

Quantifying the Uncertainty of a Conceptual Herbicide Transport Model With Time-Dependent, Stochastic Parameters

Journal Article**Author(s):**

Ammann, Lorenz; Stamm, Christian; Fenicia, Fabrizio; Reichert, Peter

Publication date:

2021-08

Permanent link:

<https://doi.org/10.3929/ethz-b-000503351>

Rights / license:

[Creative Commons Attribution 4.0 International](#)

Originally published in:

Water Resources Research 57(8), <https://doi.org/10.1029/2020WR028311>

Water Resources Research

RESEARCH ARTICLE

10.1029/2020WR028311

Key Points:

- Catchments behave stochastically at the resolution we observe them, yet they are often approximated by deterministic models
- We combine a framework for stochastic, time-dependent parameters with a conceptual model of fast herbicide transport at catchment-scale
- Stochastic parameters can increase model flexibility, improve uncertainty quantification, and inspire structural model improvements

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Ammann,
lorenz.s.ammann@gmail.com

Citation:

Ammann, L., Stamm, C., Fenicia, F., & Reichert, P. (2021). Quantifying the uncertainty of a conceptual herbicide transport model with time-dependent, stochastic parameters. *Water Resources Research*, 57, e2020WR028311. <https://doi.org/10.1029/2020WR028311>

Received 20 JUL 2020

Accepted 2 JUN 2021

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Quantifying the Uncertainty of a Conceptual Herbicide Transport Model With Time-Dependent, Stochastic Parameters

Lorenz Ammann^{1,2} , Christian Stamm¹ , Fabrizio Fenicia¹ , and Peter Reichert^{1,2} 

¹Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland, ²Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland

Abstract Small streams in catchments with agricultural land use are at high risk of diffuse pollution by herbicides. Fast transport processes can cause concentration peaks that exceed regulatory requirements. These processes have a high spatio-temporal variability and data characterizing their occurrence is often sparse. For this reason, such systems show a stochastic behavior at the resolution we observe them (same input and initial conditions lead to different output). Realistic model representations should acknowledge this pronounced apparent intrinsic stochasticity. However, a deterministic description of the physical and chemical processes at the catchment scale is state of the art in research and practice. We explore the potential of stochastic process formulations in combination with the Bayesian learning paradigm to (a) improve the quantification of the uncertainty of conceptual catchment-scale pesticide transport models and (b) gain new mechanistic insights about the system by interpreting the temporal evolution of the stochastic processes. This is done with the help of a framework for time-varying stochastic parameters. Thereby, we find that (a) the stochastic process formulation can lead to a more realistic characterization of the uncertainty of internal states and model output compared to the deterministic one, and that (b) the temporal dynamics of parameters resulting from the inference can highlight model deficits (and inspire improvements) such as a better sustained baseflow in dry periods. We also identify two key challenges: numerical difficulties in sampling the posterior and the question of where to introduce and how to constrain the additional degrees of freedom such that they are not misused.

1. Introduction

Hydrological catchments are complex systems in which various interacting processes transport water and pollutants at different spatio-temporal scales. Understanding and quantification of such processes is of high relevance to estimate the risk of diffuse pollution of headwater streams, especially in catchments dominated by agricultural land use. We can only ever observe a limited fraction of the relevant input, states and outputs of transport processes at the catchment scale. As a consequence, an important property arises: such natural systems apparently behave stochastically at the resolution we observe them, that is, they can show a different response for the same *measured* driving forces since the *true* driving forces were different. For example, for two cases of the same measured volume of precipitation, input pollutant mass and catchment state; the true output (e.g., streamflow and pollutant concentration in the stream) will not respond identically in the two cases. This is because the observations of the driving forces and of the states are highly aggregated, which means that they are compatible with many different true realizations of those quantities.

The numerical models used to characterize such complex systems can only be approximate representations and they are subject to multiple sources of uncertainty. A better understanding and quantification of the uncertainties of water quality (WQ) models has been identified as one of the most pressing issues in the field (e.g., Rode et al., 2010), which is further emphasized by the scarcity of WQ data such as pollutant concentrations compared to, for example, more abundant hydrological data such as streamflow measurements. In addition, the success of machine-learning techniques (e.g., Shen, 2018) has shown the benefit of models that are more flexible in nature (e.g., Nearing et al., 2016). The limited amount of WQ measurements makes the calibration of such data-hungry models often difficult, however. There is thus a need to add flexibility

to conventional mechanistic process-based WQ models in an interpretable way and to enable uncertainty quantification and calibration of such models with limited amounts of data.

In light of the considerations on stochasticity and flexibility in the two preceding paragraphs, let us look at the fundamental distinction between deterministic process model (DPM) and stochastic process model (SPM). As DPMs, we classify models that rely on process descriptions that yield the same result when provided with the same input, parameters, and initial states. In such models, a stochastic error term is usually added to the output of the underlying purely DPM. SPMs, on the other hand, are more flexible and account for stochasticity in internal model states (as opposed to just the output), for example, through stochastic differential equations, stochastic parameters, or stochastic input. While the latter allows to account for input uncertainty, the former two are possible approaches to consider model structural uncertainty. SPMs are a promising technique to combine the need for increased flexibility with the interpretability provided by process-based (mechanistic) approaches. A more detailed definition of DPMs and SPMs can be found in Section 2.

A large part of the previous studies dealing with uncertainty quantification of WQ models at the catchment scale fall into the category of DPMs. For example, Zheng and Keller (2007) account for input, parameter, model structure, and observational uncertainty through an additive error model in combination with the GLUE approach (Beven & Binley, 1992). Bertuzzo et al. (2013) quantify parametric uncertainty of a conceptual WQ model to quantify herbicide concentration dynamics in a small catchment. Bayesian approaches have been identified as especially a promising technique in dealing with uncertainty of WQ models (Rode et al., 2010). Previous studies have proven the value of Bayesian approaches to quantify uncertainties of WQ models (e.g., Gardner et al., 2011; Hantush & Chaudhary, 2014; Hong et al., 2005; Raat et al., 2004; Talamba et al., 2010; Wellen et al., 2014). Bayesian frameworks have also been developed (Schoups & Vrugt, 2010) or applied in WQ modeling studies (Han & Zheng, 2016; Zheng & Han, 2016). Han and Zheng (2018) present a Bayesian framework for jointly considering input, parametric and residual uncertainty based on a perturbation of input pollutant loading, acknowledging the large input uncertainties typically present in WQ modeling studies. While these developments have certainly advanced the assessment of uncertainty of WQ models, they fail to acknowledge the fundamental inherent stochasticity of the systems we deal with.

SPMs, on the other hand, have been rarely applied in WQ studies so far. Beck and Young (1976) illustrate the potential of stochastic, time-dependent parameter methods in identifying model structures of simple river WQ models, while state assimilation techniques have been used at much larger scales in river WQ modeling (e.g., Kim et al., 2014; Loos et al., 2020). SPMs are common in purely hydrological applications, where it has been argued that the assumption of deterministic behavior of a catchment is “indefensible” (Kuczera et al., 2006). The mentioned study explains in detail why this is the case and makes an argument for what we call SPMs. In hydrology, SPMs are mostly used in connection with a class of methods called “data assimilation.” These methods aim at reconciling model states, parameters, or structures with observations to obtain more accurate estimates of model states and parameters including their uncertainty (Liu & Gupta, 2007). For an overview of the large number of previous studies on this topic, the reader is referred to Liu and Gupta (2007) and Liu et al. (2012). While many approaches estimate either states or parameters separately, it is generally recognized that a joint estimate is conceptually preferable (Liu & Gupta, 2007; Vrugt et al., 2005). Various approaches for the joint estimation of parameters and states have been achieved (e.g., Moradkhani, Hsu, et al., 2005; Moradkhani, Sorooshian, et al., 2005; Vrugt et al., 2005); an interdisciplinary overview is provided by Kantas et al. (2015). However, many of the previous approaches are based on stochastic states and parameters, but fewer studies rely on the concept of *time-varying* stochastic parameters (e.g., Pathiraja et al., 2016b; Rajaram & Georgakakos, 1989; Reichert & Mieleitner, 2009; Smith et al., 2008). Even though the two approaches are technically very similar, the concept of time-varying parameters is more satisfying as it conserves mass balances and therefore respects fundamental physical principles. In this study, we rely on the approach for stochastic, time-dependent parameters suggested by Reichert and Mieleitner (2009) since it is fully compatible with the Bayesian learning paradigm by providing an approach to sample from the full joint posterior distribution of constant and time-dependent parameters. The approach taken here can be classified as a “smoothing” approach according to McLaughlin (2002), since the posterior of the quantity of interest at each time point is calculated conditional on all (past, present, and future) observations; it is therefore used offline in batch mode. Filtering approaches, which only condition on past and present obser-

vations, have been frequently applied for joint state and parameter estimation. Smoothing approaches, on the other hand, have been given much less attention (Yang et al., 2018).

The objectives of this study are the following:

- Explore the potential of SPMs of conceptual nature for the characterization of herbicide transport in small headwater catchments. Investigate whether this approach can reveal mechanisms in the underlying system that have not been accounted for so far.
- Check whether incorporating the identified patterns in time-dependent parameters into the deterministic part of the model structure improves the likelihood compared to the original model.
- Assess the uncertainty estimates provided by the SPMs for quantities that are of interest in WQ problems, like high-frequency dynamic fluctuations of herbicide concentrations in a small catchment with agricultural land use.

2. Definition of Deterministic and Stochastic Process Models

For DPMs, the states, \mathbf{y} (note the use of boldface for vectors), are given by the model, m , which is a function of the parameters θ :

$$\mathbf{y} = m(\theta) \quad (1)$$

Note that we omit the input from the notation, since statistically speaking, it is not fundamentally different from the parameters; both are quantities that affect the states and the output of the model. By model output, we mean the model's estimate of the true value of the observed state (e.g., streamflow). Let us assume that the actual observations, \mathbf{y}_{obs} , are a specific realization of the random variable \mathbf{Y}_{obs} (note the use of capital letters for random variables), which is either given by

$$\mathbf{Y}_{\text{obs}} = o(\mathbf{y}, \theta) + \mathbf{E}_o(\mathbf{y}, \psi) \quad (2a)$$

or by
$$\mathbf{Y}_{\text{obs}} = h^{-1}\left(h\left(o(\mathbf{y}, \theta)\right) + \mathbf{E}_h(\psi)\right) \Leftrightarrow h(\mathbf{Y}_{\text{obs}}) = h\left(o(\mathbf{y}, \theta)\right) + \mathbf{E}_h(\psi) \quad (2b)$$

where $o(\cdot)$ is a function mapping the model states to the outputs and \mathbf{E}_o is a stochastic term that describes the effect of all sources of errors (including measurement errors) on \mathbf{Y}_{obs} in a lumped way. Since errors are often heteroscedastic (e.g., Clarke, 1973), \mathbf{E}_o depends on \mathbf{y} and on error model parameters, ψ . Alternatively, the transformation $h(\cdot)$ (e.g., Box-Cox) may be applied (Equation 2b), which ideally leads to a homoscedastic error term, \mathbf{E}_h , that does not depend on the states anymore. However, the error tends to have other complex and unfavorable characteristics such as nonnormality (e.g., Schoups & Vrugt, 2010), and autocorrelation (e.g., Aitken, 1973; Beven & Westerberg, 2011; Kavetski et al., 2003). Accounting for these characteristics can be difficult (e.g., Ammann et al., 2019; Evin et al., 2014; McInerney et al., 2017; Schoups & Vrugt, 2010).

For SPMs, on the other hand, the states, \mathbf{Y} , are random variables obtained by the stochastic model M :

$$\mathbf{Y} = M(\theta) \quad (3)$$

and the observations are characterized as a deterministic or stochastic function, $o(\cdot)$, of the states:

$$\mathbf{Y}_{\text{obs}} = o(\mathbf{Y}, \theta) + \mathbf{Z}(\mathbf{Y}, \psi) \quad (4)$$

where \mathbf{Z} , describes the residual stochasticity, which considers the remaining uncertainty that is neither accounted for in the process model, M , nor in the function, o , which maps the states to the output. Ideally, \mathbf{Z} represents the measurement error only, which is the case if all other sources of uncertainty have been accounted for in M and o . Equation 4 has conceptual advantages over Equation 2. First, it is preferable to account for the errors where they actually originate: in the internal states and fluxes of the model. This is a more accurate description of many environmental systems, for which the uncertainty in the modeled states and fluxes is often larger than the measurement uncertainty of the output. Second, calibrated optimal temporal dynamics of internal states or parameters can be used to inspire model structural improvements

(e.g., Reichert & Mieleitner, 2009; Vrugt et al., 2005). For a more detailed discussion of the advantages and challenges of stochastic models, see Reichert and Mieleitner (2009).

3. Methods

The procedure applied in this study is largely based on the approach for stochastic models presented in (Reichert & Mieleitner, 2009; Reichert et al., 2020). After a short introduction to the study site and the data (Section 3.1), we describe in detail how that approach is applied in this study by (a) introducing stochasticity in a previously deterministic model by making some of its parameters stochastic, time-dependent quantities (Section 3.2), (b) screening for patterns in the inferred temporal dynamics of time-dependent parameters (Section 3.3), (c) based on the identified patterns, formulating model improvements to eliminate systematic model deficits (Section 3.4), and (d) subsequent prediction for uncertainty quantification (Section 3.5). Finally, Section 3.6 provides some details of the numerical implementation of the model and the sampler.

3.1. Study Site and Data

The modeled study catchment is the Ossingen experimental catchment, located in north-eastern Switzerland. It has a size of approx. 1.2 km² with predominantly agricultural land use (75%, crop production) and a subdued topography with an average slope of 4.3° (Doppler et al., 2012). While the dominating soil type in the catchment is a well-drained cambisol, poorly drained gleysols can be found in the center close to the stream (FAL, 1997). The average precipitation is around 900 mm/yr (MeteoSwiss, 2016). The catchment is rather typical for the Swiss Plateau regarding the mentioned characteristics. In the growing period of 2009, Doppler et al. (2012) conducted a controlled application of various herbicides in the Ossingen catchment and measured precipitation (10 min resolution), streamflow (5 min resolution), in-stream herbicide concentrations (15 min resolution during elevated streamflow), and various additional relevant states like groundwater levels and concentration of herbicides in the top soil layer. A more detailed overview of the experiment and the catchment characteristics is given in the mentioned study.

