Probabilistic indicators for soil and groundwater contamination risk assessment

Daniele la Cecilia[a,1], Giovanni M. Porta[b,*], Fiona H.M. Tang[a], Monica Riva[b,c], Federico Maggi[a]

[a]*Laboratory for Environmental Engineering, School of Civil Engineering, The University of Sydney, Bld. J05, 2006 Sydney, NSW, Australia.*
[b]*Dipartimento di Ingegneria Civile e Ambientale Politecnico di Milano, Piazza L. Da Vinci, 32, Milano, 20133, Italy*
[c]*Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, 85721, USA*

Abstract

Deterministic assessments of whether, when, and where environmental safety thresholds are exceeded by pollutants are often unreliable due to uncertainty stemming from incomplete knowledge of the properties of environmental systems and limited sampling. We present a global sensitivity analysis to rank the contribution of uncertain parameters to the probability, P, of a target quantity to exceed user-defined environmental safety thresholds. To this end, we propose a new index ($AMAP$) which quantifies the impact of a parameter on P and can be readily employed in probabilistic risk assessment. We apply $AMAP$, along with existing moment-based sensitivity indices, to quantify the sensitivity of soil and aquifer contamination following herbicide glyphosate (GLP) dispersal to soil hydraulic parameters. Target quantities are GLP and its toxic metabolite aminomethylphosphonic acid (AMPA) concentrations in the top soil as well as their leaching below the root zone. The global sensitivity analysis encompasses six scenarios of managed water amendments and rainfall events. The biodegradation of GLP and AMPA varies slightly across scenarios, while leaching below the root zone is greatly affected by the assumed hydrologic boundary conditions. $AMAP$ shows that, among the tested uncertain parameters, absolute permeability, air-entry suction, and porosity have the greatest impact on GLP and AMPA probability to pollute the aquifer by exceeding the aqueous concentration thresholds. Our results show that $AMAP$ is effective to thoroughly explore time histories arising from model-based predictions of environmental pollution hazards.

*Corresponding author
Email address: giovanni.porta@polimi.it (Giovanni M. Porta)
[1]Current address: Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

The proposed methodology may support informed decision making in risk assessments and help assessing ecological indicators through threshold-based analyses.

Keywords: global sensitivity analysis; uncertainty quantification; modeling; pollution; soil; groundwater; glyphosate; AMPA; environmental risk assessment

---

1. Introduction

Surface waters and aquifers are the recipients of contaminants resulting from anthropogenic activities such as agriculture, industry and waste treatment. According to the Lancet Commission on Pollution and Health, more than 140,000 synthetic molecules have been developed since 1950; of those, the 5,000 most produced molecules can be found in the environment worldwide (Landrigan *et al.*, 2018). These molecules and their metabolites can be persistent in the environment and, therefore, their detrimental effects can dramatically extend over space and time, thus harming not only humans and other living organisms but also their descendants (Kubsad *et al.*, 2019). In order to minimize and control harmful impacts, the use of these molecules has to be properly planned, managed, and regulated, thus requiring a good understanding of their dynamics in the environment.

Mathematical models are often used as decision support tools to evaluate contaminant degradation and transport (e.g., EPA, 2008; Porta *et al.*, 2018; Manheim *et al.*, 2019). Predictive models are also employed to carry out assessments on future scenarios such as climate change, land-use change, and global environmental change (e.g., Hiscock *et al.*, 2007; Armitage *et al.*, 2011; Brack *et al.*, 2015). Processes and environmental factors controlling contaminant dynamics (e.g., soil-water dynamics, contaminant sorption to soil minerals and organic matter, biochemical degradation, microbial-nutrient interactions, and soil-plant interactions) are described and coupled in models through mathematical equations with parameters often sourced from literature or estimated against laboratory or field experiments (Jackson *et al.*, 2000; Barrios *et al.*, 2019). Since environmental systems are complex and open to energy and mass flows, it is difficult to constrain and

model all controlling processes (Oreskes *et al.*, 1994). Deterministic models neglect the inherent uncertainty associated with model structure and parameters (Uusitalo *et al.*, 2015).

To overcome the shortcoming of deterministic models, uncertainty quantification and sensitivity analysis methods are used within a stochastic framework to provide decision makers with estimates of the potential outcomes of tested scenarios (e.g., Bates *et al.*, 2003; Walker *et al.*, 2003; Uusitalo *et al.*, 2015). There exist many different approaches for sensitivity analysis (e.g., Razavi & Gupta, 2015; Pianosi *et al.*, 2016; Ceriotti *et al.*, 2018). Among these, global sensitivity analysis (GSA) is one of the most comprehensive approaches because it allows (a) quantifying sensitivity across the entire parameter space and (b) accounting for the joint effects resulting from the uncertainty in diverse parameters even in nonlinear models. Output sensitivity to model parameters is commonly quantified using variance-based techniques such as the Sobol's indices (Sobol', 1993); more recently, the *AMA* family of indices was introduced (Dell'Oca *et al.*, 2017) to quantify sensitivity in terms of any statistical moment of the probability density function (pdf) of the model outputs. Available sensitivity indices do not provide a straightforward assessment of how a parameter influences the probability for a model output (e.g., the concentration of a contaminant) to exceed a user-defined threshold value. In environmental risk assessment and management, the policies and protection strategies often rely on regulatory guidelines that specify a safety limit for a certain contaminant. For example, the European Commission has set two severe safety limits to protect water resources quality from pesticide contamination: (a) the concentration of single pesticides and their relevant metabolites must not exceed 0.1 μg l^{-1}; and (b) the sum of pesticides and their metabolites concentration must not exceed 0.5 μg l^{-1} (2006/118/EC, 2006).

In this study we introduce a new global sensitivity index (*AMAP*) to rank parameters based on their impact on the probability to exceed a defined safety limit for single contaminants and mixtures. *AMAP* complements the available *AMA* moment-based indices by targeting sensitivity with respect to the exceedance probability rather than the statistical moments of the output

pdf. We demonstrate the use of *AMAP* to the case study of dispersal of the herbicide glyphosate (GLP) throughout a soil profile in an irrigated winter wheat field, using the results of la Cecilia *et al.* (2018) as a reference. The case study selection is motivated by the observation that GLP is currently the most widely used herbicide worldwide (Maggi *et al.*, 2019). Modeling GLP biodegradation pathways requires a complex network of bioreactive processes coupled to water flow and solute transport. Hence, quantitative indicators that can identify relevant parameters and processes are important to reduce the uncertainty involved in risk assessment and help constraining the whole decision-making process. Societal implications of the uncertainties underlying environmental risk assessment of GLP have been widely discussed in the recent literature (e.g., Van Straalen & Legler, 2018). The relevance of hydraulic parameters on the fate of GLP and AMPA in soil has been documented in field experiments (e.g. Soracco *et al.*, 2018; Lupi *et al.*, 2019), as well as in numerical studies (Heuvelink *et al.*, 2010). In this work we analyze the impact of uncertainty in soil hydraulic parameters on risk assessment of GLP and AMPA accumulation and leaching. We apply a number of different boundary conditions affecting transport and biodegradation processes to gain a wider understanding of how (*i*) *AMAP* informs on the effects of parameter uncertainty within pollution risk assessment and (*ii*) GLP and AMPA biodegradation predicted by the model is affected by the assumed boundary water fluxes. We designed such specific scenarios to represent managed (irrigated) and unmanaged (not irrigated) cropping. Analyses are also accompanied by specific robustness tests of the proposed *AMAP* index to show limits and advantages of its generalized application beyond the test case presented here. While we focus on soil and water contamination risk assessment, we emphasize that the proposed sensitivity index is readily applicable in other contexts, such as to assess the response of ecological and environmental systems, and particularly within threshold-based analysis of ecological indicators reported in recent literature (e.g., Libralato *et al.*, 2019; Fu *et al.*, 2019).

2. Materials and Methods

We introduce here the definition of the sensitivity indices (or metrics) as well as the approach we employ for their application within contamination risk assessments. In Section 2.1 we introduce the proposed sensitivity metrics and provide an operational framework for their application in a generic environmental problem. Next, we illustrate the application of our method to soil contamination as a result of glyphosate (GLP) dispersal. We start by presenting an overview of the kinetic model used to describe GLP contamination in a winter wheat field (Section 2.2). Target outputs for risk analysis and their corresponding safety thresholds are identified in Section 2.3. In Section 2.4, six scenarios with different ecohydrological boundary conditions are designed to be used in GSA. Finally, we select the uncertain parameters and describe the sampling methodology in Section 2.5.

2.1. Sensitivity indices and application to environmental problems

The *AMA* sensitivity indices (Dell'Oca *et al.*, 2017) quantify the impact of each uncertain parameter on the statistical moments of the pdf of the target model outputs. Let $g(\mathbf{p})$ be an output of interest, and $\mathbf{p} = (p_1, \ldots, p_i, \ldots, p_N)$ a vector gathering N uncertain parameters. The $AMAM_i$ indices quantify the impact of variability in parameter p_i on the statistical moment M_i of $g(\mathbf{p})$ (e.g., *AMAE* for the expected value E and *AMAV* for the variance V). $AMAM_i$ is defined as

$$AMAM_i = \begin{cases} \dfrac{\int_{\Gamma_i} |\mathrm{M}[g(\mathbf{p}|p_i)] - \mathrm{M}[g(\mathbf{p})]|\,\rho(p_i)dp_i}{|\mathrm{M}[g(\mathbf{p})]|} & \text{if } \mathrm{M}[g(\mathbf{p})] \neq 0 \\ \int_{\Gamma_i} |\mathrm{M}[g(\mathbf{p}|p_i)]|\,\rho(p_i)dp_i & \text{if } \mathrm{M}[g(\mathbf{p})] = 0 \end{cases} \tag{1}$$

where $\rho(p_i)$ is the pdf of p_i defined in the parameter space Γ_i. Along with Eq. (1), and to cast our work within a risk assessment framework, we introduce the new index *AMAP*, which allows quantifying the expected variation of the probability of exceedance of a threshold value *thr* as

$$AMAP_i = \int_{\Gamma_i} \left| P_{thr} - P\left[g(\mathbf{p}|p_i) > thr\right] \right| \rho(p_i)dp_i, \tag{2}$$

where $P_{thr} = P[g(\mathbf{p}) > thr]$ is the unconditional probability that the quantity $g(\mathbf{p})$ exceeds the threshold thr and $P[g(\mathbf{p})|p_i]$ indicates the same probability conditional to parameter p_i. Note that $AMAP$ provides the probability-weighted average distance between conditional and unconditional exceedance probability within Γ_i and is limited between 0 and 1. The output $g(\mathbf{p})$ and its related threshold thr can be any quantity of interest, including a contaminant concentration, and can be used in a generalized way for the purpose of, but not limited to, risk analysis as shown later in this work.

A flowchart is presented in Figure 1 to illustrate the workflow for the $AMAP$ application within model-based environmental pollution assessment. First, prior information needs to be collected to define (*i*) a model structure and reference values of model parameters, and (*ii*) target outputs of interest and the related safety thresholds, where the latter can be user-defined or taken from guidelines. Additionally, various scenarios can be selected to explore the system response in diverse conditions (e.g., diverse hydrologic or climatic regimes, socio-economic and/or legislative constraints). As environmental models typically embed a large number of parameters, a subset of these is selected to conduct sensitivity analysis. The latter may be then used to (a) rank parameter importance (b) design and prioritize experimental campaigns aimed at constraining the uncertainty of the selected output. A pdf $\rho(p_i)$ must be defined for each uncertain input to compute $AMAP_i$ through Eq. (2). This can be determined from available prior information through empirically defined frequency distributions or according to general pdf models (e.g., Gaussian or uniform distributions), thus defining a probability space for the selected parameters set. Different choices for the input pdfs $\rho(p_i)$ can be performed and the results of the analyses may depend on the chosen input distribution. Therefore, the chosen $\rho(p_i)$ should reflect available information as closely as possible. Stochastic sampling of the parameters within this set is then performed N times, rendering N values of the output $g(\mathbf{p})$. These latter are employed to evaluate the conditional and unconditional probability to exceed a given threshold needed in (2) to compute $AMAP$. The work-

flow is replicated for each of the selected scenarios, thus providing a scenario-dependent sensitivity ranking that can guide in implementing strategies to reduce uncertainty.

