

## Journal Pre-proof

The global environmental hazard of glyphosate use

Federico Maggi, Daniele la Cecilia, Fiona H.M. Tang, Alexander McBratney



PII: S0048-9697(20)30677-X

DOI: <https://doi.org/10.1016/j.scitotenv.2020.137167>

Reference: STOTEN 137167

To appear in: *Science of the Total Environment*

Received date: 4 January 2020

Revised date: 1 February 2020

Accepted date: 5 February 2020

Please cite this article as: F. Maggi, D. la Cecilia, F.H.M. Tang, et al., The global environmental hazard of glyphosate use, *Science of the Total Environment* (2020), <https://doi.org/10.1016/j.scitotenv.2020.137167>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

## The global environmental hazard of glyphosate use

Federico Maggi<sup>a,\*</sup>, Daniele la Cecilia<sup>a,1</sup>, Fiona H.M. Tang<sup>a</sup>, Alexander McBratney<sup>b</sup>,

<sup>a</sup>*Environmental Engineering, School of Civil Engineering, The University of Sydney, 2006, Sydney, NSW, Australia.*

<sup>b</sup>*Sydney Institute of Agriculture, The University of Sydney, 2006, Sydney, NSW, Australia.*

### Abstract

Agricultural pesticides can become persistent environmental pollutants. Among many, glyphosate (GLP) is under particular scrutiny because of its extensive use and its alleged threats to the ecosystem and human health. Here, we introduce the first global environmental contamination analysis of GLP and its metabolite, AMPA, conducted with a mechanistic dynamic model at 0.5×0.5 degree spatial resolution (about 55 km at the equator) fed with geographically-distributed agricultural quantities, soil and biogeochemical properties, and hydroclimatic variables. Our analyses reveal that about 1% of croplands worldwide (385,000 km<sup>2</sup>) are susceptible to mid to high contamination hazard and less than 0.1% has a high hazard. Hotspots found in South America, Europe, and East and South Asia were mostly correlated to widespread GLP use in pastures, soybean, and corn; diffuse contributing processes were mainly biodegradation recalcitrance and persistence, while soil residue accumulation and leaching below the root zone contributed locally to the hazard in hotspots. Hydroclimatic and soil variables were major controlling factors of contamination hotspots. The relatively low risk of environmental exposure highlighted in our work for a single active substance does not rule out a greater recognition of environmental pollution by pesticides and calls for worldwide cooperation to develop timely standards and implement regulated strategies to prevent excess global environmental pollution.

*Keywords:* Hazard analysis, glyphosate, AMPA, global scales, environmental modelling.

---

\*Corresponding authors

*Email addresses:* federico.maggi@sydney.edu.au (Federico Maggi), daniele.lacecilia@sydney.edu.au (Daniele la Cecilia), fiona.tang@sydney.edu.au (Fiona H.M. Tang), alex.mcbratney@sydney.edu.au (Alexander McBratney)

<sup>1</sup> Current address: Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

## 1. Introduction

Since its first commercialization in the 70s (Baird, 1971), glyphosate (GLP) has become the most used herbicide worldwide (Benbrook, 2016; Duke & Powles, 2008) with about 600 to 750 thousand tonnes used annually and an expected 740 to 920 thousand tonnes to be used by 2025 (see estimates for 2015, 2020 and 2025 in Supplementary Information S1 after (Maggi *et al.*, 2019)). In agriculture, GLP mainly serves to control weeds and clear vegetation cover before sowing crops. Depending on country-specific regulations, GLP can be used as a pre-harvest desiccant to improve threshing, or ripen some crops to prevent yield loss in wet conditions (e.g., in some European countries, EFSA 2015, and the USA, Moechnig & Deneke 2011). Crops with a GLP resistance trait or genetically modified to be GLP resistant allow post-emergence applications. On the one hand, this has likely increased the global agricultural yield in the past 20 years (FAO, 2018) and has made GLP a relatively sustainable herbicide to support monoculture farming (Baylis, 2000) because is less toxic ( $LD_{50} = 4,230$  mg/kg-body-weight) than other contemporary and legacy herbicides such as atrazine ( $LD_{50} = 2,000$  mg/kg-body-weight, banned in 37 countries), 2,4-D ( $LD_{50} = 375$  mg/kg-body-weight, banned in 3 countries) and paraquat ( $LD_{50} = 150$  mg/kg-body-weight, banned in 46 countries)<sup>2</sup>. On the other hand, GLP-treated croplands have seen the evolution of 38 GLP-resistant weeds across 37 countries and in 34 different crops including orchards, vineyards, plantations, cereals, fallow and others since 1996 (Heap & Duke, 2018). We do not have information on whether farmers have increased GLP application rates to combat resistance in GLP-resistant monoculture but this practice is likely non-effective in the longer term.

About 130,000 tonnes of GLP were used in the last USA census in 2016, that is about 16 to 19 times more than in 1992 (retrieved from the USGS/PNSP database in Baker, 2018b). As of now, we have estimated application rates up to about 6.6 kg/ha on corn, 3.9 kg/ha on soybean, 0.7 kg/ha on wheat, and 0.4 kg/ha on cotton (derived from Maggi *et al.*, 2019). At these rates,

---

<sup>2</sup> The median lethal dose  $LD_{50}$  refers to the most restrictive between oral and dermal exposure in rats as per standard procedures in toxicology and are inferred to be applicable to man. Values quoted here were retrieved from the World Health Organization (WHO, 2010). Bans were retrieved from the Pesticide Action Network organization (PAN, 2019)

measured GLP residues exceeding 0.5 mg/kg-soil in European and USA croplands (Battaglin *et al.*, 2014; Silva *et al.*, 2018), between 0.01 and 500  $\mu\text{g/L}$  in ditches, drains, lakes, ponds, and wetlands in the USA (e.g., Battaglin *et al.*, 2014), and between 0.02 and 5200  $\mu\text{g/L}$  in large streams and groundwater (Székács & Darvas, 2018) substantiate a global risk of environmental contamination caused not only by GLP but also its recalcitrant metabolite aminomethyl-phosphonic acid (AMPA, Tang *et al.*, 2019), which has recently been measured at between 0.05 and 1.92 mg/kg-soil in European croplands (Silva *et al.*, 2018). This evidence reflects risks similar to legacy agrochemicals such as atrazine banned within the European Union in 2004 but still found along with metabolites in most groundwater monitoring wells (Paris *et al.*, 2016; Vonberg *et al.*, 2014). Both the Food and Agriculture Organization (FAO) of the United Nations and the Lancet Commission on Pollution and Health have recently indicated that environmental pollution by pesticides can impact the ecosystem and public health at unprecedented scales (Eugenio *et al.*, 2018; Landrigan *et al.*, 2018), and glyphosate formulations appear potential contributors to this according to the WHO and FAO joint committee on pesticide residues (FAO & WHO, 2016). Hence, whether GLP in its current and future formulations is going to retain its effectiveness is uncertain, but whether further increases in application rates may become not just a contamination problem but a global-scale pollution case are more realistic in light of knowledge of GLP as an ubiquitous and unpreventable environmental contaminant.

Here, we conduct a global-scale analysis of GLP and AMPA environmental contamination hazard to provide information over geographic regions of interest. The most important working method for our study is to deploy a suite of time- and space-resolved mechanistic descriptors of GLP and AMPA dynamics within the environment. Our objectives are to: (1) identify geographic regions of GLP and AMPA biodegradation recalcitrance in the root zone, residue accumulation in the root zone, potential runoff and leaching below the root zone to groundwater, and persistence in the root zone; and (2) combine the above information to identify hotspots of contamination hazard and the controlling anthropogenic and environmental factors using a multi-criteria framework.