The model input and observation data we use for calibration is a preprocessed subset of the data obtained by Doppler et al. (2012) and it is the same that was used by Ammann et al. (2020). The observations used for calibration extend from March 1 to October 12, 2009 and involve measurements of streamflow, in-stream concentrations of the two herbicides atrazine and terbuthylazine, and soil/water distribution coefficients of atrazine. High-frequency in-stream concentration measurements are available during five major precipitation events that differ considerably in volume, intensity, and dominant runoff processes that they cause (Doppler et al., 2012). The variability in the characteristics of the rainfall events is expected to increase the general validity of the results. The overall resolution of the preprocessed input data (precipitation, potential evapotranspiration) is 15 min. A more detailed description of the data can be found in Ammann et al. (2020).

3.2. Probabilistic Model Based on Stochastic, Time-Dependent Parameters

We base our analysis on the deterministic conceptual model of herbicide fate and transport that has been suggested based on experimentalist knowledge for the Ossingen catchment. The model is described in detail in Ammann et al. (2020), where it is denoted as “MexpH4.” Let $m(t, \theta)$ be the model function, which defines how the model states, \mathbf{y} , evolve in time depending on the parameters θ (Table 1). The states consist of multiple reservoirs that represent fast and slow responding elements of the catchment in terms of water and herbicide transport. For example, the “groundwater” reservoir (representing an ensemble of slow transport processes) defined by the state y_{S_g} (for ease of notation we use X instead of y_X in the following) is given by:

$$\frac{dS_g}{dt} = Q_{g,in} - \theta_{k_g} S_g^{\theta_{\alpha_g}} \quad (5)$$

$$o(\mathbf{y}, \theta) = \theta_{k_g} S_g^{\theta_{\alpha_g}} + \dots \quad (6)$$

Table 1
Parameters, θ_i , of the Transport Model, Which are Calibrated in This Study

Parameter index i	Units	Description
ϕ_c	-	Multiplication factor for the potential evapotranspiration
P_{ex}	mm/h	Precipitation intensity threshold for the onset of infiltration excess runoff
D	-	Splits the flux between the fast and the slow reservoirs
$S_{t,max}$	mm	Maximum water content of the topsoil reservoir
$S_{u,max}$	mm	Maximum water content of the lower unsaturated zone reservoir
β_u	-	Exponent of the nonlinear relation of the unsaturated zone reservoir
k_i	h^{-1}	Coefficient of the constitutive function of the imperivous areas reservoir
k_c	h^{-1}	Coefficient of the constitutive function of the connected areas reservoir
k_d	$mm^{1-\alpha_d} h^{-1}$	Coefficient of the constitutive function of the drained areas reservoir
α_d	-	Exponent of the constitutive function of the drained areas reservoir
k_g	$mm^{1-\alpha_g} h^{-1}$	Coefficient of the constitutive function of the groundwater reservoir
α_g	-	Exponent of the constitutive function of the groundwater reservoir
λ	h^{-1}	Degradation rate (1st order) of the herbicides
$S_{t,z1}$	mm	Extent of completely mixed dead zone of the topsoil reservoir (fast sorption)
$S_{t,z2}$	mm	Extent of partially mixed dead zone of the topsoil reservoir (slow sorption)
r_s	h^{-1}	Kinetic coefficient of the slow sorption process

Note. For each inference setting, one of them is turned into a time-dependent stochastic process and fitted jointly with all the others.

where $Q_{g,in}$ is the inflow to that reservoir, θ_{k_g} and θ_{α_g} are part of θ , $o(\cdot)$ is the function that maps the states to the observations, and “...” denotes that this function depends on other states and parameters not shown here. Note that, because the model is conceptual in nature, its states and parameters are not directly measurable quantities but they are related to key states and processes in the catchment. For example, S_g represents the water content of the part of the saturated zone that is connected to the stream. See Ammann et al. (2020) for a more detailed explanation of the meaning of the individual parameters and states of the model.

In order to turn m into a stochastic model that can be written in the form of Equations 3 and 4, we make use of a previously developed framework for time-dependent parameters (Reichert & Mieleitner, 2009) which turns (a subset of) model parameters that were previously constant in time into a time-dependent stochastic process, $\Theta_s(t, \xi_s)$. This leads to

$$Y(\theta_{-s}, \xi_s) = m(\theta_{-s}, \Theta_s(\xi_s)) \quad (7a)$$

$$Y_o(\theta_{-s}, \xi_s) = o(Y(\theta_{-s}, \xi_s)) \quad (7b)$$

$$Y_{obs}(\theta_{-s}, \xi_s, \psi) = Y_o(\theta_{-s}, \xi_s) + Z(Y_o, \psi) \quad (7c)$$

where Y are the model states, Y_o is the model output that is comparable to the observations, Y_{obs} , is a random variable characterizing the observations, s is the index of the stochastic parameter, and $\theta_{-s} = \{\theta_i\}$ for all $i \neq s$. The probability distribution of Y_o describes our knowledge about the system's states represented by the model output variables. This knowledge is uncertain due to incomplete (epistemic) knowledge of model parameters as well as due to stochastic (aleatory) system behavior. As we describe both of these types of uncertainty probabilistically, we can combine them in a Bayesian approach despite the fact that this includes uncertainty of frequentist nature (Reichert et al., 2014).

Note that, to limit the computational effort and to avoid potentially excessive identifiability problems, only one parameter is time-dependent in each inference setting. For the stochastic process, $\Theta_s(t, \xi_s)$ and its realization $\theta_s(t)$, we use an Ornstein-Uhlenbeck (OU) processes (Uhlenbeck & Ornstein, 1930),

$$d\theta_s(t) = -\frac{1}{\tau_s}(\theta_s(t) - \mu_s) + \sqrt{\frac{2}{\tau_s}}\sigma_s dW(t), \quad (8)$$

which is defined by parameters $\xi_s = \{\mu_s, \sigma_s, \tau_s\}$ (mean, asymptotic standard deviation, and characteristic correlation time, respectively), and where $W(t)$ is a Wiener process (i.e., random walk in continuous time).

The residual error model, Z , is formulated for the different types of measurements (indexed by j) to which the models are calibrated; streamflow data, concentration measurements, and soil water distribution coefficients (see Ammann et al., 2020 for detailed information on the data). For the error of a measurement of type j at time t , we assume that

$$Z_j(t) \sim N(0, \sigma_{Z,j}(t)) \quad (9)$$

where N is the normal distribution and $\sigma_{Z,j}(t)$ is its standard deviation:

$$\sigma_{Z,j}(t) = a_j y_{0,j} \left(\frac{Y_{o,j}(t)}{y_{0,j}} \right)^{c_j} + y_{0,j} b_j \quad (10)$$

where $\psi_j = \{a_j, b_j, c_j\}$ are the parameters of the residual error model for observations of type j and $y_{0,j}$ is a scaling constant. Equation 10 is equivalent to the formulation used in Ammann et al. (2020) with the difference that we assume no temporal autocorrelation in $Z(t)$, since it should mostly represent the random component of the measurement error. Note that we cannot specifically account for systematic measurement errors such as the ones caused by faulty rating curves.

In our case, the model states, \mathbf{Y} , do not require a probabilistic formulation since they are a deterministic function of the parameters. Assuming independence of the priors of ξ_s and ψ , the joint distribution of parameters and potentially observed states, \mathbf{y}_{obs} , is given by:

$$f_{\text{joint}}(\theta_{-s}, \theta_s(t), \xi_s, \psi, \mathbf{y}_{\text{obs}}) = f_{Y_{\text{obs}}}(\mathbf{y}_{\text{obs}} | \theta_{-s}, \theta_s(t), \psi) f_{\text{pri}}(\theta_{-s}, \psi) f_{\text{OU}}(\theta_s(t) | \xi_s) f_{\text{pri}}(\xi_s) \quad (11)$$

where $f_{Y_{\text{obs}}}$ is the likelihood conditional on the time-dependent and constant parameters, which arises from the assumed statistical properties of the measurement error, \mathbf{Z} (Equation 9). $f_{Y_{\text{obs}}}$ is based on the likelihood framework proposed by Ammann et al. (2019) and its details for the specific application to the Ossingen data are given in Ammann et al. (2020). $f_{\text{OU}}(\cdot | \xi)$ is the probability density of the discretized OU-process given parameters ξ , and f_{pri} is the prior distribution. The latter consists of independent distributions for most of the individual parameters and we use hard boundaries on all of the marginals (see Appendix A). Regarding the prior for ξ_s , note that τ_s is fixed for all s due to theoretical difficulties in estimating that parameter (Roberts & Stramer, 2001). Liu and Gupta (2007) rightly argue that temporal variations of the parameters should be slower than those of the states, which reduces the peril of producing output response to input solely by the temporal dynamics of the parameter and not the states. This is why we fix τ_s at a relatively large value of ca. 10 days. μ_s and σ_s are inferred. Out of ψ , a is inferred and the other elements are fixed (see Appendix A).

According to the Bayesian inference paradigm, we learn about the parameters by conditioning Equation 11 on the variable \mathbf{y}_{obs} and by substituting it with actual observations:

$$f_{\text{post}}(\theta_{-s}, \theta_s(t), \xi_s, \psi | \mathbf{y}_{\text{obs}}) = \frac{f_{\text{joint}}(\theta_{-s}, \theta_s(t), \xi_s, \psi, \mathbf{y}_{\text{obs}})}{f_y(\mathbf{y}_{\text{obs}})} \propto f_{\text{joint}}(\theta_{-s}, \theta_s(t), \xi_s, \psi, \mathbf{y}_{\text{obs}}) \quad (12)$$

where the normalizing factor $f_y(\mathbf{y}_{\text{obs}})$ is constant w.r.t. the parameters. The procedure used to sample from f_{post} is described in Section 3.6.

Table 2
Model States and External Influence Factors Used as Features in Local Regression to Approximate the Maximum Likelihood Estimate of the Time-Dependent Parameter

Feature	Description
S_t	State of the topsoil reservoir
S_u	State of the unsaturated reservoir
S_d	State of the drainage reservoir
S_g	State of the groundwater reservoir
E_{pot}	Potential evapotranspiration
Precip.	Precipitation

Note. The humidity, precipitation and wind speed are based on data recorded by Doppler et al. (2012).

In order to keep the time-dependent parameters in reasonable ranges, and for reasons of numerical efficiency of the sampler, we redefine the time-dependent parameter as:

$$\theta_s(t) := u_s(\bar{\theta}_s + \theta_s^*(t)) \quad (13)$$

where $u_s(\cdot)$ is a scaled and shifted logistic function for $s = D$, and equal to the exponential function otherwise, $\bar{\theta}_s$ is a time-constant parameter and $\theta_s^*(t)$ is a dimensionless time-dependent parameter. The logistic function is used to asymptotically reach the relatively narrow bounds within which the prior for θ_D is nonzero (see Table A1). Through Equation 13, we hope to gain numerical efficiency, since the time-constant $\bar{\theta}_s$ becomes part of θ_{-s} , which allows for $\theta_s(t)$ to be additionally updated in the Metropolis sampling step in Equation 17, as opposed to only in Equation 19. The mean of $\theta_s^*(t)$ is kept fixed at 0 for the inference.

3.3. Finding Patterns With Local Regression and Cross-Correlation

Once the posterior distribution of $\theta_s(t)$ has been obtained for each parameter s , potential temporal patterns in $\theta_s(t)$ are of interest, since they could indicate possibilities to improve our deterministic model. In other words, such patterns would indicate that the introduced stochasticity has been misused to compensate for systematic model deficiencies. Such deficiencies would need to be corrected before a meaningful interpretation of the predictive uncertainty can be made. For example, if the evaporation multiplication factor was found to correlate with temperature, the parameterization of the potential evaporation would have to be re-evaluated.

Let us define $\hat{\theta}_s^* = \{\hat{\theta}_s^*(t_1), \dots, \hat{\theta}_s^*(t_n)\}$ as the component $\theta_s^*(t)$ of the posterior sample $\{\theta_{-s}, \theta_s^*(t), \zeta\}$ for which $f_{Y_{\text{obs}}}$ is maximal. Let us further assume that

$$\hat{\theta}_s^*(t_i) = g_s(\mathbf{x}(t_i), \mathbf{y}_g(t_i)) + \epsilon_i \quad (14)$$

where g_s is a function of external influence factors, \mathbf{x} , and of a vector \mathbf{y}_g that contains only a subset of all the states \mathbf{y} (see Table 2). ϵ_i are random deviations. If g_s can approximate $\hat{\theta}_s^*$ well, we have a chance of improving our model m by (i) using the empirical model g_s as is (i.a) or in a parameterized form (i.b) to replace some constant model parameters by their estimated temporal dynamics, or by (ii) reasoning about the deficits in physical process representation that cause the pattern g_s and adapt the process representation of m accordingly. In this paper, we choose approach (i.a). An example for (ii) would be the extension of the model structure by adding a new reservoir after identifying deficits in the constitutive function of an existing reservoir.

