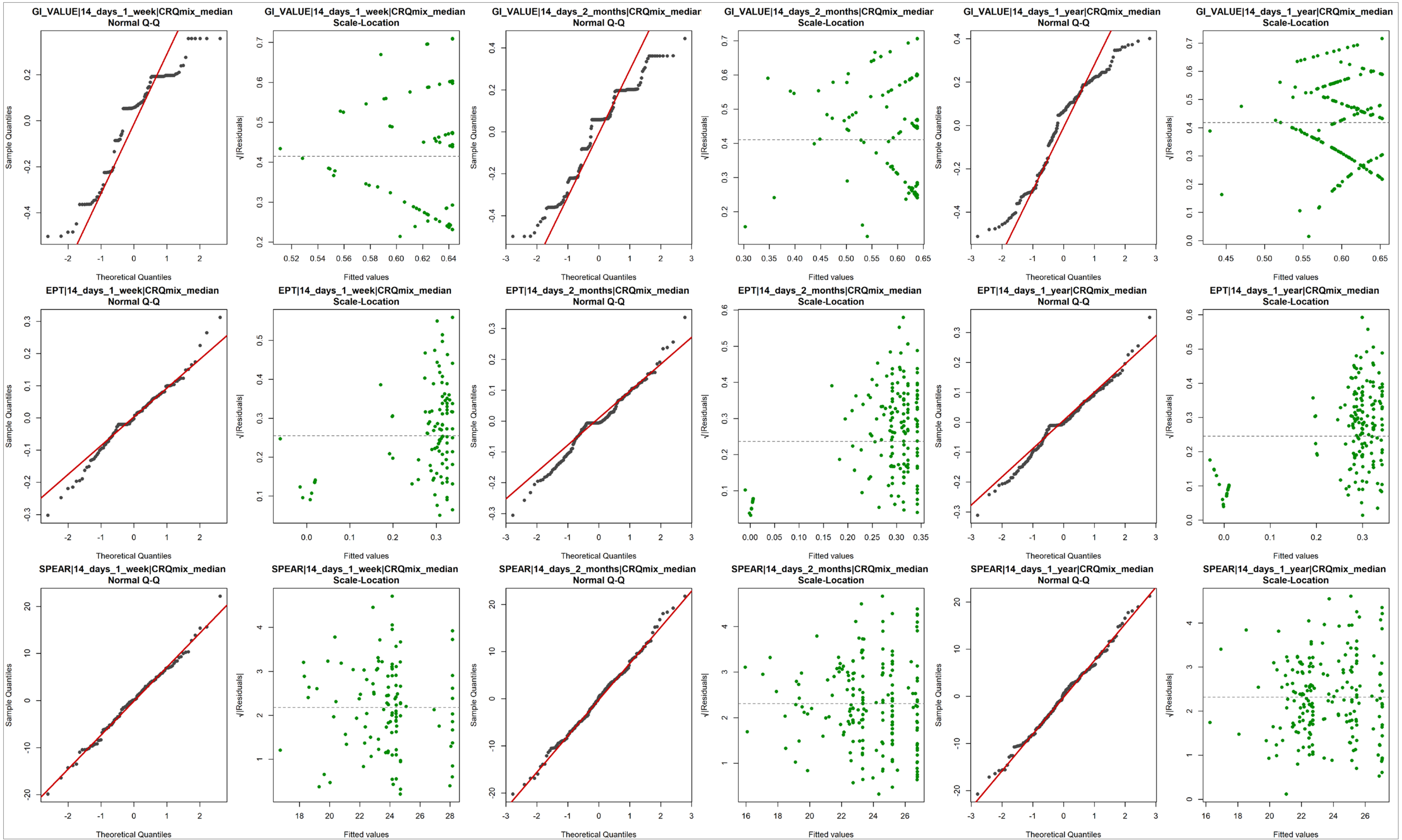


*Figure S 4. QQ and Scale-Location plots for GI, EPT, and SPEARpesticides (response variables) and chemical metrics (predictors) in 1-week, 2-week, and 1-year time windows for linear mixed models in 14 days sampling period.*