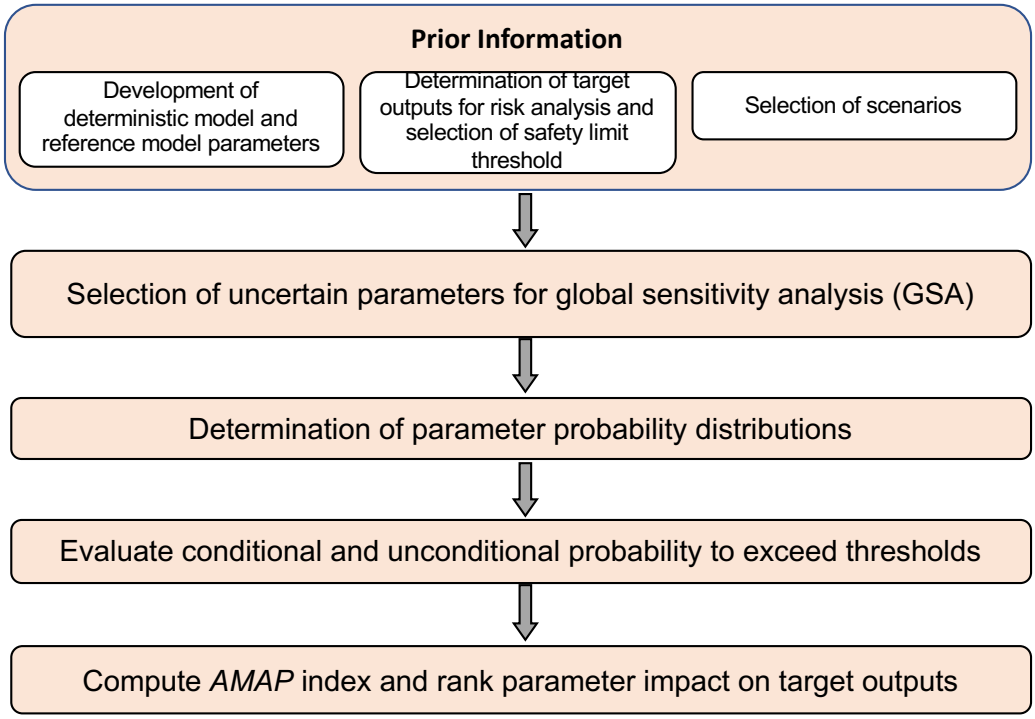

Figure 1: Flowchart of the steps used in this study for the application of *AMAP* index.

2.2. *Reference site and modeling description*

We apply the framework introduced in Sec. 2.1 to the case of GLP biodegradation considered in la Cecilia *et al.* (2018). The reference winter wheat field is located in the Modena Municipality, Italy (44°40′57″N; 10°57′48″E). The soil is a typical alluvial deposit of the Po Valley region characterized by a silt loam and loam layers (SGSS, 2016). Two regions of interest were identified along the soil profile: the root zone (RZ) with thickness $h_{RZ} = 1$ m and the soil below RZ (BRZ) with thickness $h_{BRZ} = 4$ m.

Rainfall data in the period 2006-2016 were collected (Arpae-Simc, 2016) and post-processed to compute the water infiltration after assuming a 20% rainfall interception by the crop. The actual

crop evapotranspiration was calculated from data in Arpae-Simc (2016) with the time-varying crop coefficient K_C in Allen *et al.* (1998). Irrigation was estimated to match groundwater table depth observations in Chiari *et al.* (2016) as described in la Cecilia *et al.* (2018). The 11 year precipitation, actual evapotranspiration, and irrigation time series were repeated periodically to build 50 years of daily boundary conditions. GLP was applied annually at rate 1.2×10^{-3} mol m^{-2} (i.e., 2 kg ha^{-1} year^{-1}) in a single application event. An interception fraction of 0.2 and a drift fraction of 0.2 for post-emergence pesticide application on winter wheat (Trevisan *et al.*, 2009) were accounted for as losses of the applied GLP rate, thus resulting in a net application rate $A = 7.2 \times 10^{-4}$ mol m^{-2} (i.e., 1.2 kg ha^{-1} year^{-1}).

The GLP reaction network includes GLP and AMPA biodegradation and neglects chemical degradation because it has been shown to only occur in soils rich in Mn oxides (e.g., birnessite mineral, la Cecilia & Maggi, 2018). Inhibition of the reaction by heavy metals (e.g. Cu^{2+})(Barrett & McBride, 2005; la Cecilia & Maggi, 2018; la Cecilia *et al.*, 2018; Li *et al.*, 2015) is also not considered. GLP and AMPA degradation is tightly coupled with the nitrogen (N) cycle and a pool of soil organic matter, which releases ammonium (NH_4^+), ortophosphate (PO_4^{-3}), and monomeric organic carbon (represented by CH_2O) for microbial metabolic purposes (Maggi *et al.*, 2008). Six microbial functional groups describe the soil microbial ecology and include: GLP and AMPA hydrolizing and oxidizing bacteria (B_{HyO}), and aerobic and anaerobic bacteria (B_{AER} and B_{ANAER}) that consume organic carbon (la Cecilia & Maggi, 2018; la Cecilia *et al.*, 2018), and NH_4^+ and NO_2^- oxidizing bacteria (B_{AOB} and B_{NOB}) that mediate a two-steps nitrification, and denitrifying bacteria (B_{DEN}) that perform a three-step NO_3^- denitrification reduction to N_2 (Maggi *et al.*, 2008). Dynamic stability of soil microbial ecology resorts to group-specific biomass background recovery rates after Porta *et al.* (2018) showed that some functional groups can be outcompeted for some parameter combinations. Inhibition on various reactions include O_2 effects to anaerobic reactions and pH below 6 and above 8 for microbial activity (Boon & Laudelout, 1962). Protection of

aqueous species, including GLP and AMPA, to the mineral phase is modeled as a linear equilibrium process. Biodegradation neglects the protected phase because it is assumed to not be accessible to exoenzymes as suggested in Riley *et al.* (2014). Benchmarking of the reaction network has been performed in Maggi *et al.* (2020) against field measurements of GLP and AMPA concentrations reported in the literature. All the details of the biodegradation reaction network used in this work are available in la Cecilia *et al.* (2018).

Deterministic simulations of GLP biodegradation were conducted using the BRTSim-v3.1a general-purpose solver for reaction-advection-diffusion processes in variably saturated soils (Maggi, 2019). BRTSim numerically resolves the mass, momentum and energy conservation laws, biochemical kinetics, and equilibrium reactions using hybrid explicit-implicit finite volumes solvers, which are described in detail in the User Guide and Technical Manual (Maggi, 2018). Under the assumption that the gas phase undergoes negligible pressure gradients, advection was neglected here for gaseous species while diffusion of gaseous species was explicitly included. A steady temperature profile was assigned linearly changing from 20 °C at the top soil to 14 °C at 5 m depth.

In the following we describe the key equations used to describe flow and transport processes, i.e. where the investigated uncertain parameters are directly involved (see Section 2.5). We indicate variables dimensions using the notation [M,L,T,Θ] for mass, length, time and temperature, respectively. The mass and momentum conservation laws for water in a one dimensional variably saturated soil with constant porosity ϕ can be written as (Richards, 1931)

$$\phi \frac{\partial S_l}{\partial t} = -\frac{\partial q}{\partial z} - ET(z,t) - \Delta B(z,t), \qquad \text{with} \qquad q = -\frac{\rho_l g k}{\mu} k_r(S_l)\left(\frac{\partial \psi(S_l)}{\partial z} - 1\right), \tag{3}$$

where t [T] is time, S_l [−] is water saturation, z [L] is the vertical coordinate, k [L^2] and k_r are the absolute and relative permeabilities, g [LT^{-2}] is the gravitational acceleration, q [LT^{-1}] is the Darcy's velocity, ψ [L] is the water potential, and ρ_l [ML^{-3}] and μ [$ML^{-1}T^{-1}$] are the water density

and viscosity, respectively. Terms $ET(z,t)$ and $\Delta B(z,t)$ [T^{-1}] describe contributions to the soil water saturation due to actual evapotranspiration and immobilization in the microbial biomass, respectively. Precipitations $P(t)$ and irrigation $I(t)$ [LT^{-1}] per unit planar surface are assigned at the soil surface. Eq. (3) was solved assuming the following empirical formulations for k_r and ψ proposed by Brooks & Corey (1964)

$$k_r = S_{le}^{2b+3}, \tag{4a}$$

$$\psi = \psi_s S_{le}^{-b}, \tag{4b}$$

where b [-] is the pore volume distribution index, ψ_s [L] is the air-entry suction at water saturation, and S_{le} is the effective water saturation defined as

$$S_{le} = \frac{S_l - S_{lr}}{1 - S_{lr} - S_{gr}}, \tag{5}$$

S_{lr} and S_{gr} indicate the water and gas residual saturations, respectively.

Mass balance of chemical species is expressed by

$$\phi \frac{\partial X^i}{\partial t} = \frac{\partial J_{X^i}}{\partial z} + \hat{r}^i \qquad J_{X^i} = -qX^i + \phi D_m \frac{\partial X^i}{\partial z} \tag{6}$$

where X^i [ML^{-3}] is aqueous concentration, D_m [L^2T^{-1}] is molecular diffusion, J_{X^i} [$ML^{-2}T^{-1}$] is the advective diffusive solute flux, $\hat{r}^i$ [$ML^{-3}T^{-1}$] lumps the contribution from all chemical and biochemical reactions.

Chemical and biochemical kinetic reactions are assumed to occur only in the aqueous phase. The modelling approach employed to represent the feedbacks between water stress, temperature and bioreactive processes is detailed in Appendix A.

The microbial biomass was initialized with a simulation of 100 years to allow the water flow

and microbial processes related to GLP and the N cycle reaching a stationary state. The initialized system was then used repeatedly for the purpose of GSA to simulate a period of $T_s = 50$ years, setting GLP and AMPA initial concentrations to zero in each realization. As a consequence, this initialization presumed the ecological capability to degrade GLP, i.e., the microbial functional groups have already adapted to degrade GLP when they receive the first application. Recent experiments show that adaptation may imply a time lag in the order of months before soil microorganisms can effectively degrade GLP (Tang *et al.*, 2019a), or in the order of hours due to catabolite repression mechanisms triggered by substrate preference and memory of previous growth conditions (la Cecilia *et al.*, 2019). As these observed adaptation times are significantly smaller than the considered time window of 50 years we neglect these effects in our analysis. This assumption may generally be reasonable because agricultural soils have typically been exposed to a wide suite of xenobiotics before the application of GLP.

2.3. Target quantities and thresholds for pollution assessment

The target quantities selected for risk analysis are (see also Figure 2):

- the depth-averaged aqueous concentration of GLP and the mixture MXT = GLP + AMPA in BRZ (1 to 5 m depth), labeled as $C_{\mathrm{GLP}}|_{BRZ}$ and $C_{\mathrm{MXT}}|_{BRZ}$. The threshold values used are 0.1 μg l^{-1} for GLP and 0.5 μg l^{-1} for MXT, as prescribed by the Directive 2006/118/EC (2006) for the tolerable contamination of groundwater.

- GLP and AMPA mass in the top 30 cm of soil. These variables are indicated as $M_{\mathrm{GLP}}|_{TOP}$ and $M_{\mathrm{AMPA}}|_{TOP}$. We used the ecotoxicological concentration (i.e., LC_{50}) of GLP and AMPA mass fractions to earthworms as thresholds, which are set to $M_{thr,\mathrm{GLP}} = 5,600$ mg kg-soil^{-1} and $M_{thr,\mathrm{AMPA}} = 1,000$ mg kg-soil^{-1} for GLP and AMPA, respectively (Lewis *et al.*, 2006).

