

# Characterizing fast herbicide transport in a small agricultural catchment with conceptual models

Lorenz Ammann<sup>a,b,\*</sup>, Tobias Doppler<sup>c</sup>, Christian Stamm<sup>a</sup>, Peter Reichert<sup>a,b</sup>, and Fabrizio Fenicia<sup>a</sup>

<sup>a</sup>Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

<sup>b</sup>ETH Zurich, Department of Environmental Systems Science, Zurich, Switzerland

<sup>c</sup>VSA, Swiss Water Association, Glattbrugg, Switzerland

\*Corresponding author; e-mail: [lorenz.ammann@eawag.ch](mailto:lorenz.ammann@eawag.ch), address: Überlandstrasse 133, 8600 Dübendorf, Switzerland

April 29, 2020

## Abstract

Herbicide pollution in headwater streams due to agricultural practices is a major environmental concern and is characterized by episodic peak concentrations from fast transport paths. We rely on previous experimental studies in a small (1.2 km<sup>2</sup>) agricultural catchment in the Swiss Plateau, to model dynamic diffuse herbicide pollution with emphasis on fast transport paths in a conceptual modelling framework at the catchment scale. We show how experimentalists' understanding of the fate of herbicides (perceptual model) can be translated into conceptual models considering sorption, degradation, and fast transport of water and chemicals facilitated by impervious surfaces, tile drains and artificial shortcuts. Different types of experimental data (streamflow, high-frequency concentration measurements, and soil-water distribution coefficients) are used in a joint Bayesian inference of model parameters. We assess the ability of different spatial configurations of hydrological response units in explaining observed heterogeneity in transport behaviour of two corn herbicides. Thereby, we find that (1) relatively simple conceptual models can provide a realistic description of herbicide fate in small agricultural catchments, (2) accounting for spray drift onto hard surfaces is necessary to avoid a severe model bias during the first rainfall event after application, and (3) including catchment-specific experimentalist knowledge about important elements like artificial shortcuts and tile drains leads to a reduction in

uncertainty of 30 % compared to the more conventional conception of the proximity to the stream as the dominant risk factor.

Keywords: Pesticide transport; Experimentalist knowledge; Controlled application; Conceptual model; High-frequency concentration data; Bayesian inference

## 1 Introduction

Diffuse pollution of water bodies by micropollutants has been a growing environmental concern during the last few decades. In small headwater catchments dominated by agricultural land use, peak herbicide concentrations caused by fast transport during precipitation events are especially concerning [e.g. Munz et al., 2012, Doppler et al., 2012, Leu et al., 2005, Meyer et al., 2010]. Investigation of fast transport processes in headwater catchments is therefore important for assessing and mitigating the risk of diffuse pollution of streams by herbicides and other micropollutants. Since peak concentrations can only be measured in costly monitoring campaigns involving high-frequency measurements, modelling is a valuable and cost-effective complementary tool to quantify and investigate the herbicide pollution of headwater streams.

The processes involved in herbicide mobilization and transport have been investigated in many studies at a large range of spatial scales. At lab and field scales, mostly physically based models like the pesticide root zone model (PRZM) [Carsel et al., 1985] have been shown to successfully predict water and pollutant fluxes through the upper soil layers [e.g. Singh and Jones, 2002, Ma et al., 2004, Fox et al., 2007]. A major finding of previous studies is the paramount importance of preferential flow and tile drains in transporting pollutants from the top soil layer to streams [e.g. Traub-Eberhard et al., 1994, Zehe and Flüßler, 2001, Gärdenäs et al., 2006, Ghafoor et al., 2013].

At the scale of headwater catchments in the order of 1 km<sup>2</sup>, the states, internal fluxes, and output of interest are likely to be affected by severe heterogeneity. For example, it has been shown that a substantial part of the herbicide load in a creek can originate from only a small fraction of the catchment's area [e.g. Freitas et al., 2008], the so-called critical source areas [Pionke et al., 1996]. Doppler et al. [2014b] even observed a temporal variability of those areas between successive rainfall events. However, data sets with controlled input of pesticides [e.g. Doppler et al., 2012] are lacking at this scale, impeding the generation of knowledge about critical source areas. Since any catchment-scale transport model can only be a simplification of the highly complex underlying system, we are faced with the difficult task of retaining in the model the most important features regarding its emergent behaviour.

One common simplification is the use of relatively large grid cells in combination with equations describing physical processes, in order to reduce the computational and input data demand. Examples include Kannan et al. [2006], who apply the SWAT model [Arnold et al., 1998] to reproduce daily average concentrations of four pesticides. Frey et al. [2009] show that critical source areas can be located through a topographical connectivity analysis (requiring a high-resolution DEM) and that observed herbicide loss rates are spatially correlated with the modelled amount of surface runoff to the stream and preferential flow to tile drains on different experimental fields. The ZIN-AgriTra model Gassmann et al. [2013] was successfully used to simulate high-frequency pesticide loads in an agricultural catchment. Villamizar and Brown [2017] extend the MACRO model [Jarvis et al., 1991] from the field to the catchment scale and compare it to the SPIDER model [Renaud et al., 2008] based on data of daily resolution. Morselli et al. [2018] apply the DynAPlus model in a larger catchment (130 km<sup>2</sup>) containing apple orchards to characterize high-frequency concentration dynamics of three pesticides.

The application of such physically based models at the catchment scale has been criticised for the excessive degrees of freedom compared to what observations can constrain and for being based on an “implicit upscaling premise” [Kirchner, 2006]. Data to inform their effective parameters is often lacking, even at smaller scales [e.g. Marín-Benito et al., 2014] and their high computational cost tends to prohibit comprehensive assessments of their large uncertainty. Models based on residence time distributions have also been applied to characterize daily to weekly average streamflow and concentrations of herbicides [Bertuzzo et al., 2013]. However, it seems challenging to reflect experimentalist knowledge about specific processes with such approaches.

Conceptual models, on the other hand, combine some degree of physical interpretation of the processes with simplicity and short run-times. They allow for the direct description of the emergent behaviour of the system at the catchment scale, which might not follow the equations used to describe small-scale physical processes [e.g. Kirchner, 2006]. Their structure and governing equations can be adapted to include experimental information on the most important processes in a study site.

Modellers might find it difficult, however, to consider qualitative elements like experimentalists’ process knowledge, which may be based on their interpretation of site-specific quantitative or qualitative data [e.g. Seibert and McDonnell, 2002] or general scientific knowledge of the experimentalists as experts. In fact, applied models have been found to be strongly dependent on a modeller’s background and experience [e.g. Holländer et al., 2009]. Since the recognition of this deficit, there has been some effort to propose more systematic ways of including experimental data and knowledge in the model building process [e.g. Seibert and McDonnell, 2002, Fenicia et al., 2008, McMillan et al., 2010, Wrede et al., 2014]. Specifically, the latter study shows how the construction of a per-

ceptual hydrological model [Beven, 2012] can be informed by field measurements in the first step of a modelling study. On one hand, perceptual models are well suited to accommodate qualitative knowledge of experimentalists, on the other hand they facilitate the use of such knowledge in a modelling study. This study investigates how experimentalist knowledge and data can be considered in a modelling study through the use of perceptual and conceptual models of herbicide transport in agricultural catchments, which has been insufficiently done so far.

Considering the undeniably large uncertainties of water and pollutant transport models at the catchment scale [Rode et al., 2010], the assessment of the uncertainty of such models is important. The high simulation cost of the largely physically-based models [e.g. Frey et al., 2009, Doppler et al., 2014a, Gassmann et al., 2013] used in this field of research pose a technical challenge for uncertainty assessment. Moreover, there is a lack of established approaches for rigorous stochastic description of the output uncertainty of deterministic transport models. The GLUE methodology [Beven and Binley, 1992] has been used in previous studies [Zheng and Keller, 2007, Bertuzzo et al., 2013, e.g.] for this purpose. However, GLUE lacks statistical validity when used with an informal likelihood measures [Mantovan and Todini, 2006, Beven et al., 2007, Mantovan et al., 2007, Stedinger et al., 2008]. Other studies successfully use more rigorous approaches [Raat et al., 2004, Gardner et al., 2011, e.g.], but neglect potential correlation of the errors. Hantush and Chaudhary [2014], Wellen et al. [2014], Zheng and Han [2016], Han and Zheng [2016] account for correlation, but refrain from sampling the full joint posterior distribution of error model and transport model parameters, which would be conceptually preferable [Evin et al., 2014]. Previous work which does sample the joint posterior distribution [e.g. Hong et al., 2005, Talamba et al., 2010, Han and Zheng, 2018] is based on data resolutions in the order of hours to years. It was shown that considering temporal correlation of residuals can be much more problematic when working with hourly than with daily data [Ammann et al., 2019]. In this study, we explore the applicability of such techniques to data resolutions in the order of 15 min, profiting from the time-dependent correlation approach suggested by Ammann et al. [2019].

This study tries to address the aforementioned shortcomings concerning the characterization of intermittent peak herbicide concentrations, the lack of research on how to translate perceptual models to conceptual models, and the uncertainty quantification approaches used in water quality modelling. Consequently, the goals are the following:

- 1) Show how experimentalist knowledge and data can be used to inform and constrain the construction of a simple, yet realistic, conceptual transport model to assess in-stream fluctuations of herbicide concentration.
- 2) Assess which degree of experimentalist knowledge and model complexity is

needed for an accurate representation of the available data by simplifying the model developed for goal 1).

- 3) Quantify the uncertainty of the predictions by using a recently presented framework for likelihood functions, considering heteroscedastic and correlated errors.

Section 2 gives a brief overview of the chosen study site and past experimental work carried out there, which provides the extensive experimentalist knowledge and data required for goals 1) and 2). Section 3 summarizes the pre-existing perceptual model of the catchment (Sect. 3.1) and how it is translated to conceptual models (Sect. 3.2). The approach to uncertainty quantification is described in Sect. 3.3 including the error model (Sect. 3.3.1), the inference techniques (Sect. 3.3.2), and the prior distribution (Sect. 3.3.3). Section 4 summarizes the results of the parameter inference for all the tested models, including the uncertainty of the model output. Section 5 discusses the implications of the results and Sect. 6 makes some concluding remarks.

## 2 Study Site and Data

The Eschibach catchment is located in the northeastern part of Switzerland (Fig. 1) near the town of Ossingen. For a detailed description of the catchment, the reader is referred to Doppler et al. [2012, 2014b,a]. It has a temperate climate typical for a mid-latitude region. Meteorological measurements at 11 km distance from the study site indicate an average yearly precipitation of 900 mm (without strong seasonal pattern) and an average temperature of 9.4 °C (ranging from a monthly average of 0 °C in January to 19 °C in July) between 1981 and 2010 [MeteoSwiss, 2016]. The catchment has a size of 1.2 km<sup>2</sup> and a subdued topography (the elevation ranges from 423 to 473 m a.s.l.). The land use in the catchment is predominantly agriculture (75 %) and forest (13 %), with some impervious areas (paved roads, field roads, and settlements) that partially drain into the stream. The agricultural area is used for crop production (corn, sugar beet, winter wheat, and rape seed) and roughly half of it is tile drained (Fig. 1), with pipes located around 1.0 - 1.5 m below ground surface. Some of the roadside drains are connected to this tile drainage system.