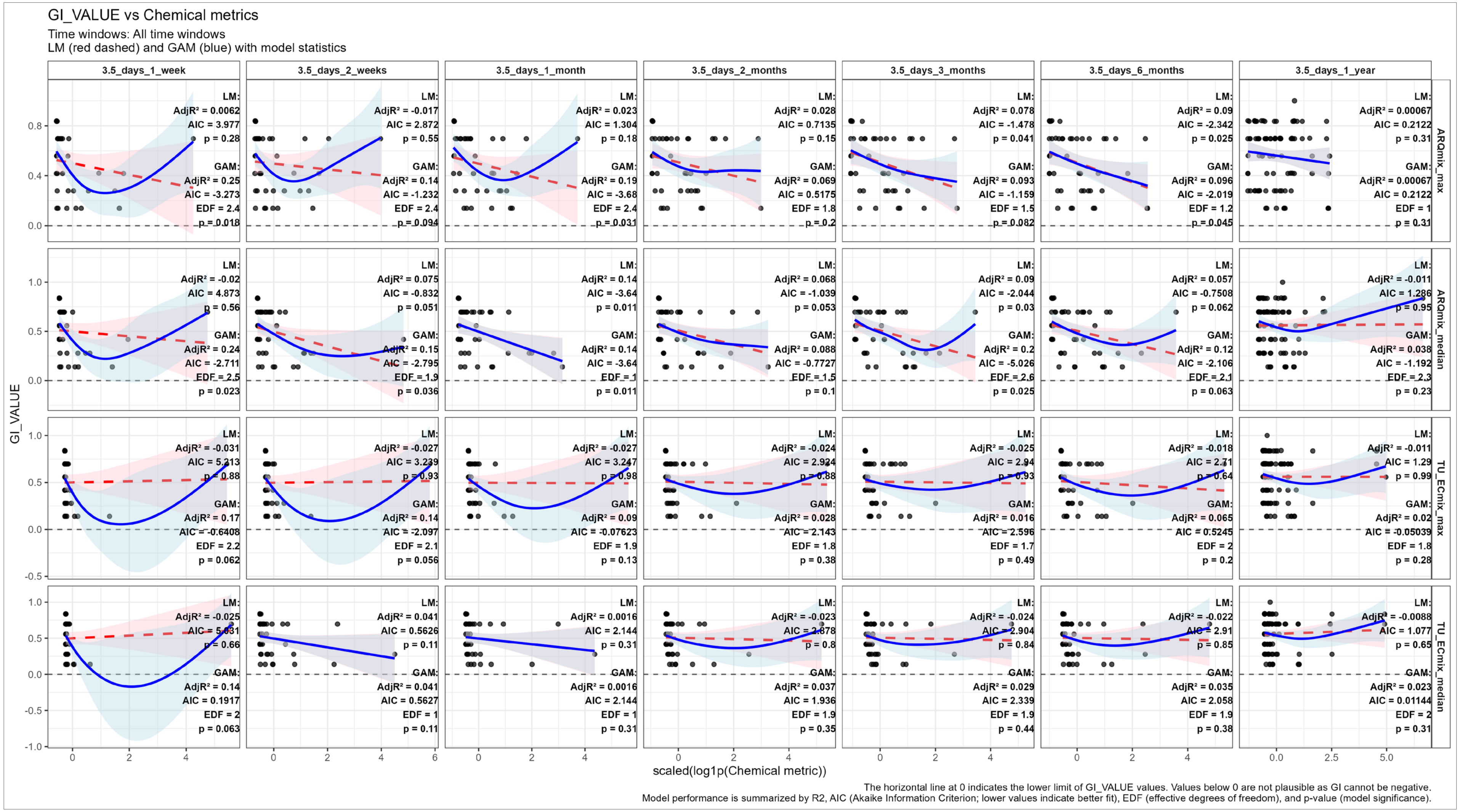


*Figure S 5. Comparison of Generalized Additive Models (blue) and Linear Models (red) fitted across different time windows using GI as response variable. When the estimates or uncertainties for the bioindicator fall below zero, it is considered an ecological implausible value.*

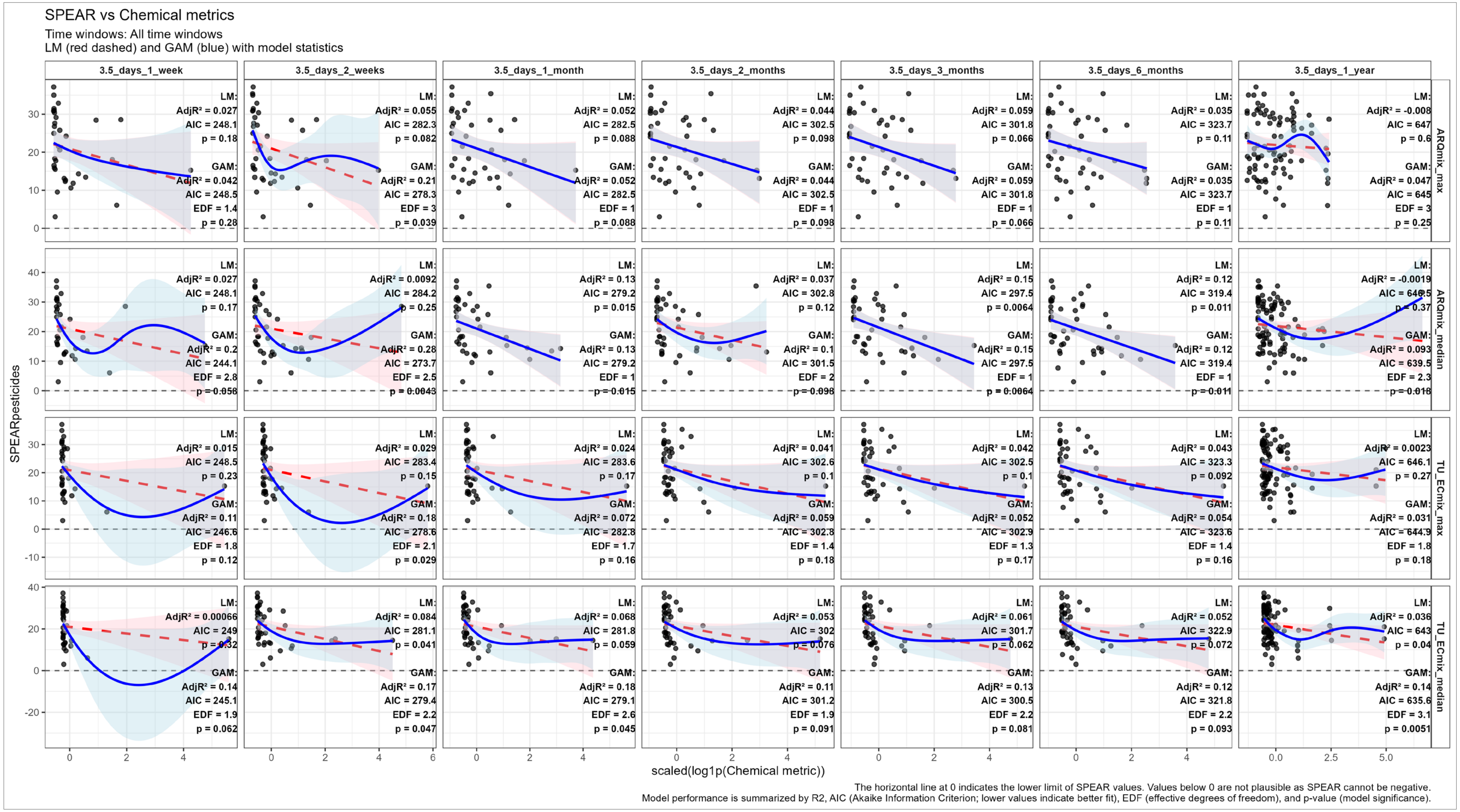


*Figure S 6. Comparison of Generalized Additive Models (blue) and Linear Models (red) fitted across different time windows using SPEARpesticides as response variable. When the estimates or uncertainties for the bioindicator fall below zero, it is considered an ecological implausible value.*

