A Supplementary data

A.1 Data pre-processing

The calculation timestep of the model was chosen to be 15 min, which is equal to the resolution of the precipitation measurements in the catchment. The amount of sprayed substance was given per field [Doppler et al., 2012] and had to be converted to a concentration in the precipitation falling on each HRU on the day of application, since substance enters the catchment via precipitation in SUPERFLEX. Daily data of potential evapotranspiration were obtained through services of MeteoSwiss [2016] for the Schaffhausen weather station, which is located 11 km north of the study catchment. The data was disaggregated to 15 min resolution by assuming a sinusoidal shape from sunrise to sunset on each day.

The concentrations measured in the stream remained unprocessed and the model output was linearly interpolated to the time points at which the samples were taken. Since some measurements consisted of a mixture of 3 time-proportional samples during high-flow conditions [Doppler et al., 2012], the predicted concentration was interpolated to the sampling times and then averaged to ensure comparability with the corresponding measured concentration for the inference process. The streamflow observations, generally available at 5 min resolution or finer, were aggregated to 15 min resolution by taking the average of all the point observations within each 15 min interval to match the calculation time step. The distribution coefficients measured by Camenzuli [2010] showed long-term increasing trends on some of the fields. The kinetically limited sorption process included in the model, on the other hand, is meant to represent processes that have an equilibration time in the order of days, which means that a long-term increase in apparent distribution coefficients cannot be modelled. Therefore, the positive trend in the observations was removed by assuming a constant value, which was the mean of the measured distribution coefficients during the first two weeks after application.

A.2 Modelling sorption and degradation in SUPERFLEX

A.2.1 Literature review

There have been a multitude of studies on sorption processes of organic pollutants in soils. A comprehensive, yet concise, review of the sorption models used to describe pesticide sorption processes is given by Wauchope et al. [2002]. General agreement exists that the kinetic rate of sorption of organic compounds can be clustered into a fast and a slow group [Karickhoff and Morris, 1985]. The former is considerably faster than the usual time scale of interest (say faster than 1 day), hence the reaction is often observed to be in equilibrium, while the latter is considerably slower and equilibration times are in the order of months. Mathematical descriptions of the kinetics of the slower reaction can be classified into first-order models [e.g. Fortin et al., 1997] or spherical diffusion approaches [e.g. Altfelder and Streck, 2006]. The latter have been shown to lead to rate

constants that were more robust and less time-dependent than those determined with a fist-order model [Altfelder and Streck, 2006]. However, first-order models are conceptually considerably simpler and have been widely applied, often yielding good results, which is why the sorption model of choice for this study is the first-order model. We furthermore assume a linear relationship between the sorbed and the dissolved concentration (Freundlich parameter $n = 1$) for the slow sorption process. Oftentimes, n was found to be somewhat smaller than unity, typically between 0.7 and 1 [Wauchope et al., 2002].

The sorption speed and the equilibrium distribution coefficient depend on the chemical properties of the substances, as well as soil characteristics like organic matter content, soil moisture, and pH [e.g. Grathwohl, 1990, Brusseau and Rao, 1991, Rutherford et al., 1992, Gaillardon, 1996]. Except for the soil moisture, those factors are assumed to be constant in time and space over the studied catchment. The effect of soil moisture on sorption speed and equilibrium is believed to be dependent on the hydrophobicity of the substance [Roy et al., 2000], but is still insufficiently understood [Roy et al., 2000].

A.2.2 Modelling sorption and degradation with partially mixed reservoirs

This section specifies how the fast and the slow sorption processes are implemented in SUPERFLEX via completely and partially mixed zones in the generic reservoir. Figure A.1 shows the generic reservoir with a total of three zones: the hydrologically active zone and two hydrologically inactive zones, which are completely and partially mixed. They are denoted as h, z1 and z2, respectively. Zones h and z1 are completely mixed, ensuring that $C \equiv C_{\mathrm{h}} \equiv C_{\mathrm{z1}}$. Zone z1 enables the conceptual modelling of fast sorption processes. With the addition of zone 2, which is only partially mixed with the upper two zones, we allow for the modelling of slow sorption processes. This means that in general, $C \neq C_{\mathrm{z2}}$.

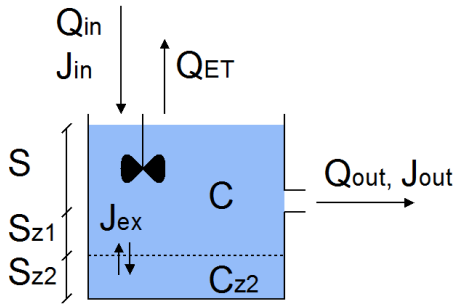

Figure A.1: Schematic representation of a partially mixed reservoir, which is a basic element in SUPERFLEX. Fluxes of water (Q) and tracer (J), water levels (S), and concentrations (C) are also indicated.

$M_{\mathrm{h,z1}}$ is the mass in the upper (completely mixed) zone of the reservoir and

M_{z2} is the mass in the lower (partially mixed) zone:

$$M_{h,z1} = (S + S_{z1})C \tag{A.1}$$

$$M_{z2} = S_{z2}C_{z2} \tag{A.2}$$

Their temporal evolution is described by the following differential equations:

$$\frac{dM_{h,z1}}{dt} = Q_{in}C_{in} - Q_{out}C - J_{ex} - \lambda M_{h,z1} \tag{A.3}$$

$$\frac{dM_{z2}}{dt} = J_{ex} - \lambda M_{z2} \tag{A.4}$$

where

$$J_{ex} = r_s(S + S_r)(C - C_{z2}) \tag{A.5}$$

is the flux of substance exchanged between the upper and the lower zone and λ is the first order degradation rate.

In the following, we express the hydrological variables above as a function of the chemical variables in Sect. 3.2.1:

$$(A.1) \leftrightarrow (4) \qquad S_{z1} = S_r + m_{soil}K_{d,f} \tag{A.6}$$

$$(A.2) \leftrightarrow (5) \qquad C_{z2} = \frac{m_{soil}}{S_{z2}}q_s \tag{A.7}$$

$$(A.4) \leftrightarrow (7) \qquad r_s = r_s^\star \rho_w K_{d,s} \tag{A.8}$$

$$S_{z2} = K_{d,s}m_{soil} \tag{A.9}$$

Note that the total distribution coefficient considering fast and slow sorption, K_d, can be obtained from Eq. (A.6) and (A.9) as

$$K_d = K_{d,f} + K_{d,s} = \frac{(S_{z1} - S_r + S_{z2})}{m_{soil}} \tag{A.10}$$

Obviously, these theoretical sorption coefficients are a function of constants only.

The concept of the apparent sorption coefficients arises from observed distribution of the substance after a limited equilibration time, in this case 24 h [Camenzuli, 2010], under static laboratory conditions. Therefore, the apparent distribution coefficient is additionally a function of the porewater and sorbed concentrations at the beginning of the experiment, the expermiental equilibration time and the kinetic rate of adsorption or desorption. The apparent distribution coefficient at any time, t, is given by:

$$K_{d,app}(t) = \frac{q_f(t) + q_s(t)}{C(t)} \tag{A.11}$$

In the model, the apparent distribution coefficient at any time is given by:

$$K_{d,app}(t) = \frac{C(t)(S_{z1}(t) - S_r) + C_{z2}(t)S_{z2}(t)}{C(t)} \frac{n_e}{\rho_{bulk}S_{max}} \tag{A.12}$$

Comparing Eq. (A.11) and (A.12) we can see that the mass of dissolved substance in the residual pore water, CS_r has to be subtracted from M_{z1} to get the mass sorbed to fast sorption sites, q_f.

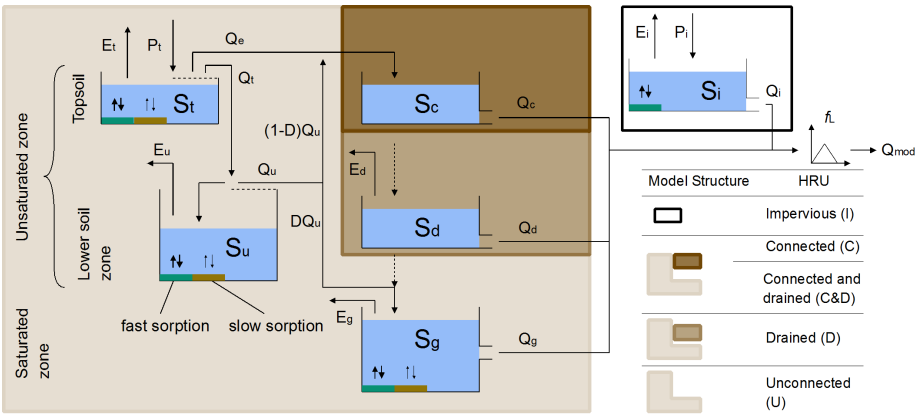

Figure A.2: Arrangement of the reservoirs and the fluxes in all the HRUs, including the variable names of the fluxes. Dashed arrows indicate fluxes that are only present in model structures in which the reservoir at their origin is absent.

A.3 Model structures and equations

This section provides detailed information on the models applied in this study. The equations of all the fluxes of water and tracers entering and leaving the specific reservoirs are provided in the following. The internal evolution of the water level, S, of each reservoir is given by Eq. (1). The differential equation of the concentration, C, is given by Eq. (6) and additionally by Eq. (A.4) if slow sorption is considered.