The experimental study by Doppler et al. [2012] took place from summer 2008 to autumn 2009 with a focus on the growing season of 2009. It consisted of a controlled application of two herbicide mixtures on 19 May 2009, each to its own set of fields. We model one substance from each mixture; atrazine (CAS no.: 1912-24-9) and terbuthylazine (5915-41-3). On 19 May, 8.3 kg of atrazine and 5.6 kg of terbuthylazine were applied. The mixture containing terbuthylazine was applied on a second occasion, the mass of which was known (4.2 kg) but the timing of application was not. Note that the controlled application of

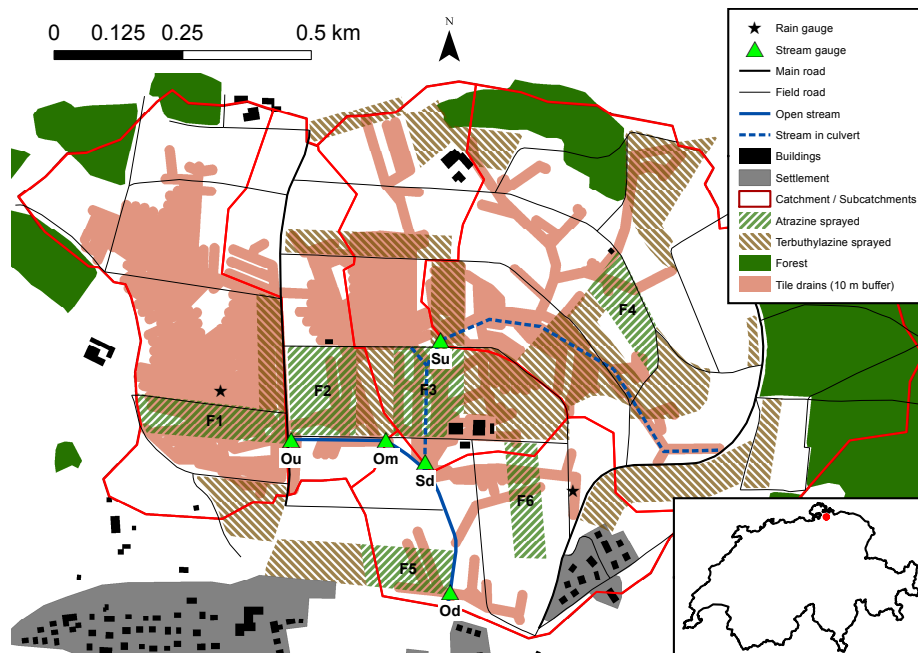


Figure 1: Overview of the study catchment of the Eschibach stream near the town of Ossingen, Switzerland. Experimental fields are denoted as F1-F6. The land use is crop production if not indicated otherwise. Source: swisstopo [2008], Gemeinde Ossingen [1995]

Table 1: Overview of the data used for calibration.

Data type	Period with data	Typical resolution	No. of data points
Streamflow	1 Apr - 12 Oct	15 min	18621
Atrazine conc.	19 May - 22 July	15 min <sup>a)</sup>	259
Terbutylazine conc.	19 May - 28 May	15 min <sup>a)</sup>	58 <sup>b)</sup>
Distribution coeff.	19 May - 18 July	5 days	6

<sup>a)</sup> During baseflow conditions, data is limited to a few grab samples.

<sup>b)</sup> Only events E1 and E2 are used for calibration w.r.t. terbutylazine concentrations, since the date of a second application of this substance after E2 is unknown.

the subsequently analyzed herbicides means that the input is known with high accuracy in both space and time, which is rarely the case at the spatial scale of this study. They measured streamflow and electric conductivity at 5 min resolution and herbicide concentrations in time-proportional samples taken every 15 min during rainfall events at 5 stream gauges within the catchment (Fig. 1). The available data also includes continuous measurements of groundwater levels and soil moisture, as well as chemical analysis of samples of ponding water. Additionally, sorption and degradation experiments have been performed [Camenzuli, 2010, Doppler et al., 2014b] with soil sampled from all the experimental fields on which atrazine was sprayed (F1-F6, Fig. 1).

Two rain gauges were active within the catchment (Fig. 1) between 19 May and 31 July 2009, by which point a large part of the sprayed herbicide mass had been transformed or degraded and the sampling was therefore stopped. The gauges operated at 15 min resolution and registered a total of 353 mm of precipitation, where 5 rain events triggered the sampling device sufficiently to lead to a meaningful representation of the dynamic fluctuations of the in-stream concentration during the event [Doppler et al., 2012]. Those events are denoted as E1-E5 in this study (see Fig. A.17). Data on potential evapotranspiration for the study period was kindly provided by MeteoSwiss [2016].

### 3 Methods

This section presents the perceptual model (Sect. 3.1), which summarizes the available knowledge about transport processes in the studied catchment and continues by showing how we translated this knowledge into numerical models that are conceptual in nature (Sect. 3.2). It continues with a section on the

chosen approach for uncertainty quantification (Sect. 3.3), which is split into a description of the probabilistic error model, the numerical technique for inference and prediction, and the chosen prior distribution. The section ends with information on the metrics chosen to assess the performance of the developed models (Sect. 3.4).

### 3.1 Perceptual model

The perceptual model of a catchment conceptualizes the main transformation and advection processes of water and substances that happen inside and across the boundaries of that catchment [Beven, 2012]. In our case, it is formulated based on experimentalist knowledge and data available from previous work at the study site. Figure 2 is a graphical summary of the elicited perceptual model, which is described in the following, focusing on aspects that are especially relevant for guiding model construction. For more details on the experimentalist understanding of the study catchment, see Doppler et al. [2012, 2014b,a].

Water enters the catchment via precipitation and leaves it via streamflow or evapotranspiration. Herbicides enter via spraying, experience sorption, are transformed or degraded (larger part), and exit via streamflow (smaller part) [Doppler et al., 2012]. Volatilization is not expected to be a major flux due to the chemical properties of the applied substances. Degradation and sorption to organic matter in the soil, on the other hand, are important processes influencing the availability and mobility of pesticides [Grathwohl, 1998, e.g.]. The topsoil contains a large part of the applied herbicide mass for a considerable time after application [e.g. Weed et al., 1995] and is known to be the primary origin of the herbicides found in the stream.

Distinctly different dynamics of in-stream concentration time series were observed for atrazine and terbuthylazine [Doppler et al., 2012]. Additionally, atrazine loss rates varied in space and time [Doppler et al., 2014b]. Data from degradation and sorption experiments [Camenzuli, 2010] show that the variability of degradation and sorption cannot explain the difference in dynamics and loss rates described above. That difference can also not be explained by the chemical properties of atrazine and terbuthylazine, which are very much alike due to the similarity of their molecular structure. Instead, the presumed reason for this difference are the dominant transport processes on the fields sprayed with either atrazine or terbuthylazine [Doppler et al., 2012]. Well-known dominant paths are macropore infiltration to tile drains [e.g. Traub-Eberhard et al., 1994] and direct overland flow to the stream [e.g. Leu et al., 2004]. In addition to these paths, we have visual evidence of overland flow that reaches the stream via artificial structures that act as shortcuts [Doppler et al., 2012]. In fact, based on the connectivity analysis of Doppler et al. [2012] in the study catchment, the area topographically connected to such shortcuts is much larger than the area



connected to the stream directly. Despite its potential impact on surface water quality, this path has rarely been considered in modelling studies. In summary, the streamflow and the in-stream concentration signal is believed to be affected by four dominant groups of transport paths, which are listed below in order of increasing travel times:

- 1) Flow on impervious areas. Runoff from paved roads and driveways, as well as from unpaved field roads<sup>1</sup> enters the stream via storm drains.
- 2) Overland flow on connected areas. Overland flow from the fields reaches the stream either directly or via artificial shortcuts such as maintenance manholes of the tile drainage system or roadside storm drains.
- 3) Macropore flow. This path conceptualizes the part of the infiltrating water that reaches the tile drains via fast vertical macropore flow through the variably saturated upper soil layers.
- 4) Flow through saturated zone. This path comprises slower trajectories that contain vertical transport through the soil matrix or lateral flow in the saturated zone or both.

Herbicides are expected to reach the stream primarily via paths 2) and 3). Path 1) and 4) are expected to be less important because the amount of substance input is small for path 1), and sorption, degradation and dilution are strong for path 4). Water, on the other hand, is expected to be predominantly transported on paths 1) during rainfall events and 4) during baseflow conditions, and to a lesser degree on paths 2) and 3).

Note that paths 1) and 3) always include a subsurface pipe on the last stretch before entering the stream. Paths 2) and 4) may or may not include such a pipe. The difference in travel times among paths 1) - 4) in the segments preceding possible pipes is assumed to be large compared to the residence time in the pipes. Therefore, the paths are distinguished based on those preceding elements.

### 3.2 Deterministic conceptual models

In the following, we show how the perceptual model developed in Sect. 3.1 is translated to conceptual models, which are implemented with the hydrological modelling framework SUPERFLEX [Fenicia et al., 2011, Kavetski and Fenicia, 2011] (Sect. 3.2.1). The flexibility of this framework facilitates the incorporation of prior experimentalist knowledge and enables systematic model comparisons.

---

<sup>1</sup>In a visual assessment of a precipitation event in the catchment, the onset of surface runoff on field roads was much earlier than on the fields and not much later than on the paved areas.

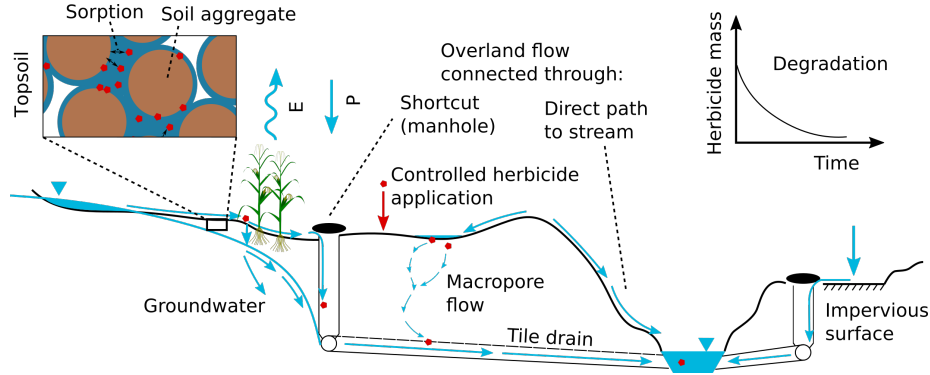


Figure 2: Graphical summary of the perceptual model of the fate of water (blue) and herbicides (red) in the study catchment.

The translation of the perceptual to the conceptual models requires decisions about how to split the landscape into response units (Sect. 3.2.2) and about the appropriate model structure and parameterization of each unit (Sect. 3.2.3).

The process of translating the perceptual model to a conceptual one was an iterative procedure that required repeated exchange between modellers and experimentalists. The deficits of initial, very simple models were identified and improved by incorporating more and more available process knowledge while still constraining model complexity. This section presents some interesting alternative hypothesis (i.e., models) emerging out of this process.

### 3.2.1 Implementation of substance transport processes within SUPERFLEX

The SUPERFLEX framework [Fenicia et al., 2011, Kavetski and Fenicia, 2011] is based on elementary building blocks that can be customized and combined in a flexible manner. They consist of reservoirs, splits and unions of fluxes, as well as lag functions [Fenicia et al., 2011, 2016]. In conceptual modelling applications, those blocks are typically intended to conceptualize the storage and release of water at the catchment scale, as well as the river routing. While previous applications of SUPERFLEX focused on water fluxes, the framework was extended to consider the fate (transport, degradation, and sorption) of pollutants for this study. Note that the fate of transformation products is not considered.

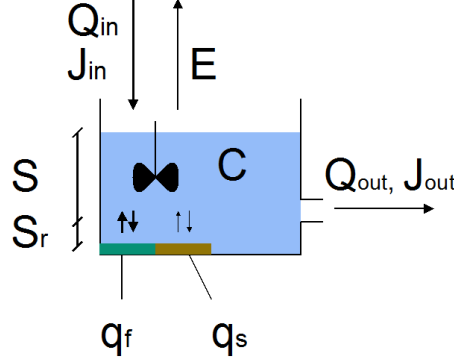


Figure 3: Basic reservoir element in SUPERFLEX, with water fluxes,  $Q$ , evapotranspiration,  $E$ , and substance fluxes,  $J$ . The exchange of substance between the dissolved and two sorbed phases is also shown; sites that are in instantaneous equilibrium,  $q_f$ , and sites for slow, kinetically limited sorption,  $q_s$ . The mixer symbol indicates complete mixing of the substance within the water volume.  $S$  denotes the variable hydrologically active water volume, and  $S_r$  is the constant volume of residual pore water.