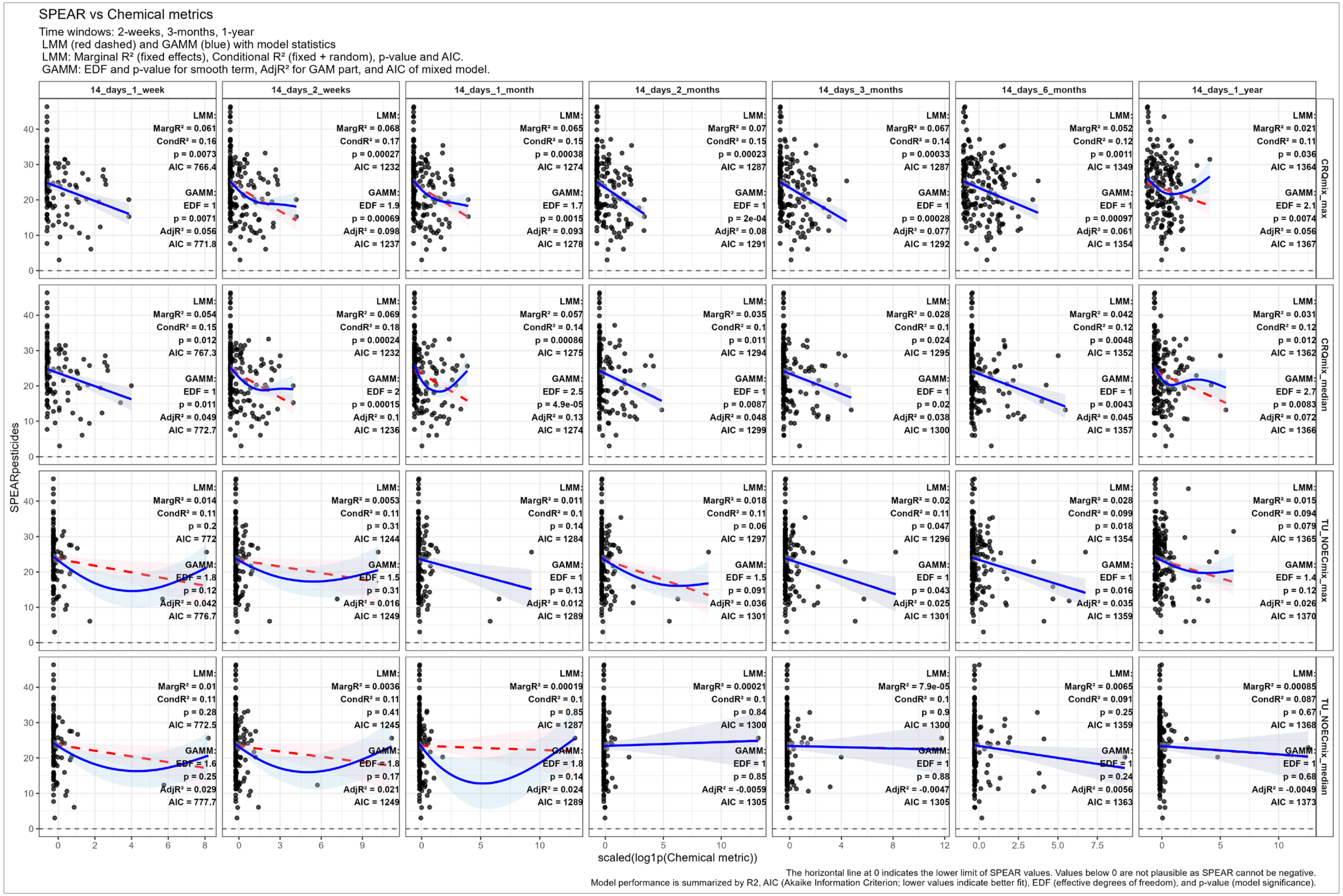


*Figure S 7. Comparison of Generalized Additive Mixed Models (blue) and Linear Mixed Models (red) fitted across different time windows using SPEARpesticides as the response variable.*