Impervious reservoir:

$$E_\mathrm{i} = \phi_\mathrm{e} E_\mathrm{pot} \left(1 - \exp\left(-\frac{S_\mathrm{i}}{m_\mathrm{e}}\right)\right) \tag{A.13}$$

$$Q_\mathrm{i} = k_\mathrm{i} S_\mathrm{i} \tag{A.14}$$

$$J_\mathrm{i} = C_\mathrm{i} Q_\mathrm{i} \tag{A.15}$$

Topsoil reservoir:

$$E_\mathrm{t} = \phi_\mathrm{e} E_\mathrm{pot} \left(1 - \exp\left(-\frac{S_\mathrm{t}}{m_\mathrm{e}}\right)\right) \tag{A.16}$$

$$Q_\mathrm{e} = \begin{cases} 0 & \text{if} \quad P_\mathrm{t} < P_\mathrm{ex} \\ P_\mathrm{t} - P_\mathrm{ex} & \text{if} \quad P_\mathrm{t} \geq P_\mathrm{ex} \end{cases} \tag{A.17}$$

$$J_\mathrm{e} = C_\mathrm{t} Q_\mathrm{e} \tag{A.18}$$

$$Q_\mathrm{t} = (P_\mathrm{t} - Q_\mathrm{e})\left(1 - \frac{(1 - \frac{S_\mathrm{t}}{S_\mathrm{t,max}})(1 + m_\mathrm{e})}{1 - \frac{S_\mathrm{t}}{S_\mathrm{t,max}} + m_\mathrm{e}}\right) \tag{A.19}$$

$$J_\mathrm{t} = C_\mathrm{t} Q_\mathrm{t} \tag{A.20}$$

Lower soil zone reservoir:

$$E_{\mathrm{u}} = (\phi_{\mathrm{e}} E_{\mathrm{pot}} - E_{\mathrm{t}}) \frac{\frac{S_{\mathrm{u}}}{S_{\mathrm{u,max}}}(1 + m_{\mathrm{e}})}{(\frac{S_{\mathrm{u}}}{S_{\mathrm{u,max}}} + m_{\mathrm{e}})} \tag{A.21}$$

$$Q_{\mathrm{u}} = Q_{\mathrm{t}} \left(\frac{S_{\mathrm{u}}}{S_{\mathrm{u,max}}} \right)^{\beta_{\mathrm{u}}} \tag{A.22}$$

$$J_{\mathrm{u}} = C_{\mathrm{t}} Q_{\mathrm{u}} \tag{A.23}$$

$$Q_{\mathrm{u,b}} = k_{\mathrm{u,b}} S_{\mathrm{u}} \tag{A.24}$$

$$J_{\mathrm{u,b}} = C_{\mathrm{u}} Q_{\mathrm{u,b}} \tag{A.25}$$

Connected overland flow reservoir:

$$Q_{\mathrm{c}} = k_{\mathrm{c}} S_{\mathrm{c}} \tag{A.26}$$

$$J_{\mathrm{c}} = C_{\mathrm{c}} Q_{\mathrm{c}} \tag{A.27}$$

Macropore flow to tile drain reservoir:

$$E_{\mathrm{d}} = (\phi_{\mathrm{e}} E_{\mathrm{pot}} - E_{\mathrm{t}} - E_{\mathrm{u}}) \left(1 - \exp\left(-\frac{S_{\mathrm{f}}}{m_{\mathrm{e}}} \right) \right) \tag{A.28}$$

$$Q_{\mathrm{d}} = k_{\mathrm{d}} S_{\mathrm{d}}^{\alpha_{\mathrm{d}}} \tag{A.29}$$

$$J_{\mathrm{d}} = C_{\mathrm{d}} Q_{\mathrm{d}} \tag{A.30}$$

Groundwater reservoir:

$$E_{\mathrm{g}} = (\phi_{\mathrm{e}} E_{\mathrm{pot}} - E_{\mathrm{t}} - E_{\mathrm{u}} - E_{\mathrm{d}}) \left(1 - \exp\left(-\frac{S_{\mathrm{g}}}{m_{\mathrm{g,e}}} \right) \right) \tag{A.31}$$

$$Q_{\mathrm{g}} = k_{\mathrm{g}} S_{\mathrm{g}}^{\alpha_{\mathrm{g}}} \tag{A.32}$$

$$J_{\mathrm{g}} = C_{\mathrm{g}} Q_{\mathrm{g}} \tag{A.33}$$

$$\tag{A.34}$$

The water and substance flux at the catchment outlet is the sum of the contributions of all the HRUs, weighted by the area of the HRUs. The contribution of each HRU, in turn, consists of the sum of the outflow of all the reservoirs that discharge into the stream:

$$Q_{\mathrm{mod}} = f_{\mathrm{L}} * \left(\sum_i w_i \sum_k Q_k^{(i)} \right) \tag{A.35}$$

$$J_{\mathrm{mod}} = f_{\mathrm{L}} * \left(\sum_i w_i \sum_k J_k^{(i)} \right) \tag{A.36}$$

$$w_i = \frac{A_i}{\sum_j A_j} \tag{A.37}$$

where i is the index of the HRUs, k is the index of the reservoirs that discharge to the stream in a given HRU, A is the surface area, and f_{L} is a convolution

operator conceptualizing routing processes in the catchment. We choose a very simple operator: $f_{\mathrm{L}}(\tau) = \delta(\tau - t_{\mathrm{L}})$, where δ is the Dirac delta function, i.e. $(f_{\mathrm{L}} * x)(t)$ just shifts x forward in time by t_{L}. Q_{mod} can be directly compared to the observed streamflow Q_{obs} and the modelled concentration at the outlet of the catchment is simply

$$C_{\mathrm{mod}} = \frac{J_{\mathrm{mod}}}{Q_{\mathrm{mod}} \sum_j A_j} \tag{A.38}$$

which can be compared to C_{obs}.

In order to obtain model output for $K_{\mathrm{d,app}}$ that is comparable to observed values, we reproduce the experimental procedure of Camenzuli [2010] in the model. Starting from the state of the topsoil reservoir at the time points when the soil samples were taken, the soil is saturated up to the point where $(S + S_{\mathrm{r}})\frac{\rho_{\mathrm{w}}}{m_{\mathrm{soil}}} = 0.8(S_{\mathrm{max}} + S_{\mathrm{r}})$ by adding tracer-free water, where $S_{\mathrm{max}} + S_{\mathrm{r}}$ represents the water holding capacity. Subsequent equilibration of the dissolved and the sorbed phases take place for 24 hours with no fluxes into and out of the soil sample (i.e. no inflow and outflow of water and tracers and no evapotranspiration, S remains constant). The resulting apparent distribution coefficient, $K_{\mathrm{d,app}}$ (Eq. A.12), after 24 hours can then be compared to the one obtained in the lab experiment, i.e. the observed one. This allows us to include the observed distributions of the herbicides beween the pore water and adsorbed phase in the calibration procedure. Equivalent to Camenzuli [2010], we assume $\rho_{\mathrm{bulk}} = 1.2\ \mathrm{g/cm}^3$. In addition, we assume that $n_{\mathrm{e}} = 0.4$ and that $S_{\mathrm{r}} = 0.06 S_{\mathrm{z1}}$.

A.4 Prior distribution and fixed parameters

We choose independent prior distributions for most of the inferred parameters (Table A.1). For $S_{\mathrm{t,1}}$ and $S_{\mathrm{t,2}}$, we use a 2-dimensional truncated lognormal distribution with $\mu = (3, 3.8)$ and covariance matrix:

$$\Sigma = \begin{bmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{bmatrix}$$

in the log-space. The distribution is truncated at 0 and 7 (in the log-space); outside of this interval, it is zero.

In some reservoirs, the parameters regarding the degree of sorption and degradation are estimated based the assessment of internal concentrations in the reservoirs in connection with expert knowledge. Those parameters are kept fixed during the calibration procedure. We assume a very small amount of fast sorption sites ($S_{\mathrm{i,1}} = 1$ mm, Table A.1) on impervious areas, so that spray drift is not washed off by insignificantly small precipitation events. A large amount of fast and slow sorption sites is chosen for the lower soil zone and the groundwater reservoirs (Table A.1), to consider the extended contact of water with soil and thus with sorption sites in those reservoirs. Sorption is neglected in the fast reservoirs of the HRUs representing connected and drained areas (Fig. 5). Degradation in the groundwater reservoir is considered with a

fixed degradation rate, λ_{g} (Table A.1), which is relatively small due to reduced abundance and activity of microbes in this layer [Rodríguez and Harkin, 1997, Schwab et al., 2006]. The chosen value corresponds to a half-life time of 100 days, which is in agreement with experimental evidence [Schwab et al., 2006].

Most of the error model parameters are fixed (see Table A.1) based on the expected magnitude of the error as a function of model output and based on the authors' experience with the applied likelihood framework [Ammann et al., 2019].

