Improving risk assessment by predicting the survival of field gammarids exposed to dynamic pesticide mixtures

Maricor J. Arlos\textsuperscript{a,b,c,*}, Andreas Focks\textsuperscript{d}, Juliane Hollender\textsuperscript{b,d}, and Christian Stamm\textsuperscript{b*}

\textsuperscript{a}Department of Civil and Environmental Engineering, University of Alberta, 9211-116 St. NW, Edmonton, Alberta, T6G 1H9, Canada

\textsuperscript{b}Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

\textsuperscript{c}Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, 8092 Zürich, Switzerland

\textsuperscript{d}Wageningen Environmental Research, P.O. Box 47, 6700 AA, Wageningen, The Netherlands

*Corresponding authors:
Maricor Arlos/Christian Stamm

Email: maricor.arlos@ualberta.ca; christian.stamm@eawag.ch

Keywords:
Toxicokinetics-toxicodynamics, survival, pesticides, GUTS, invertebrate, environmental quality standards
Abstract

Exposure assessment of pesticides has substantially improved over time, with methods that now include a combination of advanced analytical techniques and fate/transport models to evaluate their spatio-temporal distribution. However, current regulatory environmental risk assessment considers thresholds from laboratory studies completed under standardized conditions that do not reflect environmental dynamics. Using the GUTS model framework, we predicted the impact of time-varying pesticide exposures on the survival of gammarids in a small agricultural stream. LP50 values were used as an additional metric for assessing risks (defined in GUTS as the multiplication factor applied to the concentration time series to induce 50% mortality by the end of exposure). Although real-case exposures to individual pesticides were predicted to produce little to no impact on survival, the LP50 values indicate acute (LP50≤100) and/or chronic (LP50≤10) toxicities for azoxystrobin, chlorpyrifos, diazinon, and imidacloprid, while risk to propiconazole exposure was considered very low (LP50≫100). Finally, the model was extended to reflect mixture toxicity via concentration addition. It predicted risks under acute and chronic exposures to organophosphates and neonicotinoids. Given that gammarids are simultaneously exposed to multiple chemicals and other stressors throughout their lifetime, a decline in survival probabilities due to chemical stress can likely influence their overall fitness. We recognize that some assumptions require validation, but our work included a level of realism that can assist risk managers when evaluating the cumulative consequences of chemical exposure.
Graphical Abstract

How do dynamic field exposures to pesticides and their mixtures impact gammarids survival over time?
1. Introduction

Pesticide use in agriculture has offered a wide array of benefits, including improvements in crop production (i.e. reduce crop failure) and provision of fresh goods that are available to consumers at reasonable rates. However, the use of these chemicals has been associated with several environmental issues, including biodiversity loss in terrestrial and aquatic environments. Therefore, many jurisdictions have established regulatory standards to limit the unintended consequences of pesticide exposure. These regulations are typically based on environmental risk assessment that incorporate information on the fate and transport (exposure) and ecotoxicological effects of a chemical on exposed organisms.

Over the last decade, tremendous advances in analytical methodology and instrumentation improved the acquisition of reliable exposure data in various environmental matrices. Together with well-established environmental models, the occurrence and distribution of pesticides and other trace organics have been assessed in many watersheds around the globe. However, the effects data remained focused on laboratory exposure studies that are typically completed under standardized conditions (i.e., single contaminant, constant concentration, fixed exposure duration, and a few model organisms). Assessment factors are then employed to address some of the uncertainties in lab exposures, including species specificity and lab-to-field extrapolation. However, the assessment factor values are somewhat arbitrary and do not explicitly account for the dynamics in exposure conditions that have been made available via measured (i.e., analytical detections) or predicted (i.e., fate/transport models) data sets.

Merging toxicokinetics (TK) and toxicodynamics (TD) models was suggested as a potential tool that can improve the environmental risk assessment of trace organic chemicals. TK describes the time course of concentration within the whole organism or at the target site of
action resulting from exposure to fluctuating levels in the environment, and TD describes the
damage dynamics associated with this exposure\textsuperscript{9}. Combining these two "modules" can offer the
possibility of linking the impacts of exposure to relevant effects endpoints over time\textsuperscript{3}, and this
approach is particularly advantageous for assessing the effects of pesticide exposures in
receiving environments. Pesticide concentrations are especially more dynamic than wastewater-
derived substances (e.g., pharmaceuticals and personal care products) due to a significant
interplay between climatological conditions and agricultural management practices that impact
their fate and transport in the environment\textsuperscript{10}.