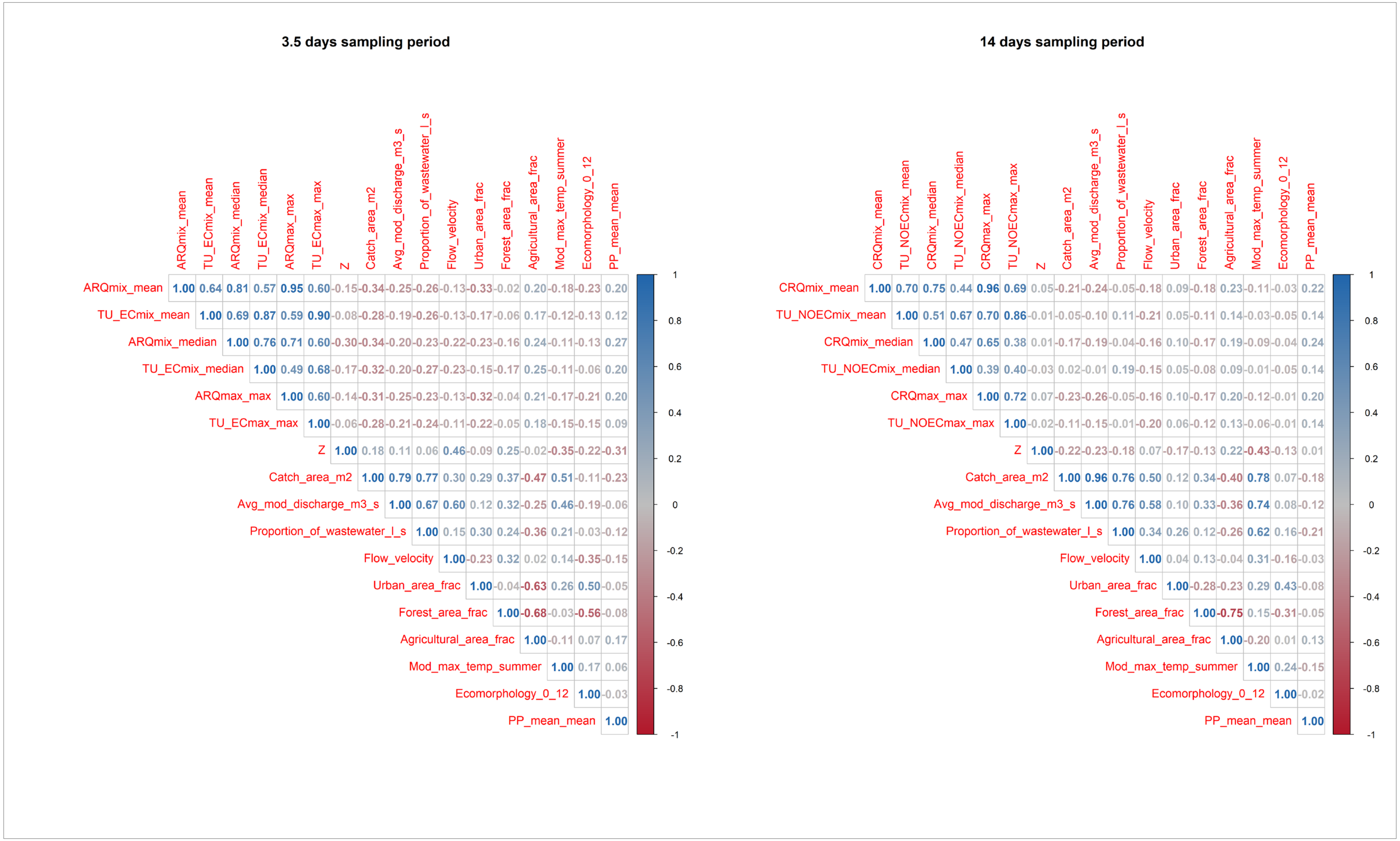


*Figure S 8. Correlogram of abiotic factors at 3.5-days and 14-days sampling periods. For practicality, only the 1-year dataset was used for this representation, as correlation values and directions were similar across time windows.*

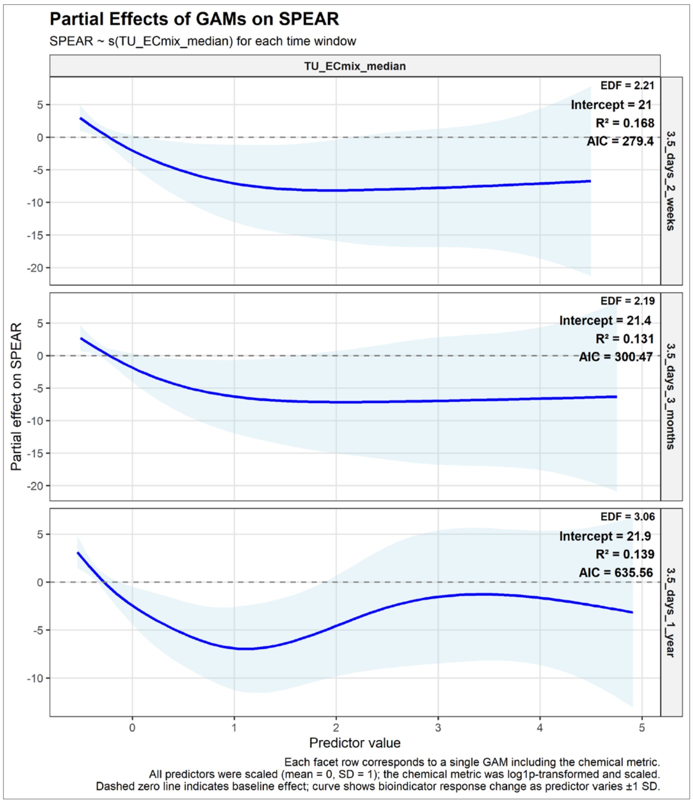


*Figure S 9. Partial effect of TU(mix, median) on SPEARpesticides across selected time windows using GAMs.*