A.5 Likelihood function and prediction

The likelihood function used in this study is adopted from Ammann et al. [2019]. More detailed information and the derivation of the likelihood function. By substituting the arbitrary distribution D_Q in Eq. (7) in Ammann et al. [2019] by a normal distribution, we obtain the following conditional probability (or probability density) of observing output $y(t_{\mathrm{i}})$ given the observation at the last time step, $y(t_{\mathrm{i}-1})$ and the parameters of the hydrological, θ, and the error model, ψ:

$$
\begin{aligned}
&\text{if } y(t_{i-1}) > 0: \\
&p_i\big(y(t_i) \mid y(t_{i-1}), \boldsymbol{\theta}, \boldsymbol{\psi}\big) \\
&\quad = \begin{cases} \sigma(t_i)^{-1} f_{\mathrm{N}\left(\eta(t_{i-1})\exp\left(-\frac{t_i - t_{i-1}}{\tau(t_i)}\right), \sqrt{1-\exp\left(-2\frac{t_i - t_{i-1}}{\tau(t_i)}\right)}\right)}(\eta(t_i)) & \text{for } y(t_i) > 0 \\ F_{\mathrm{N}\left(\eta(t_{i-1})\exp\left(-\frac{t_i - t_{i-1}}{\tau(t_i)}\right), \sqrt{1-\exp\left(-2\frac{t_i - t_{i-1}}{\tau(t_i)}\right)}\right)}(\eta(t_i)) & \text{for } y(t_i) = 0 \end{cases} \\
&\text{if } y(t_{i-1}) = 0: \\
&p_i\big(y(t_i) \mid y(t_{i-1}), \boldsymbol{\theta}, \boldsymbol{\psi}\big) \\
&\quad = \begin{cases} f_{\mathrm{N}\left(y_{\mathrm{mod}}(t_i), \sigma(t_i)\right)}\big(y(t_i)\big) & \text{for } y(t_i) > 0 \\ F_{\mathrm{N}\left(y_{\mathrm{mod}}(t_i), \sigma(t_i)\right)}(0) & \text{for } y(t_i) = 0 \end{cases}
\end{aligned}
\tag{A.39}
$$

where F and f are the cumulative distribution function (cdf) and probability density function (pdf) of the distribution given in the index, $\mathrm{N}(a, b)$ denotes the normal distribution with mean a and standard deviation b, y_{mod} is the modelled output, σ is the standard deviation of the error given in Eq. (9), τ is the characteristic correlation time, and $\eta(t_i) = (y(t_i) - y_{\mathrm{mod}}(t_i))/\sigma(t_i)$. The likelihood function for a certain type of observation (e.g. streamflow) is then obtained by multiplying the individual terms given by Eq. (A.39):

$$
f_y\big(y(t_0), y(t_1), \ldots, y(t_n) \mid \boldsymbol{\theta}, \boldsymbol{\psi}\big) = p_{\mathrm{N}(y_{\mathrm{mod}}(t_0), \sigma(t_0))}\big(y(t_0)\big) \prod_{i=1}^{n} p_i\big(y(t_i) \mid y(t_{i-1}), \boldsymbol{\theta}, \boldsymbol{\psi}\big)
\tag{A.40}
$$

Table A.1: Prior distributions used for the hydrological and error model parameters. The distribution types are abbreviated as follows; LN: lognormal, U: uniform, Exp: exponential, "Fixed": parameter was fixed at the respecitve value and not inferred.

Parameter	Distribution	Units	μ	σ	min	max
ϕ_{e}	LN	-	1.02e+00	2.06e-01	3.68e-01	2.72e+00
P_{ex}	LN	mm h^{-1}	4.07e+01	8.22e+00	8.06e+00	7.95e+01
D	U	-	-	-	5.00e-01	1.00e+00
$S_{\mathrm{t,max}}$	LN	mm	1.50e+01	1.50e+00	1.00e+01	5.00e+01
$S_{\mathrm{u,max}}$	LN	mm	2.77e+02	1.48e+02	3.00e+01	9.92e+02
β_{u}	LN	-	4.06e+00	5.32e+00	1.00e-01	6.05e+00
k_{c}	LN	h^{-1}	2.75e+00	1.47e+00	3.28e-01	1.09e+01
k_{d}	LN	h^{-1}	4.64e-01	1.42e-01	1.21e-01	1.47e+00
α_{d}	LN	-	1.25e+00	2.52e-01	4.97e-01	3.32e+00
k_{g}	LN	h^{-1}	1.63e-02	2.14e-02	2.46e-05	2.70e-02
α_{g}	LN	-	2.11e+00	6.46e-01	1.00e+00	6.05e+00
t_{L}	Fixed	h	5.00e-01	-	-	-
$S_{\mathrm{i,z1}}$	Fixed	mm	1.00e+00	-	-	-
$S_{\mathrm{t,z1}}$	LN	mm	2.28e+01	1.21e+01	1.00e+00	1.10e+03
$S_{\mathrm{u,z1}}$	Fixed	mm	2.20e+04	-	-	-
$S_{\mathrm{g,z1}}$	Fixed	mm	1.80e+04	-	-	-
$S_{\mathrm{t,z2}}$	LN	mm	5.07e+01	2.70e+01	1.00e+00	1.10e+03
$S_{\mathrm{u,z2}}$	Fixed	mm	2.20e+04	-	-	-
$S_{\mathrm{g,z2}}$	Fixed	mm	1.80e+04	-	-	-
r_{s}	LN	h^{-1}	3.05e-02	1.63e-02	4.00e-04	1.47e+00
λ	LN	h^{-1}	4.83e-03	2.01e-03	4.00e-04	1.47e+00
λ_{g}	Fixed	h^{-1}	2.88e-04	-	-	-
k_{i}	LN	h^{-1}	6.13e-01	3.27e-01	7.33e-02	4.00e+00
m_{e}	Fixed	-	1.00e-02	-	-	-
$m_{\mathrm{g,e}}$	Fixed	-	9.97e+00	-	-	-
a_Q	Exp	-	1.00e+00	-	-	-
b_Q	Fixed	-	2.00e-01	-	-	-
c_Q	Fixed	-	7.00e-01	-	-	-
τ_Q	LN	h	2.25e+01	2.95e+01	2.50e-01	1.01e+02
$y_{0,Q}$	Fixed	mm h^{-1}	4.00e-02	-	-	-
a_{atra}	Exp	-	1.00e+00	-	-	-
b_{atra}	Fixed	-	2.00e-02	-	-	-
c_{atra}	Fixed	-	8.00e-01	-	-	-
τ_{atra}	Fixed	h	0.00e+00	-	-	-
$y_{0,\mathrm{atra}}$	Fixed	μg l^{-1}	1.00e+00	-	-	-
a_{terb}	Exp	-	1.00e+00	-	-	-
b_{terb}	Fixed	-	2.00e-02	-	-	-
c_{terb}	Fixed	-	8.00e-01	-	-	-
τ_{terb}	Fixed	h	0.00e+00	-	-	-
$y_{0,\mathrm{terb}}$	Fixed	μg l^{-1}	1.00e+00	-	-	-
a_{K}	Fixed	-	2.00e-01	-	-	-
b_{K}	Fixed	-	1.00e-01	-	-	-
c_{K}	Fixed	-	8.00e-01	-	-	-
τ_{K}	Fixed	h	0.00e+00	-	-	-
$y_{0,K}$	Fixed	l kg^{-1}	1.00e+00	-	-	-

and by assuming that $\tau(t_0) = 0$. Multiplying the likelihood functions for streamflow (y_1), atrazine (y_2) and terbuthylazine (y_3) concentrations, as well as the soil/water distribution coefficient (y_4), we obtain the final likelihood function:

$$f_{\mathrm{L}} = \prod_{k=1}^{4} f_{y_k}\big(y_k(t_0), y_k(t_1), \ldots, y_k(t_n) \mid \boldsymbol{\theta}, \boldsymbol{\psi}\big) \tag{A.41}$$

For the forward simulations, an assumption about the timing of the second application of terbuthylazine is needed. We choose the 3$^{\mathrm{rd}}$ of June, because it is a day that is not much later than the first application and no precipitation was recorded on that day. After that, we generate 500 realizations of the full probabilistic model, i.e. we draw from the obtained posterior sample, use the drawn hydrological and chemical parameters to produce an output of the transport model and add a realization of the stochastic error term on top of that. This procedure provides a sample of the predicted state variables for which an error model was formulated (streamflow, atrazine and terbuthylazine concentrations, and the distribution coefficient). The uncertainty of the contribution of each HRU to the substance exported from the catchment is estimated indirectly based on the uncertainty of the mass exported from the whole catchment. More precisely, we assume that the ratio of the contributions of each HRU is invariant among realizations of the stochastic model, which allows us to multiply the maximum posterior estimate of the relative contribution of the HRUs with the total exported mass in each stochastic realization. Since the mass that can be exported from impervious areas is strongly supply limited, the mentioned assumption is not expected to hold for this HRU and therefore the uncertainty of its export is not calculated.