Despite the advantage of TKTD modelling, it continues to be underutilized within the
context of environmental risk assessment\textsuperscript{5} due to several reasons. As indicated earlier,
ecotoxicological assessments are firmly rooted in standardized lab studies, and the use of
mechanistic modelling within this field is not yet a common practice. In fact, guidance
frameworks associated with mechanistic effects modelling have just been made available. For
example, the European Food and Safety (EFSA) opinion on good practices in effects modelling
was only released in 2014\textsuperscript{11}, and their opinion specific to TKTD as an effective tool for
regulatory risk assessment of pesticides in aquatic organisms was provided in 2018\textsuperscript{12}. Also, the
use of TKTD models requires calibration for each compound-species combination, and until
recently, a software with efficient and standardized calibration technique was not available.
Hence, TKTD parameters are scarce compared to the fate and transport parameters associated
with chemical exposure (e.g., biodegradation/photolysis rate constants, sorption coefficients).
Similar to the way fate and transport models have evolved and gained their place as a useful tool
in risk assessment, it is logical that we start employing mechanistic effects modelling (such as
TKTD) to obtain its true benefits in ecotoxicology\textsuperscript{5,7,13}. 
In this study, we evaluated the suitability of a TKTD model, specifically the General Unified Threshold model for Survival (GUTS)\textsuperscript{14}, to address the impacts of dynamic pesticide exposure on the survival of gammarids in the field. We explored the impact of substances that have been known to contribute to "unsatisfactory" or "bad" water quality conditions based on the acute risk quotient profiles reported in small agricultural streams in Switzerland\textsuperscript{15}. The list includes the neonicotinoid insecticides imidacloprid and thiacloprid; the organophosphate insecticides diazinon, chlorpyrifos, and chlorpyrifos-methyl; and the fungicides azoxystrobin and propiconazole. Since there are no survival data in the field to validate the simulations, we explored the predictions based on different survival model representations and GUTS parameter sets derived from various laboratory studies. We further hypothesized that the simultaneous exposure to multiple substances would exacerbate the survival of gammarids; hence, we extended the model to assess the impacts of pesticide mixtures with the same toxic mode of action using a simple application of the concentration addition concept. Finally, we compared the model predictions to current water quality standards for aquatic ecosystem protection in Switzerland and in Europe.

2. Materials and Methods

2.1. Exposure data and compound selection

The availability of well-resolved concentration profiles is an essential requirement when predicting impacts of pesticide exposure on the survival of gammarids in the field. In 2015, the exposure profiles of pesticides in five small streams (mainly agricultural) were assessed from March to August based on time-proportional composite samples collected every 12 hours\textsuperscript{15}. Out of these streams, Eschelisbach was selected as a catchment representing intensive agriculture,
and we have also observed a higher detection frequency of our target compounds compared to other sites.

The monitoring program analyzed >200 substances for target screening. However, it appeared that stream concentrations of diazinon, chlorpyrifos, chlorpyrifos-methyl, thiacloprid, nicosulfuron, and azoxystrobin showed potential adverse effects on aquatic invertebrates at Eschelisbach. Conveniently, most of these compounds have raw gammarids survival data available that were employed to calibrate the GUTS parameters (reference studies listed in Table S1). Chlorpyrifos-methyl and thiacloprid did not have their unique survival data, but information for other compounds with the same toxic mode of action was available. No survival data is available for nicosulfuron nor other compounds with a similar toxic mode of action and was therefore not included in the study. Finally, we included propiconazole due to the availability of raw survival data for gammarids. It also provides additional insights into our analyses by confirming that the environmental concentrations of propiconazole were indeed unlikely to cause adverse effects in invertebrates as what was initially determined for Eschelisbach.

2.2. **GUTS model – a brief background**

The GUTS model predicts the survival of an organism after accruing and recovering from the damage (i.e., TD) due to the bioaccumulation, distribution, biotransformation, and elimination (i.e., TK) of chemicals in the organism. Its mathematical framework was first published in 2011 and exemplifies the two major assumptions employed in survival modelling, the (1) stochastic death (SD) and individual tolerance (IT) concepts. The SD assumption suggests that individuals are identical and have a chance to die upon chemical stress, and this probability increases with increasing internal exposure once some threshold damage has been exceeded. The IT approach assumes that individuals have differences in their sensitivity to chemical stress, and
there will be mortality when the damage exceeds the individual’s thresholds. The representation
of thresholds (a constant value in SD vs. a model distribution within a cohort of individuals in
IT), and mortality rates (as a probability in the SD model vs. a specific event in the IT model) are
the primary differences in their mathematical formulations (see Jager and Ashauer 14 for further
details).

The TK and TD processes are modelled separately in GUTS. However, since most of the
exposure data available do not include the whole-body concentrations (i.e. internal
concentrations) that are required to calibrate the TK parameters, GUTS is also available in its
reduced form (GUTS-RED) that lumps TK with damage. A user-friendly software (openGUTS)
that incorporates this simplification has been made publicly available recently and was used for
the simulations (see https://openguts.info).