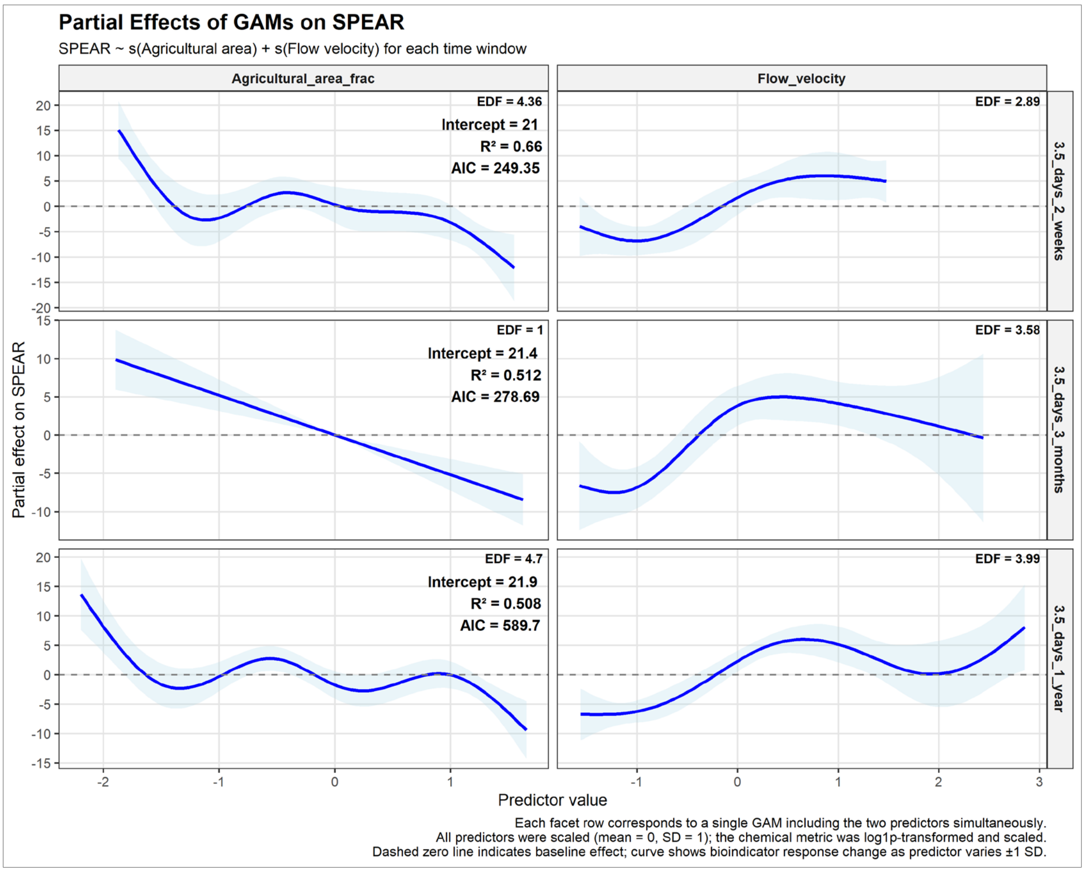


*Figure S 10. Partial effects of proportion of agricultural land use and flow velocity on SPEARpesticides across selected time windows using GAMs.*

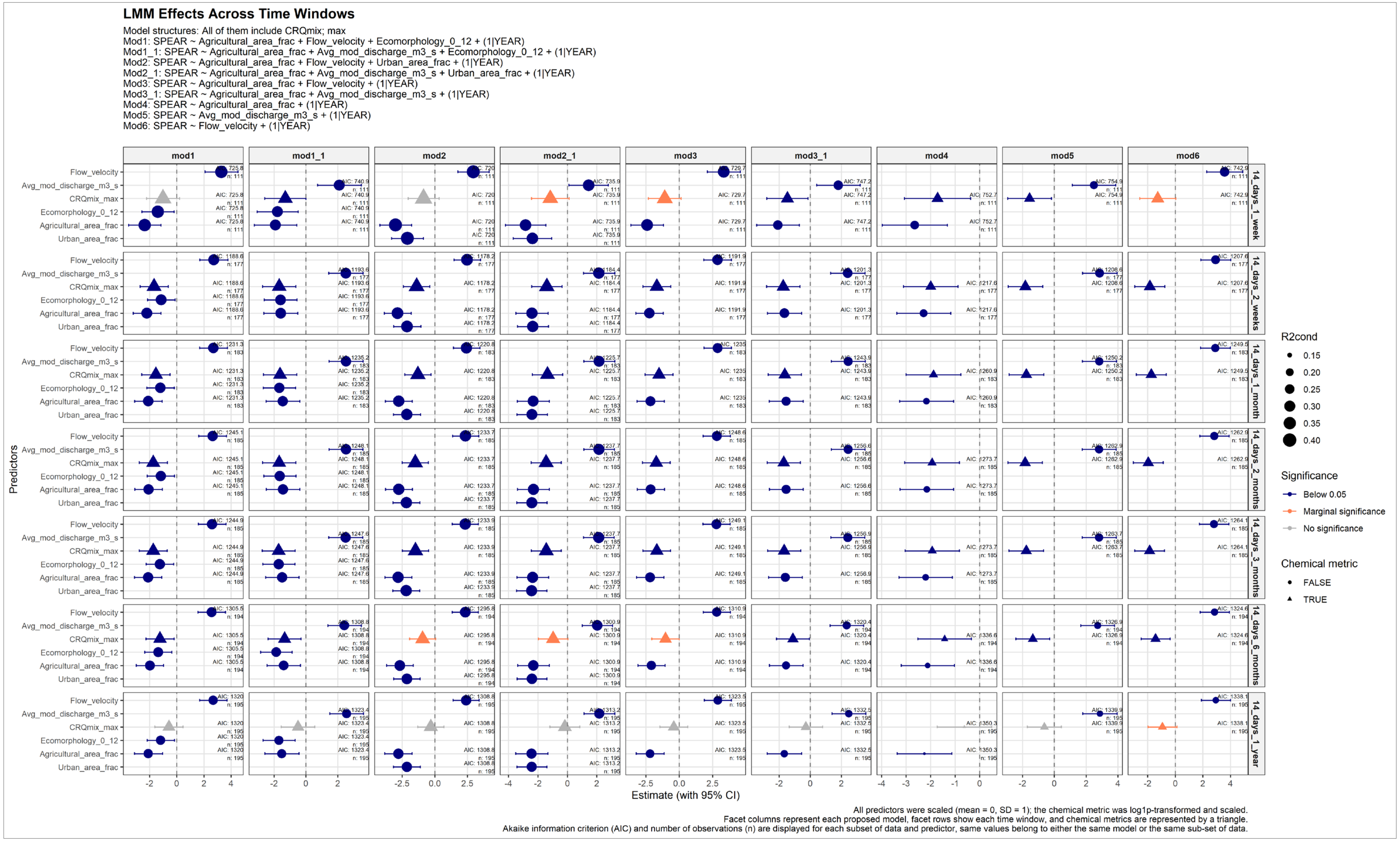


*Figure S 11. Comparison of nine linear mixed models fitted to 14-days sampling data. Model 2 was selected for further analysis and is described in the main text. Marginal significance was set as p-value < 0.1.*