A.6 Performance metrics

A.6.1 Reliability

The reliability metric used in this study is analogous to the one proposed by McInerney et al. [2017] to measure the quality of probabilistic predictions in a hydrological modelling setting. It quantifies to which degree the observations are consistent with being realizations of the predictive distribution. It was adapted by Ammann et al. [2019] so that higher values mean higher reliability, and it is defined as:

$$\Xi_{\mathrm{reli}} = 1 - \frac{2}{n}\sum_{i=1}^{n} |F_{y_{\mathrm{mod}}(t_i)}(y_{\mathrm{obs}}(t_i)) - F_{\Psi}(F_{y_{\mathrm{mod}}(t_i)}(y_{\mathrm{obs}}(t_i)))| \tag{A.42}$$

where

$$\Psi = \{F_{y_{\mathrm{mod}}(t_i)}(y_{\mathrm{obs}}(t_i)) | i \in \mathbb{N}, 1 \leq i \leq n\} \tag{A.43}$$

and n is the number of time points at which observations, y_{obs}, and the corresponding model output, y_{mod}, are available. F_{Ψ} is the empirical cumulative

distribution function of Ψ and $F_{y_{\mathrm{mod}}(t_i)}$ is the empirical cumulative distribution function of the predicted streamflow at time t_i. The reliability can take values between 0 and 1, where values close to 0 mean that many observations are far outside our predictive distribution and values close to 1 mean that the predictive distribution captures all the observations. Note that this measure is not punishing excessively broad predictive distributions; the wider the distribution, the closer Ξ_{reli} will be to unity.

A.6.2 Relative spread

The relative spread quantifies the width of the predictive distributions over all time points relative to the magnitude of the observations. We adopt the metric that was introduced by McInerney et al. [2017] as "precision", which is defined as follows:

$$\Omega_{\mathrm{spread}} = \frac{\sum_{i=0}^{n} \sigma_{y_{\mathrm{mod}}}(t_i)}{\sum_{i=1}^{n} y_{\mathrm{obs}}(t_i)} \tag{A.44}$$

where $\sigma_{y_{\mathrm{mod}}}(t_i)$ is the standard deviation of the stochastic predictions at time t_i. Ω_{spread} is limited to positive values, where small Ω_{spread} indicate a small predictive uncertainty.

A.7 Herbicide inputs to HRUs

Table A.2 shows the masses of herbicides applied to each HRU in the two spatial arrangements.

A.8 Performance metrics of all models

Table A.3 lists the performance metrics of all models as a result of the inference process.

Table A.2: Masses of herbicides (g) applied as input to each HRU in two of the spatial configuration tested in this study.

Model	HRU	Atrazine	Terbuthylazine (19^{th} of May)	Terbuthylazine (second application)
MexpH4	Impervious	0.2	0.2	0
	Connected	885.34	2225.96	21.56
	Drained	3552.13	1406.37	1105.77
	Unconnected	3832.03	1962.18	3124.67
MprxH4	Impervious	0.2	0.2	0
	Near	2614.16	112.4	339.58
	Medium	4550.7	3090.22	2158.97
	Far	1104.64	2391.88	1753.46

Table A.3: Performance of the tested models in the calibration period measured by reliability and spread w.r.t. three out of the four target variables.

Model	**Streamflow**		**Atrazine**		**Terbuthylazine**	
	Reli.	Spread	Reli.	Spread	Reli.	Spread
MexpH1	0.867	0.788	0.854	4.13	0.71	1.37
MexpH1_nsd	0.857	1.04	0.833	4.38	0.696	1.61
MexpH2	0.725	0.483	0.727	0.308	0.641	2.14
MexpH2_nsd	0.797	0.462	0.742	1.17	0.874	2.3
MexpH3a	0.718	0.479	0.697	0.662	0.768	0.722
MexpH3a_nsd	0.777	0.456	0.795	1.14	0.666	0.688
MexpH3b	0.726	0.483	0.733	0.461	0.808	1.11
MexpH3b_nsd	0.814	0.463	0.775	1.22	0.891	1.9
MexpH4	0.725	0.482	0.757	0.39	0.867	0.32
MexpH4_nsd	0.782	0.459	0.725	1.08	0.715	0.451
MtopH4	0.727	0.483	0.773	0.371	0.904	0.742
MtopH4_nsd	0.781	0.459	0.74	1.12	0.874	1.58
MprxH4	0.722	0.483	0.671	0.797	0.904	0.495
MprxH4_nsd	0.776	0.46	0.748	1.22	0.869	1.99

A.9 Prior and posterior parameter distribution

This section provides the marginal prior and posterior parameter distributions for all the models tested in the paper.

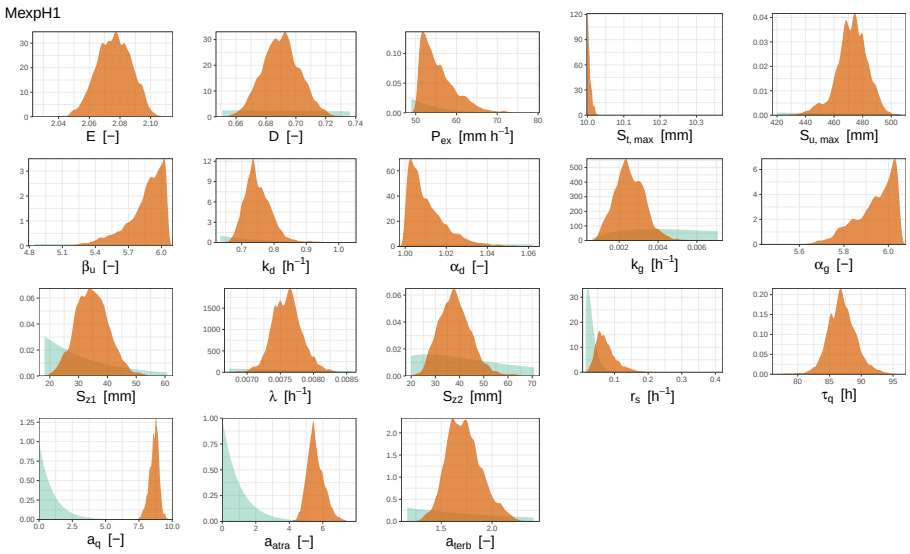

Figure A.3: Prior (blue) and posterior (red) marginal distributions for the model MexpH1.

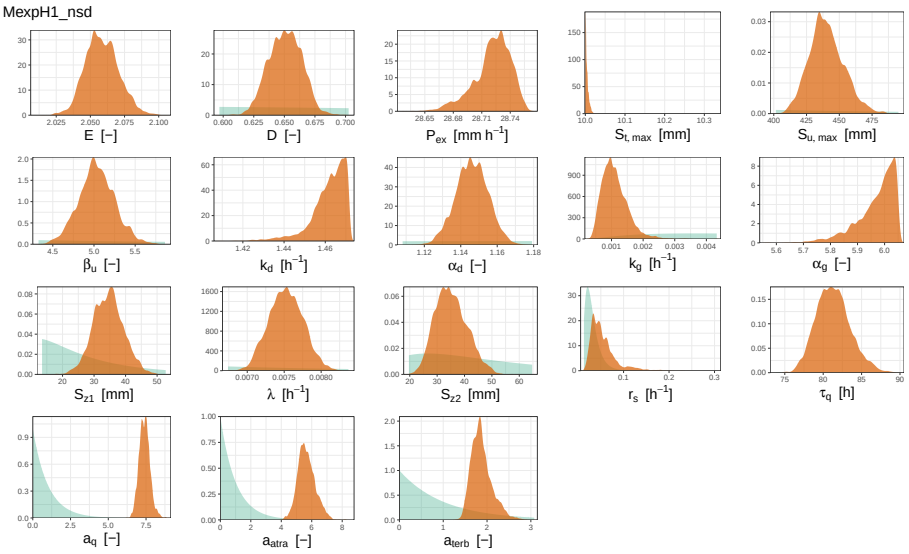

Figure A.4: Prior (blue) and posterior (red) marginal distributions for the model MexpH1_nsd.

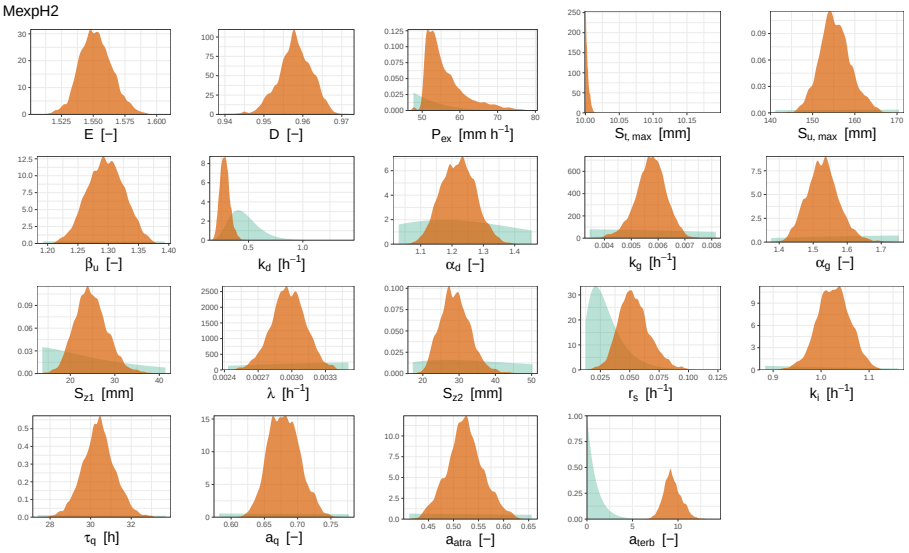

Figure A.5: Prior (blue) and posterior (red) marginal distributions for the model MexpH2.