The water balance for a generic reservoir (Fig. 3) is given by:

$$\frac{dS}{dt} = Q_{\text{in}} - E - Q_{\text{out}} \quad (1)$$

$$E = f_{\text{ET}}(S, E_{\text{pot}}, \dots) \quad (2)$$

$$Q_{\text{out}} = f_Q(S, \dots) \quad (3)$$

where  $S$  is the active storage of the reservoir,  $Q_{\text{in}}$  and  $Q_{\text{out}}$  are the inflow and outflow of water, respectively,  $E$  and  $E_{\text{pot}}$  are the actual and the potential evapotranspiration, respectively, and ' $\dots$ ' represents additional parameters of the functions  $f_{\text{ET}}$  and  $f_Q$ .  $Q_{\text{in}}$  can be precipitation or the outflow from another reservoir.

Concerning the fate of substances in the reservoir, consider the mass of dissolved substance in pore water, which is equal to  $(S + S_r)C$ , where  $S_r$  is the residual soil water content and  $C$  is the pore water concentration (Fig. 3). We distinguish fast and slow sorption; fast sorption is assumed to always be in equilibrium with a linear isotherm, i.e.  $q_f = K_{\text{d},f}C$ , where  $q_f$  is the mass of substance sorbed to fast sites per mass of soil and  $K_{\text{d},f}$  is the corresponding distribution coefficient. Let  $M_1$  be the sum of the dissolved and the fast-sorbed mass, i.e. the total mass that is in equilibrium:

$$M_1 = (S + S_r)C + m_{\text{soil}}K_{\text{d},f}C \quad (4)$$

Its counterpart, the mass that is sorbed via slow sorption processes, is given by:

$$M_2 = m_{\text{soil}}q_s \quad (5)$$

where  $q_s$  is the specific sorbed mass on slow sorption sites (mass of substance per mass of soil). The differential equations of the two compartments are given by:

$$\frac{dM_1}{dt} = Q_{\text{in}}C_{\text{in}} - Q_{\text{out}}C - J - \lambda M_1 \quad (6)$$

$$\frac{dM_2}{dt} = J - \lambda M_2 \quad (7)$$

$$J = m_{\text{soil}} r_s^* \theta (K_{\text{d},s} C - q_s) \quad (8)$$

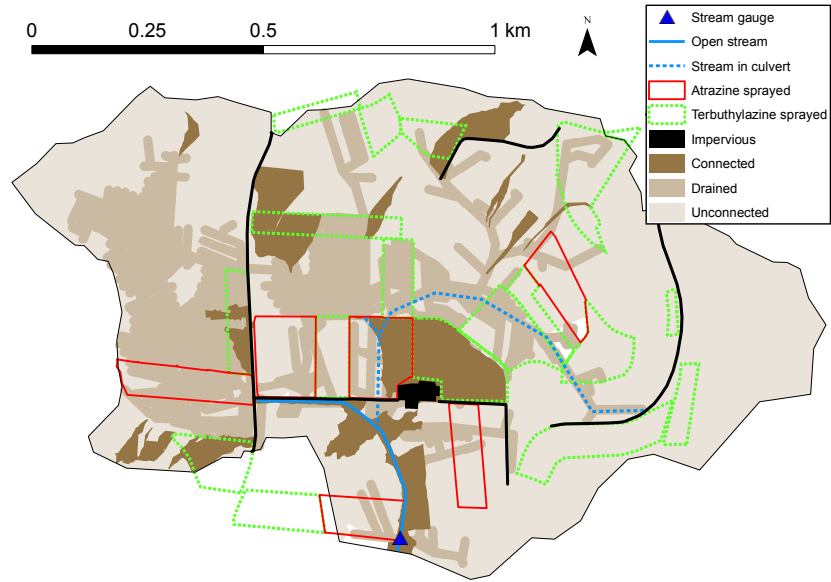
where  $J$  is the flux of substance between the equilibrium and the slowly sorbed fraction,  $r_s^*$  is the kinetic rate of sorption,  $K_{\text{d},s}$  is the distribution coefficient of the slow sorption reaction, and  $\lambda$  is the degradation rate. Note that the sorption speed is dependent on  $\theta$ , the total relative water content of the soil:  $\theta = \frac{(S+S_r)\rho_w}{m_{\text{soil}}}$ , based on findings by Gaillardon and Dur [1995] and similar to Christophersen and Wright [1981]. Note also that we assume a first order degradation of substances (Eq. 6 and 7) in the dissolved and the sorbed state in all the soil compartments (reservoirs). Based on Eq. (6) to (8) and Eq. (A.12), we construct a set of output vectors  $y_{\text{mod}} = (y_{\text{mod},1}, \dots, y_{\text{mod},4}) = (Q_{\text{out}}, C_{\text{atra}}, C_{\text{terb}}, K_{\text{d},\text{app}})$ .

Note that all the quantities that would naturally be extensive (masses  $M$  and  $m_{\text{soil}}$ ; fluxes  $Q$ ,  $E$ , and  $J$ ; volumes  $S$  and  $S_r$ ) are turned into intensive quantities through division by the area of the part of the catchment which they represent. These similarly behaving parts of the catchment are called hydrological response units (HRUs), see Sect. 3.2.2.

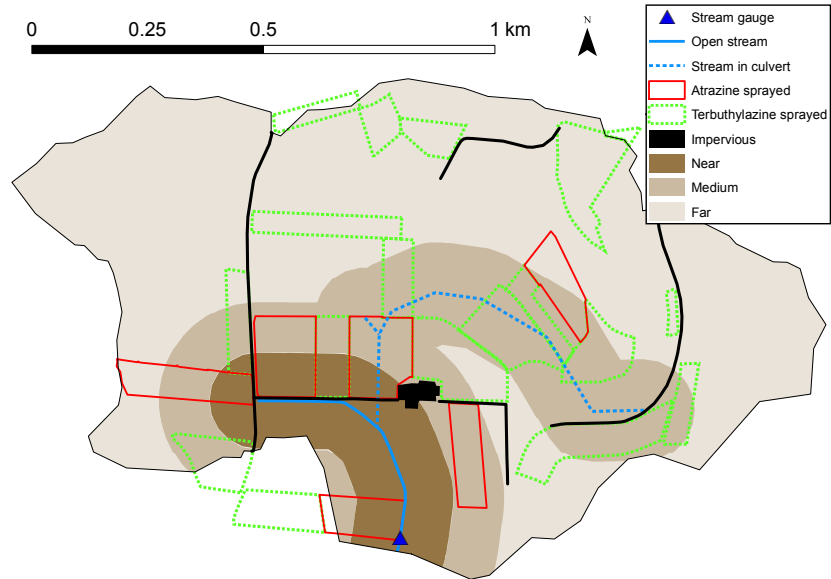
### 3.2.2 HRU delineation

Since two of the three important transport paths elicited in Sect. 3.1 are linked to spatial elements (tile drains and shortcuts), some degree of spatial distribution of the model is required. HRUs are suitable to represent the routing of fluxes through those elements, which would be more difficult to achieve with a grid-based approach.

To reflect the experimentalist knowledge, the splitting of the landscape is directly derived from the four main transport paths listed in Sect. 3.1. Paths 1) - 3) can be located in the catchment with some degree of confidence; the landscape is split according to their presence and absence (Fig. 4a), resulting in the following HRUs: impervious areas (I), areas topographically connected to a shortcut or the stream (C), drained areas (D), areas that are both connected and drained (C&D), and all the remaining “unconnected” areas (U). Path 4) (groundwater flow) is expected to be relevant in all of the HRUs except for the impervious areas; instead of influencing the division of the landscape, it is included as a parallel flow path in the HRUs listed above.



(a) Delineation of the catchment into HRUs based on experimentalist understanding of the most important transport paths (see Sect. 3.1).



(b) HRUs based on the proximity to the stream as alternative hypothesis.

Figure 4: Two sets of HRUs used as alternative hypothesis of spatial elements controlling transport processes. Note that only the fields sprayed at the first application of terbutylazine are shown. Source: swisstopo [2008], Gemeinde Ossingen [1995]

The impervious areas were delineated in ArcGIS (version 10.2.2) based on spatial vector data [swisstopo, 2008]. Only roads, settlements, and field roads that are known to drain into the stream based on spatial data or visual observations at the study site were considered as impervious areas. For the areas connected to shortcuts, we make use of the areas resulting from the topographical connectivity analysis conducted by Doppler et al. [2012]. Based on visual observations of overland flow in the catchment, we modify the original version of the connected areas by assuming that overland flow does not cross field roads and by changing the catchment polygons of some shortcuts. The resulting shortcut areas are visible in Fig. 4a. The HRU for the drained areas is obtained by applying a 10 m buffer around each tile drain, which results in a homogeneous coverage, given the typical distance of ca. 20 m between two pipes.

The second set of HRUs represents the hypothesis that spatial proximity to the stream is the most important driver for the risk of herbicide export to the stream, which is the basis for current Swiss legislation [FOAG, 2018]. The resulting set of HRUs (Fig. 4b) consists of impervious areas, which are the same as in the other set of HRUs, and areas that are near, medium and far from the stream. In order to guarantee comparability to the previously introduced set of HRUs, we use the same total coverage for the near, medium, and far areas as we did for the connected, drained, and unconnected areas, respectively. Note that the stream is partly open and partly in a culvert; the near areas are assumed to be close to the open stream only, where fast transport is possible, while the medium areas are a buffer around the near areas and the subsurface part of the stream (Fig. 4b). By selecting a buffer of width 100 m around the open channel to define the near area, we assure that it has the same total coverage than the connected areas. In analogy, a buffer of 106 m around the near area and the culvert part of the stream results in the same total area coverage of the medium and the drained areas.

Precipitation and evapotranspiration are assumed to be uniform over the catchment. The input of herbicides to each HRU is obtained by overlaying the sprayed fields with the HRUs (Fig. 4a and 4b). The resulting amount of herbicide input to all HRUs in both spatial configurations can be found in Table A.2. The input to the impervious areas (spray drift) is the same for both sets of HRUs; it is the load calculated from the observed streamflow and concentrations during event E1 (0.2 g, see Fig. 7). The effect of including or not including this mass as input is assessed within the model comparison performed in this study.

### 3.2.3 Model structure selection

While the above considerations determine how we split the landscape, this section describes the model structure applied in each of those landscape units. The full equations of all the models are given in Appendix A.3).

For the impervious areas, a simple structure with one single linear reservoir is chosen (Fig. 5). For all other areas, we differentiate between the saturated and the unsaturated zone, where the latter is split into two reservoirs representing the topsoil and the lower soil zone. This split is motivated by the large vertical gradient of substance concentrations commonly found in the top 10-20 cm [e.g. Weed et al., 1995] and by soil moisture measurements at different depths [Doppler et al., 2012] following recommendations of McMillan et al. [2010]. Motivated by visual evidence of both, infiltration excess and saturation excess overland flow in the study catchment those fluxes are considered in the basic model structure. The groundwater reservoir represents the slow hydrological response processes and is assigned a large sorption capacity (Table A.1) to reflect extended contact of water with the soil matrix along this path and, consequently, its limited contribution to in-stream concentrations (supported by baseflow samples). Due to the shallow groundwater level in parts of the catchment indicated by measurements, evapotranspiration from groundwater is enabled if all other reservoirs are empty. In addition to this basic configuration, we add a fast reservoir, the parameterization of which depends on the HRU (Fig. 5). For the connected areas, we add a linear reservoir, whereas for the drained areas we allow a non-linear response to consider that the wetting of the macropores before efficient transport might lead to a more threshold-like behaviour. The areas that are connected *and* drained are assigned the same model structure as the connected areas, assuming that the overland flow enters the shortcut or the stream directly and is not available for infiltration to macropores and tile drains anymore.