To realise our scope, we used the BRTSim mechanistic dynamic model to achieve an explicit space-time description of GLP and AMPA dynamics at  $0.5 \times 0.5$  degree resolution globally (about 55 km grid cell size at the equator) from the top soil to below the root zone over a 50-year time scale. GLP and AMPA dynamics were described using the biogeochemical reaction network developed earlier in Cecilia *et al.*, (2018), which includes several abiotic and biotic

processes and interactions with soil nutrients, minerals, water, and temperature. The model was fed with median GLP application rates obtained from our PEST-CHEMGRIDSv1 database (Maggi *et al.*, 2019) and a massive set of geographically-distributed data of agricultural and agronomic parameters, hydroclimatic variables, soil properties, and biogeochemical kinetics. The study area extends over 6 prevalent crops (corn, wheat, soybean, cotton, rice and alfalfa) and 4 aggregated crops (orchards and grapes, vegetables and fruits, pasture and hay, and others) but excludes urban areas and private uses such as in gardens and dwellings. The bounding box ranges over 180°E-180°W; 56°S-84°N. The modelled 38.6 million km<sup>2</sup> of agricultural area (about 29,000 grid cells) involve 98 of the UN193 member countries as of 2019 and is living space for 6.58 billion people. Modelling was tested for reliable outputs by benchmarking against existing field and distributed georeferenced data. Two scenarios (REF, reference for "as is" conditions; and BAN for a global GLP ban) were investigated here. Our hazard analyses were conducted by combining multiple criteria that include policies and safety limits proposed here. Details on source data, modelling approaches, and analytical contents are provided in Section 2 and Supplementary Information. Modeling benchmarking is provided in Section 3 and Supplementary Information. The following analyses and interpretations in Section 3 are a selection of the most important outcomes of a comprehensive ensemble of results partially available for consultation in Supplementary Information or upon request.

## 2. Material and methods

### 2.1. *Acquired and newly-developed georeferenced global data*

Data used here consist of constant and dynamic quantities detailed in Supplementary Information S2, Table T1.

Constant quantities include: crop type and area (Monfreda *et al.*, 2008, after Maggi *et al.* 2019), soil physical properties (i.e., porosity, bulk density, and textural sand, silt and clay fractions each comprising 7 layers from the soil surface to 2 m depth from SoilGrids, Hengl *et al.*, 2017), soil hydraulic and thermal properties (i.e., the pore volume distribution index and air-entry suction of the Brooks and Corey model, permeability, and heat capacity and conductivity, each comprising 8 layers from the soil surface to 2.3 m depth Dai *et al.*, 2019), soil residual saturation (Zhang *et al.*, 2018), soil thickness (Pelletier *et al.*, 2016), and equilibrium water table depth (Fan *et al.*, 2013).

Constant quantities also include GLP-, AMPA- and nutrient-specific biogeochemical kinetic parameters (i.e., chemical and biochemical reactions, reaction rate constants, Michaelis-Menten half-saturation concentrations, biomass yields, biomass mortality rate, and inhibition and competition constants) used to describe the GLP biodegradation reaction network and its interactions with the C and N nutrient cycles and soil mineral interaction via sorption (la Cecilia & Maggi, 2018; la Cecilia *et al.*, 2018). These interactions depend on local nutrient availability explicitly accounted for in our work but kinetic parameters were not geographically varied. In addition to the above, we developed new specific global datasets for the maximum and average crop root density distribution using data in Allen *et al.* (1998), USDA (2016) and Fan *et al.* (2016) conditioned to the corresponding crops in Monfreda *et al.* (2008) and pastures and hays in Ramankutty *et al.* (2008) (Supplementary Information S3). We also developed the GLP and AMPA logK of linear equilibrium adsorption to soil minerals using adsorption experiments in Autio *et al.* (2004) and Sidoli *et al.* (2016) and the georeferenced soil pH in Hengl *et al.* (2017) (Supplementary Information S4).

Dynamic quantities include: daily rainfall (NOAA/OAR/ESRL PSD, 2019), daily atmospheric and land surface temperature (Menne *et al.*, 2012), monthly actual evapotranspiration (Zhang *et al.*, 2016), and 8-day net solar radiation (FLASHFlux team, 2019), which were used as boundary conditions. Additional datasets were developed: specifically, crop calendars in Sacks *et al.* (2010) were re-interpreted to calculate the daily gridded normalized crop calendar and derive the daily GLP application rates, irrigations water volume, and N-NH<sub>3</sub>NO<sub>2</sub> and P-PO<sub>4</sub><sup>-3</sup> fertilizations using annual estimates in PEST-CHEMGRIDSv1 (Maggi *et al.*, 2019), crop water security indicators Thenkabail *et al.*, 2016), and fertilization rates (Potter *et al.*, 2011a,b), respectively (Supplementary Information S5).

Corollary georeferenced administrative borders of countries (Sandvik, 2009) and watershed boundaries (Gassert *et al.*, 2014) were used for analyses.

All acquired and newly-developed georeferenced layers were harmonized to 0.5×0.5 degree resolution using the "imresize" Matlab 2018b built-in function before modelling. New datasets of potential use by third parties are distributed as described in Section 2.7.

## 2.2. GLP and AMPA biogeochemistry

We accounted for GLP and AMPA abiotic and biotic processes essential in metabolic and

cometabolic reactions, substrate affinity, competition and inhibition, microbes dynamics, and mineral sorption as described in Cecilia & Maggi (2018) and adapted to controlling factors accounted for in this work as represented in Supplementary Information S6, Fig. F7. GLP biodegradation occurs along three aerobic pathways that depend on carbon (C) and phosphate ( $\text{PO}_4^{3-}$ ) availability: the first two (P1R1 and P1R1s) liberates AMPA during GLP cometabolism and metabolism, and the third (P2R1s) liberates sarcosine via a cometabolic reaction inhibited by  $\text{PO}_4^{3-}$ . Cometabolic AMPA biodegradation (P1R2s) is inhibited by  $\text{PO}_4^{3-}$ . Both GLP and AMPA degradation take place via hydrolysing soil microorganisms ( $\text{B}_{\text{HYO}}$ ). GLP and AMPA chemodegradation by manganese oxide at the surface of some minerals (e.g., birnessite, Li *et al.*, 2016; Paudel *et al.*, 2015) was excluded here because other ions largely available in field conditions (e.g.,  $\text{Ca}^{2+}$ ) compete with GLP and AMPA for adsorption (Barrett & McBride, 2005). We excluded photo-degradation because this is known to be relatively weak and its occurrence in natural environments is uncharacterized.