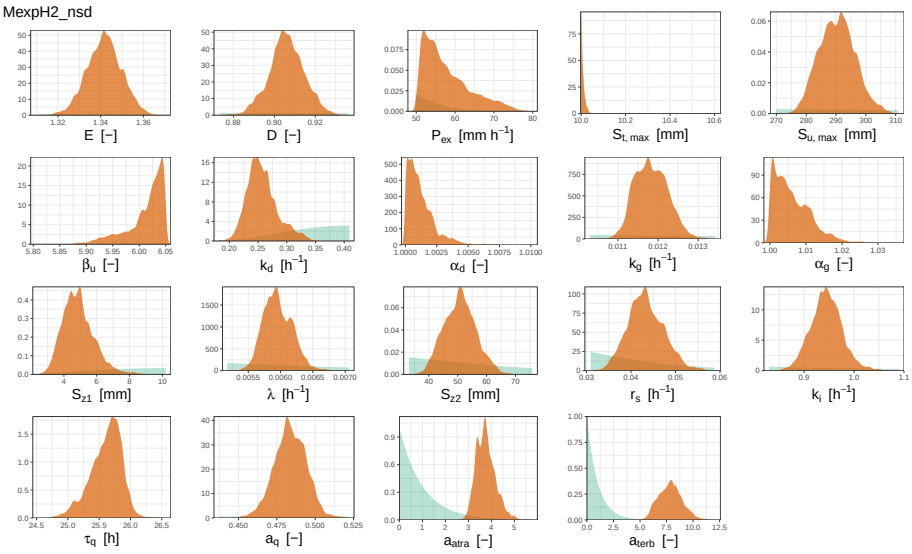

Figure A.6: Prior (blue) and posterior (red) marginal distributions for the model MexpH2_nsd.

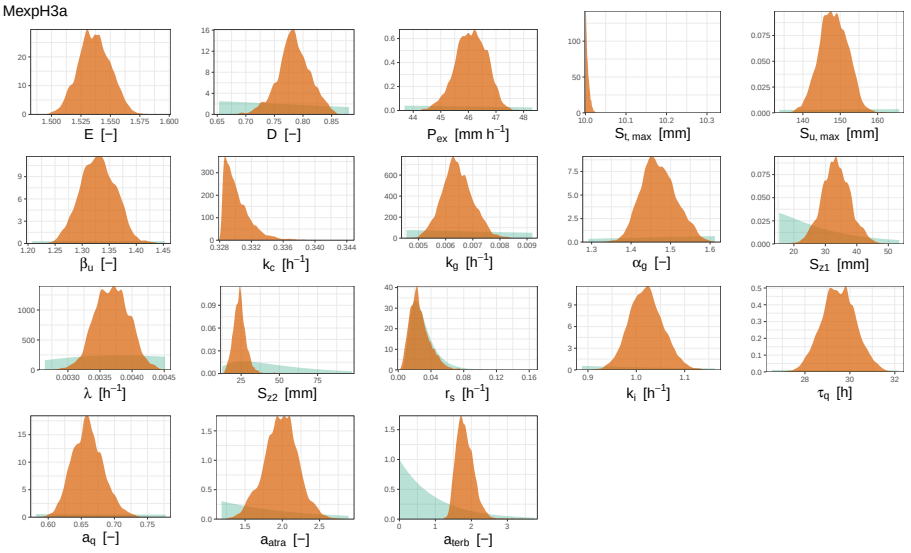

Figure A.7: Prior (blue) and posterior (red) marginal distributions for the model MexpH3a.

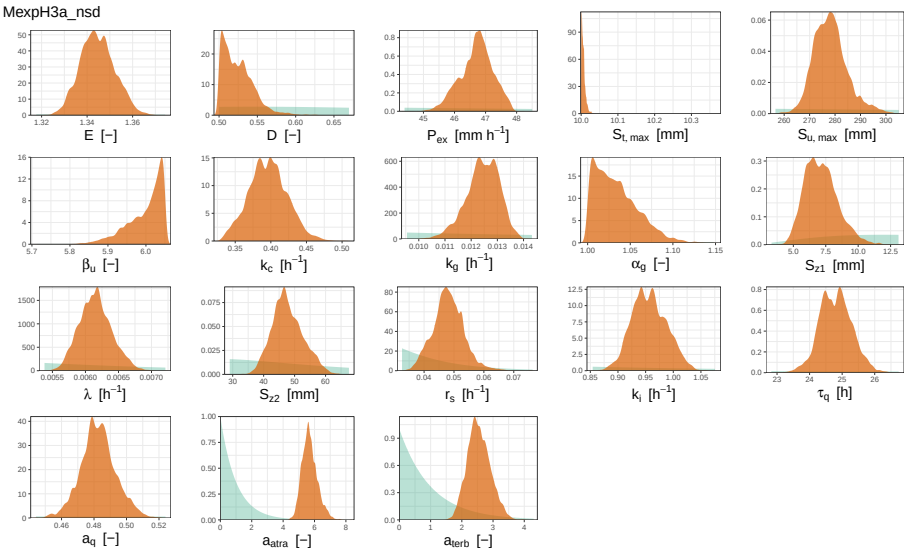

Figure A.8: Prior (blue) and posterior (red) marginal distributions for the model MexpH3a_nsd.

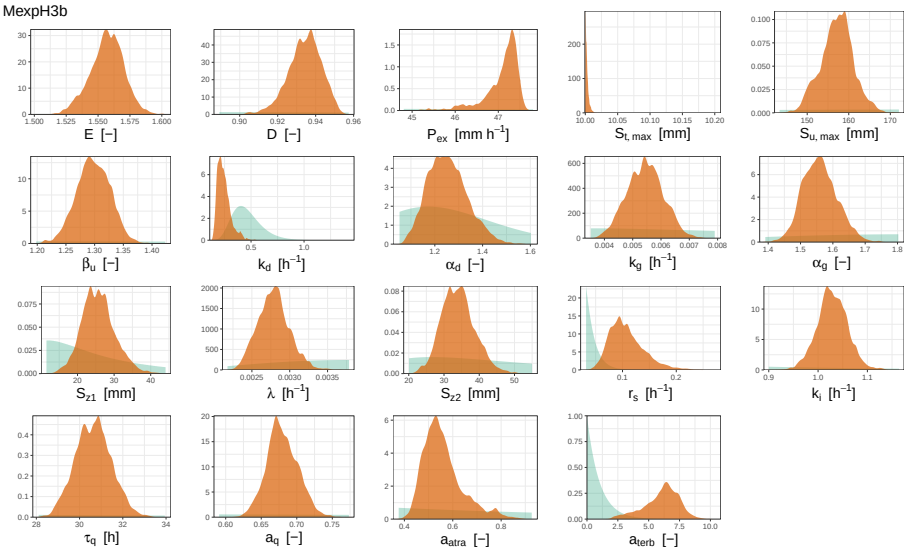

Figure A.9: Prior (blue) and posterior (red) marginal distributions for the model MexpH3b.

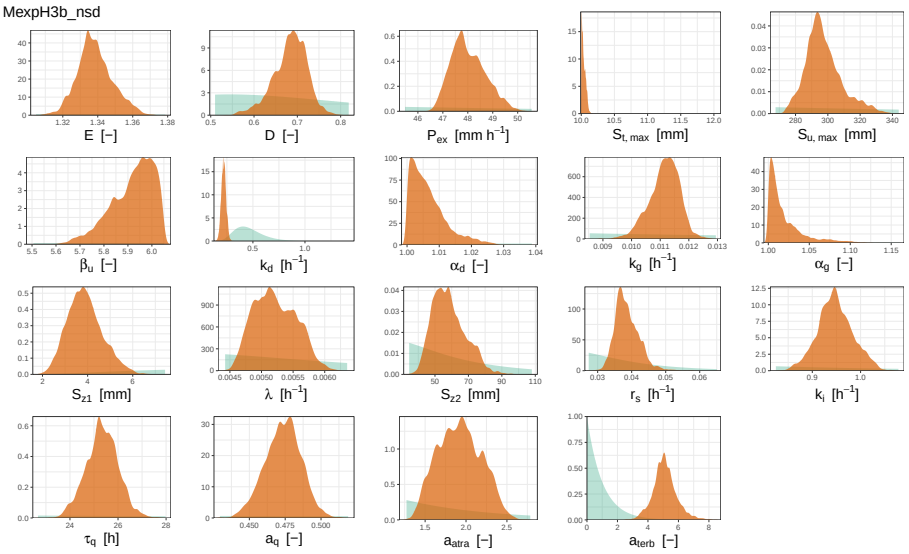

Figure A.10: Prior (blue) and posterior (red) marginal distributions for the model MexpH3b_nsd.