The parameters of the topsoil, the lower unsaturated zone, and the groundwater reservoirs are chosen to be identical in the different HRUs, which means that those reservoirs behave exactly the same in all the HRUs in which they are present. This is motivated by the fact that the topsoil texture is rather homogeneous [Doppler et al., 2014b, FAL, 1997] and a uniform underlying moraine material and Freshwater molasse formation span the whole catchment [Einsele, 2000, swisstopo, 2007]. Furthermore, agricultural practice does not show any systematic differences across space and the spatial heterogeneity of sorption and degradation rates was shown to be rather limited in the study catchment [Doppler et al., 2014b] so that there is no obvious reason to assume systematic differences in soil properties between different HRUs. Linking parameters between HRUs by introducing global parameters is an efficient and reasonable way of reducing the size of the parameter space in semi-distributed hydrological models [e.g. Fenicia et al., 2016].

Table 2 lists the tested models, which are based on combinations of the HRUs delineated in Sect. 3.2.2 and the model structures derived in this section. The naming convention is as follows: “exp” and “prx” stand for the HRU delineation based on catchment-specific experimentalist knowledge (Fig. 4a) and proximity (Fig. 4b), respectively. “top” follows the experimentalist approach, but directly uses the connected areas resulting from the topographical connec-

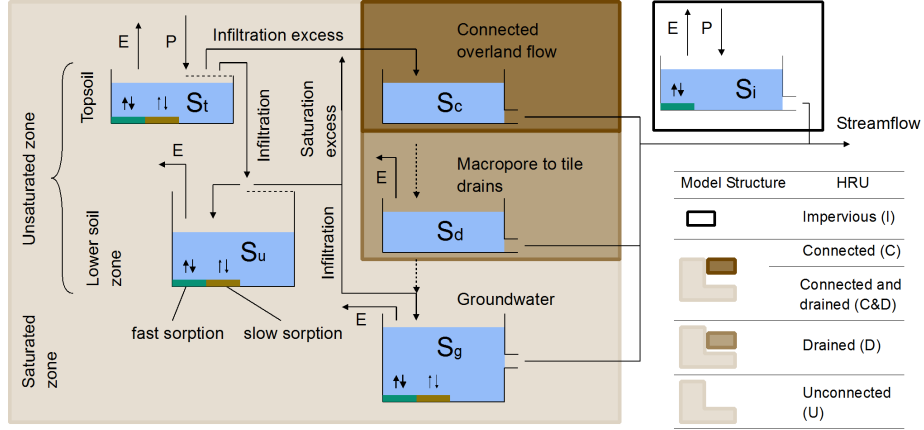


Figure 5: Combination of reservoirs that form the model structure of each HRU. All reservoirs of a certain type (e.g.  $S_u$ ) have the same parameterization in all the HRUs in which they are present. The flows indicated by the dashed arrows are only active if the reservoir at the arrow's origin is not present; in this case the flux is forwarded to the next reservoir. Dashed lines on top of reservoirs mean that a fraction of the incoming flux is directly forwarded, depending on the state of the reservoir.

Table 2: Overview of the tested model configurations. The abbreviations for the HRUs are: U: unconnected, I: impervious, C: connected, D: drained, N, M, F: areas that are near, medium and far from the stream, respectively.  $U \rightarrow (C, D, U)$  means that the model structure of U (Fig. 5) is applied to the combined HRUs C, D, and U.

Model <sup>a)</sup>	No. of parameters	Mod. Struct. $\xrightarrow{\text{applied to}}$ HRUs
MexpH1	14	$D \rightarrow (I, C, D, U)$
MexpH2	15	$I, D \rightarrow (C, D, U)$
MexpH3a	14	$I, C, U \rightarrow (D, U)$
MexpH3b	15	$I, D, U \rightarrow (C, U)$
MexpH4	16	$I, C, D, U$
MtopH4	16	$I, C^b), D, U$
MprxH4	16	$I, C \rightarrow N,$ $D \rightarrow M, U \rightarrow F$

<sup>a)</sup> Each of the listed models has an equivalent version in which spray drift is neglected, which is denoted by appending “\_nsd” to the model name.

<sup>b)</sup> This uses the connected areas originally presented in Doppler et al. [2012].



tivity analysis of Doppler et al. [2012] without modifications. The number of HRUs (e.g. “H2”) is also given in the model name. With MexpH1-4, we test different degrees of complexity of the spatial configuration based on experimentalist knowledge (Fig. 4a) in terms of number of HRUs considered. MexpH1 is fully lumped, hypothesizing that all the areas in the catchment behave similarly and can be described with the model structure of the drained areas (see Fig. 5). MexpH2 contains a minimal degree of spatial distribution by distinguishing impervious and pervious areas. MexpH3a and MexpH3b additionally consider the connected or the drained areas, respectively, as separate HRUs. MexpH4 uses the full complexity of the experimentalist HRU arrangement (Fig. 4a). MprxH4, on the other hand, follows the more conventional alternative hypothesis that spatial proximity to the stream (Fig. 4b) is the most important driver for the risk of fast transport. It only differs from MexpH4 in the spatial arrangement of the HRUs and therefore in the herbicide input to each HRU (precipitation is assumed to be homogeneous over the catchment); the model structures in all HRUs are the same as for MexpH4. All the models above are run with and without spray drift; the latter case is denoted by appending `_nd` for “no spray drift” to the model name.

### 3.3 Uncertainty assessment

#### 3.3.1 Probabilistic model formulation

We construct a probabilistic model for the observations by adding a stochastic error term to the output of the deterministic transport model. This stochastic error term accounts for all the uncertainty in the model output in a lumped way. For this, we use an approach that was recently developed in a hydrological setting [Ammann et al., 2019], but which can naturally be extended to water quality modelling. It is based on specifying the marginal distribution,  $D_{Y_k}$ , for the observation of type  $k$  at time  $t$ ,  $Y_{\text{obs},k}(t)$ , which depends on the model output,  $y_{\text{mod},k}(t)$ , and some additional parameters of the distribution,  $\psi_k$ , so that  $Y_{\text{obs},k}(t) \sim D_{Y_k}(y_{\text{mod},k}(t), \psi_k)$ . For the detailed theoretical description and an example of application, see Ammann et al. [2019].

We choose  $D_{Y_k} = \text{N}$  (where N is the Gaussian distribution) for all model outputs  $k$  (streamflow, atrazine, terbuthylazine and soil/water distribution coefficients), which is characterized by the mean,  $\mu_k$ , and the standard deviation,  $\sigma_k$ . The mean is set to be the output of the deterministic model, i.e.  $\mu_k(t) = y_{\text{mod},k}(t)$ , and the standard deviation is a function of the model output in order to consider heteroscedasticity:

$$\sigma_k(t) = a_k y_{0,k} \left( \frac{y_{\text{mod},k}(t)}{y_{0,k}} \right)^{c_k} + y_{0,k} b_k \quad (9)$$

where  $y_0$  is a typical value of the modelled quantity, and  $a$ ,  $b$  and  $c$  are error model parameters.  $y_0$  is introduced for practical reasons to make  $a$  and  $b$  unitless and independent of  $c$ , and thus more universal. In addition, we consider autocorrelation of the stochastic model deviations by a random process that has one single parameter,  $\tau$ , as described in Ammann et al. [2019]. Therefore, the parameters of our error model are  $\psi_k = (a_k, b_k, c_k, \tau_k, y_{0,k})$ , reflecting our different expectation of the uncertainty [Reichert and Schuwirth, 2012] w.r.t. the four target variables used for calibration (Table 1). For streamflow, we use the error model denoted as “E3” in Ammann et al. [2019], whereas for the other target variables, we use “E1”. Table A.1 shows which of the error model parameters are inferred and which are kept fixed. The complete likelihood function is given in Appendix A.5.

### 3.3.2 Inference procedure

Inference and forward prediction is performed according to the procedure described in Ammann et al. [2019]. For each of the models listed in Table 2, we perform Bayesian inference using the observations listed in Table 1, where we used streamflow and concentrations recorded at the most downstream gauging station only (Fig. 1). Data of the remaining stations is used to assess the spatial distribution of predicted and measured loss rates on the individual experimental fields. The calibration period extends from 1 April to 12 October 2009. As a burn-in, we add one month before that period. A local sensitivity analysis reveals that the influence of the initial condition of the groundwater reservoir completely decays within one month. The lower soil zone reservoir is initialized at 90 % of its capacity based on inspection of soil moisture data [Doppler et al., 2012]. Due to the short concentration time series, which is limited to a few rainfall events, we refrain from validation in an independent time period.

We apply an affine-invariant ensemble Markov Chain Monte Carlo technique [Foreman-Mackey et al., 2013] to obtain a sample from the full posterior distribution. An ensemble of 50 particles is started from independently sampled initial values for each parameter. The initial range is chosen wide enough to cover a reasonable spectrum of the parameters. Convergence of the chains is checked visually, the burn-in is removed, and the final sample size is 25,000 for each posterior distribution. The procedure for prediction is described in more detail in Appendix A.5.

### 3.3.3 Prior distribution

Appendix A.4 lists the independent one- and two-dimensional prior distributions for all the parameters used in this study. Note that some of the parameters related to a specific reservoir are only present if that reservoir is present in the

model configuration (see Fig. A.2 for which reservoir is present in which HRU and Appendix A.3 for which parameter belongs to which reservoir). However, if a parameter is present in a model, it always has the prior distribution given in Appendix A.4.

The priors for the hydrological parameters are largely based on the authors’ modelling experience with the kind of hydrological models applied in this study. For example,  $\phi_e$  is likely to be close to unity, since it is a correction factor for the potential evapotranspiration. A uniform distribution is used to reflect the lack of knowledge about the split parameter,  $D$ , but a lower limit of 0.5 is chosen to prevent potential identifiability problems between the fast and the slow reservoirs. The priors for the reservoir release rates  $k_c$ ,  $k_d$ , and  $k_g$  reflect our expectation of a typical time scale of the respective reservoirs response time.

The parameters with a more chemical nature are based on literature values, our understanding of the chemical processes in the study catchment, or lab measurements. For example, we neglect the difference in chemical properties of atrazine and terbuthylazine by assuming only one set of chemical parameters for both of the substances (similar degradation rates and sorption properties are reported e.g. in [Lewis et al., 2016]). The layer that feeds fast transport processes (topsoil layer), is expected to be 5 ( $\pm 5$ ) cm thick, which has to be multiplied by a typical effective porosity of 0.4 ( $\pm 0.1$ ) to get the maximal water content of this layer,  $S_{t,max}$ . As a prior for the sorption parameters  $S_{t,1}$  and  $S_{t,2}$  we choose a 2-dimensional lognormal distribution to reflect a dependence between the two parameters, motivated by Altfelder and Streck [2006], who found a typical value of approx. 1/3 for the fraction of fast sorption sites and by Karickhoff and Morris [1985], who found similar values between 0.2 and 0.38. For the kinetic rate of sorption,  $r_s$ , we use a relatively wide prior, but we limit it to the region given by a typical time scale of a few days or weeks that we expect for equilibration to happen.

### 3.4 Performance evaluation

The developed models are assessed through qualitative visual inspection, as well as quantitative performance measures. The first is applied to the hydrograph, chemographs, as well as internal states of the models, which are compared to our understanding of the catchments behaviour. Thereby we focus on the deterministic model output obtained with the maximum posterior parameter estimates.

The quantitative performance metrics, on the other hand, provide information about the quality of the full predictive distribution of a model. More specifically, they quantify to which degree the observed data are compatible with a model’s predictive distribution (henceforth called reliability) [e.g. Bröcker, 2009, McInerney et al., 2017] and how broad this distribution is (spread) [e.g. Gneiting et al., 2007, McInerney et al., 2017]. In combination, those metrics provide

concise and meaningful information about the quality of a sample of stochastic predictions in the form of time series. Here, we use the slightly adapted and renamed versions presented in Ammann et al. [2019]. The reliability is large if a large part of a model’s predictions are centered at the observations. The spread, on the other hand, quantifies the width of the predictive distribution relative to the magnitude of the observations. Good performance is therefore indicated by a large reliability *in combination* with a small spread. For more detailed information on the performance metrics, see Appendix A.6.