To effectively employ the model, the parameters must be calibrated first using lab-derived
survival data. Note that the raw survival data for chlorpyrifos, diazinon, imidacloprid, and
propiconazole have already been used to fit GUTS parameters in previous studies 16-20. However,
they were rarely utilized to predict gammarids survival in real environments except for a study by
Ashauer, et al. 20 on a single compound (diazinon). Also, the GUTS model formulations employed
within these studies either have slightly different mathematical representations or the parameter
confidence intervals were not provided. To harmonize the process of parameter calibration,
including the provision of confidence intervals, we re-fitted the parameters for all compounds
using the optimization algorithms embedded within the openGUTS in Matlab (updated Dec 2019).

2.2.1. GUTS model parameterization

Specific equations that represent both GUTS-RED-SD/IT formulations are not included
here as they are described in detail in Jager and Ashauer 14 or http://openguts.info. Briefly, the
scaled damage, $D_w(t)$ (ng/L) is referenced to the external concentration $C_w(t)$ (ng/L), and the overall dynamics is dominated by a single rate constant, $k_d$ (1/d). $D_w(t)$ is then used to calculate survival probabilities under SD or IT conditions.

For GUTS-RED-SD formulation, the hazard rate ($h_{SD}(t)$) is first calculated, which is a function that increases proportionally with the killing rate constant, $b_w$ (1/d), once the threshold $m_{w,SD}$ (ng/L) is exceeded. The hazard rate is then used to estimate the survival probability ($S_{SD}$) by integrating $h_{SD}$ over the exposure duration. The GUTS-RED-IT estimates the probability of an individual with a threshold value, $m_{w,IT}$ (ng/L), to survive until a particular day within the exposure period. This threshold follows a log-logistic distribution described by the spread-factor, $F_s$ (-). The background hazard ($h_b$) was pre-fitted to the control data for both SD and IT simulations and was considered constant over time.

The raw survival data from previously published studies were used to calibrate the parameters $k_d, m_{w,SD}, b_w, m_{w,IT}$, and $F_s$ (Table S1). The software has its built-in optimization methods that combine grid-search, genetic, and likelihood profiling algorithms that not only ensure a robust parameterization but also generate confidence intervals on model predictions and allow visual inspections of the parameter space. The calibrated parameters are then used to estimate the survival probabilities of gammarids in the field using the monitoring data as inputs for $C_w(t)$ (data from 15). The calibrated parameters for chlorpyrifos and imidacloprid were employed for chlorpyrifos-methyl and thiacloprid respectively due to the lack of available survival data specific for these compounds. At present, it is not possible to validate the calibrated parameter predictions for all compounds due to the lack of independent survival data. However, two sets of survival data from two independent studies for chlorpyrifos are available (Table S1).
For this compound, we explored the differences in model parameters when they were calibrated using (1) each dataset separately or (2) when two datasets were combined.

Under real-case exposure scenarios, it is likely that the GUTS predictions will show no impact on survival as the field concentrations tend to be lower than the acute concentrations employed in laboratory studies that measures gammarids survival. A useful information provided by openGUTS is the derivation of an LPx value defined as the multiplication factor applied to the concentration profile before an x % reduction in survival is observed at the end of the exposure. Here, we examined the LP50 values to be consistent with the EFSA recommendations. EFSA further suggests LP50 values of 100 or 10 as reasonable thresholds for acute or chronic risks, respectively (Tier 2C1 methodology 12).

2.3. Survival predictions of pesticides in mixtures

Within the GUTS framework, mixture toxicity can be approached via (1) body-residue addition, (2) damage addition, or (3) survival probability multiplication 14. The first two approaches reflect the classical “concentration addition” (CA) concept in mixture toxicity assessment, whereas survival probability multiplication mirrors “independent action”. In body-residue addition, chemicals are assumed to act on the same target site and create the same amount of damage. Hence, the internal concentrations (body-residue) of the chemicals can be expressed with respect to each other. Damage addition assumes that chemicals produce the same form of damage, but the degree of this damage varies relative to each other. For example, organophosphates and carbamates both inhibit acetylcholinesterase but the magnitudes of the overall damage may be different. In this case, a damage potency factor is used as a weighting factor based on the level of damages produced by two different chemicals. In survival multiplication approach, the chemicals of interest act on different target sites and hence, different
forms of damages are expressed. The survival probabilities are assumed independent and can then be multiplied.