For the function g_s , we choose a weighted local regression model according to Cleveland et al. (1992) under the assumption of independent Gaussian deviations, ϵ_i , with a polynomial of order two, and a smoothing parameter equal to unity. As arguments of the function g_s , we use the features in Table 2 individually and in all possible combinations of two. We refrain from combining three or more features due to the limited benefit of going from one to two features (see Section 4.1). Weights, w_i , are introduced for the least squares regression to account for the fact that the identifiability of the parameters (and thus the information content in the posterior) varies over time. The weights are given by $w_i = 1 / \max\{\sigma_{\theta_s^*}^2(t_i), \sigma_{\min}^2\}$, where $\sigma_{\theta_s^*}^2(t_i)$ is the standard deviation of the posterior distribution of the dimensionless time-dependent parameter θ_s^* at time point t_i (visible e.g., in Figure 1) and σ_{\min} is a minimal standard deviation introduced to prevent overly large weights of time points where the parameter is at the boundary. For the purposes of this study, we set σ_{\min} equal to 5% of the mean of the posterior of σ_s . The weighted R^2 ,

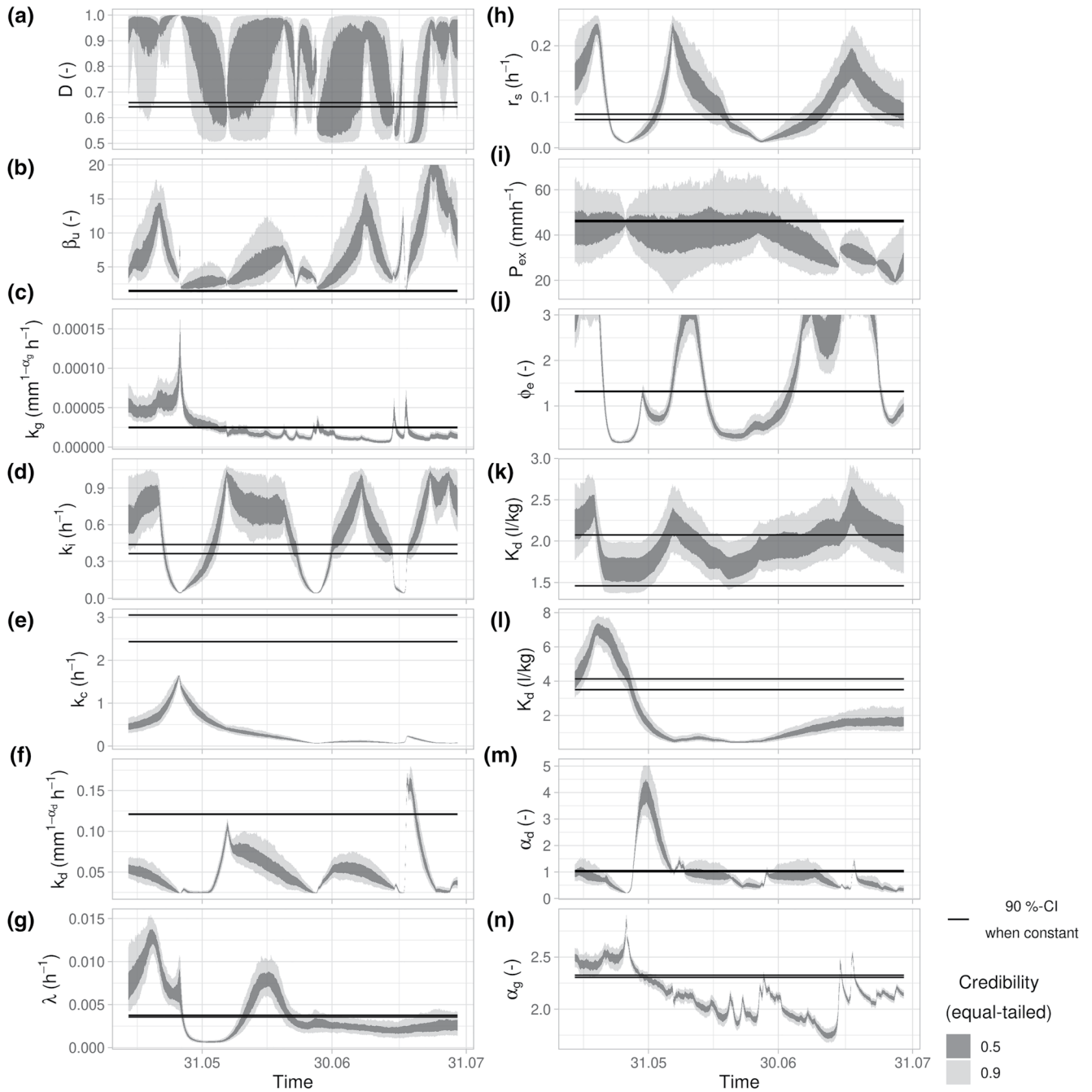


Figure 1. Posterior distribution (indicated by 50%- and 90%-credibility intervals) of $\theta_s(t)$ for all indices s (y-axes of panels a–n) during the study period in 2009. In panel (k), the soil/water distribution coefficient, K_d , is based on time-dependent $\theta_{s,t,z1}$ and constant $\theta_{s,t,z2}$, while it is vice versa for panel (l). The 90% credible interval inferred with all parameters being constant (i.e., with the deterministic model $m(\theta)$) are given by the horizontal lines for reference.

$$R^2 = 1 - \frac{\sum_i w_i \epsilon_i^2}{\sum_i w_i \left(\hat{\theta}_s^*(t_i) - \frac{\sum_j \sqrt{w_j} \hat{\theta}_s^*(t_j)}{\sum_j \sqrt{w_j}} \right)^2} \quad (15)$$

is used to quantify the goodness of fit of the regression.

In addition, we scan for potentially time-shifted dependencies between $\hat{\theta}_s^*(t)$ and $\{\mathbf{x}(t), \mathbf{y}_g(t)\}$, by calculating

$$\rho_{\theta_x}(\tau, k) = \text{corr}\left(\hat{\theta}_s^*(t), x^{(k)}(t - \tau)\right), \quad \rho_{\theta_{y_g}}(\tau, k) = \text{corr}\left(\hat{\theta}_s^*(t), y_g^{(k)}(t - \tau)\right), \quad -t_m < \tau < t_m$$

for each element k of the external influence factors and the states, where $\text{corr}(\cdot, \cdot)$ denotes the correlation and t_m is 2 weeks.

3.4. Model Improvement Based on Patterns Found in Time-Dependent Parameters

The deterministic empirical relationships between $\hat{\theta}_s$ and the influence factors, \mathbf{x} , and model states, \mathbf{y}_g , are used to improve the predictions made by model m by constructing a “hybrid” model:

$$\mathbf{y}_{\text{hyb}} = m(\boldsymbol{\theta}_{-s}, \tilde{\theta}_s^*(t)); \quad \tilde{\theta}_s^*(t) = g_s\left(\mathbf{x}(t), \mathbf{y}_{g, \text{hyb}}(t)\right) \quad (16)$$

where g_s is the optimal empirical model (identified as described in Section 3.3) approximating the temporal dynamics of $\hat{\theta}_s^*(t)$, and $\mathbf{y}_{g, \text{hyb}}$ consists of a subset of the states, \mathbf{y}_{hyb} , of this model. As an alternative and more flexible approach, one could choose another function that is based on g_s but has additional parameters, which could then be fitted together with $\boldsymbol{\theta}_{-s}$.

The posterior distribution of $\boldsymbol{\theta}_{-s}$ of the hybrid model is then obtained according to the procedure in Section 3.6. Note, however, that the time-dependent parameters of the hybrid model, $\tilde{\theta}_s^*(t)$, are not stochastic, but given by $g_s\left(\mathbf{x}(t), \mathbf{y}_{g, \text{hyb}}(t)\right)$ and therefore need not be inferred (the sampling steps in Equations 18 and 19 are not necessary). Remember that the multiplication factor, $\bar{\theta}_s$, is part of $\boldsymbol{\theta}_{-s}$, to still allow for an overall adjustment of the level of the parameter with index s . During inference, $\mathbf{y}_{g, \text{hyb}}$ might reach values that were not reached in connection with $\hat{\theta}_s^*(t)$ (based on which the function g_s was fitted). In that case, $\mathbf{y}_{g, \text{hyb}}$ is truncated so that it agrees with the support of g_s .

3.5. Prediction

The prerequisites for a meaningful interpretation of the uncertainties estimated with time-dependent parameters is that the stochasticity provided by $\hat{\theta}_s^*(t)$ is not misused to compensate for systematic deficits of the model. In addition to the procedure described in Sections 3.3 and 3.4, this should be verified by pseudo-/predictive cross validation, which is described in the following paragraphs.

The estimated uncertainty of the state variables of interest are obtained by propagating the joint posterior distribution obtained for $\boldsymbol{\theta}_{-s}$, $\boldsymbol{\xi}$, and $\boldsymbol{\psi}$ through the model. In the following, we distinguish between “intrinsic” and “residual” stochasticity. The former means all the uncertainty that is accounted for within the process model (i.e., through the constant parameters $\boldsymbol{\theta}_{-s}$ and the time-dependent parameters $\Theta_s(\boldsymbol{\xi})$ in Equation 7a), including all the uncertainty this causes in the internal states and the output. Residual uncertainty, on the other hand, is characterized by the term $\mathbf{Z}(\boldsymbol{\psi})$ (Equation 7c), which includes the effect of all additional sources of uncertainty not accounted for within the process model, such as measurement uncertainty of the output or some remaining structural uncertainty. For visualization, we estimate intrinsic stochasticity on one hand, and combined intrinsic *and* residual uncertainty on the other hand by propagating the posterior distributions of $(\boldsymbol{\theta}_{-s}, \boldsymbol{\xi})$ and $(\boldsymbol{\theta}_{-s}, \boldsymbol{\xi}, \boldsymbol{\psi})$, respectively, through the model given in Equation 7.

Note that in this study, due to the uniqueness of the data set and the limited period of high-frequency concentration measurements, the prediction period is the same as the calibration period (see Section 3.1), leading to the term “pseudo-prediction,” which is explained in more detail below. This prevents us from quantifying the prediction error to be expected for an independent time period. Nonetheless, the approach is suitable to avoid overfitting and to confirm that the additional degrees of freedom are not used to reproduce variability that should be produced by the deterministic part of the model. When “pseudo-predicting,” we neglect the posterior knowledge about the temporal dynamics of $\hat{\theta}_s^*(t)$ (even though it would be available in our pseudo-prediction period) to mimic prediction during a time period that is independent of the

calibration period. Therefore, like in a true prediction setting, the $\theta_s^*(t)$ used for the pseudo-prediction are random realizations of an OU-process with parameters ξ taken from the posterior sample. This means that the many degrees of freedom introduced through the time-dependent parameters, which constitute the main part of the danger of overfitting, are not used for prediction. Furthermore, we focus on the characterization of the resulting predictive uncertainty and its division into intrinsic and residual uncertainty, which is independent of whether we predict during the calibration or a validation period.

3.6. Numerical Implementation of Model and Sampler

The original model, $m(t, \theta)$, as well as the adapted model, $m(t, \Theta_s(t))$, which accepts time series of parameters, are implemented with the SUPERFLEX framework for conceptual hydrological models (Fenicia et al., 2011). In SUPERFLEX, the evolution of the internal states are governed by differential equations $\dot{\mathbf{y}} = f(t, \mathbf{y}, \theta)$ and the states, $\mathbf{y}(t_i)$, are obtained based on $\mathbf{y}(t_{i-1})$ with the implicit Euler method on the grid $t_i = t_0 + i\Delta_t$ with a fixed step size Δ_t of 15 min in our case. In the modified version with time-dependent parameters, $\theta_s(t_i)$ is assumed to be constant from t_{i-1} to t_i . The output for the observation time points, $o(\mathbf{y}(t_{\text{obs}}))$, is obtained through linear interpolation of $o(\mathbf{y}(t_i))$.

The hybrid model is implemented as an iterative procedure where the k -th step consists of evaluating $\mathbf{y}_{\text{hyb}}^{(k-1)} = m(\theta_{-s}, \tilde{\theta}_s^{*(k-1)})$ and then calculating $\tilde{\theta}_s^{*(k)} = g(\mathbf{x}, \mathbf{y}_{\text{g,hyb}}^{(k-1)})$ for all time points. The convergence criterion was chosen to be the maximal (over all time points) relative difference in $\mathbf{y}_{\text{g,hyb}}$ from one iteration to the next of below 1%.

We sample from f_{post} (Equation 12) according to the procedure developed by Buser (2003) and Tomassini et al. (2009). The ideas underlying this approach are described in detail in Reichert and Mieleitner (2009) and Reichert et al. (2020). The approach is based on Gibbs sampling across constant model parameters, parameters of the OU process that governs the time-evolution of the stochastic parameter, and the time-course of the stochastic, time-dependent parameter. In the k th step of this Gibbs sampling approach, we sample

$$(\theta_{-s}^k, \psi^k) \propto f_{Y_{\text{obs}}}(\mathbf{y}_{\text{obs}} \mid \theta_{-s}^k, \theta_s^{*,k-1}(t), \psi^k) f_{\text{pri}}(\theta_{-s}^k, \psi^k), \quad (17)$$

$$\xi_s^k \propto f_{\text{OU}}(\theta_s^{*,k-1}(t) \mid \xi_s^k) f_{\text{pri}}(\xi_s^k), \quad (18)$$

$$\theta_s^{*,k}(t) \propto f_{Y_{\text{obs}}}(\mathbf{y}_{\text{obs}} \mid \theta_{-s}^k, \theta_s^{*,k}(t), \psi) f_{\text{OU}}(\theta_s^{*,k}(t) \mid \xi_s^k). \quad (19)$$

Sampling of the constant parameters, (θ_{-s}^k, ψ^k) and ξ_s^k , within this Gibbs sampling scheme is done by Metropolis sampling with automatic adaptation of the proposal distribution during an initializing phase that is discarded as part of the burn-in phase. The stochastic parameter, $\theta_s^{*,k}(t)$, is proposed by sampling from the OU-process on sub-intervals (i.e., arbitrary sections of the whole simulation period) conditional on the values at the endpoints of that interval and accepting or rejecting based on the observation likelihood. By adapting $\theta_s^*(t)$ only within sub-intervals, we reach a higher acceptance rate, since the new proposal for $\theta_s^{*,k}(t)$ is largely similar to $\theta_s^{*,k-1}(t)$ and only differs within one sub-interval. This leads to faster convergence. On the other hand, one has to be aware that the split into sub-intervals increases the number of model runs required to complete one step of the outer Gibbs sampling procedure. In the present application, for most parameters that we made stochastic, we chose 50 sub-intervals, which results in a reasonably high acceptance rate for most time-dependent parameters. However, for some parameters that led to a relatively low acceptance rate during inference, we chose 100 sub-intervals. Convergence is assessed visually based on the Markov chains for the constant parameters, some selected time-points of the time-dependent parameters, the observational likelihood, $f_{Y_{\text{obs}}}$, and the posterior density, f_{post} .

To improve the numerical efficiency of the sampler, all parameters in θ_{-s} are log-transformed for sampling, except for $\bar{\theta}_D$ (see Equation 13). Parameters in ψ and ξ are not transformed.

We use the R Language for Statistical Computing (R Core Team, 2019) for all the computations and visualizations, except for the hydrological process model, m , which is written in Fortran. For the sampling procedure, we use the R-package “timedepar” (Reichert, 2020).

4. Results and Discussion

Convergence was achieved in most inference cases with time-dependent parameters and in the case of all parameters being constant in time (see supporting information). Convergence problems were observed for the time-dependent $\theta_{St,max}$ and $\theta_{Su,max}$ (the results of which are not shown), and for the infiltration excess parameter $\theta_{p_{ex}}$. The latter parameter is expected to be difficult to infer due to its threshold-like behavior and the fact that its effect on the output is restricted to very limited time periods. Inference of these parameters was, however, successful in the cases in which they were part of the constant parameters estimated jointly with other time-dependent parameters. No grave convergence problems were encountered for the other cases of time-dependent parameters and when inferring all parameters as constant in time.