## 4 Results

### 4.1 Model comparison

Figure 6 shows the reliability and spread of all tested models in the calibration period for streamflow, atrazine, terbuthylazine, and the average performance with respect to those variables. Note that the two metrics are not calculated for the distribution coefficient,  $K$ , due to the limited number of measurements. Figure 7 shows the hydrographs and chemographs obtained with the maximum posterior parameter set of MexpH2, MexpH4, and MprxH4, and compares them to the observed values. The posterior parameter distributions of all models are given in Appendix A.9 and information on some selected internal states can be found in Appendix A.10.

The performance of the models regarding streamflow (Fig. 6a) varies less than the performance regarding atrazine (Fig. 6b) and terbuthylazine (Fig. 6c). For streamflow, the lumped models (1 HRU) have a large spread of over 0.75, while models with 2 or more HRUs clearly show a better performance with spreads that are all close to 0.5 (Fig. 6a). Fig. 7 confirms the similar performance of the more complex models w.r.t. streamflow. The performance (the spread in particular) of the different models is more variable for atrazine and terbuthylazine than for streamflow (Fig. 6b,c, Fig. 7).

Models considering spray drift show an overall better performance than models neglecting spray drift (Fig. 6d). In particular, they tend to have a smaller spread than models without spray drift, and the performance of each individual model improves when adding spray drift (Fig. 6c). The better overall performance is based on the herbicide concentrations (Fig. 6b,c). For streamflow, the models considering spray drift have actually a lower reliability than the ones neglecting it (Fig. 6a).

The lumped models have a poor performance that clearly deviate from the other models (Fig. 6d). Models with 2 HRUs already show a strongly improved performance w.r.t. streamflow (Fig. 6a) which is confirmed by the visual in-

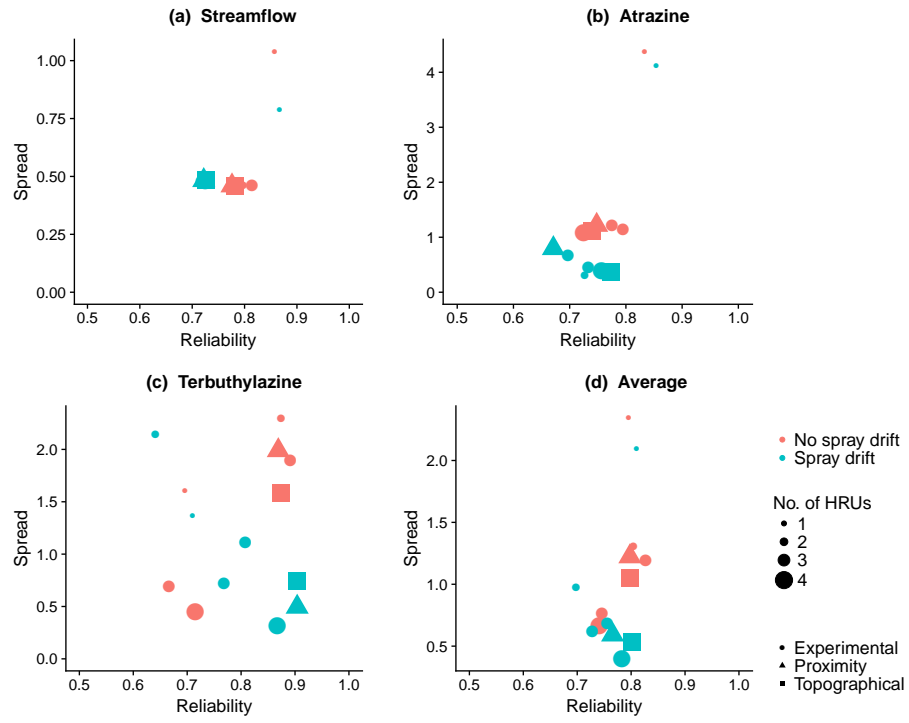


Figure 6: Performance metrics resulting from the inference of all tested models. “Average” means the arithmetic mean of the performance metrics regarding streamflow, atrazine, and terbutylazine.

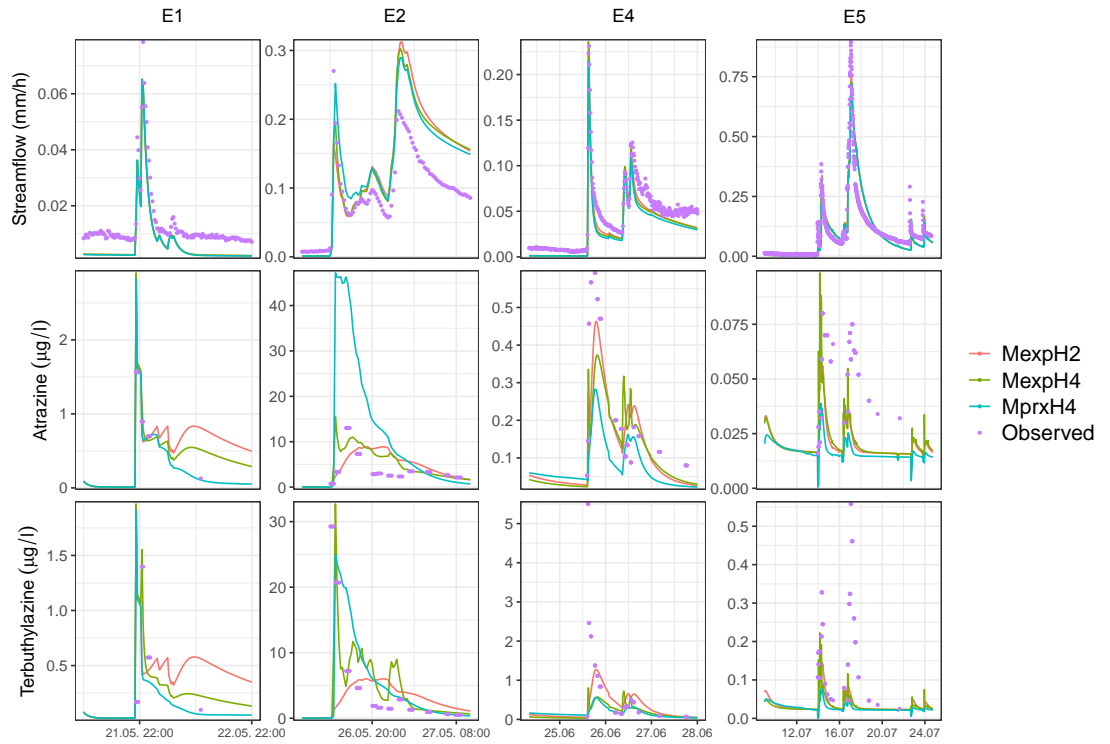


Figure 7: Comparison of the observed and modelled streamflow, atrazine, and terbutylazine concentrations during events E1, E2, E4, and E5. Note that the application of terbutylazine is not known exactly after event E2, which is why this data was not used in the calibration process.

spection of hydrographs (Fig. 7). MexpH2 fits the atrazine chemograph rather well, but shows substantial deficits w.r.t. terbuthylazine concentrations (Fig. 7, Fig. 6c). MexpH2 produces very similar chemographs for atrazine and terbuthylazine, failing to reproduce the measured earlier and sharper peak of terbuthylazine compared to atrazine during event E2. The performance of the “H3” models strongly increase w.r.t. terbuthylazine (Fig. 6c), but not w.r.t streamflow and atrazine compared to the models with 2 HRUs (Fig. 6a,b). The models with 4 HRUs show the best overall performance (Fig. 7) with relatively high reliability and low spread (Fig. 6d), given that they consider spray drift.

Regarding the different spatial configurations of the HRUs, the “experimentalist” and the “topographical” setup tend to show a better performance than the “proximity” setup in terms of herbicide concentrations (Fig. 6b,c). Model MprxH4 overestimates the atrazine concentrations during event E2 by a factor of 3-4 (Fig. 7). In terms of streamflow, MprxH4 leads to similar results as MexpH4 (Fig. 7).

While the marginals of the posterior parameter distribution of all the tested models can be found in the Appendix, we provide a summary in the following. The inferred value of the hydrological parameters are similar for all the tested models except for MexpH1. This is in agreement with the very similar performance metrics regarding streamflow for those models (Fig. 6a). Typical values of the evaporation multiplier,  $\phi_e$ , are around 1.5, the infiltration capacity,  $P_{ex}$ , is ca. 47 mm/h and the split between the groundwater and the fast reservoir,  $D$ , is ca. 0.95 in all the mentioned models. The maximal water content of the topsoil layer,  $S_{t,max}$ , and the lower unsaturated zone,  $S_{u,max}$ , are around 10 mm and 150 mm, respectively. Note that the former is also the lower bound we set for that parameter (see Table A.1). Regarding the parameters of chemical nature, the fitted degradation rate,  $\lambda$ , corresponds to a half-life time of ca. 10 days in the case of MexpH4 and is within 5 - 10 days for the other models. The sorption parameters,  $S_{t,1}$  and  $S_{t,2}$ , both result in values of around 20 mm, which corresponds to a distribution coefficient of  $K_{d,app} \approx 1.9$  l/kg according to Eq. (A.10), which compares well with the measured distribution coefficients of 0.7 - 3 l/kg in the first two weeks of the study period [Camenzuli, 2010]. The parameters that mainly cause the different performances for atrazine and terbuthylazine among the tested models are the ones of the fast responding reservoirs in the HRUs representing connected and drained areas,  $k_c$  (Eq. A.26),  $k_d$ , and  $\alpha_d$  (Eq. A.29).

## 4.2 Output uncertainty and loss rates

Figure 8 shows the output uncertainty to be expected for predictions in streamflow, atrazine, and terbuthylazine concentrations during events E2-E5 with MexpH4, based on the 90%-confidence intervals given by the stochastic predic-

tions obtained with the full probabilistic model after inference (see Sect. 3.3.2). The modelled relative export of atrazine and terbuthylazine via streamflow at the catchment outlet predicted by MexpH4 and MprxH4 is visible in Fig. 9. MexpH4 predicts that over the whole study period, the relative export amounts to between 0.2 and 0.5 %, and between 0.4 and 0.8% of the applied mass for atrazine and terbuthylazine, respectively. The former range of values agrees well with the measured overall atrazine loss rate of 0.26 % [Doppler et al., 2014b]. The contribution of the individual events to atrazine exports are very different: a large part of the atrazine mass is exported during event E2, which is the first major event after spraying, while the other events export a much smaller fraction (Fig. 9a). Note that there were two applications of terbuthylazine and the exported mass in the different events is given as a fraction of the mass applied up to that time, which is why the total exported fraction can be lower than the one for event E2 (Fig. 9b). The resulting spatial distribution of atrazine loss rates at the individual experimental fields is less biased with MexpH4 than with MprxH4 (Fig. 10a). The terbuthylazine loss rates, for which no measurements are available, are also distinctively different between MexpH4 and MprxH4 (Fig. 10b).

Figure 9c and 9d show the relative exported mass of atrazine and terbuthylazine over the whole study period split according to the HRUs. The connected and the drained areas export approx. 0.4-0.8 % of the mass of atrazine applied to them, while the HRU representing unconnected areas exports less than 0.03 % of the mass applied to it. The mass of atrazine exported from impervious areas according to the maximum posterior model output amounts to ca. 0.14 g or 70 % of the assumed amount of spray drift.

## 5 Discussion

### 5.1 Model building and inference

A large part of the available prior knowledge was incorporated in the division of the landscape (Sect. 3.2.2) and the choice of the model structure (Sect. 3.2.3); only a small part of the prior knowledge was used for the definition of the actual prior distribution of parameters. It is important to be aware of this, in particular when applying more complex physically based models “off the shelf”, in which case the splitting of the landscape and the choice of model structure are potentially more implicit.