We coupled the GLP reaction network to a soil organic matter pool (SOM) that releases dissolved organic carbon ( $\text{CH}_2\text{O}$ ) used cometabolically by  $\text{B}_{\text{HYO}}$ , and  $\text{NH}_4^+$  and  $\text{PO}_4^{3-}$  used by competing soil microorganisms. The GLP reaction network was also coupled with the soil N cycle in Maggi *et al.* (2008), which includes three microbial functional groups transforming  $\text{NH}_3 \rightarrow \text{NO}_2^-$  ( $\text{B}_{\text{AOB}}$ ),  $\text{NO}_2^- \rightarrow \text{NO}_3^-$  ( $\text{B}_{\text{NOB}}$ ), and  $\text{NO}_2^- \rightarrow \text{NO} \rightarrow \text{N}_2\text{O} \rightarrow \text{NO}_2^-$  ( $\text{B}_{\text{DEN}}$ ). Coupling with N and P nutrients was essential to include anthropogenic effects caused by fertilizations on competitive and inhibitory ecological feedbacks regulating microbial biodegradation. GLP and AMPA linear sorption was accounted for as a function of soil pH using our new georeferenced layers of the logK values as described above. GLP and AMPA soil biochemistry in our modelling also accounted for nonlinear effects caused by soil moisture, temperature, and pH on microbial activity, water immobilization and remobilization by microbial biomass, and stressors associated with competition for electron donors, acceptors, and space (Maggi, 2019b). The kinetic parameters describing all above processes are listed in Supplementary Information S6, Table T5 and T6.

### 2.3. *BRTSim* modelling

Modelling was conducted within the *BRTSim* v4.0a computational environment (based on Maggi, 2019a). *BRTSim* solves for non-isothermal continuity and conservation laws for water, gas, and heat flows, for transport of aqueous and gaseous species, and for kinetic and equilibrium reactions

amongst defined biogeochemical species. The computations use hybrid explicit-implicit numerical techniques within finite volumes largely detailed in Maggi (2019b). BRTSim was deployed on a 3-dimensional global soil grid resolved at  $0.5 \times 0.5$  degree horizontally and extended vertically over 3 soil layers (30 cm, 30 cm, and 40 cm thick, respectively) within the first meter of the root zone, and a layer with variable thickness depending on the distance to the equilibrium water table or bedrock. The grid also included 2 atmospheric layers to allow for heat exchange between soil and atmosphere and for water ponding. The 1-dimensional solvers in BRTSim simulated ecohydrological, biogeochemical and transport processes along the soil vertical in each cell of the global grid. In total, about 29,000 grid cells describing the agricultural land were modelled in this work. Methods of benchmarking of key modelled environmental variables and GLP and AMPA residues is provided in Section 3.

To help the reader capturing dynamic variations over the seasonal time scale in our modelling, we compiled animations of GLP and AMPA residue mass fraction in soil at various soil depths and their leaching below the root zone (Section 2.7).

#### **2.4. Scenario analyses**

We investigated two GLP application scenarios. The reference (REF) scenario assumed 50 years of GLP annual application rates as of 2015 (Maggi *et al.*, 2019). The second scenario (BAN) assumed a global GLP ban after 30 years. The hydrometeorological boundaries over the 50-year simulations were generated by cycling 6 years of complete datasets of daily precipitations, monthly evapotranspiration, and 8-day net solar radiation available from 2007 to 2012 (Supplementary Information S2, Table T1) and were unchanged in the two scenarios. We did not include scenarios of climate change or changes in agricultural practices in our analyses, which are a plan for future developments of this work; hence, analyses using the last 5 years of simulations are considered representative of current conditions.

#### **2.5. GLP and AMPA contamination calculations**

Modelling outputs include all hydraulic, thermal, and biogeochemical variables. Using the instantaneous GLP and AMPA total aqueous and equivalent sorbed concentration ( $C$ ), we calculated the instantaneous mass fraction



$$M_X(t)_\Omega = \frac{\int_{\Omega} C_X(t, z) m_X V_w(t, z) dz}{\int_{\Omega} M_s(z) dz},$$

where  $X$  is either GLP or AMPA,  $m$  is the molar mass,  $V_w$  is the water volume,  $M_s$  is the soil mass, and  $\Omega$  is either the top soil (TS, upper 30 cm), the root zone (RZ, upper 100 cm), or below the root zone (BRZ, depth below 100 cm). Instantaneous leaching below the root zone was calculated as

$$L_X(t) = \frac{\partial [C_X(t, BRZ) m_X V_w(t, BRZ)]}{\partial t}.$$

The biodegradation efficiency in the time interval  $\Delta t$  was calculated as

$$B_X(\Delta t) = \frac{\int_{\Delta t} \int_{\Omega} R_X(t, z) m_X V_w(t, z) dz dt}{\int_{\Delta t} PR(t) dt},$$

with  $R_X$  the reaction velocity of GLP or AMPA degradation, and  $PR$  either the GLP application rate or the GLP degradation rate feeding AMPA degradation. For GLP,  $B_{GLP}$  was calculated for each degradation pathways in the reaction network. Finally, persistence was calculated as the half-life

$$t_{1/2} | M_X(t)_\Omega \leq 0.5 M_X(t_0)_\Omega,$$

with  $t_0$  the time since the last GLP application. Time averages of  $M_X(t)$ ,  $L_X(t)$  and  $B_X(\Delta t)$  were calculated over the period from year 45 to 50 of the 50-year simulations. Other calculations such as residues by crop, country, and watershed used the same approaches distributed over the solving grid but masked with the corresponding data layers.

## 2.6. Multi-criteria hazard analysis

Hazard analysis revolves around two steps. The first identifies regions of hazard and detects hotspots, while the second ranks the controlling factors that discern hotspots from other regions. Hazard was calculated for 4 markers (target variables) upon satisfaction of a specific criterion. The first hazard refers to biodegradation of either GLP or AMPA when its efficiency is below 50%. There is currently no established standard on this marker even if the half-life of individual active substances is used in environmental risk analyses. We note however that the half-life of pesticides measured in laboratory conditions generally differs from environmental conditions and those

measured in the field generally refer to a sample possibly too small for statistical reliability as compared to the environmental variability, thus justifying the aid of mechanistic models that account for multiple processes and environments contributing to microbial dynamics in this assessment. The second hazard applies when the total GLP and AMPA residues accumulation (aqueous and adsorbed) is greater than or equal to  $1 \text{ mg/kg}_{\text{dry-soil}}$  for at least 50% of the time in TS. In this case, we used the ecotoxicity to earthworms, that is, the median lethal concentration  $\text{LC}_{50} = 5,600 \text{ mg/kg}_{\text{drysoil}}$  for GLP and  $\text{LC}_{50} = 1,000 \text{ mg/kg}_{\text{drysoil}}$  for AMPA Lewis *et al.* (2016). As a cautionary approach, we used  $\text{LC}_{50}$  of AMPA and we applied an assessment factor 1,000 according to EU (2003), which returns a safe value of  $1.0 \text{ mg/kg}_{\text{dry-soil}}$ . For this work, we also used the 50-th percentile for time exceedence to associate intensity to time. The third criterion for hazard is when leaching below the RZ of combined GLP and AMPA exceeds 5% of applied pesticide for at least 50% of the time. The 5-th percentile reflects the procedure in Cecilia *et al.* (2020) but we also included the 50-th percentile for time exceedence to associate intensity to time. The fourth criterion for hazard refers to when either GLP or AMPA in the TS are persistent, that is, their half-life since last GLP application is in excess of 6 months as defined in the Stockholm Convention (UNEP, 2001). Hazard of individual target markers are additively used to rank the hazard from 0 (lowest) to 4 (highest). Geographic grid cells with an hazard of 3 or 4 were considered "hotspots".

The second step links the target markers to the hazard and the controlling factors, and was ultimately used to isolate the controlling factors that mostly affect hotspots *versus* no- or low-risk regions. We calculated the correlation (determination) coefficient  $R_{i,j}^{HS}$  and  $R_{i,j}^0$  in hotspots and no risk regions, respectively, for marker  $i$  and controlling factor  $j$ .  $R$  values with significance  $p > 0.05$  were neglected. The strength  $S_j$  of a controlling factor  $j$  was next calculated as

$$S_j = \frac{|\sum_i R_{i,j}^{HS} - \sum_i R_{i,j}^0|}{df_j},$$

where  $df_j = \sum_i (p_{i,j} \leq 0.05)$  (Boolean sum) are the degrees of freedom for the controlling factor  $j$  on the target marker  $i$  and are such that  $S_j \in [0,1]$ .