- the yearly cumulative leaching rate of GLP and MXT between RZ and BRZ soil, F_{GLP} and F_{MXT}, corresponding to advective-diffusive fluxes between the two soil regions, defined pos-

itive for downward fluxes (i.e., from RZ to BRZ). As threshold fluxes we set 0.02 mg m^{-2} y^{-1} and 0.1 mg m^{-2} y^{-1} for GLP and MXT, respectively. These values correspond to the 0.01% and 0.05% of the gross GLP application rate A_{GLP} = 2 kg ha^{-1} y^{-1}, following the rationale employed in 2006/118/EC (2006). We explore here the impact of using a threshold mass rate to reflect the risk of aquifer contamination and pollution. This definition is motivated by the possibility of direct comparison with application and biodegradation rates, and detachment from water saturation-dependent assessment indicators.

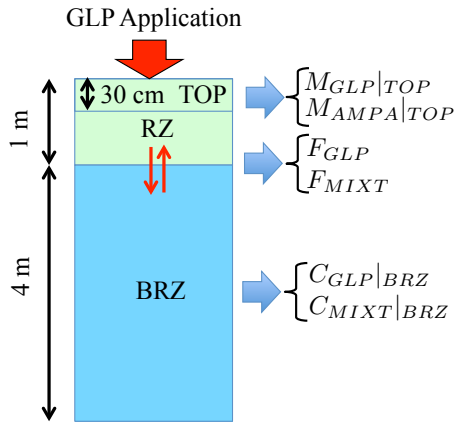

Figure 2: Graphical representation of the outputs considered for pollution and contamination risk analysis

In the following, we employ the term 'contamination' for a non-negligible concentration of a molecule in an environmental compartment where it should not exist, and 'pollution' when a given quantity exceeds a prescribed safety threshold following Rodriguez Eugenio *et al.* (2018). Therefore, contamination refers here to any positive level of concentration or flux detected below the root zone.

$AMAP_i$ indices were calculated considering all the selected outputs within the simulation time window T_s = 50 y. Results obtained at year 15 are emphasized in the discussion because 15 y is the maximum approval period for pesticides in the EU. Note that while concentrations and top soil masses of contaminants are considered as continuous functions of time, fluxes are annual averages

used to detect persistent contaminant leaching.

Finally, we calculated the annual biodegradation efficiency in RZ of GLP

$$E_{\mathrm{GLP}}(n) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{A\Delta t}\int_{h_{RZ}}\int_{(n-1)\Delta t}^{n\Delta t}\left[R_{\mathrm{GLP}}^{\mathrm{Oxi}}(z,t) + R_{\mathrm{GLP}}^{\mathrm{Hyd}}(z,t)\right]dzdt \tag{7}$$

as the ratio of biodegraded GLP through oxidation $R_{\mathrm{GLP}}^{\mathrm{Oxi}}$ and hydrolysis $R_{\mathrm{GLP}}^{\mathrm{Hyd}}$ calculated as in Eq. A.1 relative to the GLP application rate A in year n. Similarly, the biodegradation efficiency of AMPA is defined as

$$E_{\mathrm{AMPA}}(n) = \frac{1}{n}\sum_{j=1}^{n}\int_{h_{RZ}}\int_{(n-1)\Delta t}^{n\Delta t}\frac{R_{\mathrm{AMPA}}(z,t)}{R_{\mathrm{GLP}}^{\mathrm{Oxi}}(z,t)}dzdt \tag{8}$$

where R_{AMPA} is biodegraded AMPA. In both (7)-(8) we set $\Delta t = 1$ year. Note that Eq. (8) measures the ratio between biodegraded and produced AMPA, where the latter is a metabolite of GLP biodegradation through oxidation, see la Cecilia & Maggi (2018) for details. The biodegradation efficiencies in Eqs. (7)-(8) are not used for *AMAM* or *AMAP* analyses because there is no legislation that recommends threshold biodegradation efficiencies in field applications. However, Eq. (7)-(8) were used to interpret how parameter uncertainty affected GLP and AMPA biodegradation across different scenarios.

2.4. Selection of scenarios

We consider six hydrometeorological scenarios (Table 1). The first scenario corresponds to the reference (REF) case study described in la Cecilia *et al.* (2018), where the precipitation $P = P_{\mathrm{REF}}$ and actual evapotranspiration $ET = ET_{\mathrm{C,\,REF}}$ are used to estimate irrigation rates $I = I_{\mathrm{REF}}$ necessary to match locally-measured water table depths. This case study is representative of managed agricultural crops. In the second and third scenarios, P is decreased and increased by 20% of P_{REF} to simulate drier (DRY) and wetter (WET) conditions, respectively, while irrigation is maintained as in the REF scenario. In the fourth scenario, steady state (SS) boundary conditions are set equal

to time-averaged fluxes $P = \overline{P_{\mathrm{REF}}}$, $ET = \overline{ET_{\mathrm{C,\,REF}}}$, $I = \overline{I_{\mathrm{REF}}}$ calculated over the whole simulation time and applied as constant boundary flows. In this case, the GLP application rate is also time-averaged and applied throughout the simulated time period as $A = \overline{A_{\mathrm{REF}}}$. DRY and SS unmanaged scenarios without irrigation are also considered.

The managed and unmanaged SS scenarios are tested to investigate the extent to which accounting of time-resolved as compared to constant hydrological fluxes can influence the prediction of GLP and AMPA biodegradation.

Boundary conditions	Managed (REF)	(DRY)	(WET)	(SS)	Unmanaged (DRY)	(SS)
P	P_{REF}	$P_{\mathrm{REF}} \times 0.8$	$P_{\mathrm{REF}} \times 1.2$	$\overline{P_{\mathrm{REF}}}$	P_{REF}	$\overline{P_{\mathrm{REF}}}$
ET	$ET_{\mathrm{C,\,REF}}$	$ET_{\mathrm{C,\,REF}}$	$ET_{\mathrm{C,\,REF}}$	$\overline{ET_{\mathrm{C,\,REF}}}$	$ET_{\mathrm{C,\,REF}}$	$\overline{ET_{\mathrm{C,\,REF}}}$
I	I_{REF}	I_{REF}	I_{REF}	$\overline{I_{\mathrm{REF}}}$	0	0
A	A_{REF}	A_{REF}	A_{REF}	$\overline{A_{\mathrm{REF}}}$	A_{REF}	$\overline{A_{\mathrm{REF}}}$

Table 1: Boundary conditions applied in the tested scenarios. REF, DRY, WET and SS indicate the reference scenario, dry and wet scenarios, and steady state scenarios, respectively.

2.5. Selection of uncertain parameters

To illustrate the use of the sensitivity metrics defined in Section 2.1, we analyze here the impact of uncertainty related to soil hydraulic properties on the target quantities defined in Section 2.3. We consider a total of six uncertain parameters indicated in Eqs (3)-(4), i.e. k, b, ψ_s, ϕ, S_{lr} and S_{gr}. In principle, permeability k can be expressed as a function of the other parameters by means of empirical or semi-empirical correlations (e.g. Brutsaert, 2000). However, we have included k within the set of uncertain parameters because of the reported non-exact dependence between ϕ and k (e.g., Maggi & Porporato, 2007). These parameters are assumed to be mutually independent and uniformly distributed within given ranges $\delta(\mathbf{p})$. Our choice of using a uniform pdf to characterize the uncertain model inputs rests on the idea of assigning equal weight to each value of the distribution (i.e., an equal prior probability). The range $\delta(p_i)$ associated with k, b, ψ_s and ϕ (Table 2) is retrieved from naturally occurring soils, i.e., from hydrothermal properties database by Dai

et al. (2013) and the SoilGrids database in Hengl *et al.* (2017). Measured residual liquid saturation S_{lr} range from 0.046 to 0.31 (Ghanbarian-Alavijeh *et al.*, 2010), while S_{gr} observations vary between 0.092 and 0.22 (Smith & Browning, 1943; Peck, 1969). Based on these observations, S_{lr} and S_{gr} are here considered as uniformly distributed parameters with values comprised between 0.05 and 0.2 (Table 2). Our analysis neglects vertical heterogeneity of soil properties. We solve one-dimensional flow and transport along a vertical soil column. This choice is consistent with models used for assessing pesticide leaching in regulatory frameworks (Jene, 1998; Carsel *et al.*, 1985; Van den Berg *et al.*, 2012; Carsel *et al.*, 1985; 2006/118/EC, 2006). The assumed uncertainty in the soil hydraulic parameters can be used to assess the impact of their spatial variability, such as rendered by geo-referenced databases discussed in Heuvelink *et al.* (2010).

Sampling of the parameters space was conducted using a quasi Monte Carlo (QMC) technique (Sobol', 1998). In total, 5,000 parameter realizations were generated and applied to each scenario. Upon performing a forward modeling run for each of the selected sampling points, we obtain a QMC ensemble of our target outputs of interest that is next used to conduct the sensitivity analysis. We verified the convergence of the QMC samples in terms of the outputs sample pdfs, see Appendix B. The CPU time for each QMC flow and transport simulation is 600 s (Intel Xeon Platinum 8160 @ 2.10 GHz).

k	b	ψ_s	ϕ	S_{lr}	S_{gr}
$\times 10^{-13}$ [m^2]	[-]	[m] water	[-]	[-]	[-]
(0.50, 10)	(3, 7)	(-0.6, -0.1)	(0.4,0.5)	(0.05, 0.2)	(0.05, 0.2)

Table 2: Parameter value ranges.

3. Results and discussions

In the following we apply the global sensitivity indices *AMAM* and *AMAP* introduced in Section 2 to the reference scenario and we then analyze the impact of the considered hydrological and management regime on the system response. Finally we consider the impact of the selected scenar-

ios on the biodegradation efficiency, to assess the relevance of the hydrologic boundary conditions on the GLP biodegradation reaction network.

3.1. Global sensitivity indices in the reference (REF) scenario

Figure 3 reports $AMAE_i$, $AMAV_i$ and $AMAP_i$ computed for all uncertain parameters p_i and for all target quantities evaluated at the time corresponding to the maximum approval period for pesticides in the EU, i.e. at t = 15 y.

$AMAE_i$ (Figure 3a-d) suggests that the soil permeability k has the greatest influence on the sample average of all target quantities analyzed followed by (a) ψ_s and ϕ for the concentration targets and (b) by the gas residual saturation S_{gr} and ψ_s for the flux targets. Other parameters display moderate to minor effects. Similar results have been obtained for $AMAV_i$ for the two concentrations (Figure 3e-f), while the fluxes variances are also greatly influenced by the pore volume distribution index b (Figure 3g-h) .

$AMAP_i$ (Figure 3i to l) shows that the probability of C_{GLP} and (to a lesser extent) F_{MXT} to exceed their thresholds are impacted by the variability of the uncertain parameters at t = 15 y. Otherwise, $AMAP_i$ obtained for $C_{MXT}|_{BRZ}$ and F_{GLP} are negligible. This result is explained observing that the threshold MXT concentration/GLP flux is never or always exceeded in the investigated sample regardless of the parameters' values (see also Sections 3.4 and 3.3). The investigated parameters have then a negligible influence on the probability of exceeding the threshold, while they still influence the outputs mean and variance as shown by the corresponding $AMAE_i$ and $AMAV_i$ values.

Overall Figure 3 suggests that an accurate characterization of k, ψ_s, ϕ should be prioritized to predict agrochemicals' concentrations in the aquifer. In addition, estimating agrochemical fluxes from the root zone to the aquifer would benefit from an accurate knowledge of S_{gr} and b.