Although the experimentalist knowledge provides guidance in the delineation of HRUs in this study, the exact extent of each HRU is believed to be associated with uncertainty; it likely depends on time-varying factors like the wetness of the catchment, crop growth stages, plowing practice, microstructures, etc., that



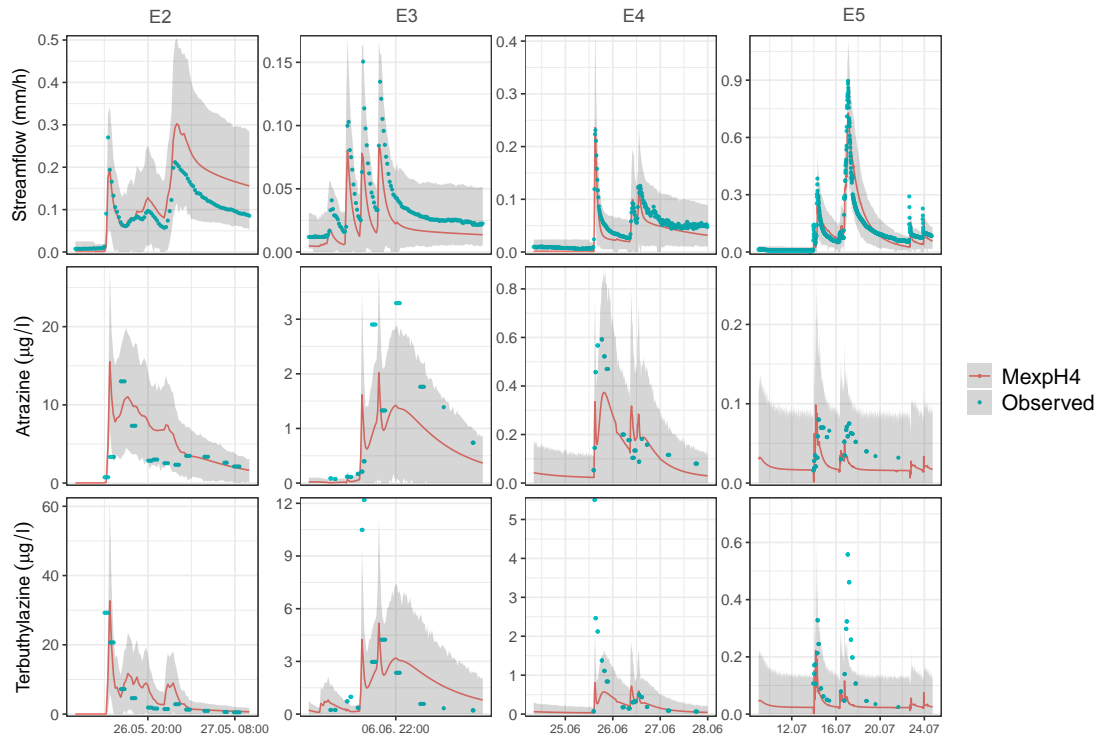


Figure 8: Maximum posterior model output (red line) and 90%-confidence bands for streamflow, atrazine, and terbutylazine during events E2-E5. Note that the terbutylazine concentrations of events E3-E5 were not used for calibration due to uncertain input.

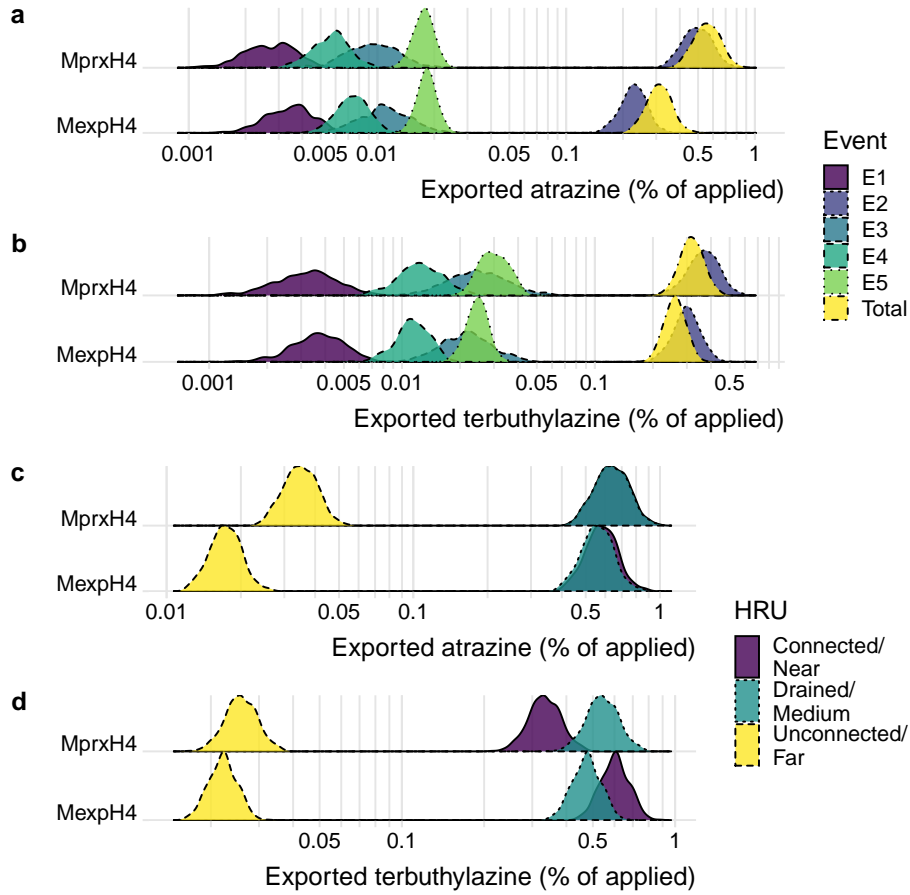


Figure 9: Modelled exported fraction of atrazine and terbuthylazine. a, b: during events E1-E5 and during the total study period. c, d: split according to HRUs. Note the log-transformed x-axis. The y-axis shows the density of the logarithm of the exported amount. The terbuthylazine fraction is given relative to the amount applied in the first application (events E1, E2) and relative to the sum of the two applications (events E3, E4, E5, Total).

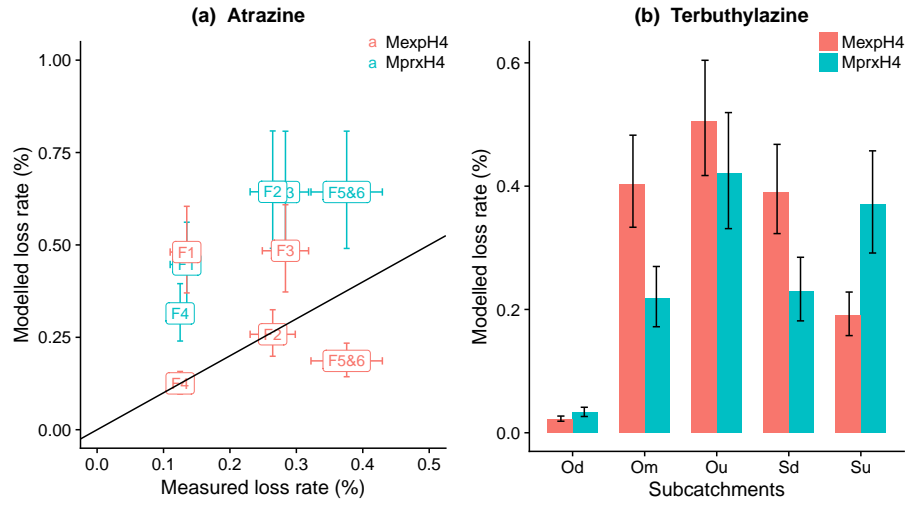


Figure 10: Relative herbicide loss rate to the stream as a fraction of the mass applied to the different experimental fields. (a): observed [Doppler et al., 2012] and modelled loss rate of atrazine at experimental fields F1-F6, (b): modelled loss rate of terbutylazine in the subcatchments defined by the gauging stations. Error bars for modelled and measured loss rates indicate 90 %-confidence intervals and the uncertainty of the load calculation in Doppler et al. [2012], respectively.

can potentially alter flow paths. This requires us to make some assumptions on the spatial extent of the active processes that are difficult to test. However, the rather similar performance of MtopH4 and MexpH4 (Fig. 6d), which use two different versions of connected areas, indicate that the results might be robust w.r.t. the delineation of the catchment of artificial shortcuts, which is a promising result for up-scaling of the experimentalist informed HRUs that consider artificial shortcuts. Soil thickness and slope potentially affect hydrological and transport processes. Nonetheless, their spatial variation was not considered when delineating the HRUs, in order to ensure model parsimony. Another assumption to be aware of is the absence of fluxes across the boundaries of HRUs, which is not very realistic at the small spatial scale of this study. In fact, Doppler et al. [2012] found evidence for substance fluxes via overland flow across the HRU boundaries delineated in this study in at least one instance. However, these fluxes are not believed to be excessive compared to the fluxes from the HRUs to the stream and the chosen approach is seen as a necessary compromise to ensure model parsimony.

## 5.2 Model comparison

The results in Sect. 4.1 highlight the need to distinguish the impervious areas directly connected to the stream as a separate HRU, with a characteristically different hydrological behaviour, and to acknowledge the limited amount of substance input to those areas. The improved performance with increasing spatial resolution is in agreement with a previous study [Ghirardello et al., 2014]. The results also indicate that the heterogeneity of fast paths relevant for herbicide transport require rather detailed information about a catchment like the location of potentially active artificial shortcuts and tile drains. The models that consider only one of those elements also show reduced performance, in particular for terbuthylazine (Fig. 6c), which supports the assumption that both of those elements are important transport paths.

## 5.3 Uncertainty and extrapolation

Note that the output uncertainty presented in Sect. 4.2 is conditional on the assumptions related to the model structures of MexpH4 and MprxH4, the prior distribution of the parameters, and the characteristics of the error model (in particular the fixed exponent,  $c$ , Eq. 9). Furthermore, the applied likelihood function assumes that the errors in one output variable are independent of the other output variables, which is not necessarily the case, but is a common assumption in multi-objective modelling problems [e.g. Hong et al., 2005, Talamba et al., 2010, Han and Zheng, 2016, 2018]. The uncertainty originating from the unknown application date of the second load of terbuthylazine is not considered in this study.

Limited availability of additional data sets similar to the one used in this study prevented a strict validation of the developed models. Therefore the validity of a spatio-temporal extrapolation remains to be tested in future work. Still, the chosen site is in many regards typical for the Swiss Plateau; i.e. topography, climate, land use, tile drained area, and presence of artificial shortcuts. The structure of MexpH4 (and MtopH4) is therefore believed to be suitable to describe the behaviour of other small catchments in the Swiss Plateau. The findings indicate, however, that high-resolution digital elevation models and site-specific knowledge about artificial shortcuts and tile drains might be necessary to achieve accurate representation of high-frequency chemographs. These are not readily available in many catchments, which impedes the delineation of HRUs similar to the ones in this study. Parameter regionalization could also be hampered by factors that are catchment-specific like soil type and organic matter content, soil management, slope, erosion risk, etc. The extrapolation to other substances is believed to be less problematic, since the model’s parameters for sorption and degradation properties can be related to measurable quantities like distribution coefficients (Sect. A.2.2) and half-life times. Validation with independent data would certainly be needed for a more accurate estimate of the extrapolation error.

## 5.4 Methodological pitfalls

We encountered some methodological pitfalls that we believe possess some generality and are therefore of potential interest for future water quality modelling studies. First, pollutant concentrations and loads can easily span many orders of magnitude. In that case, a careful choice of the error model parameters (see Eq. 9) is necessary in order to accurately reflect the expected errors in the whole range of possible model outputs. If we require a reasonable relative accuracy also for small modelled concentrations, the model can become sensitive to “missed” concentration peaks, which have very large relative errors. In our case, we could adapt the input in the form of spray drift to correct for an otherwise missed peak, but in other cases more heavy tailed distributions than the Gaussian might be necessary [e.g. Schoups and Vrugt, 2010, Ammann et al., 2019] for a reduction of the sensitivity to very strong outliers. Certainly, more research is needed on realistic uncertainty estimates with probabilistic models that allow for a robust inference process w.r.t. water quality data.