In openGUTS, the two forms of addition options cannot be differentiated since the reduced form of GUTS is implemented. Hence, we opted for the addition of exposure concentrations for assessing mixture toxicity (i.e., concentration addition). First, we grouped the compounds according to their toxic mode of action (organophosphates or neonicotinoids) and represented the $C_w(t)$ of each compound in terms of their toxic equivalence relative to a reference compound. Hence, in the mixture simulation, only a single set of GUTS parameters is employed (i.e., that of a selected reference compound). Chlorpyrifos and imidacloprid were used to represent the organophosphate and neonicotinoid groups, respectively. The toxic equivalence of a compound relative to the reference was calculated using the following equation:

$$EF_i = \frac{EC_{50_{geo,ref}}}{EC_{50_{geo,i}}} \quad \text{Eqn. 1}$$

$$C_{w,i}^* = EF_i \times C_{w,i} \quad \text{Eqn. 2}$$

where $EF$ (-) is the equivalence factor, $EC_{50_{geo,ref}}$ and $EC_{50_{geo,i}}$ are the geometric means of EC50s (µg/L) of the reference compound and compound $i$, respectively. EC50s were obtained from the database generated by De Zwart \(^{21}\) which contains information on the toxicity test endpoints, toxic mode of action, and general descriptors of toxicity tests of >80,000 chemicals (Tables S5/S6). Data specific to *Gammarus spp* were selected, except for chlorpyrifos-methyl, imidacloprid, and thiacloprid where no gammarid-specific data were available. For these compounds, acute EC50s reported for arthropods were used instead. EF is then multiplied by the measured stream concentration of compound $i$ to obtain its toxic equivalence ($C_{w,i}^*$). If the compounds have similar modes of action, then the total concentration can be calculated using concentration addition (Eqn. 3).
We also explored an alternative to estimating toxic equivalence (EF) using the 96h-LC$_{50}$ estimates provided by openGUTS. For organophosphates (chlorpyrifos and diazinon), we found a two-fold difference between EFs calculated using Eqn. 1 and De Zwart$^{21}$ database compared to EFs estimated using GUTS LC$_{50}$ (Table S6). Hence, there is a potential to use GUTS for estimating relative potencies, but not all target chemicals will have raw survival data required to run the GUTS model. The suitability of the concentration addition approach using Eqn. 1 was tested for diazinon by comparing the survival predictions (1) when its uniquely calibrated GUTS parameters are employed and (2) when it is expressed as chlorpyrifos-equivalence.

3. **Results and Discussion**

The goal of this paper is to provide predictions on the likely impacts of mixtures on gammarids survival based on real-case exposures stemming from pesticide pollution. Many steps were taken before we achieved the predictions we deemed necessary to address our study objectives. The results of using the GUTS modelling framework are described subsequently.

We first investigated the predictions when the individual and combined survival data were used to calibrate the parameters for chlorpyrifos. It was evident that the two independent survival datasets for chlorpyrifos produced different survival predictions, albeit the model uncertainty intervals overlap (Fig. 1a). Note, however, that the values for all calibrated parameters are similar except for the damage threshold ($m_{w,SD}$) with magnitudes differing by ~2 orders of magnitude (Table S1). This result suggests that the toxic dynamics of chlorpyrifos towards gammarids from both studies are similar (i.e., dominant and killing rates constants differed by 1.19-2.3-fold only), but their sensitivities vary widely. This is not a rare observation as gammarids collected at the same site but at different periods (April vs. October) can be ~31-
fold different in their sensitivities towards the same chemical \(^{22}\). A similar finding was encountered in other macroinvertebrates such as mayflies \(^{23}\). Our observation is further supported by a 10-fold variation in the 48h acute EC50s reported by Rubach, et al. \(^{24}\) and Ashauer, et al. \(^{17}\).

**Figure 1.**

**a)** GUTS-SD Survival probability predictions for chlorpyrifos using the calibrated parameters derived from the individual and combined survival data of Ashauer, et al. \(^{17}\) and Rubach, et al. \(^{25}\). The stream concentrations were measured at Eschelisbach from March to August 2015 (see Spycher, et al. \(^{15}\)).

**b)** LPx curves describe the required increase in exposure time series to produce a certain effect on survival. Depending on which GUTS parameter sets were employed, increasing the concentrations by a factor of <10-20 results to a survival probability of 50%.
Although both model calibration results are considered satisfactory, it appears that the parameters derived when using the survival data from Ashauer, et al. 17 were better than that of Rubach, et al. 25 based on the model performance evaluators recommended by EFSA 12 (Tables S1/S2, Fig. S2). The use of log-likelihood values or other classical criteria based on the loglikelihood (e.g., Akaike Information Criterion (AIC)) to compare model adequacy in representing the survival data sets are not applicable here since their values depend on number and absolute quantities of the underlying data. Given that the Ashauer, et al. 17 data were from varying exposures (pulsed) and Rubach, et al. 25 performed their study under constant exposures, one might be convinced to use the parameters derived from the former, especially when predicting the effects under fluctuating exposures 26. However, we argue that this choice may underestimate the variabilities that are inherent among laboratory studies, including differences in age, sex, diet, and natural habitat of the test organisms.

A primary advantage of the GUTS model framework is its ability to combine different survival data from different studies. However, the calibrated internal threshold parameter derived using the combined survival datasets appeared to be driven more by the Ashauer, et al. 17 data, likely due to their more intensive experimental design (i.e., survival data collected over 20 days vs. 4 days in Rubach, et al. 25). Nevertheless, we deem this approach (i.e., combined survival data) to be more robust and appropriate since the variabilities in the methodologies associated with each lab study are well-integrated. It can also be observed that the confidence range of the calibrated threshold value was considerably wider for the combined data than the one resulting from the single data sets (Table S1).