## 2.7. Data availability

Raw datasets used here are publicly available from the sources provided in Supplementary Information S2, Table 1. Newly developed gridded data of crop root density distribution are available at <https://figshare.com/s/806748599be0d4c1fbcf>. Newly developed gridded data of various adsorption parameters for GLP and AMPA are available at <https://figshare.com/s/711baaab58725b35549b>. Animations of dynamic modelling are available as visual aid to the reader at <https://figshare.com/s/f36b91fb51b2f0bc751f> but specific screen shots can be requested to the authors. Other results such as maps presented here or in Supplementary Information are also available upon request.

### 3. Model benchmarking and data quality calculations

Prior to present our selection of results, we introduce the modelling benchmarking carried out against key environmental variable and the total GLP and AMPA residues.

We benchmarked the modelled soil wetness in TS (upper 30 cm of soil) against global data of long-term monthly averages estimated in the CPC Soil Moisture dataset by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (Fan & van den Dool, 2004). Global anomalies (map, scatter, and time series in 20 randomly selected grid cells) show that modelling represented hydroclimatic processes affecting soil water availability qualitatively well ( $R \approx 0.79$ ,  $p \leq 0.01$ , Supplementary Information S7, Fig. F8). We found sparse spatial anomalies as compared to the benchmark values; some of these anomalies were related to irrigation, which we presume were not explicitly included in the NOAA/CPC data and are highlighted in our analyses. We recall also that CPC data are outputs of modelling and therefore are an indication for our assessment work but cannot be used as exact calibration set for hydraulic parameters and hydroclimatic boundary conditions in our work.

The temperature dynamics in BRTSim were benchmarked against long-term daily average land surface temperature reanalysis from Menne *et al.* (2012), and was qualitatively well described ( $R \approx 0.94$ ,  $p \leq 0.01$ , Supplementary Information S7, Fig. F9). Seasonally, BRTSim appeared to underestimate the temperature excursion but we recall that our estimates refer to the top 30 cm of soil while data in the NOAA/NCEI dataset refer to the 5 cm above land surface and hence are subject to a greater variability as compared to the mineral soil.

Benchmarking of pH against values in the top soil in SoilGrids (Hengl *et al.*, 2017) was also satisfactory ( $R \approx 0.76$ ,  $p \leq 0.01$ , Supplementary Information S7, Fig. F10) with the exception of a small fraction of grid cells departing from the benchmark values in spite of the background  $H^+$  recovery described by reactions  $R_{H^+,P}$  and  $R_{H^+,R}$  (Supplementary Information S6, Table T5). We have verified whether anomalies were related to fertilizations but we found that N and P amendments did not cause anomalies and we explain therefore those outliers as due to nonlinearities controlling  $H^+$  in our reaction network. The quality of benchmarking in soil wetness, temperature, and pH were used in the data quality metrics described below.

We also benchmarked the estimated GLP and AMPA residue mass fractions in the top 30 cm of soil against residue measured in agricultural land reported in 18 datasets from various European countries and Argentina available in Aparicio *et al.* (2013); Peruzzo *et al.* (2008); Laitinen *et al.* (2009); Bergström *et al.* (2011); Simonsen *et al.* (2008); Napoli *et al.* (2016) and Silva *et al.* (2018). Our estimates were determined from the last 5 years of the 50-year reference (REF) scenario. The modelled GLP and AMPA soil residues include both aqueous and adsorbed phases and matched reasonably well (i.e., within the same order of magnitude) the values observed in agricultural sites (Supplementary Information S8, Table T7). We slightly underestimated the residues but we note that we considered the top 30 cm of soil while field sampling was commonly performed at the top 5 to 20 cm.

We provided a measure of quality for our hazard analysis that integrates various sources of uncertainty and data reliability. The hazard quality index  $Q$  was calculated as the weighted average

$$Q = \frac{3 \cdot Q_{APR}}{6} + \frac{(Q_{SI} + Q_T + Q_{pH})}{6},$$

where  $Q_{APR}$  is the data quality of GLP applications throughout all crops retrieved from Mahhi *et al.* (2019), and  $Q_{SI}$ ,  $Q_T$  and  $Q_{pH}$  are data quality relative to modelled soil wetness, temperature, and pH and were calculated as the absolute errors against the benchmark data in Fan & van den Dool (2004), Menne *et al.* (2012), and Hengl *et al.* (2017). Before aggregation into  $Q$ , each data quality index was normalized in  $[0,1]$ ;  $Q = 1$  and  $Q = 0$  indicate high and poor data quality, respectively. Maps of each quality index and the overall data quality map are shown in Supplementary Information S9, F11.

## 4. Results

### 4.1. *GLP and AMPA natural biotic removal*

We focus our first analysis on GLP and AMPA natural removal by microbial degradation, while we excluded other abiotic processes (see Section 2.2). Hence, we tracked the mass flow through the three biotic reactions that remove GLP relative to net GLP applications (i.e., we accounted for 20 crop interception and 20 wind drift after Trevisan *et al.*, 2009), and AMPA relative to AMPA produced by GLP degradation (Supplementary Information S6). GLP removal mostly ranged between 70 and 95% (Fig. 1a) and was geographically distributed fairly homogeneously showing sparse but circumscribed degradation recalcitrance in North and South America, North Europe, and South East Asia. An efficient GLP biodegradation was likely due to the combined effect of the three GLP degradation reactions (cometabolic reactions P1R1s and P2R1s are responsible of about 46% degradation, see inset pie chart in Fig. 1), the environmental conditions such as low inhibition of GLP hydrolysis by  $\text{PO}_4^{-3}$  in P2R1s, and the relatively important GLP degradation pathway into AMPA (P1R1), which removed about 53% of GLP. Other factors can also be invoked such as pH-dependent adsorption, carbon (C) availability for cometabolic reactions, and microbial biomass dynamics (processes explicitly accounted for, see Section 2.2). In contrast, AMPA biodegradation was relatively heterogeneous and ranged from about 15 to 90% depending on the geographic location (Fig. 1b) likely because only one slow pathway is known to remove AMPA by cometabolic oxidation, which requires available C and is inhibited at low  $\text{PO}_4^{-3}$  concentration (P1R2s, Supplementary Information S6). This analysis gives an important picture of residue prevalence and long-range, and long-term contamination and potential impact on non-target organisms.