Second, the comparison of models with and without spray drift (Fig. 6) shows the sensitivity to herbicide input onto fast-responding elements like impervious areas, which possess a high risk for mobilization and transport, at least during the first event after spraying. Properly accounting for the uncertainty related to this input would require more elaborate techniques. While some methods to consider input uncertainty have been established in hydrology [e.g. Kavetski et al., 2003, Göttinger and Bárdossy, 2008], they have only recently been trans-

lated to the large input uncertainties of water quality modelling [Han and Zheng, 2016]. Note that in our study, due to the controlled herbicide application, the input uncertainties are much smaller than in conventional settings.

Third, reservoirs containing high concentrations of substances (i.e. the top soil) are supported by field observations [Doppler et al., 2012] but can be problematic in the modelling procedure. Especially when the model simulates very low streamflow, the simulated in-stream concentration is highly sensitive to the outflow of reservoirs that have high internal concentrations of a compound. In a bucket-model it is important to limit the transport of tracer along that path during baseflow conditions, which might require special parameterizations depending on the property of the reservoirs. In our case, this required a strongly non-linear constitutive function of the top soil reservoir (Eq. A.19), so that it does not contribute to the streamflow during baseflow conditions and small precipitation events.

## 6 Conclusions

In this study, we showed how a perceptual model of herbicide fate in a small agricultural catchment, which is based on prior expert knowledge, can inform and constrain the construction of a conceptual transport model in combination with a Bayesian uncertainty quantification setting. Model deficits were identified and improved based on a discussion between modellers and experimentalists. Some alternative hypothesis, in the form of different spatial model configurations were tested to assess the degree of complexity of site-specific knowledge needed. Based on this, we reach the following conclusions:

1. The available experimentalist knowledge mainly informed the division of the landscape into HRUs and in the design of the model structure. Its use in the definition of the prior parameter distribution was less pronounced. Modellers experience was mainly used as complementary information in the model structure selection and in the definition of the prior distributions of model parameters.
2. In the translation of the perceptual to the conceptual model, we found that the model bias w.r.t. the data could strongly be reduced when including spray drift on impervious areas. This is a strong indication of spray drift during the controlled herbicide application experiment in the study catchment, even if the estimated amount (0.2 g) was rather small. This also shows that experimentalist knowledge can be useful in identification and improvement of model deficits in an iterative procedure.
3. While completely lumped models fail to jointly reproduce observed streamflow and concentration time series, a minimal degree of spatial distribution

by separate treatment of impervious areas is sufficient to capture stream-flow and atrazine concentrations. However, the large heterogeneity of dominant transport processes, which is revealed by the markedly different chemograph of terbuthylazine, is more elusive. It can be partially captured once detailed knowledge about the location of tile drains and potentially active shortcuts is used.

4. The experimentalist-informed division of the landscape into connected and drained areas reduces the uncertainty by 30 % and results in a better spatial distribution of loss rates compared to the conventional way that is based on proximity to the stream. Together with finding 2 this indicates that site-specific knowledge of elements like artificial shortcuts and tile drains might be important for accurate estimations of exposure risks in catchments with sizes similar to the one in this study.
5. Regarding uncertainty quantification, the standard deviation of the model output of MexpH4 in the calibration period is in the order of 50 % of the observed values for streamflow, atrazine, and for terbuthylazine (Fig. 6a,b,c). The uncertainty of herbicide input to fast responding areas is a major factor that would ideally be accounted for separately, preferably with independent validation of the model.

## Acknowledgements

The authors would like to thank Reynold Chow for the detailed and valuable feedback he provided to this manuscript at multiple stages, and Volker Prasuhn and Sebastian Stoll for the feedback they provided to the general modelling approach chosen in this study. Bogdan Caradima kindly provided language help for this manuscript. Furthermore, we want to thank the Federal Office of Meteorology MeteoSwiss and the Federal Office of Topography swisstopo for the data they provided.

Funding: This work was supported by the Swiss National Science Foundation [grant number 200021\_163322].

## References

- S. Altfelder and T. Streck. Capability and limitations of first-order and diffusion approaches to describe long-term sorption of chlortoluron in soil. *Journal of Contaminant Hydrology*, 86(3-4):279–298, aug 2006. doi: 10.1016/j.jconhyd.2006.04.002.
- L. Ammann, F. Fenicia, and P. Reichert. A likelihood framework for deterministic hydrological models and the importance of non-stationary autocorrela-

- tion. *Hydrology and Earth System Sciences*, 23(4):2147–2172, apr 2019. doi: 10.5194/hess-23-2147-2019.
- J. G. Arnold, R. Srinivasan, R. S. Muttiah, and J. R. Williams. Large area hydrologic modelling and assessment; part I: model development. *Journal of the American Water Resources Association*, 34(1):73–89, feb 1998. doi: 10.1111/j.1752-1688.1998.tb05961.x.
- E. Bertuzzo, M. Thomet, G. Botter, and A. Rinaldo. Catchment-scale herbicides transport: Theory and application. *Advances in Water Resources*, 52:232–242, feb 2013. ISSN 03091708. doi: 10.1016/j.advwatres.2012.11.007.
- K. Beven and A. Binley. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6:279–298, 1992.
- K. J. Beven. *Rainfall-Runoff Modelling: The Primer*. Wiley-Blackwell, 2012. ISBN 978-0-470-71459-1.
- K. J. Beven, P. Smith, and J. Freer. Comment on “Hydrological forecasting uncertainty assessment: incoherence of the {GLUE} methodology” by {P}ietro {M}antovan and {E}nzio {T}odini. *Journal of Hydrology*, 338(3-4):315–318, 2007.
- J. Bröcker. Reliability, sufficiency, and the decomposition of proper scores. *Quarterly Journal of the Royal Meteorological Society*, 135(643):1512–1519, jul 2009. doi: 10.1002/qj.456.
- L. Camenzuli. *Degradation of Herbicides in Soils, Kinetics and Limiting Factors in Field and Laboratory*. PhD thesis, Swiss Federal Institute of Technology (ETH), Zürich, 2010.
- R. F. Carsel, L. A. Mulkey, M. N. Lorber, and L. B. Baskin. The Pesticide Root Zone Model ({PRZM}): A procedure for evaluating pesticide leaching threats to groundwater. *Ecological Modelling*, 30(1-2):49–69, 1985. doi: 10.1016/0304-3800(85)90036-5.
- N. Christophersen and R. F. Wright. Sulfate budget and a model for sulfate concentrations in stream water at Birkenes, a Small forested catchment in southernmost Norway. *Water Resources Research*, 17(2):377–389, apr 1981. doi: 10.1029/wr017i002p00377.
- T. Doppler, L. Camenzuli, G. Hirzel, M. Krauss, A. Lück, and C. Stamm. Spatial variability of herbicide mobilisation and transport at catchment scale: insights from a field experiment. *Hydrology and Earth System Sciences*, 16(7):1947–1967, jul 2012. doi: 10.5194/hess-16-1947-2012.
- T. Doppler, M. Honti, U. Zihlmann, P. Weisskopf, and C. Stamm. Validating a spatially distributed hydrological model with soil morphology data. *Hydrology and Earth System Sciences*, 18(9):3481–3498, sep 2014a. doi: 10.5194/hess-18-3481-2014.



- T. Doppler, A. Lück, L. Camenzuli, M. Krauss, and C. Stamm. Critical source areas for herbicides can change location depending on rain events. *Agriculture, Ecosystems & Environment*, 192:85–94, jul 2014b. doi: 10.1016/j.agee.2014.04.003.
- G. Einsele. *Sedimentary Basins: Evolution, Facies, and Sediment Budget*. Springer, 2000. ISBN 978-3-540-66193-1. URL <https://www.amazon.com/Sedimentary-Basins-Evolution-Facies-Sediment/dp/354066193X?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=sm2&camp=2025&creative=165953&creativeASIN=354066193X>.
- G. Evin, M. Thyer, D. Kavetski, D. McInerney, and G. Kuczera. Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity. *Water Resources Research*, 50(3):2350–2375, 2014. doi: 10.1002/2013wr014185.
- FAL. Bodenkarte Kanton Zürich 1:5000, 1997.
- F. Fenicia, J. J. McDonnell, and H. H. G. Savenije. Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resources Research*, 44(6), jun 2008. doi: 10.1029/2007wr006386.
- F. Fenicia, D. Kavetski, and H. H. G. Savenije. Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11):1–13, nov 2011. doi: 10.1029/2010wr010174.
- F. Fenicia, D. Kavetski, H. H. G. Savenije, and L. Pfister. From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions. *Water Resources Research*, 52(2):954–989, feb 2016. doi: 10.1002/2015wr017398.
- FOAG. Weisungen betreffend der {M}assnahmen zur {R}eduktion der {R}isiken bei der {A}nwendung von {P}flanzenschutzmitteln, 2018. URL <https://www.blw.admin.ch/blw/de/home/nachhaltige-produktion/pflanzenschutz/pflanzenschutzmittel/nachhaltige-anwendung-und-risikoreduktion.html>.
- D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. {emcee}: The {MCMC} hammer. *Publications of the Astronomical Society of the Pacific*, 125:306–312, 2013.
- G. A. Fox, G. J. Sabbagh, R. W. Malone, and K. Rojas. Modeling Parent and Metabolite Fate and Transport in Subsurface Drained Fields With Directly Connected Macropores1. *{JAWRA} Journal of the American Water Resources Association*, 43(6):1359–1372, sep 2007. doi: 10.1111/j.1752-1688.2007.00116.x.

- L. G. Freitas, H. Singer, S. Müller, R. Schwarzenbach, and C. Stamm. Source area effects on herbicide losses to surface waters—A case study in the Swiss Plateau. *Agriculture, Ecosystems & Environment*, 128(3):177–184, nov 2008. doi: 10.1016/j.agee.2008.06.014.
- M. P. Frey, M. K. Schneider, A. Dietzel, P. Reichert, and C. Stamm. Predicting critical source areas or diffuse herbicide losses to surface waters: Role of connectivity and boundary conditions. *Journal of Hydrology*, 365:23–36, 2009.
- P. Gaillardon and J. C. Dur. Influence of soil moisture on short-term adsorption of diuron and isoproturon by soil. *Pesticide Science*, 45(4):297–303, 1995. doi: 10.1002/ps.2780450402.
- A. I. Gärdenäs, J. Šimůnek, N. Jarvis, and M. van Genuchten. Two-dimensional modelling of preferential water flow and pesticide transport from a tile-drained field. *Journal of Hydrology*, 329(3-4):647–660, 2006. doi: 10.1016/j.jhydrol.2006.03.021.
- K. K. Gardner, B. L. McGlynn, and L. A. Marshall. Quantifying watershed sensitivity to spatially variable N loading and the relative importance of watershed N retention mechanisms. *Water Resources Research*, 47(8), aug 2011. doi: 10.1029/2010wr009738.
- M. Gassmann, C. Stamm, O. Olsson, J. Lange, K. Kümmerer, and M. Weiler. Model-based estimation of pesticides and transformation products and their export pathways in a headwater catchment. *Hydrology and Earth System Sciences*, 17(12):5213–5228, 2013. doi: 10.5194/hess-17-5213-2013.
- Gemeinde Ossingen. Drainagekarte 1:1000, 1995.
- A. Ghafoor, J. Koestel, M. Larsbo, J. Moeyes, and N. Jarvis. Soil properties and susceptibility to preferential solute transport in tilled topsoil at the catchment scale. *Journal of Hydrology*, 492:190–199, jun 2013. doi: 10.1016/j.jhydrol.2013.03.046.
- D. Ghirardello, M. Morselli, S. Otto, G. Zanin, and A. D. Guardo. Investigating the need for complex vs. simple scenarios to improve predictions of aquatic ecosystem exposure with the {SoilPlus} model. *Environmental Pollution*, 184: 502–510, jan 2014. doi: 10.1016/j.envpol.2013.10.002.
- T. Gneiting, F. Balabdaoui, and A. E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, apr 2007. doi: 10.1111/j.1467-9868.2007.00587.x.
- J. Götzinger and A. Bárdossy. Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research*, 44(12), nov 2008. doi: 10.1029/2007wr006691.