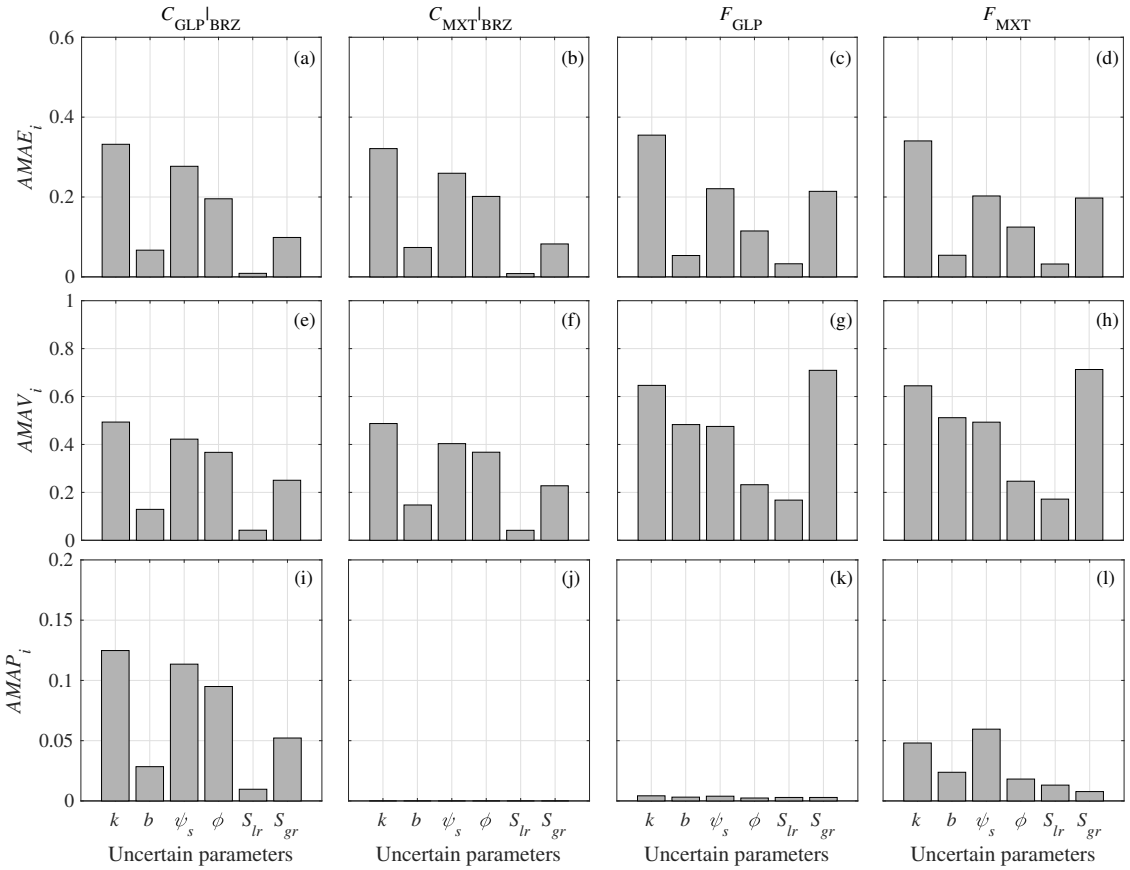

Figure 3: Global sensitivity indices (a)-(d) $AMAE_i$; (e)-(h) $AMAV_i$; and (i)-(l) $AMAP_i$ evaluated for GLP and MXT aqueous concentrations and fluxes from RZ to BRZ. Analyses are relative to REF scenario at time t = 15 years.

3.2. Range Impact Analysis - RIA

In this section, we quantify the variability of $AMAP_i$ with respect to the level of uncertainty assumed for each parameter p_i upon considering GLP concentration $C_{\mathrm{GLP|BRZ}}(t = 15\mathrm{y})$. We keep the average value for each parameter probability distribution constant and we increase/decrease the ranges of variability in $\delta(p_i)$ by a prescribed factor comprised between 0.7 and 1.1. This allows testing the robustness of parameter ranking upon maintaining uniformly distributed parameters and without violating physical constraints (i.e., positive permeability and porosity comprised between zero and one).

Figure 4 shows that the $AMAP_i$ indices smoothly vary. The ranking of parameters importance is also consistent for all the investigated ranges with k, ψ_s and ϕ chiefly influencing the system response, while the effect of b and S_{lr} and S_{gr} appears negligible. Increasing values of $AMAP_i$ are obtained for increasing parameters ranges, which reflects the increase of assumed uncertainty in the parameter values.

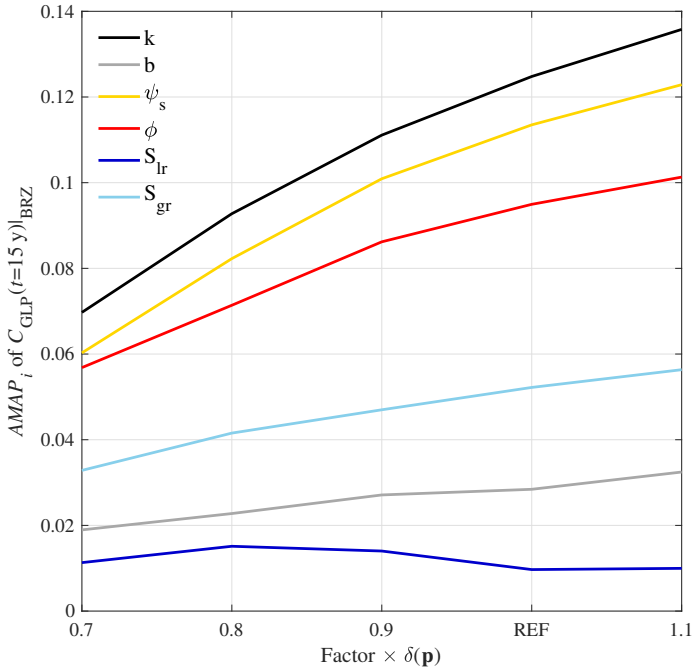

Figure 4: $AMAP$ - Range Impact Analysis for $C_{\mathrm{GLP|BRZ}}(t = 15\mathrm{y})$.

3.3. Analysis of contamination and pollution

Figure 5a-b displays the temporal evolution of the expected values (or sample-averages) $\mathrm{E}[M_{\mathrm{GLP}|_{TOP}}]$ and $\mathrm{E}[M_{\mathrm{AMPA}|_{TOP}}]$, respectively. Both quantities are significantly smaller (by three orders of magnitude) than their ecotoxicological threshold (i.e., LC_{50} for earthworms) and the probability to exceed the thresholds is negligible across the whole sample. Managed and unmanaged steady state scenarios result in the largest $\mathrm{E}[M_{\mathrm{GLP}|_{TOP}}]$ and $\mathrm{E}[M_{\mathrm{AMPA}|_{TOP}}]$ values, which are likely caused by re-

duced flushing as compared to scenarios where intense precipitations caused fast GLP and AMPA transport to BRZ.

The sample-averaged aqueous concentrations $\mathrm{E}[C_{\mathrm{GLP}}|_{BRZ}]$ and $\mathrm{E}[C_{\mathrm{MXT}}|_{BRZ}]$ (Figure 5c and d) increase over time in the aquifer. Wet scenarios (REF and WET) lead to faster increase in $\mathrm{E}[C_{\mathrm{GLP}}|_{BRZ}]$ and $\mathrm{E}[C_{\mathrm{MXT}}|_{BRZ}]$ as compared to dry scenarios (DRY). Our results also suggest that pollution may occur at substantially longer time scales in steady state conditions, i.e., SS scenarios do not cause any significant contamination within the investigated 50-year time period.

Figure 5e and f show the relative frequency (or sample probability) of the exceedance time $\hat{t}$, i.e., the time at which $C_{\mathrm{GLP}}|_{BRZ}$ and $C_{\mathrm{MXT}}|_{BRZ}$ exceed the corresponding threshold concentrations. All transient scenarios showed more than 95% probability for $C_{\mathrm{GLP}}|_{BRZ}$ and $C_{\mathrm{MXT}}|_{BRZ}$ to exceed the threshold concentrations within a time frame of 50 years. On the other hand, considering the EU maximum approval period for pesticides of 15 years, the probability to exceed the threshold concentrations is smaller than 20% for $C_{\mathrm{GLP}}|_{BRZ}$ and negligible for $C_{\mathrm{MXT}}|_{BRZ}$. This result is consistent with Figure 3j, showing negligible $AMAP_i$ values (for all parameters) for $C_{\mathrm{MXT}}|_{BRZ}$. We further note that the probability distributions of $\hat{t}$ display heavier right tails for DRY than for WET scenarios, i.e., GLP and MXT arrival times to BRZ are characterized by larger uncertainty in DRY than in WET scenarios. This result quantifies a delay in the occurrence of water pollution BRZ in DRY conditions.

The mean GLP and MXT leaching rates from RZ to BRZ (Figure 6a-b) vary significantly across all scenarios. As expected, WET scenarios lead to higher leaching rates than DRY ones. Soil BRZ can undergo pollution after 4 years since the first GLP application in both WET and DRY scenarios. Figure 6a and b also show that upward (negative) fluxes occur from BRZ to RZ in dry scenarios. These instances are driven by particularly dry periods and elevated ET_{C}, which result in high water suction in TOP and RZ from BRZ. Upward fluxes are consistent with previous observations of herbicide transport during capillary driven groundwater rise (Arjoon *et al.*, 1998). This result may

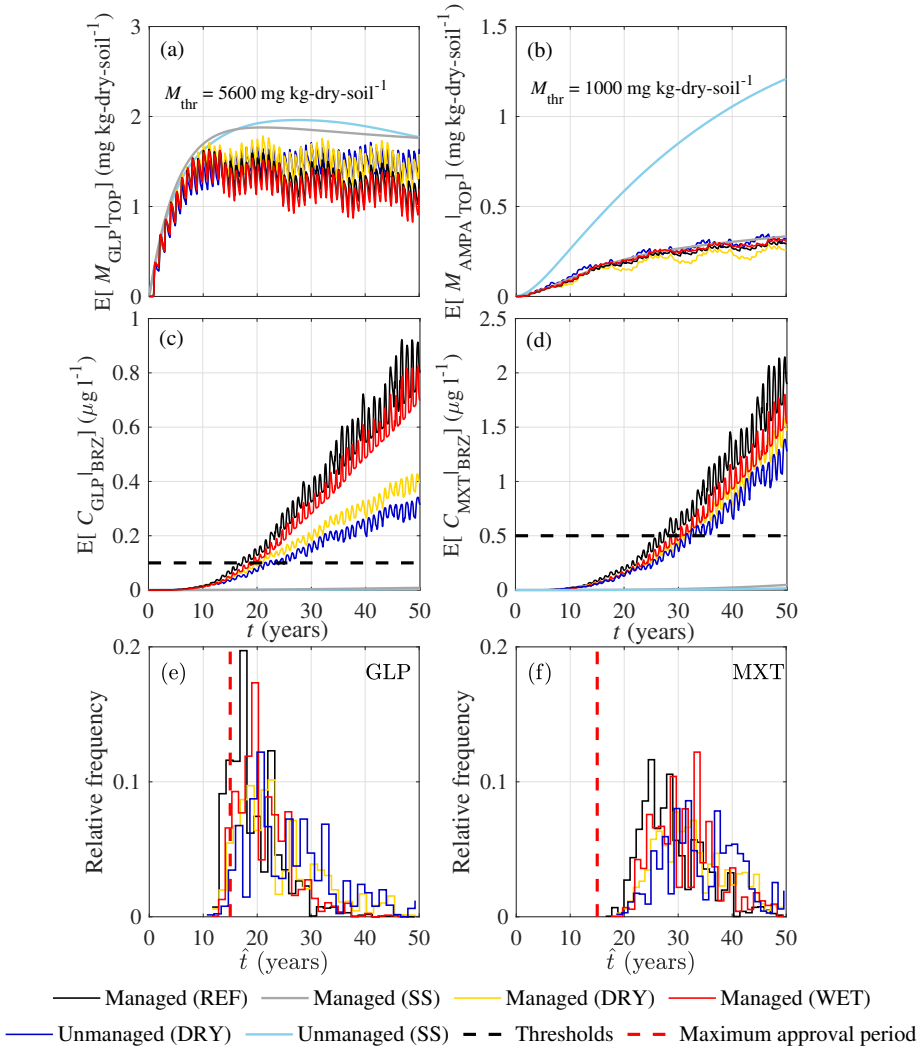

Figure 5: Average GLP and AMPA mass in RZ (a-b) and aqueous concentrations in BRZ (c-d); (e)-(f) represent the corresponding probability distribution of exceedance time $\hat{t}$, i.e., the time at which contaminant concentrations exceeded threshold concentrations.

raise awareness for pollutants accumulation at soil depths that may be reached during water table fluctuations, where there may be a lack of active biodegraders. The recontamination does not occur in the SS scenarios because the saturation profile is constant in time. This result confirms our interpretation that unsteady water inputs due to precipitations can cause rapid contaminant flushes (i.e., positive fluxes) to BRZ as compared to SS scenarios.