3.1. **LP$_{50}$ analysis and influence of exposure duration and magnitude on predictions**
Based on LP$_{50}$, the exposure profile of chlorpyrifos even at relatively low concentration can potentially induce acute and chronic toxicities to gammarids, similar to what was described by Spycher, et al.\textsuperscript{15} about the same site using the risk quotient approach. Specifically, increasing the concentrations by only a factor of $<10$-20 (depending on which GUTS parameters are employed) will reduce the average survival of gammarids to 50% at the end of the exposure period (180 days) (Fig. 1b). Interestingly, it is during the lower concentration but relatively more prolonged pulsed exposure that mostly contributed to the decline in gammarids’ survival chances (Fig. 1a). It is possible then that short-duration spills (0.5 d here) which tend to occur at higher magnitudes, may not be as damaging to gammarids compared to longer duration exposure (up to 6 days in this study). This rapid exposure does not likely provide sufficient time for the chemical to be bioaccumulated in gammarids as what has been estimated when the internal concentrations of wastewater-derived micropollutants were modelled at various spill conditions.\textsuperscript{27} Hence, high magnitude but short-duration spikes in river concentrations will result to lower body residue and little tissue damage as predicted in this study. Unfortunately, we do not have the measured internal concentrations of gammarids in the field to confirm this hypothesis. However, we predicted the whole-body burden using a separate TK model with uptake and rate constants derived from a previous study (Fig. S1, Table S8). Indeed, our TK predictions showed that exposure to low concentrations but at a longer duration produced higher internal concentrations in gammarids than high concentration-short duration exposure. Of course, the combination of exposure duration and the magnitude of exposure ultimately determines the total damage (as illustrated in Fig. S6), but our analysis is precisely the type of assessment that can be achieved via the TKTD approach. By explicitly linking the predicted effects to the fluctuations in pesticide concentrations, one can develop hypotheses on the likely hazardous exposure conditions. More
importantly, this process incorporates effects as a function of time, which is often missed in traditional risk assessment.

3.2. Survival predictions upon exposure to other pesticides – SD vs IT

Up to this point, we have only presented the survival predictions using the GUTS-RED-SD approach. We additionally explored the primary differences in IT and SD predictions for other compounds. The openGUTS predictions of gammarids' survival upon exposure to varying concentrations of azoxystrobin and diazinon are illustrated in Fig. 2 (see Fig. S3 for imidacloprid and propiconazole). As laid down in the ESFA opinion, the suitability of SD and IT in predicting gammarid survival would ideally be tested with validation data through experimentally observed survival of gammarids under time-variable exposure. Since such validation data were not available for the chemicals analysed in this study, the second-best option is to describe the suitability of SD and IT predictions based on the quality of the calibrated parameters (Tables S1/S2, Fig. S2). Recall that we re-fitted these parameters to the survival data from published studies and have reached the same conclusions: for diazinon, the SD parameters are better constrained than IT parameters, the opposite was observed for imidacloprid, and either representation appears to be suitable for propiconazole. The GUTS parameters for azoxystrobin are reported for the first time (raw survival data in Table S7), and it appeared that the SD represents the survival slightly better than IT (Fig. S2, Tables S1). Note, however, that the survival data for azoxystrobin are not as extensive as other pesticides where GUTS parameters have been calibrated, so caution must be exercised as large prediction uncertainties are expected.
Figure 2. GUTS-RED-SD/IT predictions for azoxystrobin and diazinon. Recently approved acute and chronic environmental quality standards (EQS) for pesticides in Swiss surface water are indicated along with the regulatory acceptable concentrations (RAC) from Swiss pesticide registration dossiers. Uncertainty in the LP50 predictions results from the propagation of the parameter uncertainty. Acute and chronic risks are expected when LP50 values are <100 and <10 respectively. Predictions for imidacloprid and propiconazole are presented in Fig. S3.