4.1. Inferred Dynamics of Time-Dependent Parameters and Identified Patterns

For each inference approach with time-dependent parameter s , we obtain a posterior uncertainty estimate of $\theta_{-s}, \theta_s^*(t)$, and σ_s . The resulting temporal dynamics of $\theta_s(t)$ (obtained from $\theta_s^*(t)$ acc. to Equation 13) show distinctly different characteristics for the different time-dependent parameters s (Figure 1). While some parameters show rather high-frequency fluctuations (Figure 1; e.g., panels a, c, n), others vary on longer time-scales (Figure 1; e.g., panels g, h, k, l). They also differ in the relative width of their distribution at each time point compared to the amplitude of the temporal fluctuations. For example, panels (i) and (k) in Figure 1 show parameters with a wide uncertainty band relative to the amplitude of their temporal dynamics. Also, different parameters show different relative identifiability (indicated by the width of their distribution) during different time periods. This is especially pronounced for the parameter θ_D (Figure 1a), which controls the split of the fluxes into a fast and a slow response; this parameter is only identifiable during precipitation events and reverts to the prior otherwise. Similarly, the release rate θ_{k_i} of the reservoir representing the impervious areas shows a smaller uncertainty (and also lower values) during three distinct time periods (Figure 1d). Most of the parameters result in similar ranges of the marginal posteriors when time-dependent compared to the case of all constant parameters (Figure 1), which means that the interpretation regarding process representation of the individual parameters is consistent for the case of constant and time-dependent parameters. Exceptions are $\theta_{k_c}, \theta_{k_d}$, and θ_{β_u} . For the first two, the reason for this is that $\bar{\theta}_s$ goes to the lower boundary of the prior, which is the same as for the corresponding parameter in the case of all-constant parameters. Lower values for $\theta_s(t)$ are then reached by shifting $\theta_s^*(t)$ to values predominantly below 0.

For the case of $s = k_g$, Figure 2 shows the corresponding uncertainties in model output (streamflow and atrazine concentration), internal states (level of the groundwater reservoir, S_g) and the temporal dynamics of $\theta_{k_g}(t)$ in the calibration setting. Of the two parameters (θ_{k_g} and θ_{α_g}) that lead to the largest increase in the likelihood (see Figure 3a), we choose to show θ_{k_g} , since the interpretation is more intuitive. The spread of the posterior estimate of $\theta_{k_g}(t)$ does not vary strongly between inter-event and event periods (Figure 2d); the parameter is identifiable during both periods since the groundwater reservoir doesn't fall dry during the study period. Propagating the posterior distribution of $\{\theta_{-k_g}, \theta_{k_g}(t), \psi\}$ through the model leads to a very small uncertainty of the groundwater reservoir level, which is barely visible in Figure 2c. This outcome is expected during calibration.

There is a pronounced difference in the increase of the observational likelihood, $f_{Y_{obs}}$, obtained when making different parameters time-dependent compared to the reference case where $\theta_s(t) = \text{const.}$ (Figure 3a). The largest increase in $f_{Y_{obs}}$ is observed for the parameters related to the groundwater reservoir, θ_{k_g} and θ_{α_g} (for their meaning, see Equation 5). The strong increase of the likelihood is confirmed by the very good agreement of modeled and observed streamflow (Figure 2a). A smaller increase is found for the parameters that control the chemical processes ($\theta_\lambda, \theta_{\tau_s}, \theta_{St,z1}$, and $\theta_{St,z2}$). Obviously, the strong skewness in data abundance w.r.t. the different data types (large amount of streamflow data, limited number of measurements of

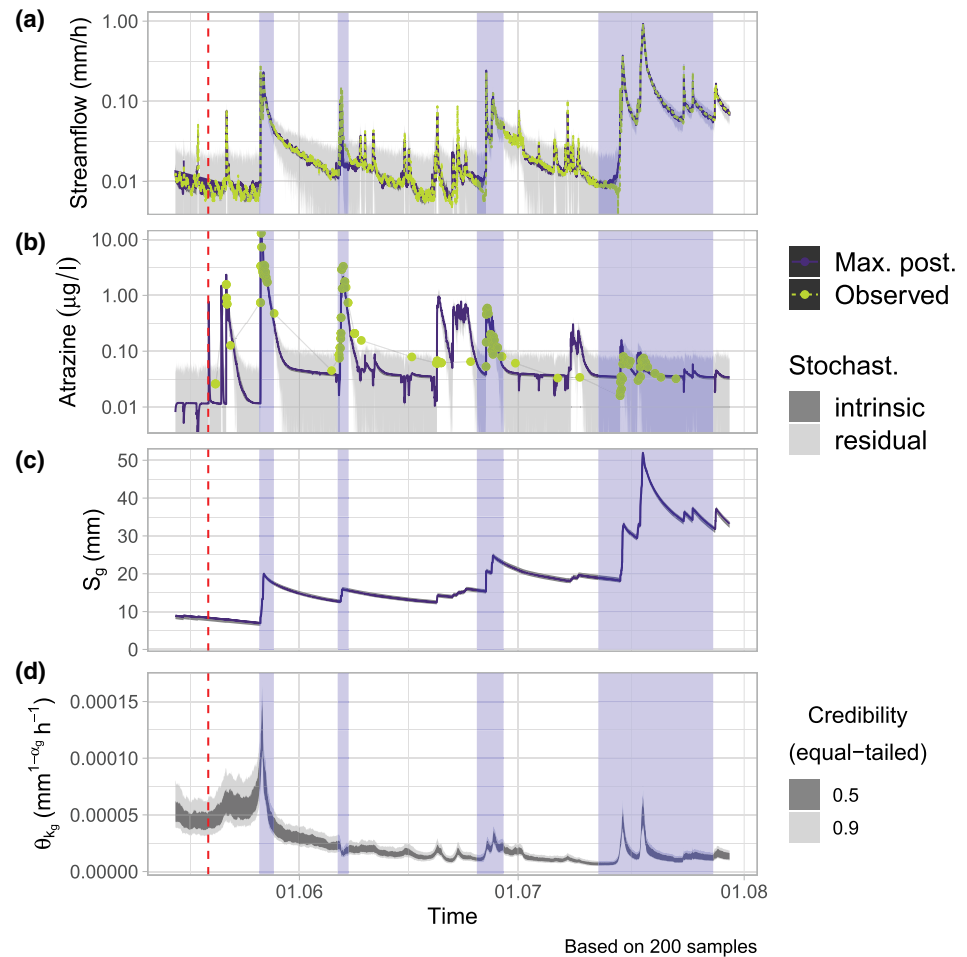


Figure 2. Posterior calibration uncertainty of model output (a and b), the level of the groundwater reservoir (c), and the time-dependent parameter θ_{k_g} (d) during the study period in 2009. The 90%-uncertainty bands are given for the intrinsic and residual stochasticity (a–c), while the 50% and 90% credible intervals are shown for the time-dependent parameter (d). The dashed red line indicates the application of herbicides and periods shaded in blue are the events zoomed in Figures 5 and 6.

concentration and soil/water distribution coefficients) is clearly reflected in the sensitivity of the observational likelihood to the degrees of freedom in the individual parameters.

The inferred temporal evolution of θ_{k_g} (Figure 2d) has very similar characteristics as the one of θ_{α_g} (see supporting information). In particular, both time series show a reasonable amplitude of the fluctuations. This is opposed to most other parameters controlling the hydrological processes (see Figure 1), for which the variability (in terms of σ_s) is larger than expected to a degree that must be seen as unreasonable. Note that the two parameters without exaggerated variability are the ones that cause the largest increase in likelihood when being made time-dependent (Figure 3a). This indicates that, in case of structural model deficits, additional degrees of freedom introduced w.r.t. θ_s can be misused (resulting in a large temporal variability of θ_s) to compensate for deficits unrelated to θ_s .

A comparison of the resulting dynamics of the time-dependent parameters when performing inference on streamflow data only, as opposed to the full data set, can reveal further interesting insights w.r.t. the sensitivity of the parameters to the different data types at different points in time. For example, Figure C3 in the appendix shows that the parameter for the shortcut reservoir, θ_{k_c} , is mainly informed by streamflow data, except during the major precipitation event in the end of May, which was the first one after spraying

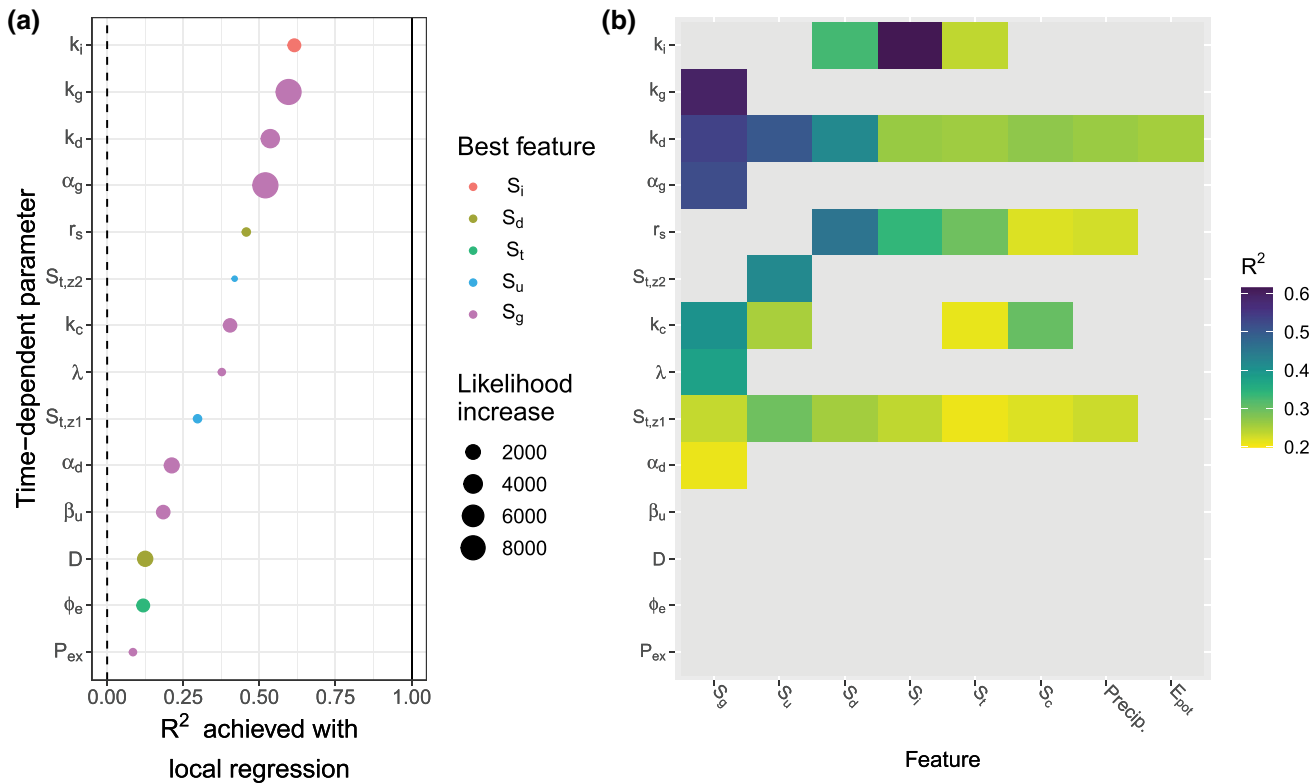


Figure 3. Values of R^2 obtained (x -axis) by approximating $\hat{\theta}_s^*(t)$ with a local regression model (g_s) when making the parameter with index s (y -axis) time dependent. (a) The color indicates which feature in the one-dimensional regression led to the highest R^2 , and the point size shows the increase in likelihood obtained compared to the case of no time-dependent parameters. (b) More detailed information on which of the features $\{\mathbf{x}(t), \mathbf{y}_g(t)\}$ (x -axis) resulted in which R^2 for which time-dependent parameter (y -axis). The y -axis is ordered according to the maximal R^2 achieved with the one-dimensional local regression, the x -axis is ordered according to which feature reaches the highest average R^2 for all parameters.

of herbicides. This shows that the model relies on this fast shortcut reservoir to produce the observed concentration peak. The equivalent comparison for all other parameters is given in the supporting information.

When approximating $\hat{\theta}_s^*(t)$ as a local regression of the features $\{\mathbf{x}(t), \mathbf{y}_g(t)\}$, we obtain values of weighted R^2 between close to 0 and 0.61 for the one-dimensional local regression (Figure 3b) and up to 0.72 for the two-dimensional local regression (Figure C2). A relatively large R^2 is found for the one-dimensional regression (0.6) for $\theta_{k_g}^*$ as a function of the level of the groundwater reservoir, S_g (Figure 4a). The identified relationship is “banana-shaped,” with higher values of the parameters coinciding with the lowest groundwater levels. The pattern for $\theta_{\alpha_g}^*$ is similar (Figure 4b). There also seems to be a tendency for θ_{k_g} and θ_{α_g} to increase again to very high groundwater levels (Figures 4a and 4b), however, this is only supported by a limited amount of data (one event only). The temporal patterns of θ_{k_g} remain largely unchanged in preliminary trials where it is inferred jointly with other time-dependent parameters. This increases confidence in the generality of the results described above. A closer look at the cause of the pattern in Figures 4a and 4b reveals that, in particular, April and August 2009 were very dry months, which caused the groundwater reservoir to fall below 10 mm (Figure 2). θ_{k_g} and θ_{α_g} increase during that period to produce the relatively large baseflow still observed in the stream. This suggests that the catchment’s baseflow is less affected by dry periods than can be represented with the static parameterization of a single nonlinear reservoir chosen in the model.