Figures 6c and d show that the probability to exceed F_{thr} is very high ($P_{thr} \approx 1$) within 15 years for both GLP and MXT in all scenarios except SS. Comparing Figures 6c and d with Figure 5e and f, one can conclude that the time scale associated to pollution observed in leaching rate (from RZ to BRZ) is significantly smaller than the time scale linked to resident agrochemical concentrations in BRZ. Therefore, measurements and modeling predictions of leaching rates would provide a more conservative indicator than concentration data within a contaminant risk assessment framework.

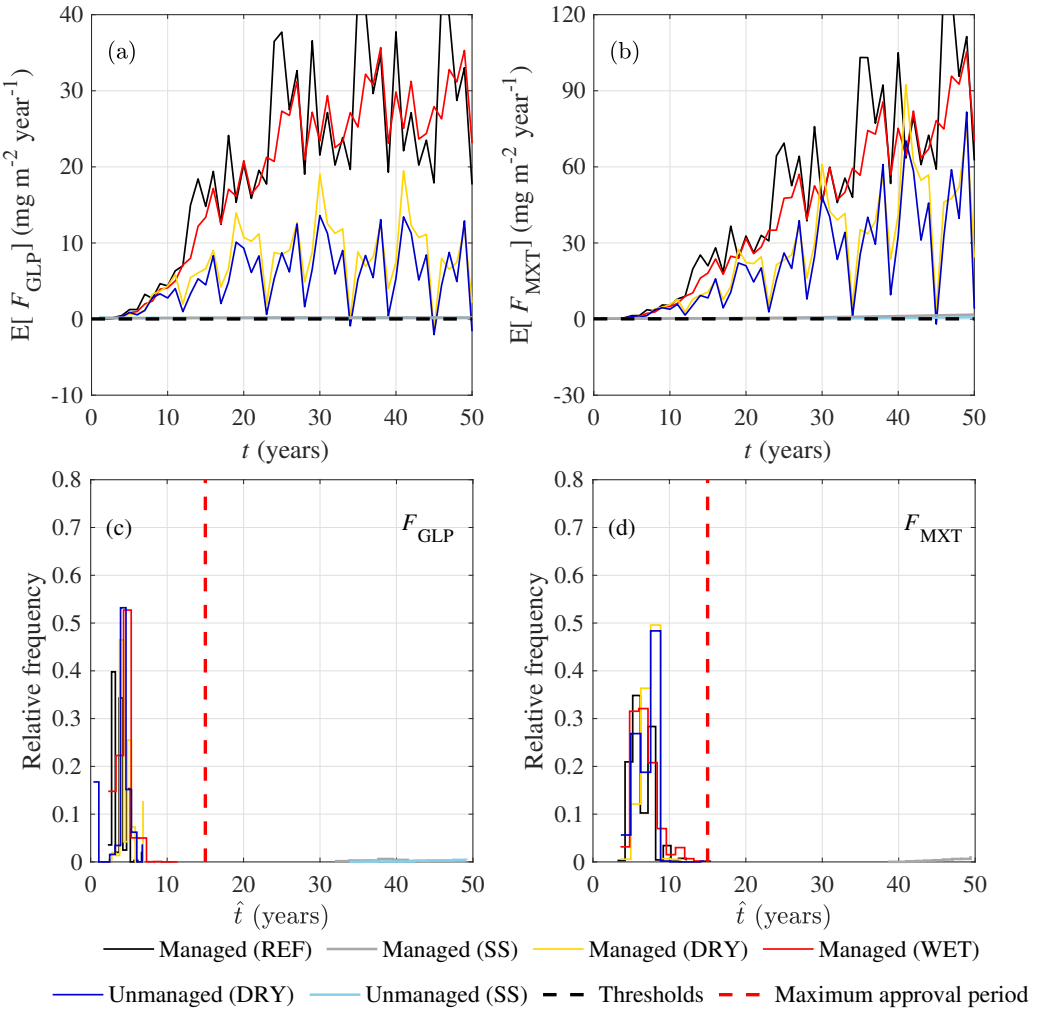

Figure 6: Sample-averaged GLP (F_{GLP}) (a) and MXT (F_{MXT})(b) leaching rate from RZ to BRZ, respectively; (c) and (d) represent the relative frequency of the exceedance time $\hat{t}$, i.e., the first time at which F_{GLP} and F_{AMPA} exceeded the threshold values.

3.4. AMAP and scenario analyses

$AMAP_i$ can be used as a time dependent sensitivity measure as illustrated in Figure 7. We observe that parameter ranking is consistent across different outputs for the same scenario because the selected outputs are inherently linked between each other. Significant differences in ranking are conversely observed across scenarios. For example, parameter b has an important influence on the probability that both $C_{\mathrm{GLP}}|_{BRZ}$ and $C_{\mathrm{MXT}}|_{BRZ}$ exceed the related thresholds in DRY scenarios (Figure 7b,f,d,h). On the other hand, the parameter ψ_s is predominant over b in assessing $C_{\mathrm{GLP}}|_{BRZ}$ and $C_{\mathrm{MXT}}|_{BRZ}$ in wet scenarios (WET and REF). This result is consistent with Eq. (4), showing that the impact of b decreases as the water saturation increases.

Figure 8 displays $AMAP_i$ values evaluated for the transient scenarios and considering $C_{\mathrm{GLP}}|_{BRZ}$ and F_{MXT} after 15 years from the first GLP application. Analogous results for C_{MXT} and F_{GLP} in the transient conditions as well as for all target quantities in the SS scenarios are not reported since negligible values of $AMAP_i$ have been obtained for all parameters.

Parameters k and ψ_s are the most influential parameters on $C_{\mathrm{GLP}}|_{BRZ}$ and F_{GLP}. These results are consistent with Eq. (3); that is, (*i*) increasing k promotes faster water flows, and thereby solute transport; (*ii*) decreasing ψ_s promotes higher water capillary rise from BRZ to RZ contrasting leaching. Figures 7-8 allow identifying which parameters should be further constrained, such as through measurement campaigns, to reduce the uncertainty associated with probabilistic groundwater or soil contamination risk assessment.

3.5. Biodegradation and flow regime

Mean biodegradation efficiencies $\mathrm{E}[E_{\mathrm{GLP}}]$ and $\mathrm{E}[E_{\mathrm{AMPA}}]$ indicate that biodegradation starts as soon as GLP is applied and increases over time (see Figure 9). Sample-averaged GLP biodegradation efficiency does not change significantly among all investigated scenarios, because $\mathrm{E}[E_{\mathrm{GLP}}]$ varies only between 0.85 and 0.9. The variability slightly increases when AMPA is considered, values of $\mathrm{E}[E_{\mathrm{AMPA}}]$ ranging between 0.24 and 0.32. Therefore, the sensitivity of microbial activity

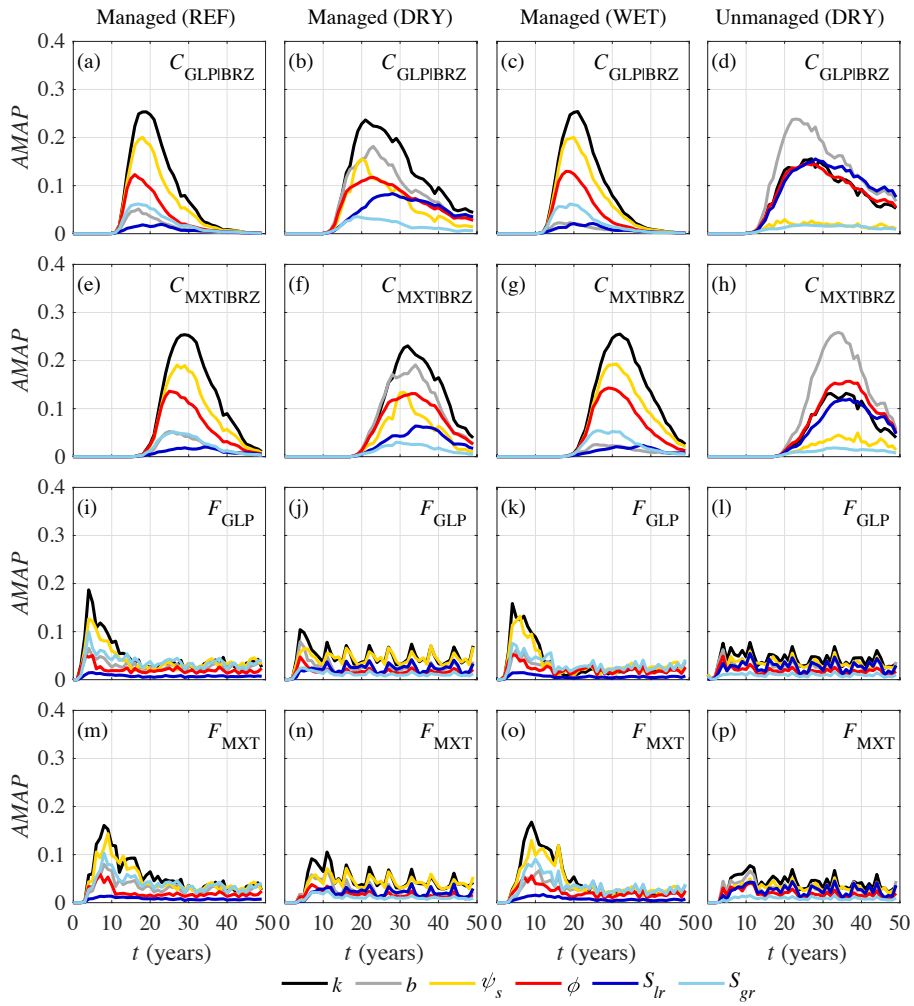

Figure 7: Time evolution of *AMAP* values: each row of plots is associated with a single output, columns distinguish the different scenarios.

and biodegradation to soil water availability is not particularly relevant in the selected scenarios even though water availability is explicitly considered in the model via Eq. (A.3).