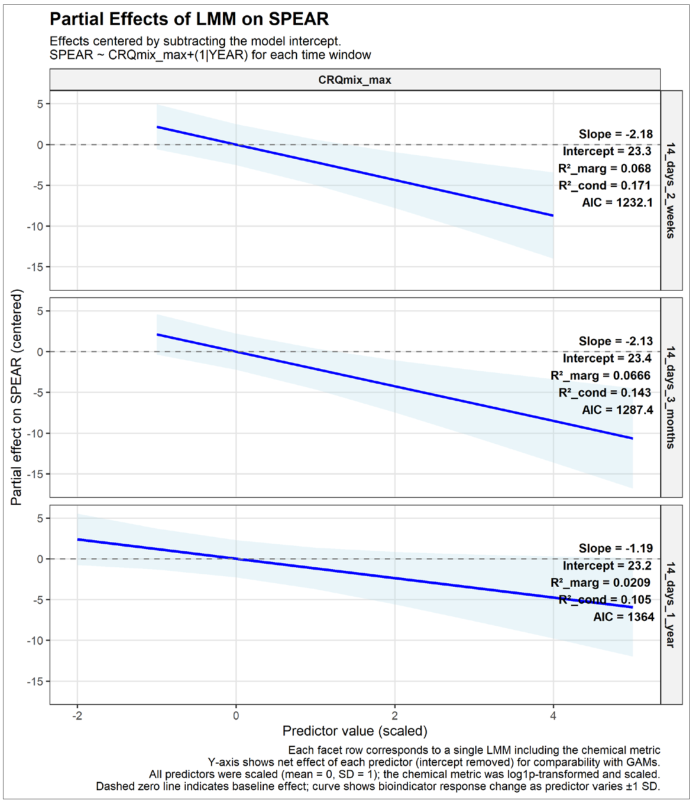


*Figure S 12. Partial effects of CRQ(mix, max) on SPEARpesticides across selected time windows using LMMs.*

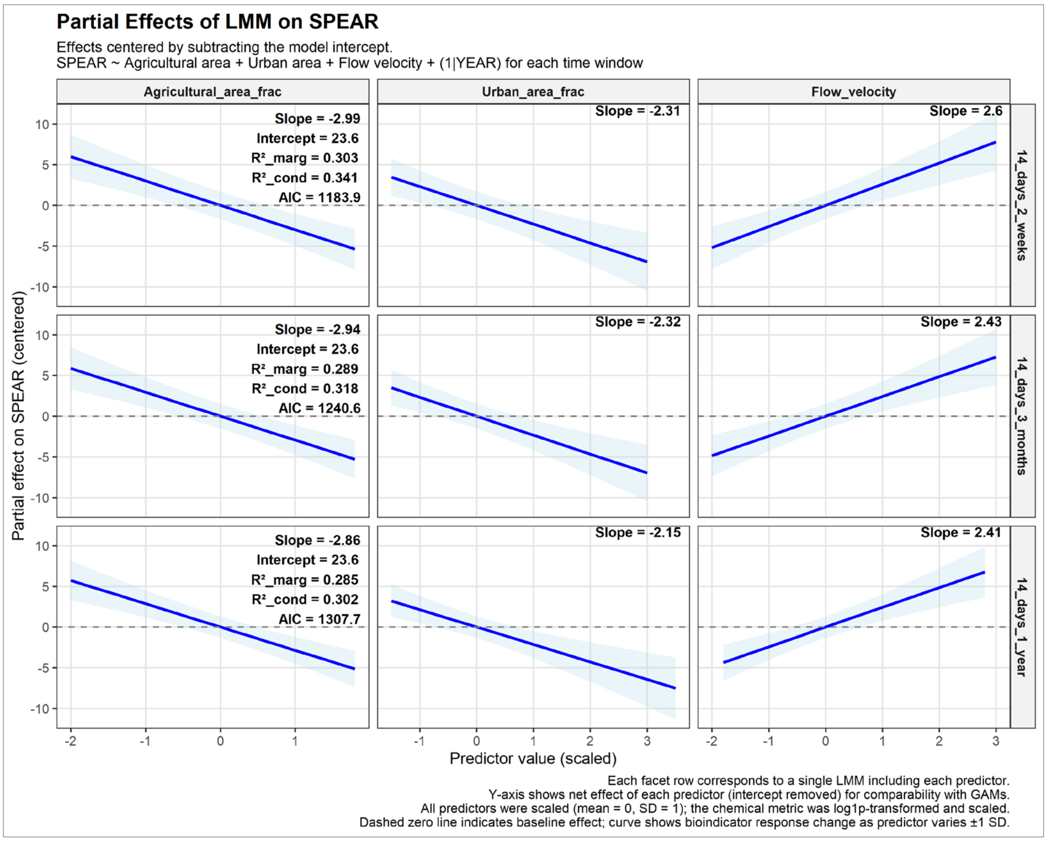


*Figure S 13. Partial effects of agricultural and urban land proportion and flow velocity on SPEARpesticides across selected time windows using LMMs.*

*Table S 1. Description of the abiotic variables included in the study (chemical metrics and environmental variables).*

| VARIABLE | DESCRIPTION | DIMENSION | TYPE | REFERENCES* |
|---|---|---|---|---|
| **Chemical metrics** | Risk Quotients (RQs) and Toxic Units (TUs) based on measured concentration of pesticides (insecticides, herbicides, and fungicides), Swiss ecotoxicological quality standards and ecotoxicological endpoints, calculated as 'maximum' and 'mixture' values, and represented by (mean, median, max) per time window. | Chemical | Dynamic | (Burdon et al., 2019; Liess et al., 2021) |
| **Catchment area ($m^2$)** | Estimation of the total area (in $m^2$) of water bodies (including lakes) upstream the sampling point. | Hydrological | Static | (BAFU, 2020; Hutter et al., 2019; Ilg & Alther, 2024) |
| **Average discharge ($m^3/s$)** | Annual mean modelled discharge (in $m^3/s$) for the watercourse 'Mittlerer modellierter Abfluss'. | Hydrological | Static | (BAFU, 2020; Burdon et al., 2016) |
| **Flow velocity (m/s)** | Modelled annual flow velocity (based on stream width and slope). | Hydrological | Static | (Khaliq et al., 2024) |
| **Stream classification** | Waterbody size classification (small 'kleines', medium 'mittel', large 'grosses') based on discharge. | Hydrological | Static | (Ilg & Alther, 2024) |
| **Proportion of urban land use** | Proportion of the total land upstream categorized as human settlements 'Siedlungsflächen.' | Land Use | Static | (Hutter et al., 2019; Ilg & Alther, 2024; Khaliq et al., 2024; Liess et al., 2021) |
| **Proportion of forest land use** | Proportion of the total land upstream categorized as forested areas, unproductive areas, artificial meadows, permanent meadows, permanent pastures, and summer grazing area. | Land Use | Static | |
| **Proportion of agricultural land use** | Proportion of the total land upstream categorized as other fruit crops, fruit crops aggregated, perennial berries, stone fruit, pears, apples, open area for production, vines, vineyards with natural biodiversity, and vines (region-specific biodiversity areas). | Land Use | Static | |
| **Ecomorphology** | Classification of the stream based on water body morphology, ranging from 0 (natural or near natural) to 12 (unnatural or artificial). | Morphological | Static | (BAFU, 2022; Hutter et al., 2019; Ilg & Alther, 2024) |
| **Altitude (m.a.s.l.)** | Measure of the altitude (m.a.s.l.) at the sampling point. | Spatial | Static | |
| **Stream identity** | Stream name where macroinvertebrate and chemical sampling occurred. | Spatial | Static | - |