Figure 1: Average percent biodegradation in the root zone (RZ, upper 100 cm) of: (a) GLP relative to net GLP applications; and (b) AMPA relative to AMPA production by GLP biodegradation. Average percent was calculated over the last 5 years of the 50-year simulation time when the RZ reaches a stationary state. Inset pie chart in panel (a) represents the prevalence in GLP biodegradation along pathways P1R1, P1R1s and P2R1s defined in the GLP biodegradation reaction network in Supplementary Information S5. (color figure, page width)

## 4.2. GLP and AMPA residues in soil

Undegraded GLP and AMPA can remain as residue in soil or be transported and dispersed in the environment. Hence, we calculated the cumulative aqueous and adsorbed GLP and AMPA residue (mass fraction) time-averaged over the last 5 years of the REF scenario. Both the top soil (TS, upper 30 cm) and root zone (RZ, upper 100 cm) show residues distributed worldwide with intensity that coarsely reflects the GLP annual application rates (see geographic map in Supplementary Information S5, Fig. F5). More than 18 and 43% of TS in croplands is affected by GLP and AMPA residues greater than or equal to  $0.01 \text{ mg/kg}_{\text{dry-soil}}$ , respectively (current detection limit by HPLC methods is 0.02 to 0.03 mg/kg, Silva *et al.*, 2018), while  $< 1$  and about 17% is affected by residues greater than or equal to  $0.1 \text{ mg/kg}_{\text{dry-soil}}$  (i.e., above detection limit, Fig. 2a and b). Residues in the RZ (data not shown) had geographic distribution similar to but with lower intensity than in TS because of dilution over the soil profile and losses via biodegradation highlighted in Section 4.1. There is currently a relatively small number of site-specific data for GLP and AMPA residues, most of which are sampled in European countries and Argentina; we have therefore collected those observations and verified that our estimates substantially capture their order of magnitude in several instances as a benchmark to our modeling work (see details in Section 3)

More than 98% of GLP and AMPA residues were found in the adsorbed phase. However, GLP sorption did not have substantial implications on biotic degradation; hence, low AMPA degradation is mainly explained by recalcitrance (i.e., the kinetics are slow as compared to GLP) and by  $\text{PO}_4^{-3}$  inhibition rather than low availability in the aqueous phase (Section 4.1). The relatively high AMPA recalcitrance can therefore have implications on its mobility and leaching through the soil (detailed in Section 4.3).

We analysed the geographic distribution of the time fraction where aqueous and adsorbed GLP and AMPA residues are present in the TS at a mass fraction equal to or greater than  $0.1 \text{ mg/kg}_{\text{dry-soil}}$  (i.e., above detection limit, Fig. 2c and d). Identified regions were weakly but significantly correlated with low biodegradation efficiency (i.e.,  $R \approx -0.38$ ,  $p \leq 0.05$  for GLP and  $R \approx -0.20$ ,  $p \leq 0.05$  for AMPA against data in Fig. 1), thus suggesting that biotic GLP and AMPA removal plays a role in controlling residue accumulation in agroecosystems. We then

calculated the product between the above time fraction relative to GLP and AMPA residues with the harvested area of each crop to highlight crops exposure to residues of at least  $0.1 \text{ mg/kg}_{\text{dry-soil}}$ ; here, we found that pastures and hays, corn, and soybean have the widest and longest exposure and the overall distribution across crops was only slightly different for GLP and AMPA residues (inset bars in Fig. 2c and d). With a similar approach applied to the absolute area of countries, we found that USA, Brasil, and China have the widest and longest exposure to both GLP and AMPA residues of at least  $0.1 \text{ mg/kg}_{\text{dry-soil}}$  (inset bars in Fig. 2c and d). In terms of country relative area, maps in Fig. 2c and d highlight a remarkably high fraction affected by AMPA residues across all European countries.

Because we accounted for time- and space-resolved biodegradation and residue accumulation, we investigated the time correlation between GLP and AMPA residues in the TS of each geographic grid cell and we found that this is nearly evenly distributed spatially with a prevalence of positive correlation and circumscribed regions of negative correlation mainly in India (Fig. 2e, only grid cells with  $p \leq 0.05$  are shown), thus suggesting a relatively small time lag between GLP applications and biodegradation leading to rapid increases in AMPA followed by slow AMPA degradation (see reaction network, Supplementary Information S6, Fig. F7). The correlation along the RZ shows that GLP and AMPA are mostly positively correlated (i.e., it is highly likely that the presence of one of the two implies the presence of the other, Fig. 2f, only grid cells with  $p \leq 0.05$  are shown). Regions of either positive or negative significant correlation in time or along the RZ did not show any evident pattern across the latitude but were clustered regionally, thus suggesting the effect of environmental factors in addition to biodegradation that are investigated in greater detail in Section 4.5.

Figure 2: Geographic distribution of: (a) and (b) time-averaged GLP and AMPA residue mass fraction in the topsoil (TS, upper 30 cm), respectively; (c) and (d) percent time during which time-resolved GLP and AMPA mass fractions exceed  $0.1 \text{ mg/kg}_{\text{dry-soil}}$  in TS, respectively; and (e) and (f) time correlation of GLP and AMPA residue in TS and depth correlation of time-resolved residue in the root zone (RZ, upper 100 cm), respectively. Residues included aqueous and adsorbed GLP and AMPA. Inset pie charts in (a) to (b) represent the residue partitioning between adsorbed and aqueous phases and the percent area with range mapped to the linear color scale in each panel. Correlations in (e) and (f) only account for grid cells where the significance is

$p \leq 0.05$ . Time of analysis covers the last 5 years of the 50-year reference (REF) scenario when the RZ reaches a stationary state. BD stands for "below detection" concentration by HPLC methods. Both panels (a) and (b), and (c) and (d) share the same color scheme. (color figure, page width)

### 4.3. *GLP and AMPA in surface water and groundwater*

Surface water flows such as runoff and stream discharge are means of GLP and AMPA residues dispersal leading to potential long-range ecosystem contamination within watersheds and along rivers. We therefore assessed the runoff load susceptibility and leaching within watersheds by intersecting our gridded GLP and AMPA residue estimates in TS with gridded maps of major watersheds available in the AQUEDUCTv2.1 database (Gassert *et al.*, 2014).

While reporting that the total residue load within a watershed was highly correlated to the watershed extension and its agricultural land (data not shown here), we analysed the load per unit watershed area and we tagged the 15 most susceptible watersheds: the Schelde (North Europe) and Dongjiang (Asia) have a relatively high runoff density potential for both GLP and AMPA but we note that a number of other European watersheds (e.g., Elbe, Weser, Loire, Rhine, Po and Seine), and North and South American watersheds (Pee Dee and Uruguay) appear susceptible to both GLP and AMPA residues (Fig. 3a and b). We also analysed the aqueous GLP and AMPA leaching below RZ as a proxy to groundwater susceptibility to contamination. Gridded average annual leaching rates highlight that potential contamination of the groundwater was generally limited to GLP fluxes up to  $10 \text{ mg/m}^2$  per year in circumscribed regions in South America and East Asia. AMPA leaching rates of at least  $100 \text{ mg/m}^2$  per year were geographically common in agroecosystems and were largely correlated to low biodegradation ( $R \approx -0.81$ ,  $p \leq 0.05$  against data in Figs. 3c and d). We remark that even if immediate leaching of aqueous residue was relatively low globally, the potential residue leaching is substantially higher considering that about 98% of residues is adsorbed onto soil minerals and therefore prone to be remobilized. For example, specific agricultural practices not explicitly accounted for in our modeling work such as tillage and plowing can lead, in combination with irrigations and soil pH corrections, to unprotection from soil aggregates and desorption from mineral surfaces, thus increasing GLP and AMPA mobility.