- P. Grathwohl. *Diffusion in Natural Porous Media: Contaminant Transport, Sorption/Desorption and Dissolution Kinetics*. Springer {US}, 1998. doi: 10.1007/978-1-4615-5683-1.
- F. Han and Y. Zheng. Multiple-response Bayesian calibration of watershed water quality models with significant input and model structure errors. *Advances in Water Resources*, 88:109–123, feb 2016. doi: 10.1016/j.advwatres.2015.12.007.
- F. Han and Y. Zheng. Joint analysis of input and parametric uncertainties in watershed water quality modeling: A formal Bayesian approach. *Advances in Water Resources*, 116:77–94, jun 2018. doi: 10.1016/j.advwatres.2018.04.006.
- M. M. Hantush and A. Chaudhary. Bayesian Framework for Water Quality Model Uncertainty Estimation and Risk Management. *Journal of Hydrologic Engineering*, 19(9):4014015, sep 2014. doi: 10.1061/(asce)he.1943-5584.0000900.
- H. M. Holländer, T. Blume, H. Bormann, W. Buytaert, G. B. Chirico, J.-F. Exbrayat, D. Gustafsson, H. Hölzel, P. Kraft, C. Stamm, S. Stoll, G. Blöschl, and H. Flühler. Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data. *Hydrology and Earth System Sciences*, 13(11):2069–2094, nov 2009. doi: 10.5194/hess-13-2069-2009.
- B. Hong, R. L. Strawderman, D. P. Swaney, and D. A. Weinstein. Bayesian estimation of input parameters of a nitrogen cycle model applied to a forested reference watershed, Hubbard Brook Watershed Six. *Water Resources Research*, 41(3), 2005. doi: 10.1029/2004wr003551.
- N. J. Jarvis, P.-E. Jansson, P. E. Dik, and I. Messing. Modelling water and solute transport in macroporous soil. I. Model description and sensitivity analysis. *Journal of Soil Science*, 42(1):59–70, 1991. doi: 10.1111/j.1365-2389.1991.tb00091.x.
- N. Kannan, S. M. White, F. Worrall, and M. J. Whelan. Pesticide Modelling for a Small Catchment Using {SWAT}-2000. *Journal of Environmental Science and Health, Part B*, 41(7):1049–1070, 2006. doi: 10.1080/03601230600850804.
- S. W. Karickhoff and K. R. Morris. Sorption dynamics of hydrophobic pollutants in sediment suspensions. *Environmental Toxicology and Chemistry*, 4(4):469–479, aug 1985. doi: 10.1002/etc.5620040407.
- D. Kavetski and F. Fenicia. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research*, 47(11):1–19, nov 2011. doi: 10.1029/2011wr010748.
- D. Kavetski, S. W. Franks, and G. Kuczera. Confronting input uncertainty in environmental modelling. In *Water Science and Application*, pages 49–68. American Geophysical Union, San Francisco, 2003. doi: 10.1029/ws006p0049.

- J. W. Kirchner. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, 42(3), 2006. doi: 10.1029/2005wr004362.
- C. Leu, H. Singer, C. Stamm, S. R. Müller, and R. P. Schwarzenbach. Simultaneous Assessment of Sources, Processes, and Factors Influencing Herbicide Losses to Surface Waters in a Small Agricultural Catchment. *Environmental Science & Technology*, 38(14):3827–3834, jul 2004. doi: 10.1021/es0499602.
- C. Leu, H. Singer, S. R. Müller, R. P. Schwarzenbach, and C. Stamm. Comparison of Atrazine Losses in Three Small Headwater Catchments. *Journal of Environment Quality*, 34(5):1873, 2005. doi: 10.2134/jeq2005.0049.
- K. A. Lewis, J. Tzilivakis, D. J. Warner, and A. Green. An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment: An International Journal*, 22(4):1050–1064, jan 2016. doi: 10.1080/10807039.2015.1133242.
- Q. Ma, R. D. Wauchope, L. Ma, K. W. Rojas, R. W. Malone, and L. R. Ahuja. Test of the Root Zone Water Quality Model({RZWQM}) for predicting runoff of atrazine, alachlor and fenamiphos species from conventional-tillage corn mesoplots. *Pest Management Science*, 60(3):267–276, feb 2004. doi: 10.1002/ps.846.
- P. Mantovan and E. Todini. Hydrological forecasting uncertainty assessment: Incoherence of the {GLUE} methodology. *Journal of Hydrology*, 330(1-2): 368–381, 2006. doi: 10.1016/j.jhydrol.2006.04.046.
- P. Mantovan, E. Todini, and M. L. V. Martina. Reply to comment by {K}eith {B}even, {P}aul {S}mith and {J}im {F}reer on “Hydrological forecasting uncertainty assessment: incoherence of the {GLUE} methodology”. *Journal of Hydrology*, 338:319–324, 2007.
- J. M. Marín-Benito, V. Pot, L. Alletto, L. Mamy, C. Bedos, E. Barriuso, and P. Benoit. Comparison of three pesticide fate models with respect to the leaching of two herbicides under field conditions in an irrigated maize cropping system. *Science of The Total Environment*, 499:533–545, nov 2014. doi: 10.1016/j.scitotenv.2014.06.143.
- D. McInerney, M. Thyer, D. Kavetski, J. Lerat, and G. Kuczera. Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors. *Water Resources Research*, 53(3):2199–2239, 2017. doi: 10.1002/2016wr019168.
- H. K. McMillan, M. P. Clark, W. B. Bowden, M. Duncan, and R. A. Woods. Hydrological field data from a modeller’s perspective: Part 1. Diagnostic tests for model structure. *Hydrological Processes*, 25(4):511–522, nov 2010. doi: 10.1002/hyp.7841.

- MeteoSwiss. No Title, apr 2016. URL <https://www.meteoschweiz.admin.ch/home/service-und-publikationen/beratung-und-service/datenportal-fuer-experten.html>.
- B. Meyer, J.-Y. Pailler, C. Guignard, L. Hoffmann, and A. Krein. Concentrations of dissolved herbicides and pharmaceuticals in a small river in Luxembourg. *Environmental Monitoring and Assessment*, 180(1-4):127–146, nov 2010. doi: 10.1007/s10661-010-1777-9.
- M. Morselli, C. M. Vitale, A. Ippolito, S. Villa, R. Giacchini, M. Vighi, and A. D. Guardo. Predicting pesticide fate in small cultivated mountain watersheds using the {DynAPlus} model: Toward improved assessment of peak exposure. *Science of The Total Environment*, 615:307–318, feb 2018. doi: 10.1016/j.scitotenv.2017.09.287.
- N. Munz, C. Leu, and I. Wittmer. Pestizidmessungen in Fliessgewässern. *Aqua & Gas*, 11:32–41, 2012.
- H. B. Pionke, W. J. Gburek, A. N. Sharpley, and R. R. Schnabel. Flow and nutrient export patterns for an agricultural hill-land watershed. *Water Resources Research*, 32(6):1795–1804, jun 1996. doi: 10.1029/96wr00637.
- K. J. Raat, J. A. Vrugt, W. Bouten, and A. Tietema. Towards reduced uncertainty in catchment nitrogen modelling: quantifying the effect of field observation uncertainty on model calibration. *Hydrology and Earth System Sciences*, 8(4):751–763, 2004. ISSN 1607-7938. doi: 10.5194/hess-8-751-2004. URL <http://www.hydrol-earth-syst-sci.net/8/751/2004/>.
- P. Reichert and N. Schuwirth. Linking statistical description of bias to multi-objective model calibration. *Water Resources Research*, 48:W09543, doi:10.1029/2011WR011391, 2012.
- F. G. Renaud, P. H. Bellamy, and C. D. Brown. Simulating pesticides in ditches to assess ecological risk ({SPIDER}): I. Model description. *Science of The Total Environment*, 394(1):112–123, 2008. doi: 10.1016/j.scitotenv.2007.11.038.
- M. Rode, G. Arhonditsis, D. Balin, T. Kebede, V. Krysanova, A. van Griensven, and S. E. A. T. M. van der Zee. New challenges in integrated water quality modelling. *Hydrological Processes*, 24(24):3447–3461, nov 2010. ISSN 08856087. doi: 10.1002/hyp.7766. URL <http://doi.wiley.com/10.1002/hyp.7766>.
- G. Schoups and J. A. Vrugt. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46:W10531, doi:10.1029/2009WR008933, 2010.

- J. Seibert and J. J. McDonnell. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research*, 38(11):1241, doi:10.1029/2001WR000978, 2002.
- P. Singh and R. L. Jones. Comparison of pesticide root zone model 3.12: Runoff predictions with field data. *Environmental Toxicology and Chemistry*, 21(8): 1545–1551, aug 2002. doi: 10.1002/etc.5620210803.
- J. R. Stedinger, R. M. Vogel, S. U. Lee, and R. Batchelder. Appraisal of the generalized likelihood uncertainty estimation ({GLUE}) method. *Water Resources Research*, 44(12), nov 2008. doi: 10.1029/2008wr006822.
- Swisstopo. Geologischer Atlas der Schweiz, Bundesamt für Landestopographie (Art.30 Geo IV): 5704 000 000, reproduced by permission of swisstopo/ JA100119, 2007.
- Swisstopo. Vector25, Bundesamt für Landestopographie (Art.30 Geo IV): 5704 000 000, reproduced by permission of swisstopo/ JA100119, 2008.
- D. B. Talamba, E. Parent, and A. Musy. Bayesian multiresponse calibration of {TOPMODEL}: Application to the Haute-Mentue catchment, Switzerland. *Water Resources Research*, 46(8), aug 2010. doi: 10.1029/2007wr006449.
- U. Traub-Eberhard, W. Kördel, and W. Klein. Pesticide movement into subsurface drains on a loamy silt soil. *Chemosphere*, 28(2):273–284, jan 1994. doi: 10.1016/0045-6535(94)90124-4.
- M. L. Villamizar and C. D. Brown. A modelling framework to simulate river flow and pesticide loss via preferential flow at the catchment scale. *CATENA*, 149:120–130, feb 2017. doi: 10.1016/j.catena.2016.09.009.
- D. A. J. Weed, R. S. Kanwar, D. E. Stoltenberg, and R. L. Pfeiffer. Dissipation and Distribution of Herbicides in the Soil Profile. *Journal of Environment Quality*, 24(1):68, 1995. doi: 10.2134/jeq1995.00472425002400010010x.
- C. Wellen, G. B. Arhonditsis, T. Long, and D. Boyd. Quantifying the uncertainty of nonpoint source attribution in distributed water quality models: A Bayesian assessment of {SWAT}’s sediment export predictions. *Journal of Hydrology*, 519:3353–3368, nov 2014. doi: 10.1016/j.jhydrol.2014.10.007.
- S. Wrede, F. Fenicia, N. Mart\`inez-Carreras, J. Juilleret, C. Hissler, A. Krein, H. H. G. Savenije, S. Uhlenbrook, D. Kavetski, and L. Pfister. Towards more systematic perceptual model development: a case study using 3 Luxembourgish catchments. *Hydrological Processes*, 29(12):2731–2750, 2014. doi: 10.1002/hyp.10393.
- E. Zehe and H. Flühler. Preferential transport of isoproturon at a plot scale and a field scale tile-drained site. *Journal of Hydrology*, 247(1-2):100–115, jun 2001. doi: 10.1016/S0022-1694(01)00370-5.

- Y. Zheng and F. Han. Markov Chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management. *Stochastic Environmental Research and Risk Assessment*, 30(1):293–308, jan 2016. ISSN 1436-3259. doi: 10.1007/s00477-015-1091-8. URL <https://doi.org/10.1007/s00477-015-1091-8>.
- Y. Zheng and A. A. Keller. Uncertainty assessment in watershed-scale water quality modeling and management: 1. Framework and application of generalized likelihood uncertainty estimation ({GLUE}) approach. *Water Resources Research*, 43(8), aug 2007. doi: 10.1029/2006wr005345.