| | | | | |
|---|---|---|---|---|
| **Maximum summer water temperature (°C)** | Modelled maximum morning summer stream temperature predicted from a linear model based on catchment area and mean catchment elevation. | Spatial | Static | (Khaliq et al., 2024) |
| **Season** | Season (spring or summer) when macroinvertebrate monitoring was performed. | Temporal | Static | (Ilg & Alther, 2024) |
| **Year** | Year of macroinvertebrate and chemical sampling. | Temporal | Static | (Khaliq et al., 2024) |
| **Average of daily precipitation (mm)** | Average of daily precipitation (mm) per site from 01-01-2017 to 31-12-2023, separated by chemical sample | Temporal | Dynamic | - |

**References that have used the same or equivalent (based on our data availability) abiotic factors to be related to the bioindicators in previous studies*

*Table S 2. Summary table of fitted GAMs across time windows in the 3.5-days sampling period.*

| | 1_week | 2_weeks | 1_month | 2_months | 3_months | 6_months | 1_year |
|---|---|---|---|---|---|---|---|
| ***Model 1*: SPEARpesticides - TU_ECmix_median + Agriculture proportion+ Flow velocity**** | | | | | | | |
| **AIC** | 222.405 | 255.692 | 254.655 | 274.785 | 273.995 | 278.718 | 600.699 |
| **R2** | 0.617 | 0.594 | 0.609 | 0.578 | 0.587 | 0.705 | 0.441 |
| **VIF** | 1.026 | 1.027 | 1.03 | 1.036 | 1.04 | 1.034 | 1.024 |
| **EDF** | | | | | | | |
| **edf_s(TU_ECmix_median)** | 1.786 | 1.752 | 2.164 | 1.873 | 1.868 | 1 | 2.535 |
| **edf_s(Agricultural_area_frac** | 1 | 1 | 1 | 1 | 1 | 4.539 | 1 |
| **edf_s(Flow_velocity)** | 3.47 | 3.393 | 3.299 | 3.718 | 3.733 | 3.578 | 3.719 |
| **F-statistic** | | | | | | | |
| **F_s(TU_ECmix_median)** | 1.87 | 3.198 | 3.49 | 2.919 | 3.226 | 2.948 | 4.848 |
| **F_s(Agricultural_area_frac)** | 22.613 | 23.248 | 23.895 | 24.981 | 24.832 | 11.436 | 14.68 |
| **F_s(Flow_velocity)** | 6.269 | 7.055 | 7.493 | 7.438 | 7.664 | 10.627 | 11.432 |
| **Significance (p_value)** | | | | | | | |
| **p_s(TU_ECmix_median)** | 0.18 | 0.0557 | 0.0434 | 0.0702 | 0.0523 | 0.0948 | 0.00351 |
| **p_s(Agricultural_area_frac)** | 5.96E-05 | 3.34E-05 | 2.75E-05 | 1.74E-05 | 1.82E-05 | 0.000007 | 0.000245 |
| **p_s(Flow_velocity)** | 0.00108 | 0.000332 | 0.000246 | 0.000209 | 0.000159 | 9.6E-06 | 0 |
| ***Model 2: SPEARpesticides - Agriculture proportion+ Flow velocity*** | | | | | | | |
| **AIC** | 211.617 | 249.352 | 249.352 | 278.686 | 278.686 | 279.87 | 589.698 |
| **R2** | 0.726 | 0.66 | 0.66 | 0.512 | 0.512 | 0.692 | 0.508 |
| **VIF** | 1.026 | 1.027 | 1.03 | 1.036 | 1.04 | 1.034 | 1.024 |
| **EDF** | | | | | | | |
| **edf_s(Agricultural_area_frac** | 4.484 | 4.36 | 4.36 | 1 | 1 | 4.651 | 4.704 |
| **edf_s(Flow_velocity)** | 2.772 | 2.889 | 2.889 | 3.583 | 3.583 | 3.384 | 3.986 |
| **F-statistic** | | | | | | | |
| **F_s(Agricultural_area_frac)** | 9.287 | 9.14 | 9.14 | 24.386 | 24.386 | 12.235 | 9.28 |
| **F_s(Flow_velocity)** | 8.952 | 9.208 | 9.208 | 6.637 | 6.637 | 9.634 | 13.169 |
| **Significance (p_value)** | | | | | | | |
| **p_s(Agricultural_area_frac)** | 0.000033 | 6.32E-05 | 6.32E-05 | 1.83E-05 | 1.83E-05 | 1.11E-06 | 2.35E-06 |
| **p_s(Flow_velocity)** | 0.000287 | 0.000153 | 0.000153 | 0.000462 | 0.000462 | 3.03E-05 | 0 |
| ***Model 3: SPEARpesticides - ARQmix_median + Agriculture proportion+ Flow velocity*** | | | | | | | |
| **AIC** | 212.585 | 250.513 | 252.353 | 276.88 | 279.798 | 282.104 | 590.491 |
| **R2** | 0.723 | 0.67 | 0.639 | 0.562 | 0.506 | 0.682 | 0.517 |
| **VIF** | 1.03 | 1.006 | 1.095 | 1.127 | 1.159 | 1.132 | 1.005 |
| **EDF** | | | | | | | |
| **edf_s(ARQmix_median)** | 1 | 2.515 | 1 | 2.207 | 1 | 1 | 2.171 |
| **edf_s(Agricultural_area_frac** | 4.487 | 3.972 | 4.198 | 1 | 1 | 4.579 | 4.646 |
| **edf_s(Flow_velocity)** | 2.665 | 2.4 | 2.766 | 3.602 | 3.249 | 3.295 | 3.703 |
| **F-statistic** | | | | | | | |
| **F_s(ARQmix_median)** | 1.106 | 2.475 | 0.463 | 2.005 | 1.86 | 0.708 | 1.562 |
| **F_s(Agricultural_area_frac)** | 8.572 | 7.21 | 8.117 | 24.964 | 18.455 | 9.229 | 6.762 |
| **F_s(Flow_velocity)** | 8.951 | 8.514 | 7.741 | 6.642 | 5.267 | 9.082 | 12.775 |
| **Significance (p_value)** | | | | | | | |
| **p_s(ARQmix_median)** | 0.303 | 0.0779 | 0.501 | 0.154 | 0.181 | 0.406 | 0.157 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **p_s(Agricultural_area_frac)** | 6.36E-05 | 0.000518 | 0.000257 | 1.75E-05 | 0.000126 | 9.66E-06 | 1.88E-05 |
| **p_s(Flow_velocity)** | 0.000333 | 0.000999 | 0.000463 | 0.000361 | 0.00203 | 6.24E-05 | 0 |
| ***Model 4: SPEARpesticides - TU_ECmix_median + Ecomorphology + Agriculture + Flow*** | | | | | | | |
| **AIC** | 223.786 | 257.849 | 256.799 | 276.883 | 264.754 | 295.774 | 584.082 |
| **R2** | 0.623 | 0.586 | 0.599 | 0.564 | 0.685 | 0.571 | 0.561 |
| **VIF** | 1.08 | 1.085 | 1.088 | 1.069 | 1.07 | 1.072 | 1.035 |
| **EDF** | | | | | | | |
| **edf_s(TU_ECmix_median)** | 1.736 | 1.863 | 2.153 | 1.847 | 1 | 1.509 | 1.536 |
| **edf_s(Ecomorphology_0_12)** | 1.709 | 1.252 | 1.227 | 1 | 1 | 1.654 | 2.241 |
| **edf_s(Agricultural_area_frac** | 1 | 1 | 1 | 1 | 4.301 | 1 | 4.584 |
| **edf_s(Flow_velocity)** | 3.481 | 3.287 | 3.265 | 3.665 | 3.683 | 3.969 | 4.082 |
| **F-statistic** | | | | | | | |
| **F_s(TU_ECmix_median)** | 1.47 | 3.173 | 3.443 | 2.852 | 3.04 | 2.59 | 2.828 |
| **F_s(Ecomorphology_0_12)** | 0.658 | 0.067 | 0.057 | 0.011 | 0.52 | 0.522 | 1.928 |
| **F_s(Agricultural_area_frac)** | 23.251 | 21.11 | 22.536 | 23.686 | 10.514 | 22.676 | 7.263 |
| **F_s(Flow_velocity)** | 5.592 | 6.089 | 6.228 | 6.342 | 7.797 | 8.185 | 11.951 |
| **Significance (p_value)** | | | | | | | |
| **p_s(TU_ECmix_median)** | 0.215 | 0.0579 | 0.0485 | 0.0755 | 0.0911 | 0.0631 | 0.0495 |
| **p_s(Ecomorphology_0_12)** | 0.572 | 0.845 | 0.917 | 0.917 | 0.476 | 0.623 | 0.143 |
| **p_s(Agricultural_area_frac)** | 6.02E-05 | 6.86E-05 | 4.78E-05 | 2.73E-05 | 7.12E-05 | 3.11E-05 | 3.63E-05 |
| **p_s(Flow_velocity)** | 0.00222 | 0.0017 | 0.000868 | 0.000811 | 0.00023 | 0.000105 | 0 |