Figure 3: Geographic distribution of: (a) and (b) watershed-aggregated runoff potential expressed



in GLP and AMPA residue in the top soil (TS, upper 30 cm) per unit area, respectively; and (c) and (d) time-averaged annual leaching rate below the root zone (RZ, upper 100 cm) per unit area. Inset bars in (a) and (b) represent the top 15 watersheds with potential contaminant runoff. Residues included aqueous and adsorbed GLP and AMPA. Time averages were calculated over the last 5 years of the 50-year simulation time when the RZ reaches a stationary state. (color figure, page width)

#### **4.4. Persistence after ban**

After identifying geographic regions susceptible to biodegradation recalcitrance, residue accumulation, and potential dispersion by runoff and leaching, we quantified the time scale for these stressors to decrease their intensity by 50% in the hypothesis that a global ban on GLP use is enacted from year 30 of the 50 simulated years (BAN scenario). That is, we assessed the half-life *a posteriori* to quantify the persistence according to the definition of the Stockholm Convention (i.e., 6-month half-life is the minimum lapse defining a persistent organic pollutant, UNEP, 2001). Hence, we tracked the GLP and AMPA residues mass fraction in TS for 20 years after the ban and the retrieved half-life was assumed to represent the time scale also for runoff and leaching to decrease by 50% for simplicity. We found that GLP was geographically distributed as a non-persistent contaminant in about 72% of croplands worldwide (Fig. 4a), while AMPA was largely distributed as a persistent contaminant in about 93% of croplands (Fig. 4b). GLP persistence (half-life) was generally not exceeding 1 to 2 years except in some regions of South East Asia and South America, while AMPA persisted in vast cropland portions up to the 20-year assessment time of the BAN scenario.

The assessed half-life in our modeling exercise shows an extraordinary greater value than what available from public repositories such as the Pesticide Properties DataBase (PPDB, Lewis *et al.*, 2016), which reports median degradation half-life  $DT_{50}$  of 15 and 23.79 days for GLP in laboratory and field conditions, respectively, by aerobic degradation at 20°C, and  $DT_{50}$  of 121.4 and 419 days for AMPA in the same conditions. We contend therefore whether complementary approaches should be used to assess the residence or life time of pesticides in soil. More importantly, our analysis of persistence suggests that AMPA is the key contaminant related to glyphosate use in agriculture and the one that requires an accurate ecotoxicity characterization, which is currently not established according to recognized standards.

Figure 4: Geographic distribution of GLP and AMPA persistence in the top soil (TS, upper 30 cm). Persistence was assessed in the scenario of global ban (BAN) applied from year 30 of the 50-year simulation and was used to determine *a posteriori* the half-life since last application. Residues included aqueous and adsorbed GLP and AMPA. (color figure, page width)

#### 4.5. Hotspots systematic identification and controlling factors

The geography of GLP and AMPA susceptibility to biodegradation recalcitrance, residues accumulation, leaching, and persistence is instrumental to eventually classify regions presenting one or more contamination hazards. We therefore used the above four quantities as the hazard markers and we integrated them in a systematic approach to identify contamination hotspots based on the following marker-specific criteria (detailed in Section 2.6): (1) the biodegradation efficiency of either GLP or AMPA in TS is lower than 50%; (2) the total residue accumulation of GLP and AMPA (aqueous and adsorbed in the TS) is greater than or equal to 1 mg/kg<sub>dry-soil</sub> for at least 50% of the time; (3) the leaching below the RZ of combined GLP and AMPA exceeds 5% of applied pesticide for at least 50% of the time; and (4) either GLP or AMPA in the TS are persistent (i.e., the half-life since the last GLP application exceeds 6 months). A grid cell was therefore classified as "no hazard" (or "hazard 0") if none of the four criteria was met, or "hazard X" when X criteria were met, the highest being "hazard 4".

Among each criterion, residue accumulation and leaching were the least impacting while biodegradation recalcitrance and persistence led to hazards over the widest cropland area (Fig. 5a to d). Within the assessment tolerance quantified by the quality index  $Q$  (see Section 3 for calculation details and data quality map in Supplementary Information S9), the combined hazard highlights minor regions subject to high and mid-high hazard (about 1%) and a wide extent subject to low or negligible hazard (about 60%, Fig. 5e). Note that we used the median GLP application rate calculated from the PEST-CHEMGRIDSv1 database; however, PEST-CHEMGRIDSv1 reports the "high" and "low" estimates for each active ingredients, including GLP. Hence, we repeated the same analyses for the two "high" and "low" application rates, each for the two REF and BAN scenarios required to retrieve the target markers. The resulting aggregated hazard maps (see Supplementary Information S10, Fig. 12) highlight minor differences mainly in South America, North Europe, and midland China. Comparisons allowed us to confidently use the scenario of median GLP application rates as a reference scenarios for hazard and for

interpretations.

Figure 5: Geographic distribution of GLP and AMPA aggregated contamination hazards relative to (a) biodegradation recalcitrance, (b) residue accumulation in the TS, (c) leaching below the RZ, and (d) persistence, respectively. (e) aggregated contamination hazard resulting from (a) to (d) (see Section 2.6). (color figure, page width)

We analysed the controlling anthropogenic and environmental factors that characterise geographic hotspots from no- or low-hazard regions. To this end, we calculated the correlation between the target markers in the 4 criteria above and several anthropogenic and environmental factors in both instances. We gained a particularly informative picture and we were able to draw explanations for the presence of contamination hazard hotspots (Fig. 6). First, agronomic practices and crop characteristics controlled markedly the hazard susceptibility in both hotspots and no-hazard regions because GLP applications are the source of contamination. Notably, N and P fertilizations (NAP and PAP labels in Fig. 6) have a weak positive correlation with biodegradation in hotspots but have a stronger positive and significant correlation with residue accumulation in no- or low-hazard regions because amended N increases denitrification and competition with GLP and AMPA degraders for C source, while amended P inhibits GLP (one pathway) and AMPA degradation (see stamp "A" in Fig. 6). However, GLP applications alone cannot be invoked to necessarily cause a contamination hazard as shown by positive and significant correlation with target markers in no- or low-hazard regions. In average, hydroclimatic conditions have an opposite effect in hotspots as compared to no- or low-hazard regions (stamp "B" in Fig. 6). For example, high precipitations and evapotranspiration (PRE and ETA labels) significantly reduce biodegradation and, to a smaller extent, also residues in hotspots but increase residues in no- or low-hazard regions. Interestingly, the effect of atmospheric temperature (ATP label) only influenced hotspots but did not have significant effects in other regions. We found that dynamic soil variables that had a significant correlation with hazard markers in hotspots were not significantly correlated to the corresponding markers in no- or low-hazard regions. However, we observed that pH (pH label) controlled hotspots with a generally positive correlation against hazard markers as compared to a range in correlations in no- or low-hazard regions. We found that GLP and AMPA degraders characterized degradation in hotspots but were not relevant in other

regions. Finally, of all soil physical, thermal, and hydraulic properties, none appeared to clearly distinguish hotspots because they were sparsely correlated in both regions with no clear pattern except the soil organic carbon (SOC label) and bulk density (BLD label), which have a significant effect on biodegradation and residues in hotspots (stamp "C" in Fig. 6). Surprisingly, SOC was negatively correlated to biodegradation in hotspots likely because the higher the SOC content, the higher the competition between GLP and AMPA degraders with other aerobes.

We ultimately developed a relatively simple additive ranking method to systematically discern controlling discriminants of target markers amongst anthropogenic and environmental factors (see Section 3). The outcome of our method (Fig. 6c) shows that the strongest discriminants distinguishing hazard hotspots from no- or low-hazard regions are soil carbon (SOC), precipitations (PRE), soil pH (pH) and bulk density (BLD).