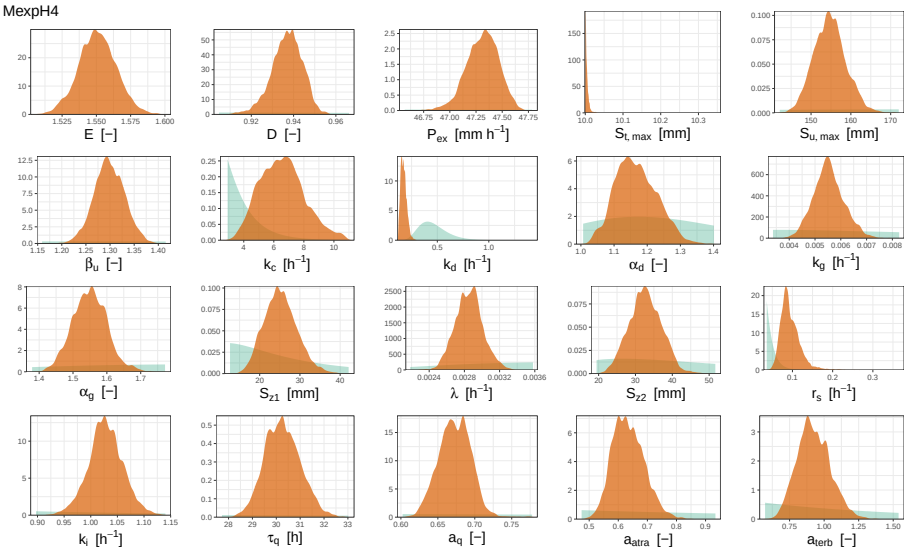

Figure A.11: Prior (blue) and posterior (red) marginal distributions for the model MexpH4.

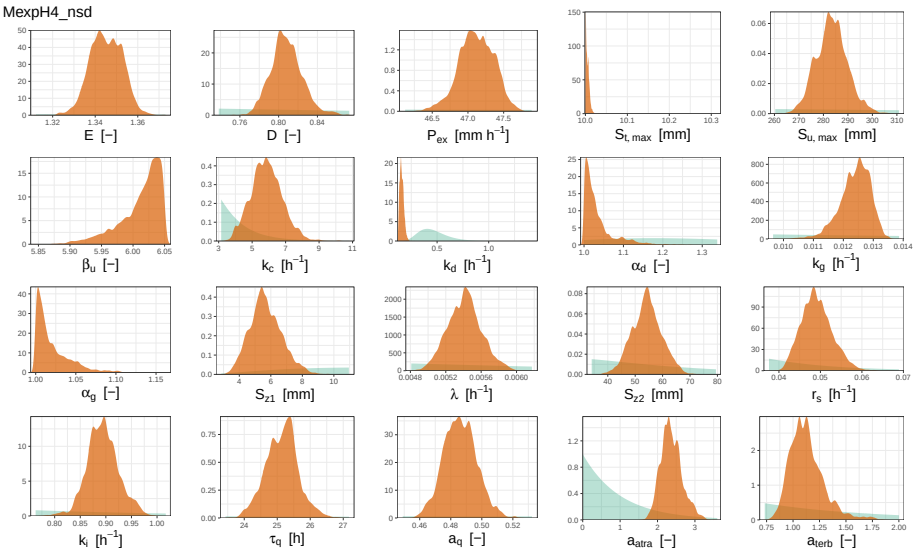

Figure A.12: Prior (blue) and posterior (red) marginal distributions for the model MexpH4_nsd.

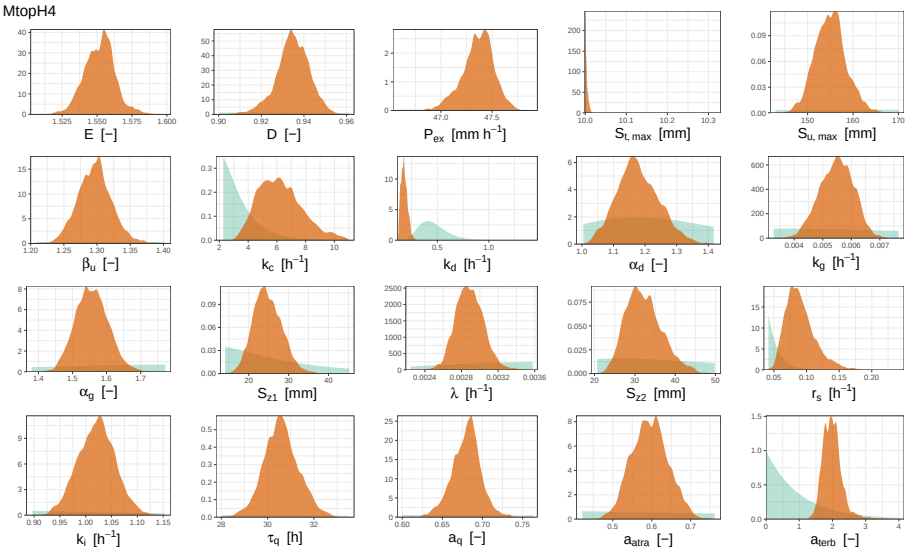

Figure A.13: Prior (blue) and posterior (red) marginal distributions for the model MtopH4.

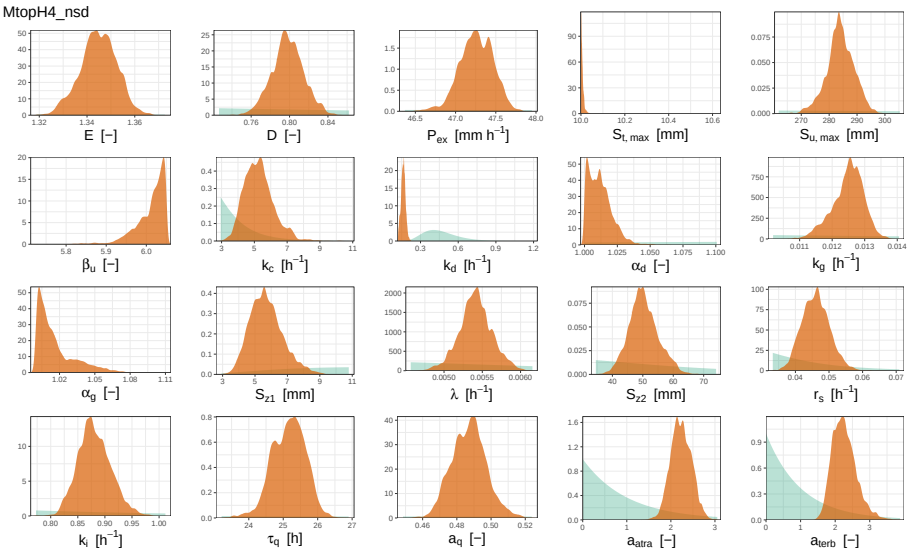

Figure A.14: Prior (blue) and posterior (red) marginal distributions for the model MtopH4_nsd.

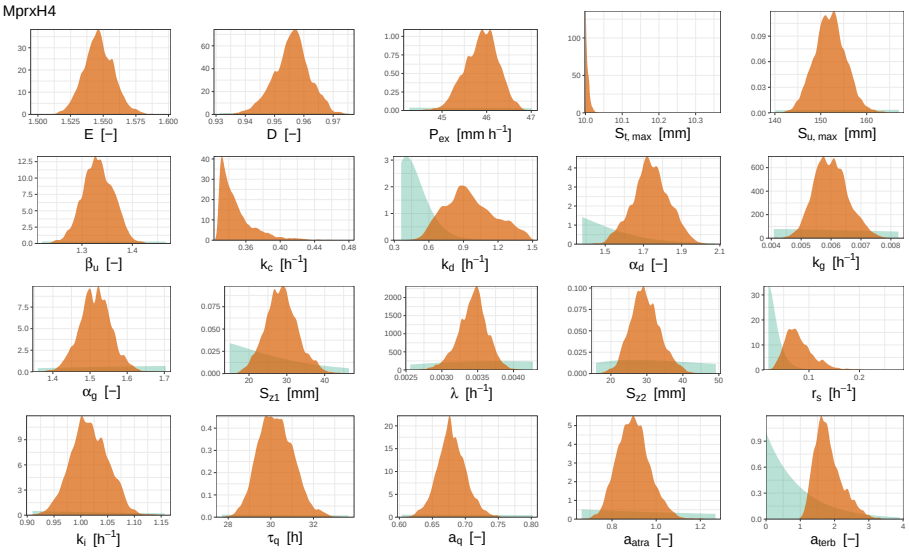

Figure A.15: Prior (blue) and posterior (red) marginal distributions for the model MprxH4.

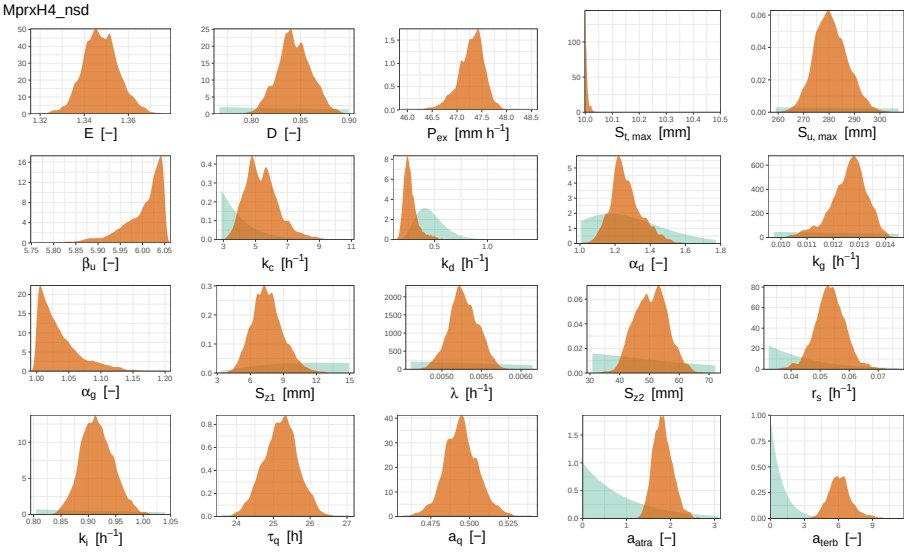

Figure A.16: Prior (blue) and posterior (red) marginal distributions for the model MprxH4_nsd.