There have been many attempts to differentiate the assumption that dominates the survival upon exposure to chemical stress. Indeed, these studies have found it challenging to identify which of the mechanisms describe survival better, and extending these hypotheses to
real environments with different confounding factors makes it even more difficult to determine
the exact mechanism of survival due to chemical stress. There have been indications in the
literature that survival is driven more by the SD process, not only at the organism level 32, 33, but
also in describing population dynamics 34. However, there is a large consensus that SD and IT
are only extreme, special cases of survival mechanisms 14, 35. Hence, it is likely that true survival
falls between these two assumptions, and if one is to decide on the dominant mechanism, the
choice must be supported with enough evidence (e.g., additional lab studies, literature) 29, 31.
Furthermore, the calculated LP_{50} curves show large differences between SD and IT
model representations (Fig. 2, Fig S3), although the calibration quality appears very similar (Fig.
S2). It is not surprising that despite similar calibration quality, SD and IT model predictions can
show these differences. This has not yet been mechanistically demonstrated for LPx values, but
within the context of model validation, it has been reported that often, one of the SD or the IT
models show a much better match to the validation data (e.g. Focks, et al. 36). For our study, we
presented the results of both models (Fig. 2) as we lack the measured data to confirm whether the
survival of gammarids upon exposure to these chemicals in the field fits the SD or IT (or neither)
assumptions.
Regardless of which models are employed, it initially appears that there would be little to
no impact on survival upon exposure to individual chemicals. However, it is important not to
conclude the GUTS analysis here, but rather extend the assessment to infer the level of exposure
needed before a certain effect may be observed (i.e., through LPx analysis). After examining the
LPx curves, some of these compounds can potentially have damaging effects on gammarids
(LP_{50} values in Table S3). For instance, the LP_{50} for azoxystrobin is <100 and <10 under SD and
IT predictions respectively, suggesting that this chemical may be both acutely and chronically
toxic to gammarids at the end of the exposure. The concentration dynamics of diazinon in the field suggest acute toxicity (LP50<100 for both SD and IT predictions), but for imidacloprid acute (LP50<100) toxicities may be expected depending on which model formulation is used (SD vs. IT, see Table S3). The risks of exposure to propiconazole appear to be very minimal as LP50 values are much greater than 100. Diazinon was banned in Switzerland as a plant protection product in 2013 but is still registered as a veterinary pharmaceutical. It is likely that the spike in diazinon concentration in Eschelisbach is driven by a rain event as the total sum of pesticides shortly after this event was highly elevated (see SI in Spycher, et al. 15). Given that diazinon degrades rapidly in soil 37, the source of diazinon might have been due to veterinary uses or from an unauthorized application in the field. Imidacloprid and propiconazole were frequently detected, but the exposure concentrations are not large enough to result in acute toxicities upon exposure.

3.3. Mixture impacts – the case of organophosphates and neonicotinoids

Our predictions showed that exposure to individual compounds may not produce substantial impacts on survival (e.g., imidacloprid, Fig. 2), but toxicity may arise once chemicals are in mixtures. There are several ways to address mixture toxicity, and a straightforward approach is via concentration addition (CA). Here, we assumed that if the chemicals have the same mechanism of action and similar slopes in classical dose-response curves, then one compound can be represented in terms of the other based on their relative potency (expressed as EF in Eqn. 1). In addition to method simplicity, CA is advantageous within the context of the GUTS modelling framework, especially for compounds with field monitoring data that do not have chemical-specific survival data to parameterize GUTS. Chlorpyrifos-methyl and thiacloprid, for instance, were deemed as relevant compounds contributing to high risk of
exposure to invertebrates at our site (see acute toxic profiles in Spycher, et al. 15), but there are no available survival data specific for gammarids. We recognize, of course, that concentration addition is a major assumption employed here, but there are no available studies to date that assess the “relative internal damage” of pesticides belonging in the same toxic mode of action so we cannot assess the possibilities of using body-residue and/or damage addition within GUTS.

To test whether the CA hypothesis is acceptable, we applied this concept to diazinon by expressing it as chlorpyrifos-equivalence. Both compounds are organophosphates that disrupt the function of acetylcholinesterase in invertebrate nervous system. We estimated the impacts on gammarid survival using the GUTS parameters for chlorpyrifos and compared the predictions when diazinon-specific GUTS parameters were employed. As shown in Fig. 3, we observed that the CA hypothesis is acceptable under the SD predictions (uncertainties overlap). The survival predictions via the IT were also similar for both presentations (Fig. S4).

**Figure 3.** a) Survival predictions for diazinon when its own GUTS parameters were employed (blue) in comparison to when the exposure concentrations were expressed as chlorpyrifos equivalence and chlorpyrifos GUTS parameters were employed (green). b) LP50 values from
both models are very similar and suggest an acute risk of exposure to diazinon ($LP_{50} < 100$).

Only the SD predictions are shown here; IT predictions are found in Fig. S4.

When ultimately evaluating the risks, the LP$x$ predictions can provide indications of the severity of the exposure. Here, expressing diazinon as chlorpyrifos equivalence and using chlorpyrifos GUTS-SD parameters for survival predictions result in similar $LP_{50}$ as when diazinon-specific parameters were employed (Fig. 3b). However, a four-fold difference was observed between GUTS-SD and GUTS-IT $LP_{50}$ values (Table S3). A previous study claims that gammarids survival after exposure to diazinon follows the SD process, but it is again challenging to select the dominant mechanism of survival, especially when examining the effects of mixtures. Nevertheless, our analysis of $LP_{50}$ under the GUTS-SD representation suggests that CA may be sufficient for the purposes of our study.