This result suggests that simplified SS models as compared to time-resolved hydrologic boundary conditions may be used to predict overall contaminant mass budgets and is in line with previous numerical results by Tang *et al.* (2019b). However, SS scenarios do not yield accurate predictions of contaminants concentrations and leaching rates, which are instead driven by hydrological fluc-

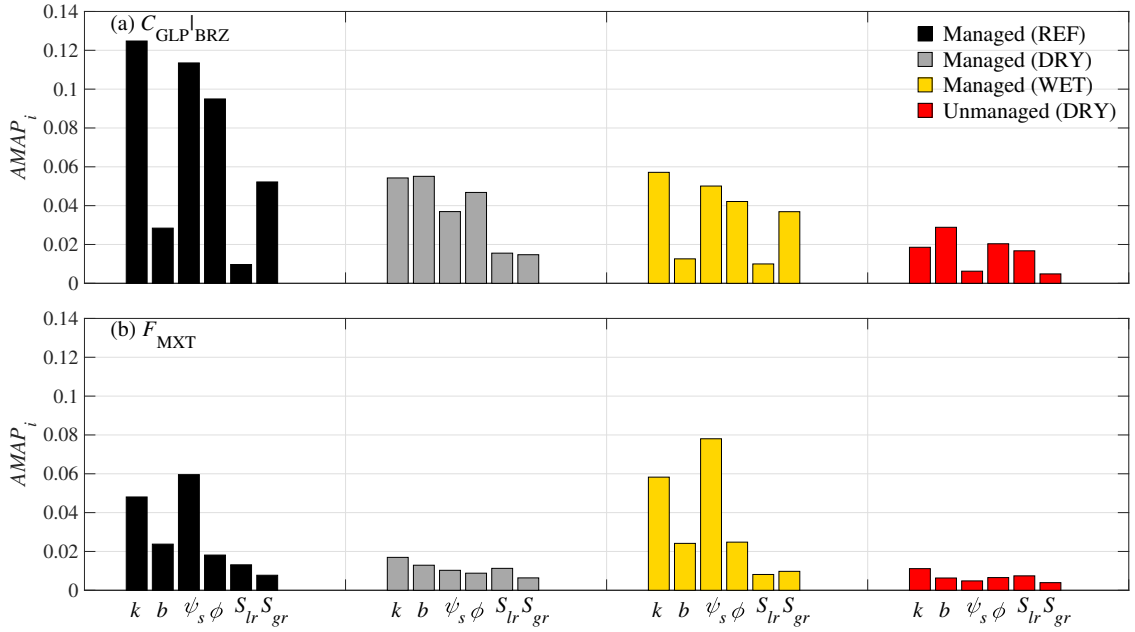

Figure 8: *AMAP* values of the uncertain soil hydraulic parameters relative to (a) GLP aqueous concentration BRZ and (b) MXT fluxes from RZ to BRZ. Analyses were carried out at time t = 15 years and results are grouped and colored according to each one of the six scenarios.

tuations, as discussed in Section 3.3.

3.6. Final remarks

Our results demonstrate that probabilistic indicators allow identifying the impact of soil hydraulic properties on pesticide contamination and leaching. In particular, we show that leaching is highly variable depending on soil properties, hydrological boundary conditions, and land management practices. In previous studies, Stenemo & Jarvis (2007) showed that the ranking of parameters may change according to soil texture, while Jury & Gruber (1989) showed that the persistence of pesticides with residence times longer than one year is more affected by soil rather than climatic variability. A number of studies have investigated uncertainty quantification and sensitivity analysis of agrochemical biodegradation and leaching (Dubus *et al.*, 2003; Stenemo & Jarvis, 2007; Heuvelink *et al.*, 2010). These modeling works assumed first order decay of pesticides in soil in

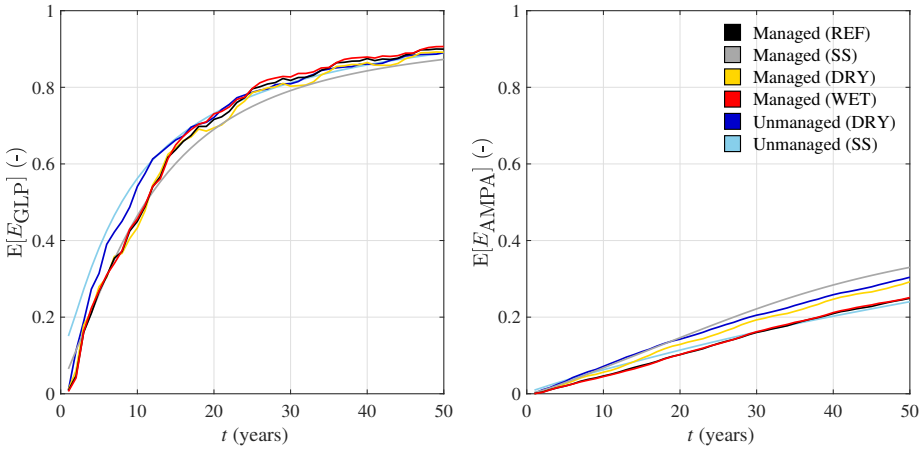

Figure 9: Mean biodegradation efficiency of (a) GLP and (b) AMPA in RZ and in each scenario.

contrast with our approach that uses Michaelis-Menten-Monod kinetics. Some of these studies suggest that soil hydraulics has a smaller impact on pesticide fate as compared to half life and sorption properties, yet our results demonstrate that soil hydraulic properties have a relevant impact on threshold exceedance probability. In particular, the *AMAP* time series demonstrate the impact of soil hydraulic parameters appears to be more persistent in dry than wet scenarios. We emphasize that the above mentioned studies typically rely on a single output statistics for parameter ranking in pesticide leaching (Heuvelink *et al.*, 2010) or quantify sensitivity by changing parameters one at a time (Dubus *et al.*, 2003). Conversely, our results advocate for the use of a suite of global sensitivity indicators to thoroughly test the system response, and are in line with recent studies (Borgonovo *et al.*, 2017; Ceriotti *et al.*, 2018). Hence *AMAP* can complement moment-based sensitivity indicators that do not account for safety threshold considered in risk assessment protocols. Likewise moment-based indicators, *AMAP* can be used with any input parameter distribution, model formulation and dimensionality. The present work considers only scenario and parametric uncertainty. Model structure and dimensionality has been identified as another possible relevant source of uncertainty. Different results may be obtained upon considering diverse flow and transport formulations as well as alternative biogeochemical reaction networks. Future work is en-

visaged to provide a formal derivation of model structure sensitivity indices within a probabilistic risk assessment framework, as recently discussed in Dell'Oca *et al.* (2020).

4. Conclusions

Our work leads to the following major conclusions:

- We test the use of a suite of sensitivity indicators for soil and groundwater environmental risk assessment. To this end, we introduce a new sensitivity index ($AMAP$) and we show its application to contamination and pollution following glyphosate (GLP) herbicide dispersion in an agricultural soil column. $AMAP$ provides a time dependent indication of the relevance of each parameter on the probability for a given output to exceed a user defined threshold and complements available moment-based sensitivity metrics. The index developed here can be readily applied to rank uncertain parameters with respect to any arbitrary threshold related to the quality of any environmental sphere (e.g., water, air, soil, and their combinations) and to inform appropriate management and restoration strategies.

- We quantify the impact of uncertainty in soil hydraulic parameters on the time required for the concentration of GLP and its toxic metabolite AMPA to exceed pollution thresholds in the top soil, below the root zone and for the fluxes measured at 1 m depth. We repeated these analyses in different scenarios of managed and unmanaged water budgets. When concentration thresholds are considered, dry scenarios result in larger uncertainty in terms of exceedance time as compared to wet ones. Measurements of contaminant fluxes reduce such uncertainty.

- Parameter ranking varies with ecohydrological scenarios of precipitations and irrigation practice. Permeability (k) and air-entry suction (ψ_s) have the greatest effect on exceedance of water quality safety limits in the reference scenario as well as in the wet scenario. The

influence of the pore volume distribution index b emerges in dry conditions, probably due to its relation with the relative permeability. Gas residual saturation may play an important role in contaminant transport in some scenarios, however it has shown a relatively minor role on pollution assessment. Porosity appears to have a minor effect on pollution risk as compared to permeability, except in the driest investigated conditions. These results suggest that *AMAP* can be used to refine uncertainty quantification in hazard assessment through measurements campaigns or to design risk management strategies, which can be specialized to local ecohydrological boundary conditions.

- Steady state scenarios do not allow assessing contamination and pollution and overpredict pollution time scales if compared with time-resolved simulations. Remarkably the sample-averaged GLP and AMPA biodegradation efficiencies show only minor differences across the tested scenarios. This result shows that steady state simulation may be able to match biodegradation efficiency yielded by time-resolved boundary conditions, but are not effective in rendering contamination hazards. Average flow conditions neglect the impact of short range fluctuations that play a predominant role in triggering contamination and pollution.

5. Acknowledgements

We acknowledge the support of the University of Sydney through the SREI2020 EnviroSphere research program, the University of Sydney Mid-career Research Award and SOAR Fellowship supporting FM. G.M. Porta and M. Riva acknowledge the EU and MIUR for funding, in the frame of the collaborative international Consortium (WE-NEED) financed under the ERA-NET WaterWorks2014 Cofunded Call. This ERA-NET is an integral part of the 2015 Joint Activities developed by the Water Challenges for a Changing World Joint Programme Initiative (Water JPI). The authors acknowledge the Sydney Informatics Hub and the University of Sydney's high performance computing cluster Artemis for providing the high performance computing resources that

have contributed to the research results reported within this paper.

Appendix A. Modelling of feedbacks between water saturation and biodegradation rates

The rate of change of an aqueous species for a given reaction in (6) is $\hat{r}^i = \mathrm{d}X^i/\mathrm{d}t = x_i R^i$, where x_i is the stoichiometric number for species i and R^i [$\mathrm{ML^{-3}T^{-1}}$] is the reaction velocity. For a generic reaction with n_O n-order kinetic products, n_{MM} Michaelis-Menten-Monod terms with concentration $X_{n_{MM}}$, n_{COM} competitive reactants ($X_{n_{COM}}$), and n_{INB} inhibition terms ($X_{n_{INB}}$), the reaction velocity R is written as

$$R^i = r f_B \prod_{n_O} X_{n_O}^{x_{n_O}} \cdot \prod_{n_{MM}} \frac{X_{n_{MM}}}{X_{n_{MM}} + K_{n_{MM}} \left(1 + \sum_{n_{COM}} \frac{X_{n_{COM}}}{K_{n_{COM}}}\right)} \prod_{n_{INB}} \frac{X_{n_{INB}}}{X_{n_{INB}} + K_{n_{INB}}}, \tag{A.1}$$

where r [$\mathrm{T^{-1}}$] is the reaction rate constant, $K_{n_{MM}}$, $K_{n_{INB}}$, $K_{n_{COM}}$ are the Michaelis-Menten half saturation, competition and inhibition constants, respectively, $f_B = 1$ if the reaction is chemical or if the reaction is biochemical.

$$f_B = \min\left\{f(S_B), f(\theta), f(S_l)/\max\{f(S_l)\}\right\}, \tag{A.2}$$

where $f(S_B)$ [-], $f(\theta)$ [-] and $f(S_l)$ [-] are the specific microbial response functions to space availability, temperature and water saturation, respectively. These terms are evaluated as

$$f(S_B) = \min\left\{1 - \frac{S_B - S_{lr}}{1 - S_{gr} - S_{lr}}, 1 - \frac{f_l S_B}{S_l}, 1 - \frac{(1 - f_l) S_B}{S_g}\right\}, \tag{A.3a}$$

$$f(\Theta) = \left(\frac{e^{\theta}}{e^{\theta_{LB}} + e^{\theta}}\right)^n \cdot \left(\frac{e^{\theta_{UB}}}{e^{\theta_{UB}} + e^{\theta}}\right)^m, \tag{A.3b}$$

$$f(S_l) = \frac{S_l}{S_{l,LB} + S_l} \cdot \frac{S_{l,UB}}{S_{l,UB} + S_l}, \tag{A.3c}$$

where f_l [-] is the biomass water fraction, S_B [-] is the biomass saturation, θ_{LB} and θ_{UB} [Θ] and $S_{l,LB}$ [-] and $S_{l,UB}$ [-] are the lower and upper temperature (in Kelvin) and liquid saturation response

parameters. The three functions in (A.3) introduce a limitation to microbially driven reactions as a function of environmental factors that may limit the bacterial growth and/or activity.