In order to allow for more sustained baseflow in dry periods, one could think of several model improvements, for example, an additional (slower) reservoir or an adaptation of the constitutive function of the existing groundwater reservoir. An advantage of the first option, is that it is effective also in time periods during which the existing reservoir is empty, while an advantage of the second option is the possibility of

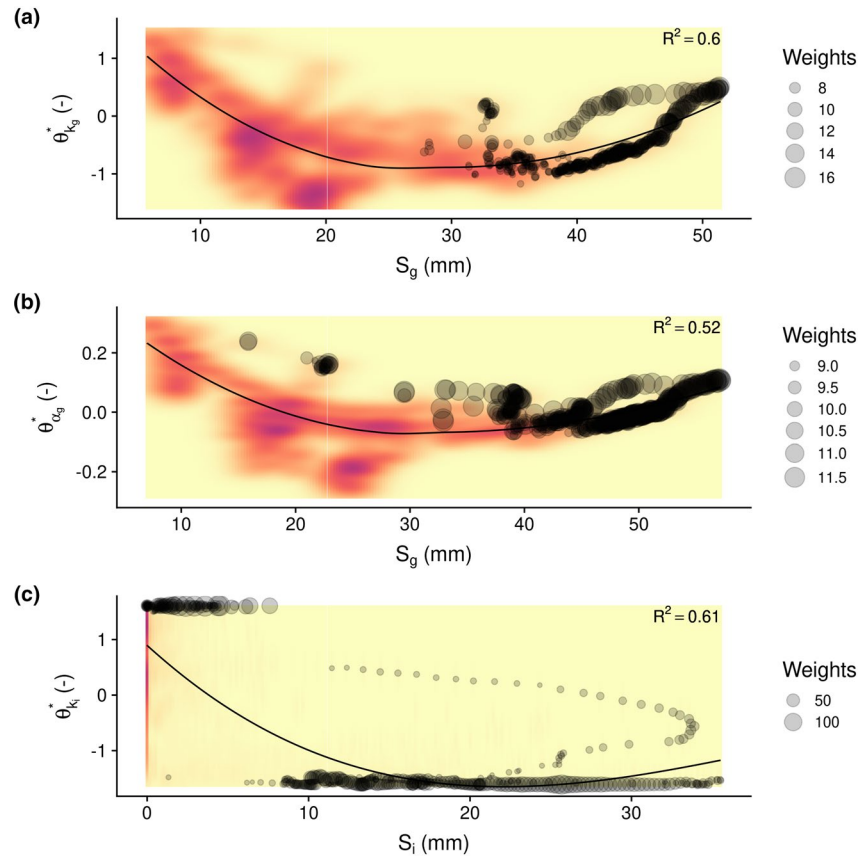


Figure 4. Weighted local regression fits (black lines) of the dimensionless form of three time-dependent parameters as a function of the features that led to the highest R^2 for those parameters. Weighted R^2 -values are also given. The weights were divided by their median for this plot and the data with the largest 1,000 weights are plotted as points. The red shading indicates the density of the remaining data. S_g is the state of the reservoir representing groundwater and $\theta_{k_g}^*$ (panel a) and $\theta_{a_g}^*$ (panel b) are parameters controlling its release rate, while S_i represents the impervious surfaces and $\theta_{k_i}^*$ (panel c) their release rate. The boundaries of $\theta_{k_i}^*$ are reached in (c).

keeping the number of additionally introduced parameters low (or even equal to zero). We refrain from introducing an additional slower reservoir, which would require at least two new parameters (a split and a release rate), since based on visual assessment, the baseflow is already reasonably well fitted even with the DPM with constant parameters (see Figure C1 in the appendix). Instead, we choose to adapt the constitutive relation of the existing groundwater reservoir as described in Section 3.4.

An elevated value of R^2 is also obtained for the parameter θ_{k_i} , which controls the response of the impervious areas. Figure 4c shows a surprisingly clear-cut threshold at $S_i \approx 8$ mm below and above which we tend to get small and large values of θ_{k_i} , respectively. Note that the limited flexibility of the regression function does not allow for a good approximation of this pattern. The pattern is interpreted as an activation of additional flow paths with a slower response time (compared to the very fast responding impervious areas) in cases where the reservoir reaches a sufficiently high level. For example, these might be paths including compacted gravel roads, which are abundant in the study catchment, and which tend to react slightly slower than completely sealed surfaces.

The other strong correlations visible in Figure 3b are judged to be of less interest, since the pattern for the pair (k_d, S_g) is dominated by a single event, $(S_{t, z2}, S_u)$ and (r_s, S_d) are judged to be coincidental, and (k_c, S_g) is similar to (k_i, S_i) . As an example for a potentially coincidental pattern, we discuss the pair $(S_{t, z2}, S_u)$ in more detail in Appendix B.

Table 3

Natural Logarithm (and Std. Dev. in Brackets) of the Observational Likelihood, $f_{y_{\text{obs}}}$, Reached During Inference With Time-Dependent Groundwater Reservoir Parameters When Approximating the Temporal Dynamics by the Local Regression Predictor g_s (Resulting in $\tilde{\theta}_s^*(t)$), and When Inferring the Temporal Dynamics of $\theta_s^*(t)$ Directly

No time-dep. parameter	Time-dep. parameter	$\tilde{\theta}_s^*(t)$	$\theta_s^*(t)$
84,042 (31)	$s = k_g$	86,410 (34)	92,836 (31)
	$s = \alpha_g$	84,919 (31)	92,905 (28)

Note. The result for the original model with no time-dependent parameters is also given for reference.

Regarding the cross-correlation analysis, we find correlations of around $\rho = 0.6$ or weaker for all the combinations of θ_s^* and $x^{(k)}$ or $y_g^{(k)}$. Many of the strongest correlations have already been discussed above or do not have an obvious interpretation that is physically meaningful. One potential exception is the exponent governing the nonlinearity of the response of the unsaturated reservoir, $\theta_{\beta_u}^*$, which is negatively correlated ($\rho = -0.6$) to the water content of the reservoir representing the unsaturated zone, S_u , 10 days earlier. However, this finding is more an example of how an expert's reasoning can be stimulated and does not prove a causal relation between the two quantities. For that, we are missing longer time series of observations, which would enable a more robust analysis.

In summary, this analysis reveals that there are meaningful patterns in the inferred temporal dynamics of the parameters that allow for interpretation by the domain expert, which can inspire potential model improvements. This confirms the findings of previous hydrological modeling studies, which have found patterns in output innovations when estimating states (Vrugt et al., 2005) or patterns in the temporal dynamics of time-dependent parameters (Pathiraja et al., 2016a), and extends them to a WQ modeling case study. However, not all the identified patterns might be physically meaningful; a potential example is the relation between $\theta_{s_{t,z2}}$ and S_u . This shows that such patterns need to be interpreted with great caution, especially if they are based on limited data that does not allow for an independent temporal validation.

4.2. Model Improvement by Incorporating Found Patterns

The patterns shown in Section 4.1 that are most promising for improving the DPM are the ones identified for θ_{k_g} and θ_{α_g} . This is due to their relatively large effect on the observational likelihood (Figure 3a) and due to the rather good interpretability of the mechanisms that cause the identified patterns. Table 3 shows an expected result; when estimating the dynamics of the time-dependent parameters through regression (with the hybrid model $m(\theta_{-s}, \tilde{\theta}_s^*(t))$, column 3) the observational likelihood is higher than the one reached with time-constant parameters (with base model $m(\theta)$, column 1), but lower than the one reached when estimating the dynamics directly from the data (with the stochastic model $m(\theta_{-s}, \Theta_s(t))$, column 4). This is true for both, $s = k_g$ and $s = \alpha_g$. In both of these cases, however, the hybrid model is closer to the deterministic than to the stochastic model, which indicates that exploiting the information contained in the temporal variation of the time-dependent parameters leads to a rather limited model improvement in this case. This means that most of the variation (in terms of likelihood improvement) of $\theta_s^*(t)$ is probably “true” stochasticity and only a minor part is caused by the systematic deficit revealed in Figures 4a and 4b. The predictive uncertainties (Section 4.3) obtained based on the inferred σ_s for those parameters are therefore valid. For an example where the revealed deficits and the inspired improvements are more significant, see Reichert et al. (2020).

Note that we opted for a static function, g_s , to improve the parameterization of the reservoir representing groundwater. More flexibility could have been traded for less accuracy by choosing a more general adaptation of the constitutive relationship of the reservoir, thereby introducing additional parameters. This might have led to a larger (or a smaller) increase in $f_{y_{\text{obs}}}$.

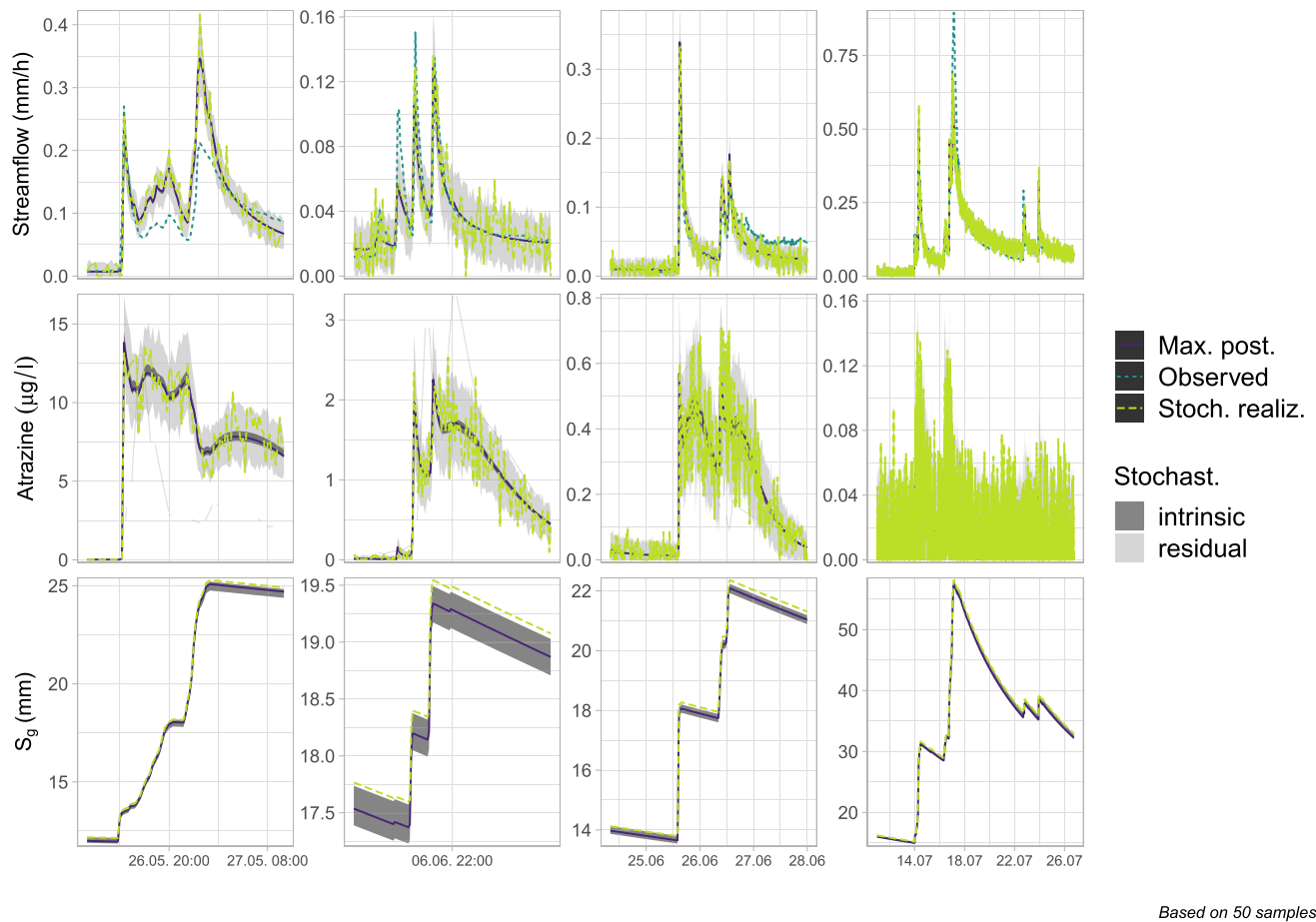


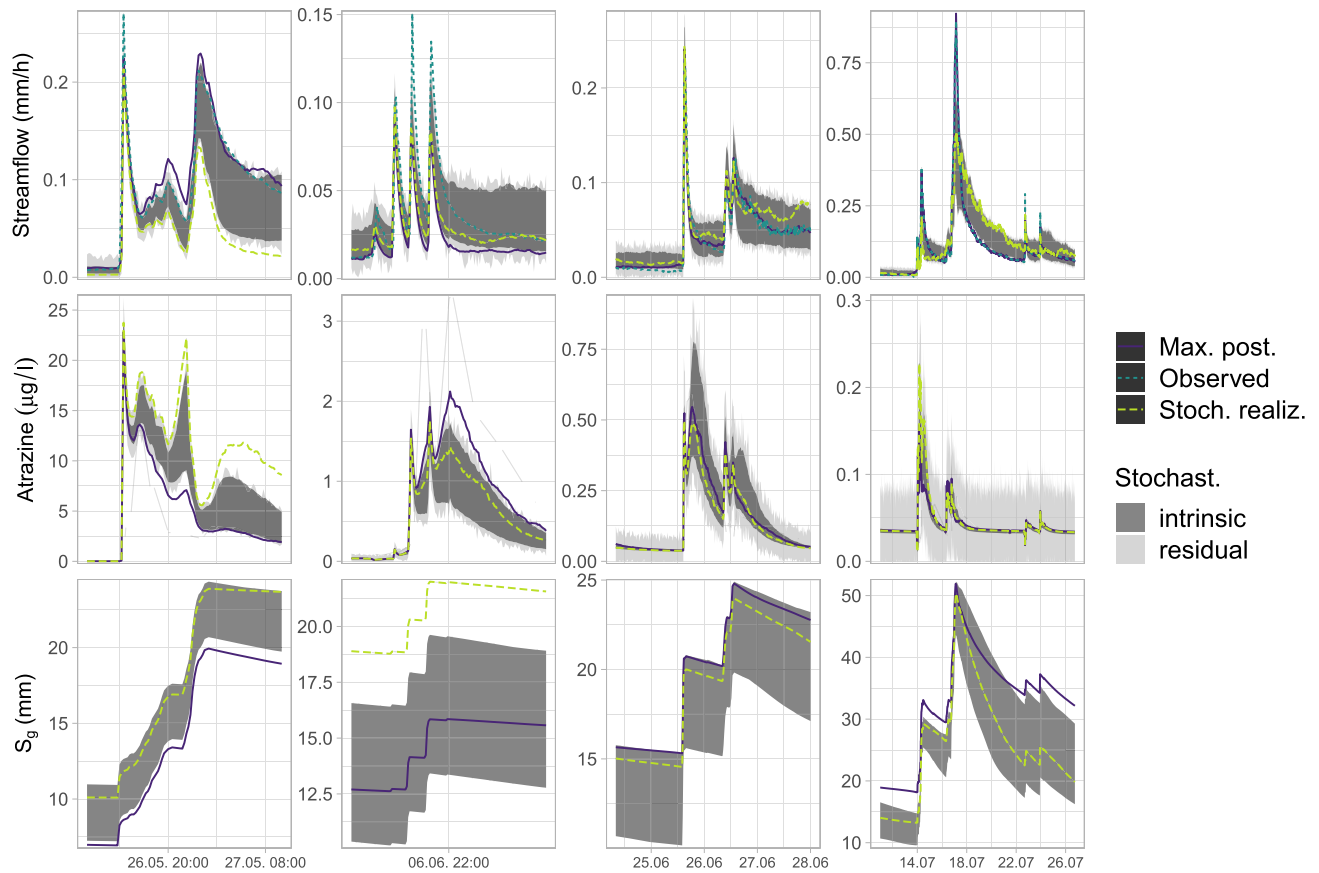
Figure 5. Parametric (intrinsic) and residual uncertainty estimated with all parameters being constant in time. The shaded areas cover 90% of the predictions.