A potential issue in this approach may be associated with the selection of a compound that represents the pesticide group of interest. For simplicity, we selected chlorpyrifos over diazinon as it employed more than one raw survival data sets available for model calibration, and it is also structurally very similar to chlorpyrifos-methyl. When we then evaluated the mixture effects of the three organophosphates, our predictions suggest some impact on gammarid survival at the end of the exposure period (Fig 4). Specifically, the mixture exposures were predicted to be both acutely and chronically toxic to gammarids, and it only requires up to $\sim 5$ times the exposure concentrations before the survival probability drops to 50% (see $LP_{50}$, Fig. 4b). This is a highly probable event as chlorpyrifos and chlorpyrifos-methyl were detected at concentrations higher than what had been observed at our site (e.g., up to 750 ng/L $^{38}$). When we further explored the changes in survival probabilities under varying exposure concentration
multiplications (MF), the overall decline in predicted survival probabilities corresponds to the spikes in stream concentrations. Obviously, higher concentration peaks led to larger drops in survival probabilities, but with mixture toxicity assessment (in time series), we can identify the compound which mostly contributed to the overall changes in survival (i.e., chlorpyrifos-methyl in the case of organophosphates).

Note that an additional uncertainty in the model prediction is associated with the calculation of toxic equivalence (EF). When EFs were estimated based on all arthropod acute EC50 data, diazinon and chlorpyrifos-methyl were found to be 0.5 and 2.1 times as toxic to chlorpyrifos (vs. EFs of 0.03 and 0.9 respectively in the current predictions). These EF values resulted to survival predictions that were more severe than what was illustrated in Fig. 4a (see Fig. S5), but obviously, gammarid-specific data were deemed more appropriate for the study. Nevertheless, this observation alludes to the flexibility for mixture modelling within the GUTS framework where EFs can be adjusted based on the availability of EC50 data associated with the species of interest.
Figure 4. GUTS predictions for a mixture of 2 groups of insecticides with different toxic modes of action. 

a) exposure concentration expressed in the as equivalent concentration. For organophosphates, the GUTS parameters for chlorpyrifos (combined survival data) were employed and the stream concentrations were converted to chlorpyrifos (chlor) equivalence based on the EC50s (see section 2.3). Diazinon (diaz) and chlorpyrifos-methyl (chlorm) are 0.03 and 0.9 times as toxic as chlorpyrifos. Imidacloprid (imida) was employed as the reference compound for neonicotinoids. The EF thiacloprid (thia) is 0.4 (Table S5). 

b) evaluation of survival probabilities when different multiplication factors (MF) were imposed on the exposure time series. With mixture toxicity assessment illustrated here, we can hypothesize the compound that mostly contributed to the cumulative drop in survival.
It has been reported that there are minor differences in the relative toxicity of neonicotinoids, and they tend to be additively or synergistically toxic together \(^3^9\). We further confirmed the applicability of the CA hypothesis for imidacloprid and thiacloprid by comparing the GUTS parameters derived by Focks, et al. \(^3^6\) for mayfly (Cloeon dipterum). When the values were compared on a molar basis, we found that the mortality-related parameters for both compounds, \(F_s\), \(m_w\), and \(b_w\), only differed by a factor of 0.4, 1.5, and 7 respectively (i.e., if toxicity are similar, these parameters should also be similar). Hence, the CA approach can be applied to gammarids when exposed to a mixture of neonicotinoids by expressing the total concentrations as imidacloprid-equivalence. Overall, it appears that only a minimal impact on survival is predicted at the end of the exposure, but the corresponding LP\(_{50}\) values indicate acute (LP\(_{50}\)<100) and chronic risks (LP\(_{50}\)<10). Although thiacloprid is not as toxic to invertebrates as imidacloprid (EF=0.4), it occurs at higher environmental concentrations and thus was observed to contribute more to the predicted risk.

We further hypothesize that the combined exposure to organophosphates and neonicotinoids have detrimental effects on gammarids. Low biodiversity and the absence of several sensitive species at Eschelisbach have already been reported, indicating that this stream is indeed heavily impacted not only in the year we made the predictions but also at another year as documented in 2017 \(^4^0\). Given that chemical exposure is just one of the stressors that gammarids experience in the field, reducing the exposure to pesticides will likely improve their overall fitness in response to other types of environmental pressures (e.g., increased predation, warming climates).

3.4. **Comparison of the predictions to current environmental quality standards**
In Switzerland, the individual environmental quality standards (EQS) for 19 pesticides have just been recently added as part of their nationwide Water Protection Ordinance (SR 814.20, Annex 2, No. 11, Par. 3). For compounds with no specific EQS values (e.g., propiconazole), a threshold of 0.1 µg/L is still in place. As expected, the compounds with frequent exceedance to acute or chronic EQS values were also considered by the model to be acutely or chronically (or both) toxic based on the analysis of their LP50 values (e.g., azoxystrobin at 17% and 40% exceedance for acute and chronic EQS respectively) (Fig. 2). This observation suggests that the GUTS model predictions are supportive of the Swiss standards and confirms that these thresholds can adequately protect the aquatic environment.