The function $f(S_B)$ implies that microbial functional groups can grow as long as there is enough free water to immobilize, or gas space available for the cell solid fraction $(1 - f_l)$ to occupy, or there is enough pore volume to host the total microbial biomass volume. Following the scheme in Maggi & Porporato (2007), the function $f(S_B)$ also implies that the total water saturation includes the free (mobile) water saturation S_l and the immobilized water saturation $S_{lB} = f_l S_B$. As a consequence, the term ΔB in Eq. (3) accounts for the rate of change in mobile water saturation S_l when the total microbial biomass increases (i.e., $\Delta B > 0$ expresses water immobilization) or decreases (i.e., $\Delta B < 0$ expresses water remobilization). Hence, Eq. (3) is subject to the constraint $S_l + S_g + S_B = 1$. Function $f(T)$ limits R when temperature is below θ_{LB} and above θ_{UB}. Finally, function $f(S_l)$ limits R when S_l is below $S_{l,LB}$ or above $S_{l,UB}$.

The response function $f(\Theta)$ appearing in (A.1)-(A.2) was calculated with $m = 0.1$, $n = 0.5$, $\theta_{LB} = 6$ °C and $\theta_{UB} = 45$ °C to return the microbial activity curve typical of mesophiles documented in (Rittmann & McCarty, 2001) (Figure A.1a), while $f(S_l)$ was implemented with $S_{l,LB} = S_{l,UB} = 0.46$ to represent typical water stresses (e.g., Moyano *et al.*, 2012; Yan *et al.*, 2018) (Figure A.1b). The response function $f(S_B)$ was implemented with $f_l = 0.85$ after Rockhold *et al.* (2005) and varies with S_B and therefore with t and over z.

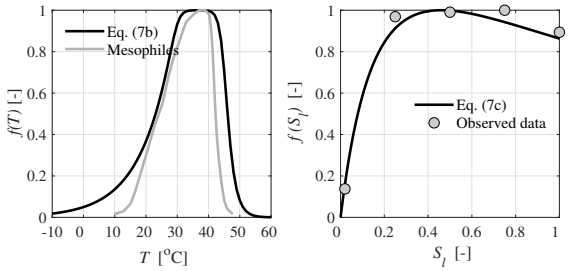

Figure A.1: Microbial response functions for (a) temperature stress $f(\Theta)$, and (b) water stress $f(S_l)$ as a function of temperature and mobile water saturation, respectively. Functions apply to all microbial functional groups accounted for in the GLP biogeochemical reaction network. The curve for mesophiles (left panel) is from (Rittmann & McCarty, 2001); experiments in the right panel are from Wickland & Neff (2008).

Appendix B. Stability of QMC sampling

Results in Figure B.2 shows that 5,000 stochastic realizations resulted in stable relative frequency in the target outputs. The results also allow appreciating that different scenarios resulted in different frequencies.

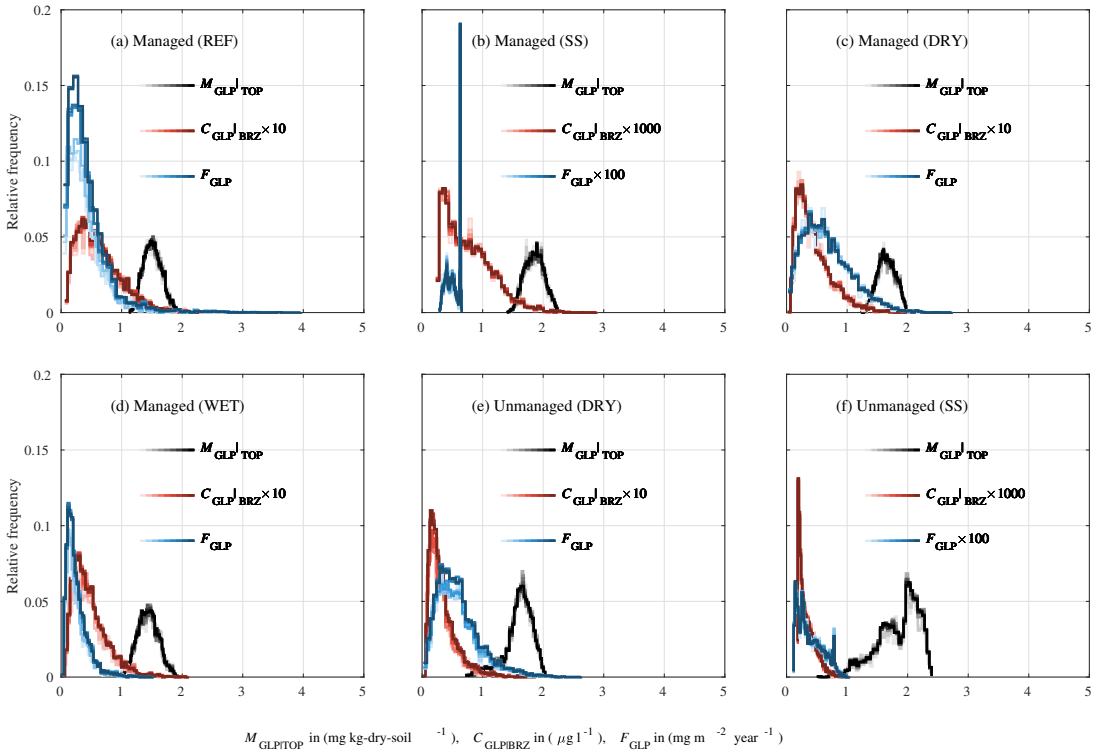

Figure B.2: Relative frequency of M_{GLP}, C_{GLP}, F_{GLP} at increasing number of QMC realizations (from 1,000 in light to 5,000 in dark, grey, red and blue, respectively).

References

2006/118/EC, Directive. 2006. European Parliament and Council, 2006. Directive 2006/118/EC of 12 December 2006 on the protection of groundwater against pollution and deterioration. **372**, 1931.

Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. 1998. *FAO - Food and Agriculture Organization of the United Nations, Natural Resources and Environment, Rome, 1998. ISBN 92-5-104219-5.* `http://www.fao.org/land-water/databases-and-software/eto-calculator/en/`.

Arjoon, D., Prasher, S.O., Clemente, R.S., Gallichanda, J., & Salehis, F. 1998. Effects of water table management on groundwater contamination from the use of prometryn in organic soils. *Canadian Water Resources Journal*, **23**, 9–20.

Armitage, James M, Quinn, Cristina L, & Wania, Frank. 2011. Global climate change and contaminants an overview of opportunities and priorities for modelling the potential implications for long-term human exposure to organic compounds in the Arctic. *Journal of Environmental Monitoring*, **13**(6), 1532–1546.

Arpae-Simc. 2016. *database accessed on 01.02.2017 at "*`http://www.smr.arpa.emr.it/dext3r/`*". Arpae emilia-romagna. Servizio IdroMeteoClima. Manuale Dext3r. Versione 0.95.*

Barrett, K. A., & McBride, M. B. 2005. Oxidative degradation of glyphosate and aminomethylphosphonate by manganese oxide. *env sci tec*, **39**, 9223–9228.

Barrios, Renys E., Gaonkar, Omkar, Snow, Daniel, Li, Yusong, Li, Xu, & Bartelt-Hunt, Shannon L. 2019. Enhanced biodegradation of atrazine at high infiltration rates in agricultural soils. *Environ. Sci.: Processes Impacts*, –.

548 Bates, Samantha C, Cullen, Alison, & Raftery, Adrian E. 2003. Bayesian uncertainty assessment
549 in multicompartment deterministic simulation models for environmental risk assessment. *Envi-*
550 *ronmetrics: The official journal of the International Environmetrics Society*, **14**(4), 355–371.

551 Boon, B., & Laudelout, H. 1962. Kinetics of nitrite oxidation by *Nitrobacter Winogradskyi*.
552 *biochem j*, **85**, 440–447.

553 Borgonovo, E., Lu, X., Plischke, E., Rakovec, O., & Hill, M.C. 2017. Making the most out of a
554 hydrological model data set: Sensitivity analyses to open the model black-box. *Water Resour.*
555 *Res.*, **53**(9), 7933–7950.

556 Brack, Werner, Altenburger, Rolf, Schüürmann, Gerrit, Krauss, Martin, Herráez, David López,
557 Van Gils, Jos, Slobodnik, Jaroslav, Munthe, John, Gawlik, Bernd Manfred, Van Wezel, An-
558 nemarie, *et al.* 2015. The SOLUTIONS project: challenges and responses for present and future
559 emerging pollutants in land and water resources management. *Science of the total environment*,
560 **503**, 22–31.

561 Brooks, R.H., & Corey, A.T. 1964. Hydraulic properties of porous media. *Hydrology Papers*, **3**.

562 Brutsaert, Wilfried. 2000. A concise parameterization of the hydraulic conductivity of unsaturated
563 soils. *Advances in water resources*, **23**(8), 811–815.

564 Carsel, R.F., Mulkey, L.A., Lorber, M.N., & Baskin, L.B. 1985. The Pesticide Root Zone Model
565 (PRZM): A procedure for evaluating pesticide leaching threats to groundwater. *Ecol Model*, **30**,
566 49–69.

567 Ceriotti, G., Guadagnini, L., Porta, G., & Guadagnini, A. 2018. Local and Global Sensitivity Anal-
568 ysis of Cr (VI) Geogenic Leakage Under Uncertain Environmental Conditions. *Water Resources*
569 *Research*, **54**(8), 5785–5802.

570 Chiari, G., Genovesi, R., Raimondi, S., Sarno, G., & Tarocco, P. 2016. *database accessed on*
571 *01.02.2017 at "http://cloud.consorziocer.it/FaldaNET/retefalda/index". CG*
572 *and GR from Consorzio di Bonifca di secondo grado per il Canale Emiliano-Romagnolo; RS*
573 *from Coop. I.ter; SG from Regione Emilia-Romagna. Servizio ricerca, innovazione e promozione*
574 *del sistema agroalimentare; TP from Regione Emilia-Romagna. Servizio Geologico, Sismico e*
575 *dei Suoli.*

576 Dai, Yongjiu, Shangguan, Wei, Duan, Qingyun, Liu, Baoyuan, Fu, Suhua, & Niu, Guoyue. 2013.
577 Development of a China dataset of soil hydraulic parameters using pedotransfer functions for
578 land surface modeling. *Journal of Hydrometeorology*, **14**(3), 869–887.

579 Dell'Oca, A., Riva, M., & Guadagnini, A. 2017. Moment-based Metrics for Global Sensitivity
580 Analysis of Hydrological Systems. *Hydrol. Earth Syst. Sci.*, **21**, 6219–6234.

581 Dell'Oca, A., Riva, M., & Guadagnini, A. 2020. Global Sensitivity Analysis for Multiple Interpre-
582 tive Models With Uncertain Parameters. *Water Resources Research*, **56**(2), e2019WR025754.

583 Dubus, Igor G, Brown, Colin D, & Beulke, Sabine. 2003. Sensitivity analyses for four pesticide
584 leaching models. *Pest Management Science*, **59**(9), 962–982.

585 EPA, US Environmental Protection Agency. 2008. *Integrated Modeling for Integrated Environmen-*
586 *tal Decision Making*. Tech. rept. EPA-100-R-08-010. Office of the Science Advisor, Washington,
587 DC.

588 Fu, C., Xu, Y., Bundy, A., Grüss, A., Coll, M., Heymans, J.J., Fulton, E.A., Shannon, L., Halouani,
589 G., Velez, L., Akoglu, E., Lynam, C.P., & Shin, Y.-J. 2019. Making ecological indicators man-
590 agement ready: Assessing the specificity, sensitivity, and threshold response of ecological indi-
591 cators. *Ecological Indicators*, **105**, 16–28.

592 Ghanbarian-Alavijeh, B, Liaghat, A, Huang, Guan-Hua, & Van Genuchten, M Th. 2010. Estima-
593 tion of the van Genuchten soil water retention properties from soil textural data. *Pedosphere*,
594 **20**(4), 456–465.

595 Hengl, T., de Jesus, J.M., Heuvelink, G.B., Gonzalez, M.R., Kilibarda, M., Blagotic, A.and Shang-
596 guan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R.,
597 MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., &
598 Kempen, B. 2017. SoilGrids250m: Global gridded soil information based on machine learning.
599 *PLoS One*, **12**(2), e0169748.