4.3. Estimated Uncertainties Based on Bayesian Posterior Distributions

When using the DPM, the predictive uncertainty for streamflow and herbicide concentrations is dominated by the residual stochasticity, which is much larger than the parametric uncertainty (Figure 5). The combination of the two yields error bands of reasonable width for the type of observations, j , for which we formulated an error model according to Equation 10. The unobserved states, for which one did not formulate a residual error model, for example, S_g , have only parametric uncertainty, which seems too small, especially compared to the uncertainty of observed states like streamflow (Figure 5). Furthermore, the stochastic realizations of model output for an observed state, say streamflow, have to be produced by sampling the residual stochasticity, since it combines multiple relevant sources of uncertainty. However, this leads to model realizations with unreasonably weak temporal short-range correlation compared to the observations (see the first and second row in Figure 5).

Regarding the predictions with the SPMs, we focus on the one with time-dependent parameter θ_{k_g} (Figure 6), since it led to the largest increase in the observational likelihood. The results regarding predictive uncertainty that we obtain for $s = \theta_{a_g}$ (see supporting information) are similar due to the similar structural function of the parameters. The predictive uncertainties obtained for all other parameters made time-dependent are given in the supporting information.

Figure 6 shows that the total output uncertainty is dominated by the intrinsic stochasticity whereas the residual stochasticity is rather small. This is opposed to the dominant role of the residual stochasticity for the deterministic model (Figure 5). Internal model states, such as the level of the groundwater reservoir, show a larger (probably more realistic) uncertainty in the case of the stochastic model (Figure 6) compared to the



Based on 50 samples

Figure 6. Parametric (intrinsic) and residual uncertainty estimated with time-dependent parameter θ_{k_g} . The shaded areas cover 90% of the predictions.

deterministic model (Figure 5). This is of practical relevance whenever we are interested in such internal states. The uncertainty that characterizes our knowledge of states, output, and observations is much smaller during calibration (Figure 2) than during prediction (Figure 6) when relying on the SPM, which is a very intuitive result. This is not the case for the DPM. Furthermore, with the stochastic model, we can easily obtain realizations of the states that have both, a reasonable autocorrelation and a reasonably large uncertainty (green dashed line, Figure 6). With DPMs, a similar result would require considering autocorrelation in the residual error term, which has been shown to be problematic (e.g., Evin et al., 2013), especially for high-frequency data (e.g., Ammann et al., 2019) like the one used in this study. Note that another study successfully considers autocorrelation in the residual error term and finds stronger model deficits (Reichert et al., 2020).

4.4. Challenges, Recommendations, and Transferability

One major issue we encountered, which we believe to be general, is related to the additional degrees of freedom introduced by making different parameters stochastic. The freedom w.r.t. some stochastic parameters led to exaggerated temporal variability in these parameters, which was misused to reduce errors originating from other sources. In particular, parameters to which the likelihood was rather insensitive suffered from this deficiency, presumably because large changes in these parameters are necessary to achieve a higher likelihood. This shows that failing to constrain the stochastic parameters through an adequate prior for their asymptotic standard deviation, σ_s , which is difficult in our experience, can lead to unreasonable behavior of the SPM. We find that an effective solution to this is to make the parameter that has the largest impact on the likelihood time-dependent, which is also conceptually reasonable, since in that case, the degrees of freedom are introduced where they are most needed. A systematic procedure for selecting which parameters to

make time-dependent is still to be developed in future studies. Identifiability issues are expected to increase with more time-dependent parameters. A preliminary analysis in which we infer two time-dependent parameters jointly (see supporting information) shows that the two fast reservoirs S_i and S_c are less misused as slow reservoirs when their release rates are fitted jointly with θ_{k_s} of the slow reservoir as compared to when they are fitted as single time-dependent parameters. This is an indication of potential benefits of inferring multiple time-dependent parameters, given that the most sensitive parameters are included.

Sampling from the posterior distribution was numerically challenging; the chosen sampler required up to 100 evaluations of the DPM for one iteration of the Markov chain in order to achieve a reasonable acceptance rate. This resulted in a rate of ca. 2,500 samples per week on a modern processor. This sampler is therefore only applicable if at least one of the following conditions is met: (a) the model is cheap to run, (b) a limited amount of information is contained in the inferred time series (i.e., a short length or a small signal-to-noise ratio), or (c) only part of the simulation period has to be re-evaluated if the value of the time-dependent parameters have changed only on a short sub-interval.

When choosing the parameter(s) to be made time-dependent, we recommend to focus on the ones to which the likelihood function is most sensitive. The degree of flexibility of the empirical model, g_s , on which the screening is based, should be reasonable in light of the amount of data; a g_s that is too simple will conceal interesting dependencies, an overly complex g_s will find many coincidental relationships that lack physical meaning. Since it is not a priori clear what the optimal complexity is, such patterns need to be interpreted with great caution. Section 3.3 gives an example of how to scan the potentially large amount of results for patterns (visualized in Figures 3 and 4), which are expected to work well for a large class of problems. Section 3.4 presents a generally applicable empirical improvement of the original model. It might be preferable to design more process-based changes, but these are much more dependent on the case study and should be formulated by the respective domain expert.

Since we explore a single case study, we can only make limited statements about the general applicability of the stochastic parameter approach to other problems or fields. On the other hand, the study involves a coupled hydrological and WQ model and joint inference to three types of data with vastly different number of observations each. This proves the transferability of the approach from purely hydrological applications (Reichert & Mieleitner, 2009; Reichert et al., 2020) to WQ models, which increases the confidence in further transferability to other models, case studies, or domains. In WQ modeling, approaches like this are especially needed due to the large uncertainties and the lack of data, which often precludes other flexible approaches that are more data-hungry, like machine-learning techniques.

5. Conclusions

This study has shown that SPMs based on time-dependent, stochastic parameters can be used at the catchment scale to assess dynamic herbicide pollution from diffuse sources on short timescales. This is one step in a potentially much broader future application of similar techniques in WQ modeling, a field in which SPMs have not received sufficient attention so far.

The analysis with time-dependent parameters revealed that the parameters controlling the hydrological processes lead to a larger increase in the observational likelihood (i.e., goodness of fit) when being made stochastic than the parameters controlling the chemical processes. Based on the temporal dynamics of the inferred parameter time series we conclude that the release rate of the slow reservoir seems to be slightly underestimated during dry conditions with the original deterministic model. After verifying that the predictive uncertainty is reasonable and after accounting for the identified deficit by adapting the DPM and repeating inference, we find that the increase in model performance in terms of observation likelihood is rather limited. Therefore we conclude that the identified deficit of the original model is not severe, which is a prerequisite for a save interpretation of the uncertainty estimates obtained with the chosen approach.

The resulting temporal dynamics of the time-dependent parameters were shown to contain information that is interpretable by the modeler or the domain expert. The spread of the posterior distribution of a parameter's time-series provides additional information about the periods during which the observed model output is especially sensitive to this parameter. This proves that the chosen method is able to disclose

systematic deficits of a model's internal process formulation that have been unaccounted for so far. It is therefore a promising tool to produce insights that stimulate experts' reasoning about the respective parts of the system and about potential model improvements. We showed that the SPM leads to a more reasonable distinction of process-related and residual stochasticity than the DPM. This allows for a better separation of the predictions of the true states of interest from their random observation error. On one hand, this leads to model output that better resembles the observations in terms of elevated autocorrelation. On the other hand, with the SPM, we obtain a more realistic uncertainty of internal model states and of the predicted true output, and a smaller residual (observation) error than with the DPM.

One major practical challenge when applying the tested approach turned out to be the control over the excessive degrees of freedom introduced with time-dependent parameters. This calls for further investigations into the possibilities of constraining the stochastic parameters, for example, through the prior distribution. Other challenges of more numerical nature were linked to the chosen sampler. These suggest that the sampler is only applicable if rather restrictive conditions are met and further research is needed to design more efficient samplers capable of inference with stochastic models. Future research should also focus on the application of SPMs to longer time series of observations that allow for proper calibration/validation splits and on better exploring the capability of such models to account for the large uncertainty in input that is often encountered in WQ settings, but which was of limited relevance in this study. Overall, there is a large potential benefit of using more SPMs to assess WQ problems at the catchment scale in the future, since we are often forced to use structurally deficient models to describe systems that obviously behave highly stochastic at the resolution we can observe them.

Appendix A: Prior Parameter Distribution

The prior distribution we use for the parameter vectors θ and ψ are independent distributions for each of the parameters contained in those vectors (Table A1). The only exceptions are $\theta_{s,t,z1}$ and $\theta_{s,t,z2}$, which are believed to be correlated and which are therefore described by a two-dimensional lognormal distribution with nonzero off-diagonal elements. The prior in this study is mostly equivalent to the prior used by Ammann et al. (2020), where more detailed information can be found.

The only changes in the prior compared to Ammann et al. (2020) are (i) the replacement of the uniform distribution for D with a lognormal distribution, (ii) the adaptation of the lower boundary for $\theta_{\alpha d}$, (iii) the adaptation of the upper boundary for θ_{k_i} , and (iv) the prior for $a \in \psi$ and the neglect of autocorrelation in the residual error model. Change (i) was made because a lognormal distribution is a more realistic approximation of our expectation of this parameter than a uniform distribution. The effect on the results is believed to be insignificant since we chose a rather large standard deviation. Change (ii) was made to acknowledge the expected increase in rate of macropore flow with wetting of the soil.

Change (iv) expresses the new interpretation of the residual stochasticity in the face of a SPM as opposed to a DPM; compare Equations 2 and 4. The parameter controlling autocorrelation, denoted as γ_j in Table A1, is fixed at zero for all types of observations j , to account for the fact that when using a SPM, the residual stochasticity is expected to characterize measurement errors only, which are assumed to be uncorrelated in time. However, considering correlation of the errors was a convenient way of achieving a reasonable weighting between observation types with a lot of data and such with little amounts of data. When neglecting correlation, we have to prevent unreasonably large a_{atra} and a_{terb} resulting from the preference of the model to fit streamflow due to the much more abundant data. We do so by choosing narrower priors for a_{atra} and a_{terb} .

The prior distribution for $\bar{\theta}_s$, which is effectively the median of the time-dependent parameter $\theta_s(t)$, are the same as for the constant version of the parameter with index s (see Table A1). The asymptotic standard deviation of the OU-process, σ_s , is equal to one fifth of σ listed in Table A1. In addition, we limit the values of θ_s^* to a range of $[-1.5, 1.5]$ to prevent overly large fluctuations of the time-dependent parameters.

Table A1
Prior Distribution of the Parameters of the Transport Model, θ_{-s} , the Residual Error Model, ψ , and the OU-Process, ξ

Parameter index	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	LN	-	7.00e-01	3.00e-01	5.00e-01	1.00e+00
$S_{t,max}$	LN	mm	1.50e+01	1.50e+00	1.00e+01	5.00e+01
$S_{u,max}$	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	1.00e+00	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
$S_{t,z1}$	LN	mm	2.28e+01	1.21e+01	1.00e+00	1.10e+03
$S_{t,z2}$	LN	mm	5.07e+01	2.70e+01	1.00e+00	1.10e+03
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
k_i	LN	h ⁻¹	6.13e-01	3.27e-01	7.33e-02	1.09e+01
a_Q	LN	-	8.22e-02	2.11e-03	5.00e-02	9.00e-01
b_Q	Fixed	-	2.00e-01	-	-	-
c_Q	Fixed	-	7.00e-01	-	-	-
γ_Q	Fixed	h	0.00e+00	-	-	-
$y_{0,Q}$	Fixed	mm h ⁻¹	4.00e-02	-	-	-
a_{atra}	LN	-	8.22e-02	2.11e-03	5.00e-02	9.00e-01
b_{atra}	Fixed	-	2.00e-02	-	-	-
c_{atra}	Fixed	-	8.00e-01	-	-	-
γ_{atra}	Fixed	h	0.00e+00	-	-	-
$y_{0,atra}$	Fixed	μg l ⁻¹	1.00e+00	-	-	-
a_{terb}	LN	-	8.22e-02	2.11e-03	5.00e-02	9.00e-01
b_{terb}	Fixed	-	2.00e-02	-	-	-
c_{terb}	Fixed	-	8.00e-01	-	-	-
γ_{terb}	Fixed	h	0.00e+00	-	-	-
$y_{0,terb}$	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	-	-	-
γ_K	Fixed	h	0.00e+00	-	-	-
$y_{0,K}$	Fixed	l kg ⁻¹	1.00e+00	-	-	-

Note. The distribution types are: LN, lognormal, U, uniform, Exp, exponential, “Fixed”: parameter was fixed at the respective value and not inferred.

Appendix B: Temporal Dynamics and Predictive Uncertainty Obtained for Time-Dependent $S_{t,z2}$

A rather high R^2 (0.42, Figure 3b) was obtained when describing the temporal dynamics of $\theta_{S_{t,z2}}$ as a function of S_u . The coinciding dynamics of the two quantities are shown in the original space in the last two rows of Figure B1. However, a dependency of the two quantities would be in contradiction of the physical understanding we have of sorption processes. Especially the peak of $\theta_{S_{t,z2}}$ close to the application day and the following sharp decrease (Figure B1) is unrealistic and probably caused by compensating the models tendency to overestimate the concentration during a small rainfall event shortly after application. On the other hand, the subsequent slow increase in sorption equilibrium during the second half of the experiment (Figure B1) is in agreement with a process called “aging” of soils. This process has been observed at many sites (Wauchope et al., 2002) and was also observed in the study catchment (Camenzuli, 2010), but it was not included in the model m to keep parsimony. The results of the two-dimensional local regression reveal a significant increase of R^2 to 0.63 when additionally including S_g as an explanatory variable for $\theta_{S_{t,z2}}^*$ (Figure C2). Note that this should be interpreted with great caution, as the concentration measurements that support the inference of $\theta_{S_{t,z2}}$ are limited and contain only a few precipitation events.