Figure 6: Controlling anthropogenic, environmental, soil dynamic variables, and soil physical properties on aqueous GLP and AMPA biodegradation, and aqueous and adsorbed GLP and AMPA residue in the top soil (TS, upper 30 cm), aqueous GLP and AMPA leaching below the root zone (RZ, upper 100 cm), and persistence in TS. Controls are calculated as correlation (determination) coefficient over the grid for constant quantities and for the 5-year time average for time-dependent variables when the RZ reaches a stationary state. Acronyms stand for: APR, glyphosate annual application rate; NAP, nitrogen annual application rate; PAP, phosphorus annual application rate; CRTD, crop median root density; PRE, 5-year mean precipitation depth; ETA, 5-year mean actual evapotranspiration; ATP, 5-year mean atmospheric temperature; pH, 5-year mean soil pH; SI, 5-year mean soil wetness; STP, 5-year mean soil temperature; BHYO, 5-year mean GLP- and AMPA-degrading bacterial concentration; SND, sand; SLT, silt; CLY, clay; SOC, soil organic carbon; BLD, soil bulk density; PHI, soil porosity; CSOL, soil heat capacity; KS, soil hydraulic conductivity at saturation; LAMBDA, pore volume distribution index of the Brooks and Corey model; KT, soil thermal conductivity;  $K_{GLP}$ , glyphosate adsorption constant;  $K_{AMPA}$ , AMPA adsorption constant. The level of significance in this analysis is  $p \leq 0.01$ . (color figure, page width)

## 5. Discussion

This work complements and extends other assessments such as the GLP and AMPA contamination mapping in surface waters and groundwater (reviewed in Székács & Darvas, 2018), the GLP and AMPA contamination mapping in European agricultural soils Silva *et al.*, 2018), as well as other studies reporting GLP soil residues at plot or field scales (Aparicio *et al.*, 2013; Peruzzo *et al.*, 2008; Laitinen *et al.*, 2009; Bergström *et al.*, 2011; Simonsen *et al.*, 2008; Napoli *et al.*, 2016). Knowledge achieved combining different sources of assessment can serve to specialize monitoring strategies in geographic regions that we have identified as hotspots and that currently lack in or there is no knowledge of monitoring operations.

Our estimates of glyphosate residue in soil suggests no acute impact on earthworms but the environmental persistence, however, contradicts the earlier perception of rapid degradation and rather indicates extended periods of time during which residues may be remobilized such as by wind erosion, runoff, and leaching and cause exposure to non-target organisms and ecosystems. A clear understanding of health consequences due to long-term and low-dose chronic exposure is currently missing and only recent attention being paid to this type of exposure has challenged the previous assumption of low GLP toxicity to non-target organisms showing oxidative stress in a model human liver cancer and potentially cause cytogenetic and primary DNA damages (Kašuba *et al.*, 2017). More importantly, little is currently known about the ecotoxicity of AMPA, which appears from our analyses to be the most persistent (long lasting) and recalcitrant metabolite (difficult to be catabolized by soil microorganisms), and the one requiring attention.

Regulatory authorities and health and environmental agencies hold controversial opinions on GLP safety to human, and AMPA seems not to be under particular investigation. This hampers the undertaking of a firm position on the use of glyphosate in agriculture and two prevalent but contrasting views have emerged: one argues that agriculture may lose yield without GLP, and this can severely affect the capability to satisfy the agricultural production demand and carry a global food security crisis (McBratney *et al.*, 2014; Danne *et al.*, 2019); the other considers unsustainable the excessive GLP use and alternative agronomic practices should be implemented to meet food and soil security criteria (Wilson & Tisdell, 2001). Our analyses, however, provide the ground for testing hypothesis that can fill in gaps between these two standpoints so to reduce the hazard in hotspots without necessarily making a case for banning glyphosate. For example, conditioning soil microbial diversity to that in no- or low-hazard regions such as facilitating the growth of hydrolyzing GLP degraders (interpreted from Fig. 6a) can enhance biodegradation and reduce

residues accumulation. Likewise, fine tuning pH and SOM amendments can reduce competitive and inhibitory enzymatic mechanisms among soil microbial functional groups and enhance GLP and AMPA degradation. Yet, this or similar approaches may not suffice alone. Finding the just intersection between profitability and sustainability should be undertaken with agreed international governance that guarantees minimal ecosystem losses and optimal agricultural production. Within such a framework, soil health is to be acknowledged for its role to maintain environmental health and new standards should therefore be established for the ecotoxicity to key soil microorganisms, the standard that is currently missing not only for glyphosate but all pesticides. In addition, regulations of pesticide use may rely on standardised modelling scenarios, which may not always be appropriate for a specific region and should rather be specialized such as highlighted by controlling factors in hotspots in our analyses. If we exclude emerging agricultural techniques such as biodynamic, biological and others, use of pesticides as well as other plant protection agrochemicals include a large number of active substances (about 500 and 1300 are registered in the USDA/NASS and the EU pesticide database European Commission (2016), Baker (2018a)) leading to residue mixtures whose combined effect may be amplified as compared to single substances. Hence, the outcome of this work, though relevant for GLP use in agriculture, may provide an incomplete picture of the greatest context. We remark therefore that the results proposed here and the discussion initiated above are an initial step into a relevant research area that is still in its infancy and requires a greater and coordinated effort.

We acknowledge that even under the constraints of available datasets used in this work, a number of degrees of freedom may have introduced a level of uncertainty not necessarily captured by the quality index that characterizes our contamination hazard analysis. For example, our work reflects the estimated current GLP application rates and does not employ projections on pesticide biotechnology such as new formulations, changes in governance such as regulations and bans, changes in climatic patterns, and changes in dietary habits and consumers attitude, all aspects that can have implications on land use and agriculture. With the development of new datasets and the possibility to interlace them to our modeling infrastructure we are confident that mechanistic, process-based modelling will become more widely used in this type of environmental assessments.

We do acknowledge some numerical and modeling limitations, though, which may be overcome with new releases of our model and further understanding of underpinning processes. For example, we did not account for lateral flow in this current modelling work but we do expect

that the hydrological information on flow direction and accumulation available in the HYDROSHED or similar databases will be of use to determine the actual transport of pesticide through runoff and aquifer flow at the spatial scales of interest to this work. Similarly, we have explicitly accounted for important agricultural practices such as irrigations and N and P fertilizations but we acknowledge that plowing and tillage as well as other operations can have an impact on pesticide dynamics, which we did not explicitly include in our assessment. We acknowledge also limitations in the representation of some processes such as the adsorption of glyphosate and AMPA, which was described in this work by a linear equilibrium and thus we excluded mineral surface saturation. We have produced estimates of nonlinear equilibrium parameters such as for the Langmuir adsorption, which we believe will be of use to a third party and in future improvement of contamination assessments.

We finally acknowledge that a level of uncertainty is present also in the publicly sourced datasets used here and, in particular, that only a few datasets of glyphosate and AMPA are available for benchmarking our estimates. In spite of our modelling work generally fitting available data quite well, we underline that new initiatives should be undertaken to provide as detailed picture as possible of the global extent of contamination by glyphosate as well as other plant protection products used in agriculture.

## **6. Concluding remarks**

We conducted a comprehensive analysis of glyphosate and aminomethyl-phosphonic acid (AMPA) potential environmental contamination hazard using a dynamic model that allowed us to achieve a time- and spatially-resolved picture of key pathways of GLP and AMPA dynamics globally. Although we retrieved a massive amount of geographically distributed outputs, we concentrated on biodegradation recalcitrance, soil residue, leaching below the root zone, and persistence as key markers to assess which environmental compartments are mostly susceptible and which regions aggregate the greatest combined hazard.