600 Heuvelink, G.B.M., Burgers, S.L.G.E., Tiktak, A., & Berg, F. Van Den. 2010. Uncertainty and
601 stochastic sensitivity analysis of the GeoPEARL pesticide leaching model. *Geoderma*, **155**(3),
602 186 – 192.

603 Hiscock, K, Lovett, A, Saich, A, Dockerty, T, Johnson, P, Sandhu, C, Sünnenberg, G, Appleton,
604 K, Harris, B, & Greaves, J. 2007. Modelling land-use scenarios to reduce groundwater nitrate
605 pollution: the European Water4All project. *Quarterly Journal of Engineering Geology and*
606 *Hydrogeology*, **40**(4), 417–434.

607 Jackson, Leland J, Trebitz, Anett S, & Cottingham, Kathryn L. 2000. An introduction to the
608 practice of ecological modeling. *BioScience*, **50**(8), 694–706.

609 Jene, B. 1998. *PELMO 3.00, Manual Extension, SLFA Neustadt, Ecology Department, Staatliche*
610 *Lehr - und Forschungsanstalt für Landwirtschaft, Weinbau und Gartenbau, Breitenweg 71, D-*
611 *67435 Neustadt, Germany.*

612 Jury, William A., & Gruber, Joachim. 1989. A stochastic analysis of the influence of soil and
613 climatic variability on the estimate of pesticide groundwater pollution potential. *Water Resources*
614 *Research*, **25**(12), 2465–2474.

615 Kubsad, D., Nilsson, E. E., King, S. E., Sadler-Riggleman, I., Beck, D., & Skinner, M. K. 2019.
616 Assessment of Glyphosate Induced Epigenetic Transgenerational Inheritance of Pathologies and
617 Sperm Epimutations: Generational Toxicology. *Scientific Reports*, **9**, 6372.

618 la Cecilia, D., & Maggi, F. 2018. Analysis of glyphosate degradation in a soil microcosm. *Envi-*
619 *ronmental Pollution*, **233**, 201–207.

620 la Cecilia, D., Tang, F.H.M., Coleman, N.V., Conoley, C., Vervoort, R.W., & Maggi, F. 2018.
621 Glyphosate dispersion, degradation, and aquifer contamination in vineyards and wheat fields in
622 the Po Valley, Italy. *Water Research*, **146**, 37–54.

623 la Cecilia, D., Riley, W. J., & Maggi, F. 2019. Biochemical modeling of microbial memory effects
624 and catabolite repression on soil organic carbon compounds. *Soil Biology and Biochemistry*,
625 **128**, 1–12.

626 Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N., Baldé, A.B., Bertollini,
627 R., O'Reilly, S.B., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Crop-
628 per, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M.,
629 Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray,
630 C.J.L, Ndahimananjara, J.D., Perera, F., Potocnik, J., Preker, A. S., Ramesh, J., Rockström, J.,
631 Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk,
632 W.A., van Schayck, O.C.P., Yadama, G.N., Yumkella, K., & Zhong, M. 2018. THE LANCET
633 COMMISSIONS. *THE LANCET COMMISSIONS*, **391**, 462–512.

634 Lewis, Kathleen, Tzilivakis, John, Green, Andrew, Warner, Douglas, *et al.* 2006. Pesticide Proper-
635 ties DataBase (PPDB).

636 Li, H., Joshi, S. R., & Jaisi, D. P. 2015. Degradation and isotope source tracking of glyphosate and
637 aminomethylphosphonic acid. *J. Agric. Food Chem.*, **64**, 529–538.

638 Libralato, S., Pranovi, F., Zucchetta, M., Anelli Monti, M., & Link, J.S. 2019. Global thresholds in
639 properties emerging from cumulative curves of marine ecosystems. *Ecological Indicators*, **103**,
640 554–562.

641 Lupi, Leonardo, Bedmar, Francisco, Puricelli, Marino, Marino, Damián, Aparicio, Virginia C.,
642 Wunderlin, Daniel, & Miglioranza, Karina S.B. 2019. Glyphosate runoff and its occurrence in
643 rainwater and subsurface soil in the nearby area of agricultural fields in Argentina. *Chemosphere*,
644 **225**, 906 – 914.

645 Maggi, F., & Porporato, A. 2007. Coupled moisture and microbial dynamics in unsaturated soils.
646 *Water resources research*, **43**, W07444.

647 Maggi, F., Gu, C., Riley, W., Hornberger, G., Venterea, R., Xu, T., Spycher, N., Steefel, C., Miller,
648 N., & Oldenburg, C. 2008. A mechanistic treatment of the dominant soil nitrogen cycling pro-
649 cesses: Model development, testing, and application. *J Geophys Res*, 1–13.

650 Maggi, F., la Cecilia, D., Tang, F.H.M., & McBratney, A. 2020. The global environmental hazard
651 of glyphosate use. *Science of the Total Environment*, **717**, 137167.

652 Maggi, Federico. 2018. *BRTSim v3.1, Release a, A general-purpose*
653 *multiphase and multispecies computational solver for biogeochemical*
654 *reaction-advection-dispersion processes in porous and non-porous media*.
655 https://www.dropbox.com/sh/wrfspx9f1dvuspr/AAD5iA9PsteX3ygAJxQDxAy9a?dl=0: User
656 guide and technical manual, first Ed., 15 August 2018, p 75.

657 Maggi, Federico. 2019. BRTSim, a general-purpose computational solver for hydrological, bio-
658 geochemical, and ecosystem dynamics. *arXiv preprint arXiv:1903.07015*.

659 Maggi, Federico, Tang, Fiona H.M., la Cecilia, Daniele, & McBratney, Alexander. 2019. PEST-

660 CHEMGRIDS, global gridded maps of the top 20 crop-specific pesticide application rates from
661 2015 to 2025. *Scientific Data*, **6**, 170.

662 Manheim, D.C., Detwiler, R.L., & Jiang, S.C. 2019. Application of unstructured kinetic models to
663 predict microcystin biodegradation: Towards a practical approach for drinking water treatment.
664 *Water Research*, **149**, 617–631.

665 Moyano, F. E., Vasilyeva, N., Bouckaert, L., Cook, F., Craine, J., Curiel Yuste, J., Don, A., Epron,
666 D., Formanek, P., Franzluebbers, A., Ilstedt, U., Katterer, T., Orchard, V., Reichstein, M., Rey,
667 A., Ruamps, L., Subke, J.-A., Thomsen, I. K., & Chenu, C. 2012. The moisture response of soil
668 heterotrophic respiration: interaction with soil properties. *Biogeosciences*, **9**, 1173–1182.

669 Oreskes, Naomi, Shrader-Frechette, Kristin, & Belitz, Kenneth. 1994. Verification, validation, and
670 confirmation of numerical models in the earth sciences. *Science*, **263**(5147), 641–646.

671 Peck, AJ. 1969. Entrapment, stability, and persistence of air bubbles in soil water. *Soil Research*,
672 **7**(2), 79–90.

673 Pianosi, Francesca, Beven, Keith, Freer, Jim, Hall, Jim W, Rougier, Jonathan, Stephenson, David B,
674 & Wagener, Thorsten. 2016. Sensitivity analysis of environmental models: A systematic review
675 with practical workflow. *Environmental Modelling & Software*, **79**, 214–232.

676 Porta, G., la Cecilia, D., Guadagnini, A., & Maggi, F. 2018. Implications of uncertain bioreactive
677 parameters on a complex reaction network of atrazine biodegradation in soil. *Advances in Water*
678 *Resources*, **121**, 263–276.

679 Razavi, Saman, & Gupta, Hoshin V. 2015. What do we mean by sensitivity analysis? The need
680 for comprehensive characterization of "global" sensitivity in Earth and Environmental systems
681 models. *Water Res. Resear.*, **51**(5), 3070–3092.

Richards, L.A. 1931. Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, **1**(5), 318–333.

Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., & Guerry, N. 2014. Long residence times of rapidly decomposable soil organic matter: application of a multi-phase, multicomponent, and vertically resolved model (BAMS1) to soil carbon dynamics. *Geoscientific Model Development*, **7**, 1335–1355.

Rittmann, Bruce E, & McCarty, Perry L. 2001. *Environmental biotechnology: principles and applications*. New York: McGrawHill.

Rockhold, Mark L, Yarwood, RR, Niemet, MR, Bottomley, Peter J, & Selker, John S. 2005. Experimental observations and numerical modeling of coupled microbial and transport processes in variably saturated sand. *Vadose Zone Journal*, **4**(2), 407–417.

Rodriguez Eugenio, N, McLaughlin, M., & Pennock, D. 2018. *"Soil Pollution: A Hidden Reality"*. Tech. rept. FAO.

SGSS. 2016. *database accessed on 01.02.2017 at "https://applicazioni.regione.emilia-romagna.it/cartografia_sgss/user/viewer.jsp?service=pedologia&bookmark=1%22". Regione Emilia-Romagna. Servizio Geologico, Sismico e dei Suoli.*

Smith, RM, & Browning, DR. 1943. Persistent Water-Unsaturation of Natural Soil in Relation to Various Soil and Plant Factors 1. *Soil Science Society of America Journal*, **7**(C), 114–119.

Sobol', I. M. 1993. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, **1**(4), 407–414 (1995).

Sobol', I. M. 1998. On quasi-Monte Carlo integrations. *Math. Comput. Simulat.*, **47**(2), 103 – 112.

Soracco, C. Germán, Villarreal, Rafael, Lozano, Luis Alberto, Vittori, Santiago, Melani, Esteban M., & Marino, Damián J.G. 2018. Glyphosate dynamics in a soil under conventional and no-till systems during a soybean growing season. *Geoderma*, **323**, 13 – 21.

Stenemo, Fredrik, & Jarvis, Nicholas. 2007. Accounting for uncertainty in pedotransfer functions in vulnerability assessments of pesticide leaching to groundwater. *Pest Management Science*, **63**(9), 867–875.

Tang, Fiona H.M., Jeffries, Thomas C., Vervoort, R. Willem, Conoley, Chris, Coleman, Nicholas V., & Maggi, Federico. 2019a. Microcosm experiments and kinetic modeling of glyphosate biodegradation in soils and sediments. *Science of The Total Environment*, **658**, 105–115.

Tang, Fiona H.M., la Cecilia, Daniele, Vervoort, R. Willem, Coleman, Nicholas V., Conoley, Chris, & Maggi, Federico. 2019b. A simple pre-factor for contaminant biodegradation potential and its application to pesticides risk assessment. *Mathematics and Computers in Simulation*.

Trevisan, M., Di Guardo, A., & Balderacchi, M. 2009. An environmental indicator to drive sustainable pest management practices. *Environmental Modelling & Software*, **24**(8), 994–1002.

Uusitalo, Laura, Lehikoinen, Annukka, Helle, Inari, & Myrberg, Kai. 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, **63**, 24–31.

Van den Berg, E., Tiktak, A., van Kraalingen, D., Van der Linden, T., & Boesten, J. 2012. *Documentation update for FOCUS-PEARL 4.4.4, Alterra, Wageningen, The Netherlands*.

Van Straalen, N.M., & Legler, J. 2018. Decision-making in a storm of discontent. Regulation of pesticides such as glyphosate needs to include societal assessment. *Science*, **360**(6392), 958–960.

Walker, Warren E, Harremoës, Poul, Rotmans, Jan, Van Der Sluijs, Jeroen P, Van Asselt, Marjolein BA, Janssen, Peter, & Krayer von Krauss, Martin P. 2003. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated assessment*, **4**(1), 5–17.

Wickland, K. P., & Neff, J. C. 2008. Decomposition of soil organic matter from boreal black spruce forest: environmental and chemical controls. *Biogeochemistry*, **87**(1), 29–47.

Yan, Z., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S., Liu, C., & Liu, C. 2018. A moisture function of soil heterotrophic respiration that incorporates microscale processes. *Nature communications*, **9**(1), 2562.