A.10 States and fluxes of the reference model

This section contains some key visualizations of internal states and output fluxes of the reference model. Some of those states are compared to field observations to assess the realism of the model.

Figure A.18 shows the mass of substance in the dissolved and the sorbed state in the topsoil reservoir of the reference model.

Figure A.18 shows the mass of substance in the dissolved and the sorbed state in the topsoil reservoir of the reference model.

Figure A.19 shows the streamflow contributions from the most important reservoirs durig the study period as predicted by the reference model. The dominance of the impervious areas and the groundwater reservoirs in producing the fast and the slow streamflow response, respectively, is clearly visible.

Comparing Figures A.20 and A.21, one can see the different contributions of the drained and shortcut areas to the export of atrazine and terbuthylazine during the major loss event E2. Note that the difference is due to the different amount of the substances applied to the drained and the shortcut areas, not due to different substance properties.

Figure A.22 shows the measured and the modelled volumetric water content of the soil. The modelled water content is calculated as follows:

$$\theta_{\mathrm{mod}} = \frac{S_{\mathrm{u}}}{S_{\mathrm{u,max}}} n_{\mathrm{e}} + n_{\mathrm{resid}} \tag{A.45}$$

where S_{u} is the reservoir level of the lower soil zone (unsaturated zone), n_{e} is the effective porosity (assumed equal to 0.3), and n_{resid} is the residual porosity (assumed equal to 0.1). For the observed water content, Figure A.22 shows the measurements at one particular location in the center of the catchment, which is believed to be most representative.

Figure A.23 shows the observed and the measured groundwater levels.

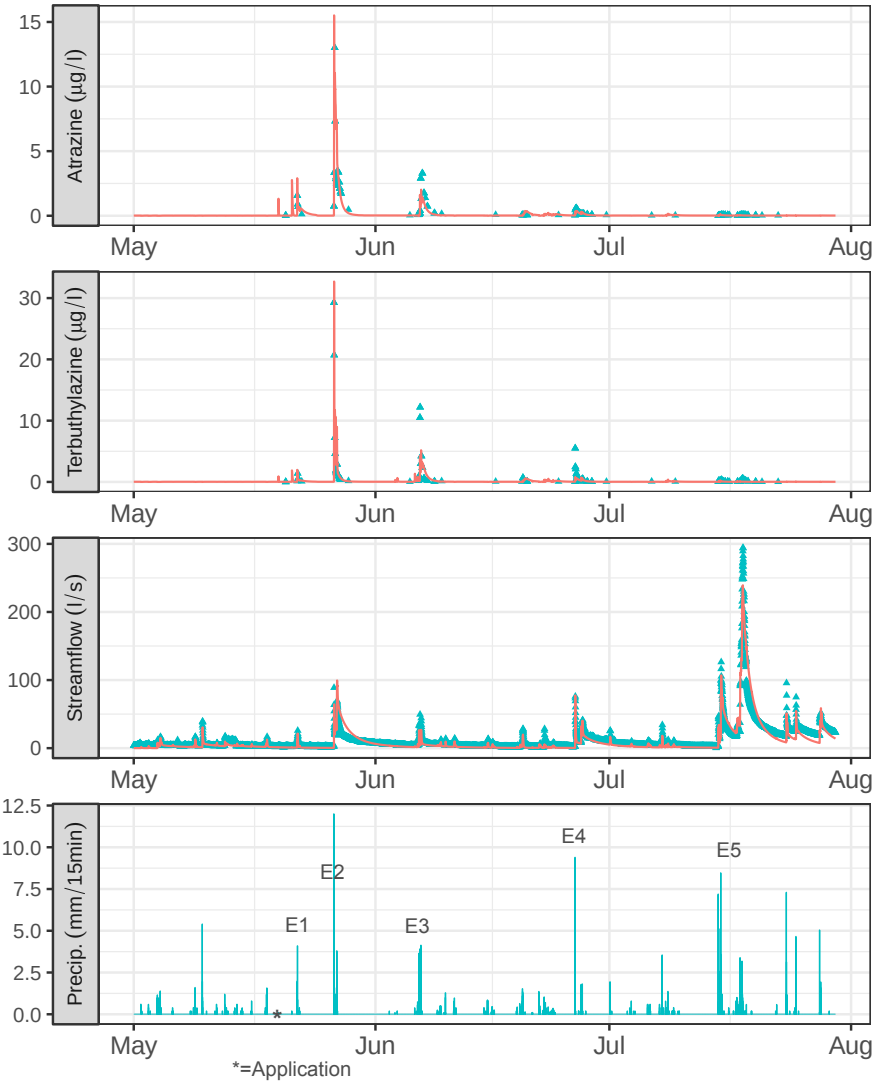

Figure A.17: Model output with the maximum posterior parameter values of MexpH4 (red lines) during the full study period, together with measured concentrations and streamflow (blue dots), and the recorded precipitation. The events are labelled in the lowermost panel.

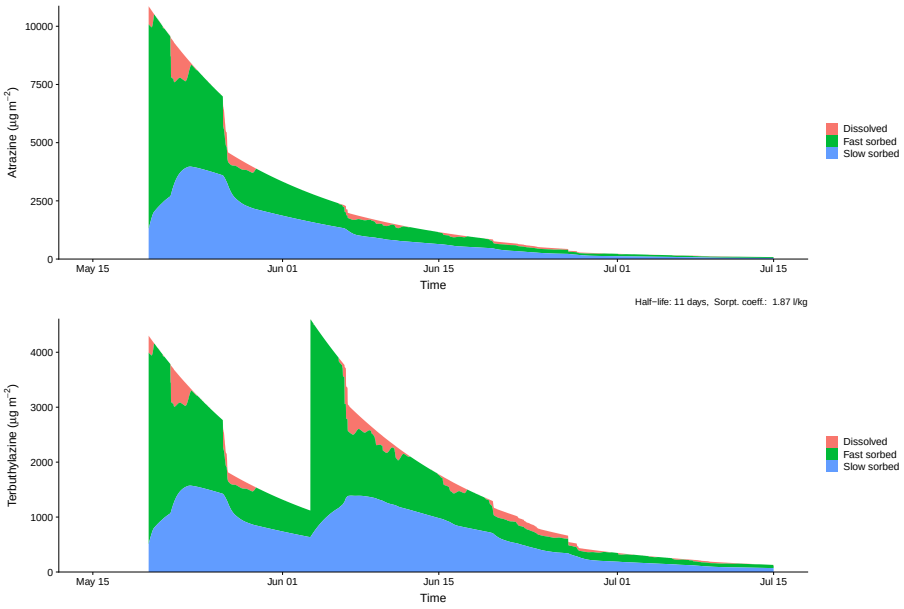

Figure A.18: The mass of atrazine and terbuthylazine that exists in the dissolved, fast, and slow sorbed state in the topsoil reservoir of the HRU representing drained areas during the whole study period within the reference model. Note that the substance is modelled as spatially homogeneous within one HRU, so the modelled concentrations do not correspond directly to soil concentrations on the sprayed fields.

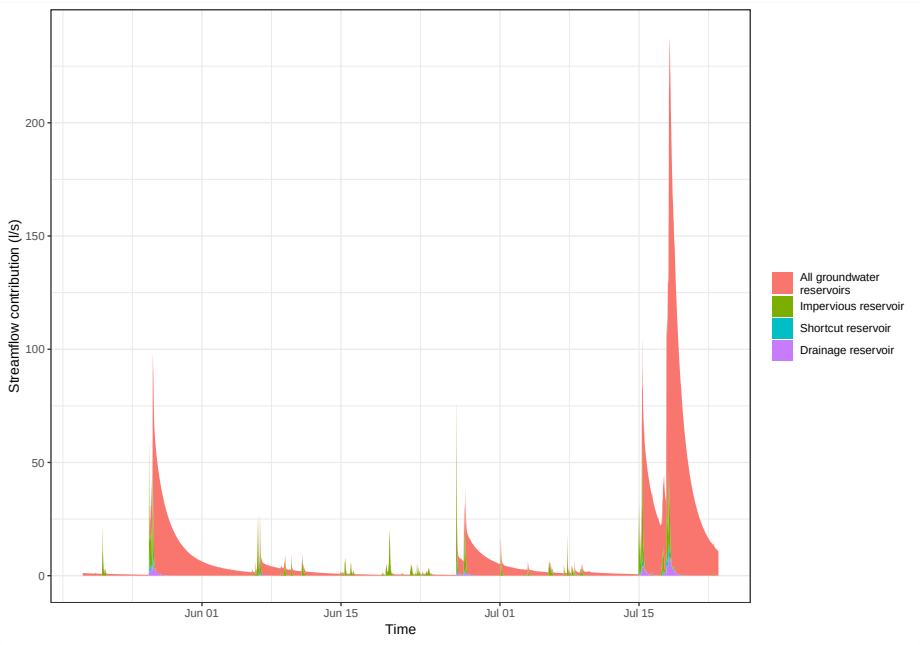

Figure A.19: Contributions of the different reservoirs to the total streamflow. Note that for this plot, all the groundwater reservoirs from the different HRUs are added. The other reservoirs are shown as individual contributions from the corresponding HRUs in which they exist.