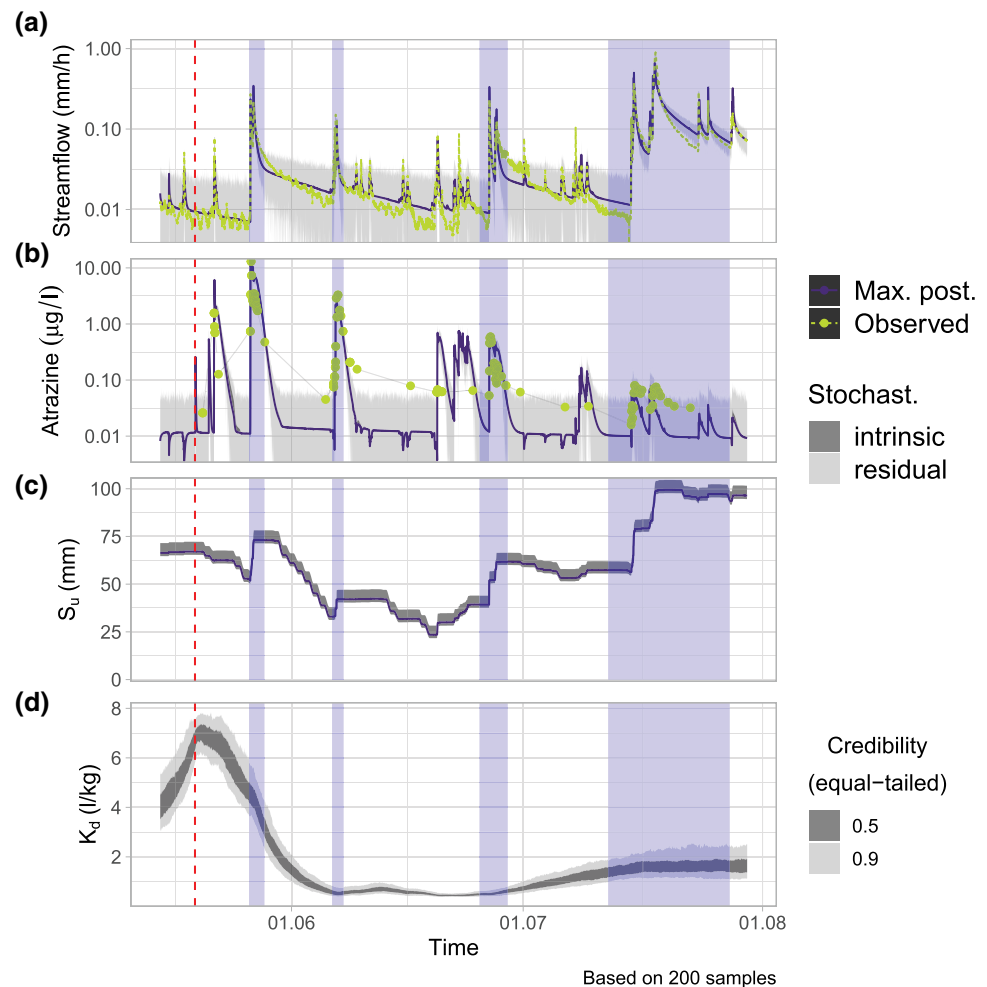
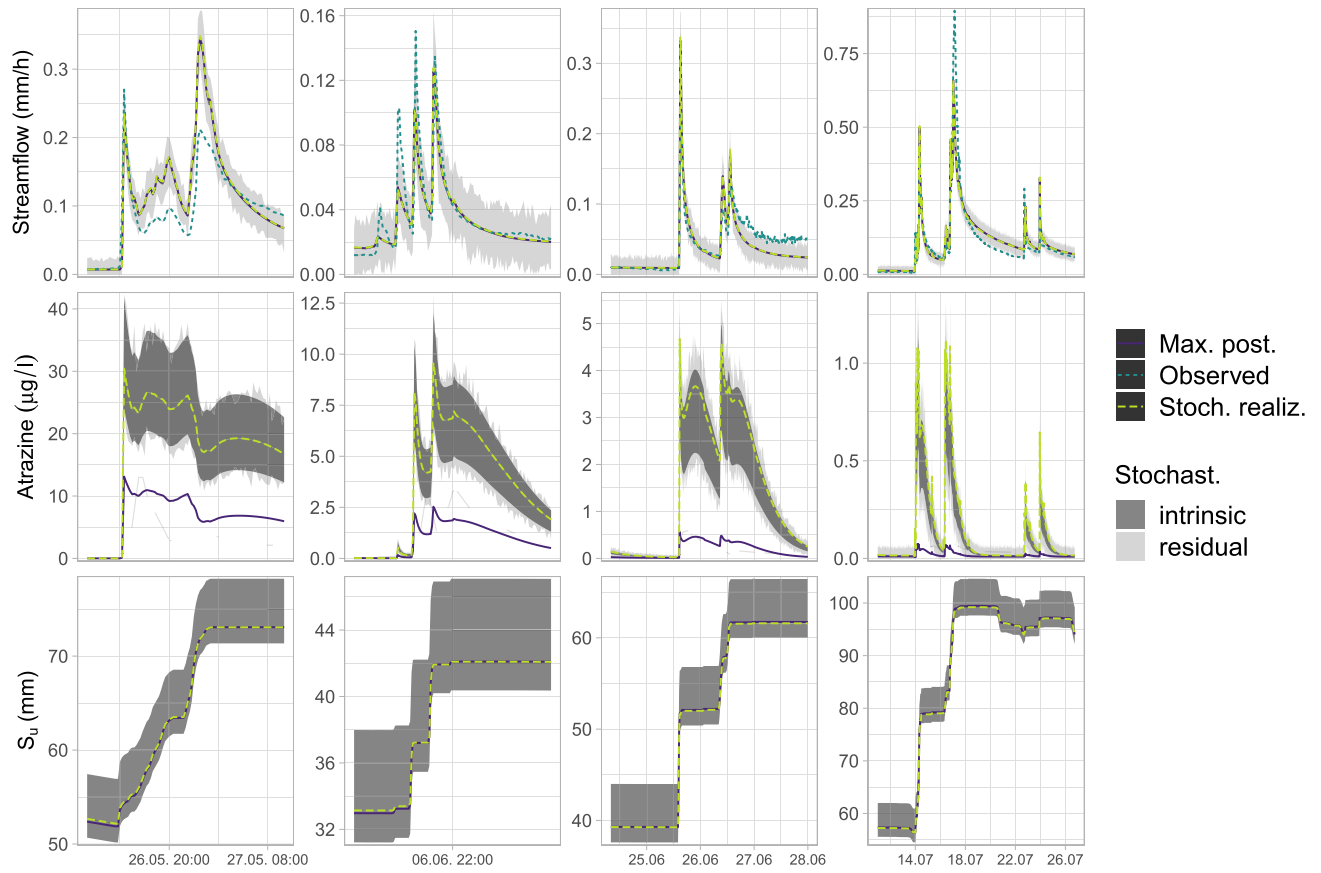


Figure B1. Posterior uncertainty estimates of model output, internal states, and $\theta_i(t)$ obtained based on inference with time-dependent parameter $\theta_{S_{t,z2}}$. For easier interpretation, panel (d) shows the soil/water distribution coefficient, K_d , calculated based on $\theta_{S_{t,z1}}$ and $\theta_{S_{t,z2}}$. The dashed red line indicates the application of herbicides. Areas shaded in blue indicate event periods zoomed in Figure B2.



Based on 50 samples

Figure B2. Parametric and residual uncertainty estimated when making $\theta_{s_{t,z2}}$ time-dependent.

Appendix C: Additional Results

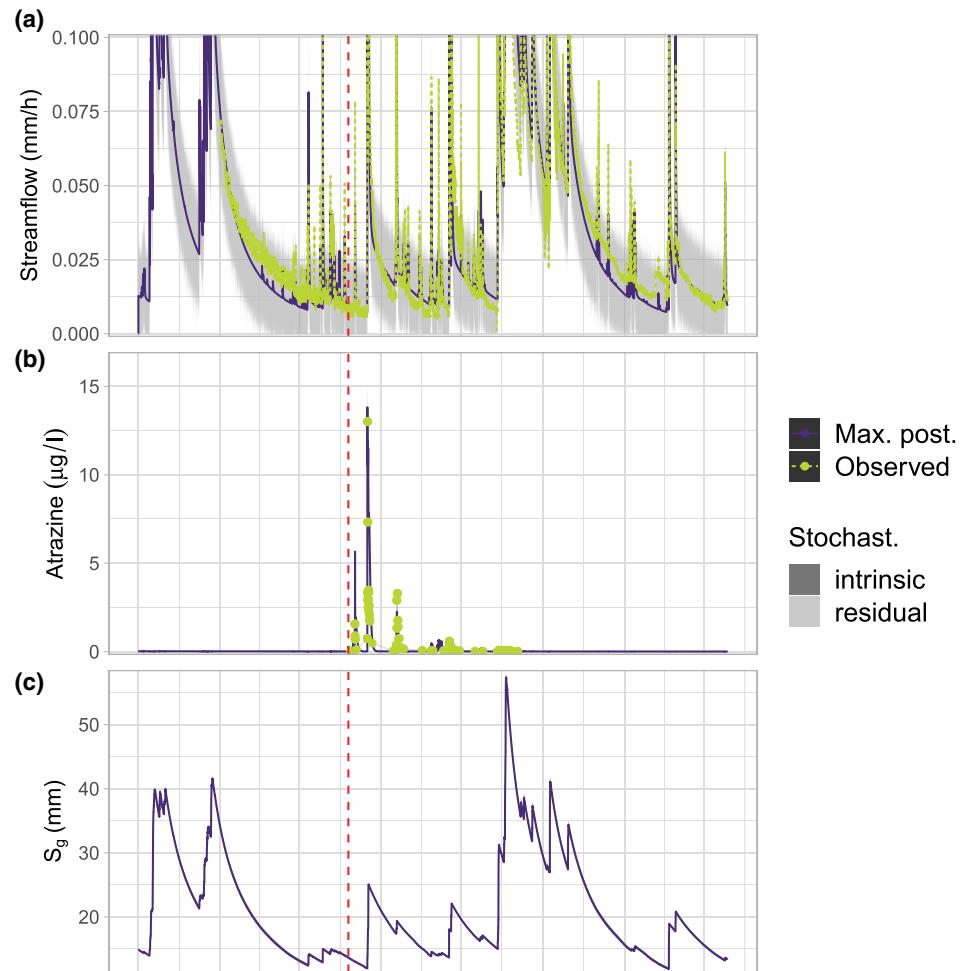


Figure C1. Maximum posterior prediction and 90% uncertainty bands obtained with the deterministic process model $m(\theta)$ with all parameters constant. The uppermost panel is zoomed to the baseflow for better visibility.

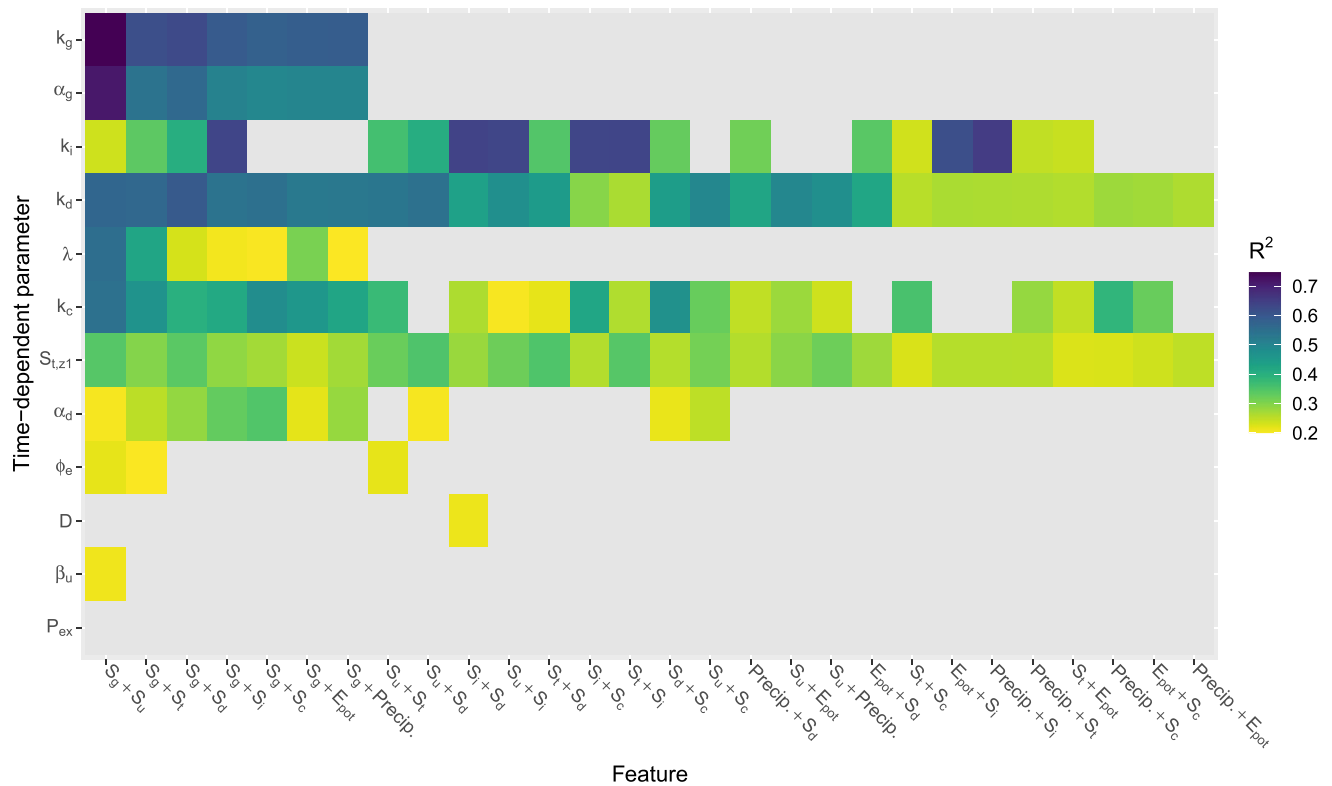


Figure C2. Results of the two-dimensional local regression. Pairwise combination of the features $\{\mathbf{x}(t), \mathbf{y}_g(t)\}$ (x-axis) and the resulting R^2 for the time-dependent parameter (y-axis). The y-axis is ordered according to the maximal R^2 achieved, the x-axis is ordered according to which combination of the features reaches the highest average R^2 for all parameters.