**Model described in the thesis, selected as the best GAM.*

### Link 1. Public Repository Structure and Access

The public repository associated with this thesis is available at: [GitHub: SPEAR PUBLIC 2025] - https://github.com/AnthonyFow/SPEAR-THESIS-2025_Public

Repository structure:

- /inputs
  - Abiotics_factors_41_streams_ds_prepared.rds
  - metrics_3.5d_0_R12_pp_ds.rds
  - metrics_14d_0_R12_pp_ds.rds
- /scripts
  - bio_chem_analysis.R
  - datasets_creation.R
- /output
- /graphics
  - (Main thesis plots + user-generated plots)
- /docs
  - (Detailed variable/column documentation)
- /README.md

Important notes:

- The datasets provided are entirely synthetic, randomly generated to closely mimic the structure, column names, and approximate number of observations and streams of the original data.
- These datasets do **NOT** contain real observations and cannot be used to reproduce the exact results of the thesis.
- All code for statistical modeling and figure generation is provided and fully reproducible with the synthetic data.
- Full variable documentation is available in the /docs folder.
- **This repository contains the primary analysis script used for the thesis. It is published to promote academic and scientific transparency.**
- For access to the original data or data preparation scripts, contact the thesis supervisor (Dr. Anne Dietzel, VSA; anne.dietzel@vsa.ch).

## THE PRESENT RESEARCH WORK HAS BEEN (PREPARED TO BE) PUBLISHED/PRESENTED IN

**This work is being prepared to be published in:**

The thesis work is scheduled to be presented at the SETAC 13$^{th}$ Young Environmental Scientists Meeting 2025, which will be held in August 2025 at the University of York, UK.

In addition, a summarized version of the work will be translated into German and published in the local technical journal “Aqua & Gas”.

Furthermore, an extended version of the thesis is being prepared for submission to an international peer-reviewed journal, potentially including Water Research (Elsevier), Freshwater Biology (Wiley) or Science of the Total Environment (Elsevier).

**The results from the recommendations will also be included in the paper.**

**PLENTZIA (UPV/EHU), JULY 2025**