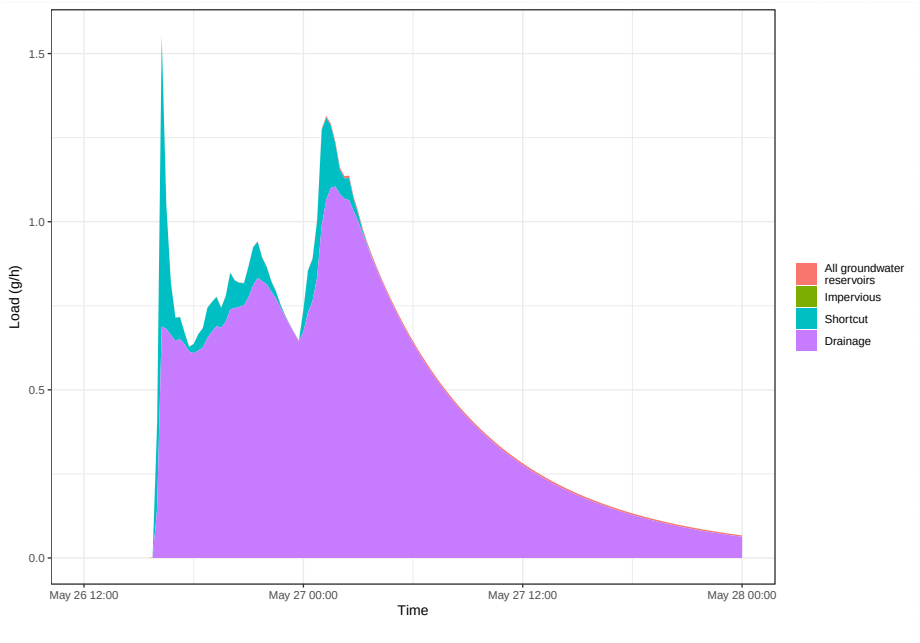

Figure A.20: Contributions of the different reservoirs to the total flux of atrazine during the major loss event E2. Note that for this plot, all the groundwater reservoirs from the different HRUs are added. The other reservoirs are shown as individual contributions from the corresponding HRUs in which they exist.

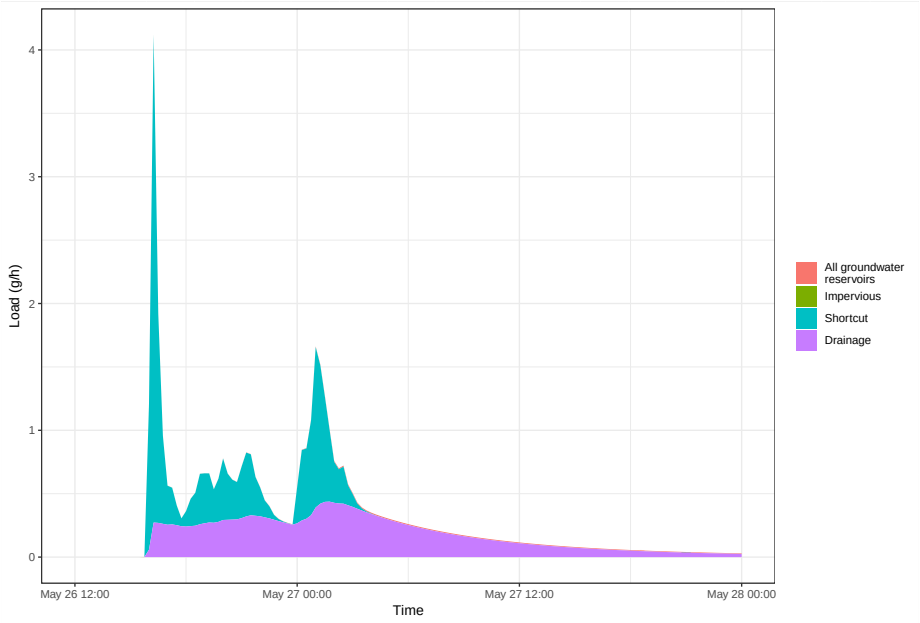

Figure A.21: Contributions of the different reservoirs to the total flux of terbuthylazine during the major loss event E2. Note that for this plot, all the groundwater reservoirs from the different HRUs are added. The other reservoirs are shown as individual contributions from the corresponding HRUs in which they exist.

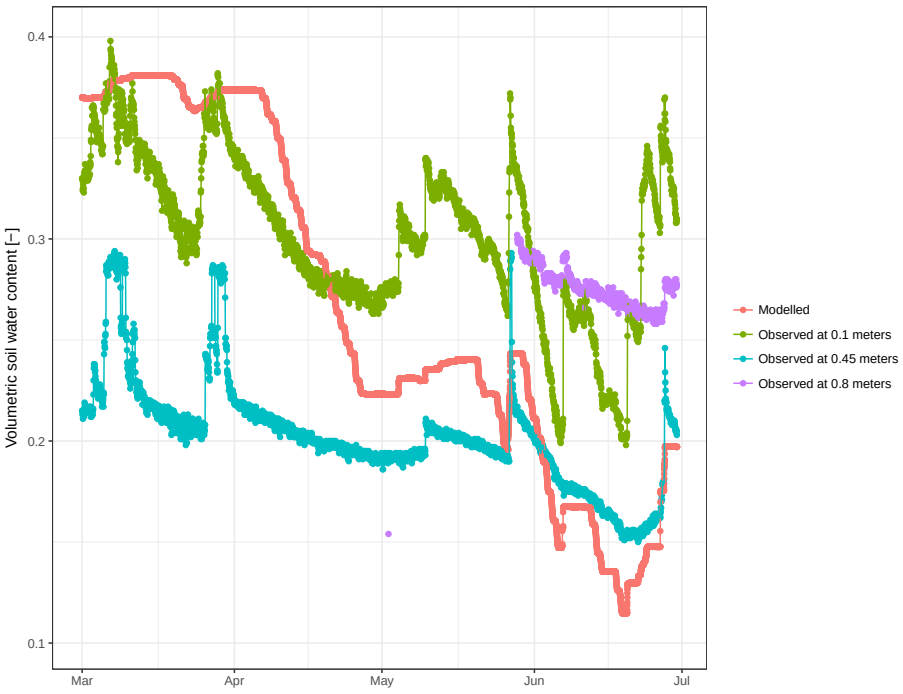

Figure A.22: Comparison of the modelled and measured volumetric water content. The measured soil water content is shown at three different depths below ground level.

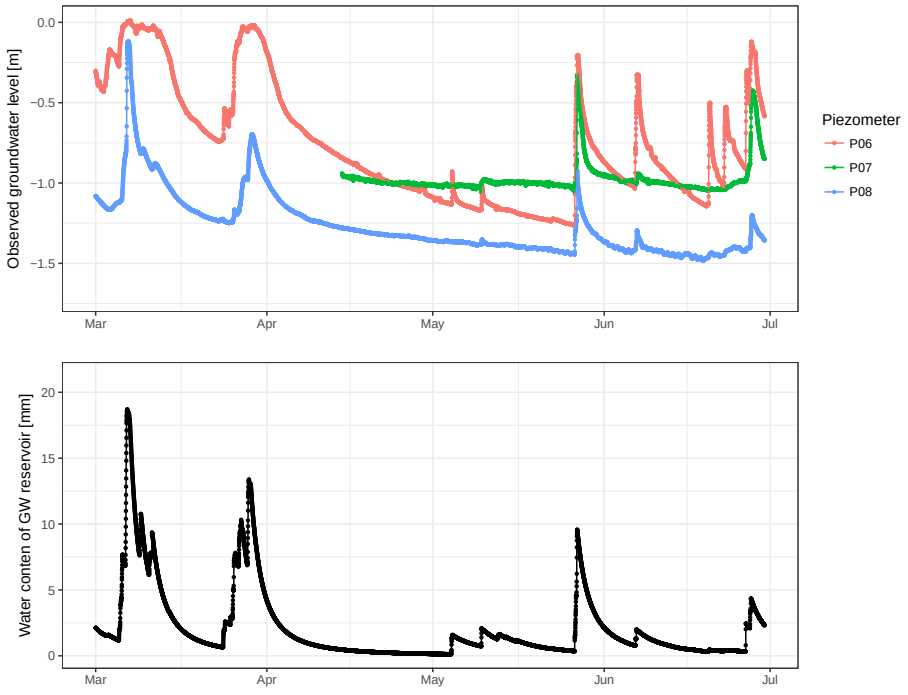

Figure A.23: Comparison of the modelled and measured groundwater levels. The measured groundwater head is given in meters below ground level at three different piezometers (P06, P07, and P08) located in the center of the catchment. These are believed to be most representative of the whole catchment area.