There are no specific environmental standards set for these substances within the European Union (EU) as only a few pesticides belonging to the priority substances list are regulated under the Water Framework Directive 2000/60/EC (e.g. atrazine, diuron). The EU Plant Protection Regulation (1107/2009) suggests the use of regulatory acceptable concentration (RAC) for surface water protection, but these values are more intended for prospective monitoring (edge-of-field exposure) within the context of pesticide (re-) registration and approvals (RAC values also shown in Fig. 3 and Table S9). Recent studies however, have used the guidelines set by the EU drinking water directive (98/83/EC)41-43, i.e., the individual pesticide concentration of pesticides should not exceed 0.1µg/L in source waters, and their total sum should not go over 0.5 µg/L. When this was applied to our monitoring data, the environmental concentrations of most substances included in this study rarely or never exceed 0.1µg/L. However, the GUTS predictions showed that low concentrations of these substances could be acutely and/or chronically toxic (e.g., chlorpyrifos and diazinon). Hence, relying on these values for water quality protection may underestimate the true risk for aquatic organisms.
4. Model Limitations and Environmental Significance

Our modelling study added a layer of realism when evaluating the risk of exposure to fluctuating concentration levels of pesticides in comparison to traditional risk assessment methods that require the calculation of risk quotients (RQ) based on maximum peak or average concentrations. Additionally, the GUTS model provided LP$_{50}$, a reasonable metric that examines the impact of time-variable exposure data (measured or modelled) while explicitly considering the uncertainty of the assessment. LP$_{50}$ values could be used as an endpoint in addition to the traditional RQs. It is important to note, however, that the LP$_{x}$ predictions between SD and IT representations may vary. If there are insufficient evidence to indicate that survival follows one form more than the other, it is important to show the LP$_{x}$ values from both model formulations.

We acknowledge that there are limitations in our modelling approach, especially the CA hypothesis within our mixture toxicity predictions (e.g., EF derivation, reference compound when calculating toxic equivalence, additivity vs. synergism). Also, only a few compounds were considered for each group and the presence of other pesticides that are less toxic but at higher concentration levels may add more to the cumulative effects on survival. Nevertheless, the model has generated some useful suggestions that can be beneficial for future mechanistic mixture effects modelling. First, the use of GUTS models in assessing the potential impacts of large numbers of chemicals as detected in environmental monitoring would require their calibration based on laboratory toxicity data. It would be useful to generate laboratory standard toxicity tests for more chemicals, ideally representing more mode-of-action groups, and to collect the GUTS parameter in open databases. Specifically, the number and quality of available survival data are not only instrumental for acceptable model calibration, but they also determine the uncertainties in parameter estimates as well as uncertainties related to extrapolation to field conditions (i.e.,
pulsed exposure vs. constant exposure). The availability of these datasets does not only allow
users to compare modelled results, but can help generate hypotheses when conducting future lab
validation or field exposures. There is also a potential to examine mixture toxicities
comprehensively in GUTS especially when assessing the mechanism (body residue, damage
addition or survival multiplication). Second, the sequence of exposure to compounds with
different toxic modes of action may matter as what has been investigated by Ashauer, et al. 44 in
the lab. It can be hypothesized further that survival may be different when gammarids were
exposed to organophosphates first before neonicotinoids and vice-versa. Given that the
application rates of pesticides vary for each pesticide, knowing the impact of exposure sequence
to survival can provide recommendations related to best management practices. Overall, the
modelling strategy presented here is a reasonable first step to address mixture toxicities, but
cautions must be exercised if the results are employed for mixture-based risk assessment purposes
as they still require validation.

5. Acknowledgments

This study received funding from the ETH-Zurich Postdoctoral Fellowships (0-20400-17,
co-funded by the European Union's Seventh Framework Programme for Research and
Technological Development under grant agreement 608881 (2015-2019). We acknowledge the
financial support of the Natural Sciences and Engineering Research Council of Canada (PDF-
517151-2018). We also thank the inputs of the journal reviewers as their comments improved
the quality of the manuscript. Finally, the figure in the graphical abstract was created with
BioRender.com.
Supporting Information. openGUTS calibrated parameter values (GUTS-RED-SD/IT) and calibration quality, LP10/LP50 values for both model variants (SD/IT), toxic equivalence derivation, raw survival data for azoxystrobin, EQS and RAC values, supplemental figures including: TK predictions, parameter space and predicted-observed plots, survival probability predictions of individual compounds, additional scenario testing, SageMath model codes for scenario testing (Python-based).
6. References