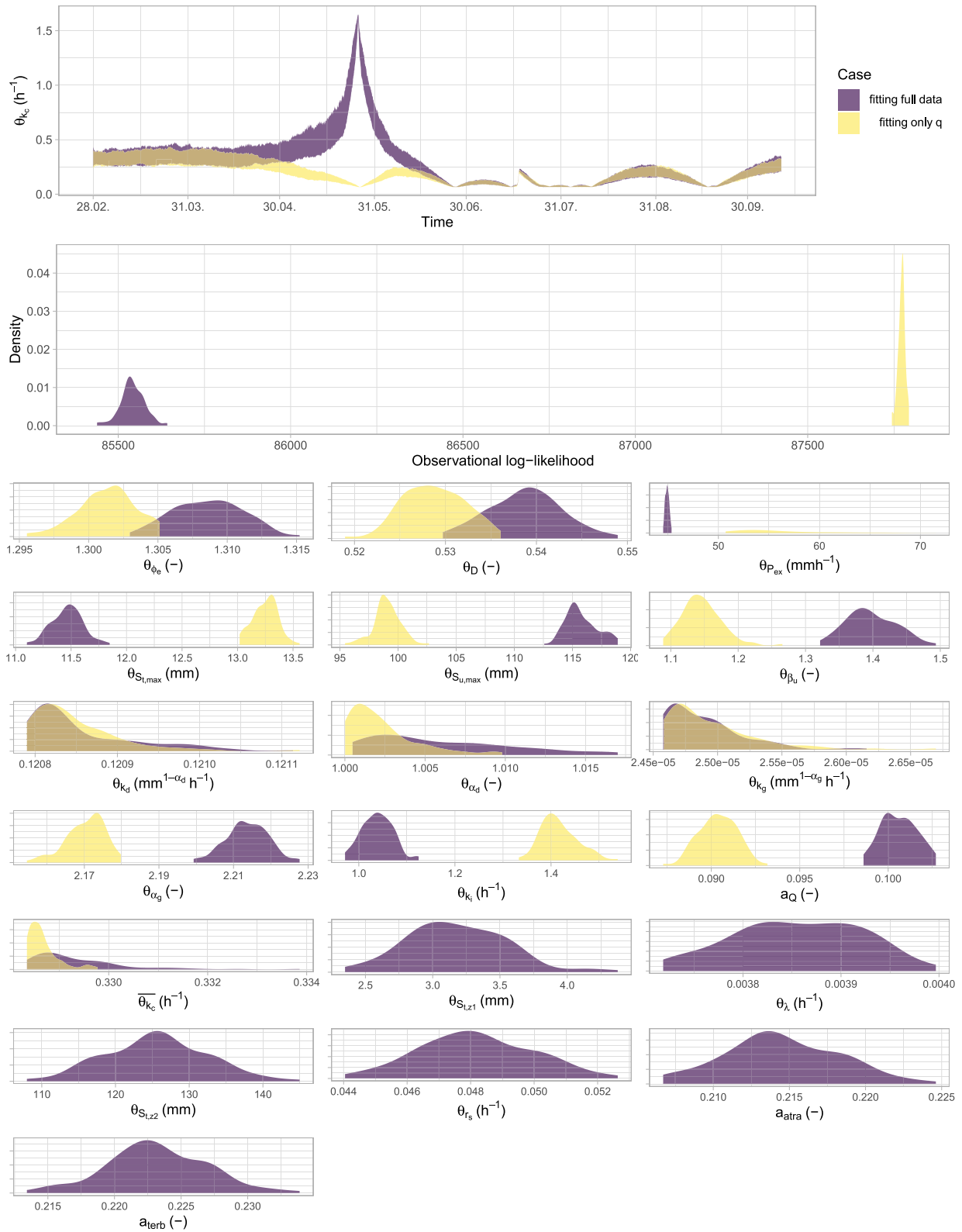


Figure C3. Comparison of posterior marginal distributions obtained when performing inference with time-dependent parameter θ_{k_c} based on streamflow only versus based on the full data set. The top two panels show the time-dependent parameter and the obtained observational likelihood, respectively. The lower panels show the marginals of the constant parameters.

Data Availability Statement

The data underlying this article will be made available upon publication through the Eawag Research Data Institutional Repository, accessible at <https://opendata.eawag.ch>.

Acknowledgments

The authors would like to thank Tobias Doppler for his valuable inputs regarding the discussion of the results, and for his feedback on draft versions of this manuscript. Jonas Sukys, Marco Bacci, and Marco dal Molin indirectly contributed to this manuscript through scientific discussions. The authors are also grateful for the data provided by the Federal Office of Meteorology MeteoSwiss. This work was supported by the Swiss National Science Foundation (grant number 200021_163322).

References

- Aitken, A. P. (1973). Assessing systematic errors in rainfall-runoff models. *Journal of Hydrology*, *20*(2), 121–136. [https://doi.org/10.1016/0022-1694\(73\)90035-8](https://doi.org/10.1016/0022-1694(73)90035-8)
- Ammann, L., Doppler, T., Stamm, C., Reichert, P., & Fenicia, F. (2020). Characterizing fast herbicide transport in a small agricultural catchment with conceptual models. *Journal of Hydrology*, *586*, 124812. <https://doi.org/10.1016/j.jhydrol.2020.124812>
- Ammann, L., Fenicia, F., & Reichert, P. (2019). A likelihood framework for deterministic hydrological models and the importance of non-stationary autocorrelation. *Hydrology and Earth System Sciences*, *23*(4), 2147–2172. <https://doi.org/10.5194/hess-23-2147-2019>
- Beck, M. B., & Young, P. C. (1976). Systematic identification of DO-BOD model structure. *Journal of the Environmental Engineering Division*, *103*(EE5), 902–927.
- Bertuzzo, E., Thomet, M., Botter, G., & Rinaldo, A. (2013). Catchment-scale herbicides transport: Theory and application. *Advances in Water Resources*, *52*, 232–242. <https://doi.org/10.1016/j.advwatres.2012.11.007>
- Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*, 279–298.
- Beven, K., & Westerberg, I. (2011). On red herrings and real herrings: Disinformation and information in hydrological inference. *Hydrological Processes*, *25*(10), 1676–1680. <https://doi.org/10.1002/hyp.7963>
- Buser, C. M. (2003). *Differentialgleichungen mit zufälligen zeitvariierenden Parametern* (Diploma thesis). ETH Zürich.
- Camenzuli, L. (2010). *Degradation of herbicides in soils, kinetics and limiting factors in field and laboratory* (Unpublished doctoral dissertation). Swiss Federal Institute of Technology (ETH).
- Clarke, R. T. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, *19*, 1–20.
- Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In J. M. Chambers, & T. J. Hastie (Eds.), *Chapter 8 in statistical models in S* (p. 608). Wadsworth & Brooks/Cole.
- Doppler, T., Camenzuli, L., Hirzel, G., Krauss, M., Lück, A., & Stamm, C. (2012). Spatial variability of herbicide mobilisation and transport at catchment scale: Insights from a field experiment. *Hydrology and Earth System Sciences*, *16*(7), 1947–1967. <https://doi.org/10.5194/hess-16-1947-2012>
- Evin, G., Kavetski, D., Thyer, M., & Kuczera, G. (2013). Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, *49*, 4518–4524. <https://doi.org/10.1002/wrcr.20284>
- Evin, G., Thyer, M., Kavetski, D., McInerney, D., & Kuczera, G. (2014). Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity. *Water Resources Research*, *50*(3), 2350–2375. <https://doi.org/10.1002/2013wr014185>
- FAL. (1997). *Bodenkarte Kanton Zürich (1:5000)*. Zürich.
- Fenicia, F., Kavetski, D., & Savenije, H. H. G. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, *47*(11), 1–13. <https://doi.org/10.1029/2010wr010174>
- Gardner, K. K., McGlynn, B. L., & Marshall, L. A. (2011). Quantifying watershed sensitivity to spatially variable N loading and the relative importance of watershed N retention mechanisms. *Water Resources Research*, *47*(8). <https://doi.org/10.1029/2010wr009738>
- Han, F., & Zheng, Y. (2016). Multiple-response Bayesian calibration of watershed water quality models with significant input and model structure errors. *Advances in Water Resources*, *88*, 109–123. <https://doi.org/10.1016/j.advwatres.2015.12.007>
- Han, F., & Zheng, Y. (2018). Joint analysis of input and parametric uncertainties in watershed water quality modeling: A formal Bayesian approach. *Advances in Water Resources*, *116*, 77–94. <https://doi.org/10.1016/j.advwatres.2018.04.006>
- Hantush, M. M., & Chaudhary, A. (2014). Bayesian framework for water quality model uncertainty estimation and risk management. *Journal of Hydrologic Engineering*, *19*(9), 4014015. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000900](https://doi.org/10.1061/(asce)he.1943-5584.0000900)
- Hong, B., Strawderman, R. L., Swaney, D. P., & Weinstein, D. A. (2005). 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). <https://doi.org/10.1029/2004wr003551>
- Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., & Chopin, N. (2015). On particle methods for parameter estimation in state-space models. *Statistical Science*, *30*(3), 328–351. <https://doi.org/10.1214/14-STS511>
- Kavetski, D., Franks, S. W., & Kuczera, G. (2003). Confronting input uncertainty in environmental modelling. In Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, & R. Turcotte (Eds.), *Water science and application* (pp. 49–68). American Geophysical Union. <https://doi.org/10.1029/ws006p0049>
- Kim, K., Park, M., Min, J. H., Ryu, I., Kang, M. R., & Park, L. J. (2014). Simulation of algal bloom dynamics in a river with the ensemble Kalman filter. *Journal of Hydrology*, *519*(PD), 2810–2821. <https://doi.org/10.1016/j.jhydrol.2014.09.073>
- Kuczera, G., Kavetski, D., Franks, S., & Thyer, M. (2006). Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *Journal of Hydrology*, *331*(1–2), 161–177.
- Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*(7). <https://doi.org/10.1029/2006WR005756>
- Liu, Y., Weerts, A. H., Clark, M., Franssen, H.-J. H., Kumar, S. V., Moradkhani, H., et al. (2012). Hydrology and Earth System Sciences Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, *16*, 3863–3887. <https://doi.org/10.5194/hess-16-3863-2012>
- Loos, S., Shin, C. M., Sumihar, J., Kim, K., Cho, J., & Weerts, A. H. (2020). Ensemble data assimilation methods for improving river water quality forecasting accuracy. *Water Research*, *171*, 115343. <https://doi.org/10.1016/j.watres.2019.115343>
- McInerney, D., Thyer, M., Kavetski, D., Lerat, J., & Kuczera, G. (2017). Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors. *Water Resources Research*, *53*(3), 2199–2239. <https://doi.org/10.1002/2016wr019168>

- McLaughlin, D. (2002). An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering. *Advances in Water Resources*, 25(8–12), 1275–1286. [https://doi.org/10.1016/S0309-1708\(02\)00055-6](https://doi.org/10.1016/S0309-1708(02)00055-6)
- MeteoSwiss. (2016). Retrieved from <https://www.meteoschweiz.admin.ch/home/service-und-publikationen/beratung-und-service/daten-portal-fuer-experten.html>
- Moradkhani, H., Hsu, K.-L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, 41(5). <https://doi.org/10.1029/2004WR003604>
- Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28(2), 135–147. <https://doi.org/10.1016/j.advwatres.2004.09.002>
- Nearing, G. S., Tian, Y., Gupta, H. V., Clark, M. P., Harrison, K. W., & Weijs, S. V. (2016). A philosophical basis for hydrological uncertainty. *Hydrological Sciences Journal*, 61(9), 1666–1678. <https://doi.org/10.1080/02626667.2016.1183009>
- Pathiraja, S., Marshall, L., Sharma, A., & Moradkhani, H. (2016a). Detecting non-stationary hydrologic model parameters in a paired catchment system using data assimilation. *Advances in Water Resources*, 94, 103–119. <https://doi.org/10.1016/j.advwatres.2016.04.021>
- Pathiraja, S., Marshall, L., Sharma, A., & Moradkhani, H. (2016b). Hydrologic modeling in dynamic catchments: A data assimilation approach. *Water Resources Research*, 52(5), 3350–3372. <https://doi.org/10.1002/2015WR017192>
- Raat, K. J., Vrugt, J. A., Bouten, W., & Tietema, A. (2004). 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. <https://doi.org/10.5194/hess-8-751-2004>
- Rajaram, H., & Georgakakos, K. P. (1989). Recursive parameter estimation of hydrologic models. *Water Resources Research*, 25(2), 281–294. <https://doi.org/10.1029/WR025i002p00281>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from <https://www.r-project.org/>
- Reichert, P. (2020). *R package for Bayesian inference with time-dependent, stochastic parameters*. Retrieved from <https://cran.r-project.org/package=timedeppar>
- Reichert, P., Ammann, L., & Fenicia, F. (2020). Potential and challenges of investigating intrinsic uncertainty of hydrological models with stochastic, time-dependent parameters. *Water Resources Research*, 57, e2020WR028400. <https://doi.org/10.1029/2020WR028400>
- Reichert, P., Langhans, S. D., Lienert, J., & Schuwirth, N. (2014). The conceptual foundation of environmental decision support. *Journal of Environmental Management*, 154, 316–332.
- Reichert, P., & Mieleitner, J. (2009). Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters. *Water Resources Research*, 45, W10402. <https://doi.org/10.1029/2009WR007814>
- Roberts, G. O., & Stramer, O. (2001). On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm. *Biometrika*, 88(3), 603–621.
- Rode, M., Arhonditsis, G., Balin, D., Kebede, T., Krysanova, V., van Griensven, A., & van der Zee, S. E. A. T. M. (2010). New challenges in integrated water quality modelling. *Hydrological Processes*, 24(24), 3447–3461. <https://doi.org/10.1002/hyp.7766>
- Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46, W10531. <https://doi.org/10.1029/2009WR008933>
- Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. *Water Resources Research*, 54(11). <https://doi.org/10.1029/2018WR022643>
- Smith, P. J., Beven, K. J., & Tawn, J. A. (2008). Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. *Water Resources Research*, 44(1). <https://doi.org/10.1029/2006WR005205>
- Talamba, D. B., Parent, E., & Musy, A. (2010). Bayesian multiresponse calibration of TOPMODEL: Application to the Haute-Mentue catchment, Switzerland. *Water Resources Research*, 46(8). <https://doi.org/10.1029/2007wr006449>
- Tomassini, L., Reichert, P., Künsch, H. R., Buser, C., Knutti, R., & Borsuk, M. E. (2009). A smoothing algorithm for estimating stochastic, continuous-time model parameters and its application to a simple climate model. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, 58, 679–704.
- Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical Review*, 36(5), 823–841. <https://doi.org/10.1103/physrev.36.823>
- Vrugt, J. A., Diks, C. G., Gupta, H. V., Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, 41(1). <https://doi.org/10.1029/2004WR003059>
- Wauchope, R. D., Yeh, S., Linders, J. B., Kloskowski, R., Tanaka, K., Rubin, B., et al. (2002). Pesticide soil sorption parameters: Theory, measurement, uses, limitations and reliability. *Pest Management Science*, 58(5), 419–445. <https://doi.org/10.1002/ps.489>
- Wellen, C., Arhonditsis, G. B., Long, T., & Boyd, D. (2014). 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. <https://doi.org/10.1016/j.jhydrol.2014.10.007>
- Yang, B., Stroud, J. R., & Huerta, G. (2018). Sequential Monte Carlo smoothing with parameter estimation. *Bayesian Analysis*, 13(4), 1133–1157. <https://doi.org/10.1214/17-BA1088>
- Zheng, Y., & Han, F. (2016). Markov Chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management. *Stochastic Environmental Research and Risk Assessment*, 30(1), 293–308. <https://doi.org/10.1007/s00477-015-1091-8>
- Zheng, Y., & Keller, A. A. (2007). 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). <https://doi.org/10.1029/2006wr005345>