Our analyses bring to light that a low contamination occurs in nearly all croplands where glyphosate is used; about 1% undergoes mid-high to high hazard and impacts about 385,000 km<sup>2</sup>. Hotspots were found spatially sparse and forming relatively small clusters. These hotspots were mainly correlated to pastures and hays, soybean, and corn agriculture, and were largely controlled

by rainfalls, pH, soil organic carbon and bulk density. Outside hotspots, glyphosate contamination was low but globally pervasive. Glyphosate was found to be a persistent contaminant at relatively low values in about 30% of global croplands but AMPA was found to be persistent in about 93% of croplands. Following our findings, we recommend that more attention should be paid to characterize in full the ecotoxicity of AMPA and develop better knowledge of its biodegradation pathways and kinetics to improve future assessments of contamination by this metabolite.

## **7. Acknowledgements**

This work was supported by The University of Sydney through the Sydney Research Excellence Initiative (SREI 2020) EnviroSphere and Voucher Programs. FM was supported by the Mid Career Research Award (MCR) and the Sydney Research Accelerator Fellowship (SOAR) by The University of Sydney. The authors acknowledge the Sydney Informatics Hub and the high performance computing cluster Artemis made available for this work.

## **8. Author contributions**

FM and DLC have conceptualized the research project, DLC and FHMT have collected data and conducted pre-modeling analyses, FM has conducted HPC modeling and post-modeling analyses, all authors have contributed to the interpretations to analyses and have written the manuscript.

## **9. Competing interests**

The authors declare no competing interests.

## **10. Code availability**

Data manipulation for pre-modelling was conducted with custom codes written in the Matlab 2018b environment and are available upon reasonable request. Dynamic modelling was conducted within the BRTSim environment, which can be accessed under Creative Commons Attribution 4.0 International License (CC-BY4.0) along with the User Manual and Technical Guide, samples of simulations, and data post-processing scripts at



<https://sites.google.com/site/thebrtsimproject/home>. All files specific to simulations presented here can be requested to the authors to replicate or modify simulations (about 840 GB).

## References

## References

- Allen, Richard G, Pereira, Luis S, Raes, Dirk, Smith, Martin, *et al.* 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO Irrigation and drainage paper No. 56, Rome*, **300**(9), D05109.
- Aparicio, V.C., De Gerónimo, E., Marino, D., Primost, J., Carriquiriborde, P., & Costa, J.L. 2013. Environmental fate of glyphosate and aminomethylphosphonic acid in surface waters and soil of agricultural basins. *Chemosphere*, **93**, 1866 – 1873.
- Autio, S., Siimes, K., Laitinen, P., Rämö, S., Oinonen, S., & Eronen, L. 2004. Adsorption of sugar beet herbicides to Finnish soils. *Chemosphere*, **55**, 215–225.
- Baird, DD. 1971. Introduction of a new broad spectrum post emergence herbicide class with utility for herbaceous perennial weed control. *In: Proceedings of the 26th North Central Weed Conference (Kansas City, USA, 7-9 December 1971)*. North Central Weed Science Society.
- Baker, N.T. 2018a. *Estimated Annual Agricultural Pesticide Use by Major Crop or Crop Group for States of the Conterminous United States, 1992–2016*. <https://water.usgs.gov/nawqa/pnsp/usage/maps/index.php>.
- Baker, N.T. 2018b. *Estimated Annual Agricultural Pesticide Use by Major Crop or Crop Group for States of the Conterminous United States, 1992–2016*. U.S. Geological Survey. Database accessed on 27.02.2019 at "<https://water.usgs.gov/nawqa/pnsp/usage/maps/index.php>".
- Barrett, K. A., & McBride, M. B. 2005. Oxidative degradation of glyphosate and aminomethylphosphonate by manganese oxide. *env sci tec*, **39**, 9223–9228.
- Battaglin, W.A., Meyer, M.T., Kuivila, K.M., & Dietze, J.E. 2014. Glyphosate and its degradation product AMPA occur frequently and widely in U.S. soils, surface water, groundwater, and precipitation. *Journal of the American Water Resources Association*, **50**.

- Baylis, Alan D. 2000. Why glyphosate is a global herbicide: strengths, weaknesses and prospects. *Pest Management Science: formerly Pesticide Science*, **56**(4), 299–308.
- Benbrook, Charles M. 2016. Trends in glyphosate herbicide use in the United States and globally. *Environmental Sciences Europe*, **28**(1), 3.
- Bergström, L., Börjesson, E., & Stenström, J. 2011. Laboratory and lysimeter studies of glyphosate and aminomethylphosphonic acid in a sand and a clay soil. *J. Environ. Qual.*, **40**, 98–108.
- Dai, Yongjiu, Xin, Qinchuan, Wei, Nan, Zhang, Yonggen, Shangguan, Wei, Yuan, Hua, Zhang, Shupeng, Liu, Shaofeng, & Lu, Xingjie. 2019. A global high-resolution dataset of soil hydraulic and thermal properties for land surface modeling. *Journal of Advances in Modeling Earth Systems*.
- Danne, M., Musshoff, O., & Schulte, M. 2019. Analysing the importance of glyphosate as part of agricultural strategies: A discrete choice experiment. *Land Use Policy*, **86**, 189–207.
- Duke, S.O., & Powles, S.B. 2008. Glyphosate: a once-in-a-century herbicide. *pest manag sci*, **64**, 319–325.
- EFSA. 2015. Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate. *EFSA Journal*, **13**, 107.
- EU, European Commission. 2003. Technical Guidance Document on Risk Assessment in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. *Ispra (IT): European Commission Joint Research Centre. EUR*, **20418**.
- Eugenio, N.R., McLaughlin, M., & Pennock, D. 2018. Soil Pollution: a hidden reality. *Rome, FAO*, 1–142.
- European Commission, EU. 2016. *PLANT, EU Pesticides database*. "<http://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/public/>".
- Fan, Jianling, McConkey, Brian, Wang, Hong, & Janzen, Henry. 2016. Root distribution by depth for temperate agricultural crops. *Field Crops Research*, **189**, 68–74.
- Fan, Ying, Li, H., & Miguez-Macho, Gonzalo. 2013. Global patterns of groundwater table depth. *Science*, **339**(6122), 940–943.

- Fan, Yun, & van den Dool, Huug. 2004. Climate Prediction Center global monthly soil moisture data set at 0.5 resolution for 1948 to present. *Journal of Geophysical Research: Atmospheres*, **109**(D10).
- FAO. 2018. *Food and Agriculture Organization of the United Nations. Database Collection of the Food and Agriculture Organization of the United Nations. Database accessed on 26.07.2019 at "http://www.fao.org/faostat/en/#data"*.
- FAO, & WHO. 2016. *Pesticide residues in food 2016. Joint FAO/WHO Meeting on Pesticide Residues. FAO Plant Production and Protection Paper no. 227. Rome. ISBN 978-92-5-109246-0*.
- FLASHFlux team, NASA Langley Research Center. 2019. *Net Radiation (8 day) dataset. NASA Earth Observations*.
- Gassert, Francis, Luck, Matt, Landis, Matt, Reig, Paul, & Shiao, Tien. 2014. Aqueduct global maps 2.1: Constructing decision-relevant global water risk indicators. *World Resources Institute*, 31.
- Heap, I., & Duke, S. O. 2018. Overview of glyphosate-resistant weeds worldwide. *Pest Management Science*, **74**, 1040–1049.
- Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagoti, Aleksandar, Shangguan, Wei, Wright, Marvin N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, G.G.B., Ribeiro, E., Wheeler, I, Mantel, S., & Kempen, B. 2017. Soil-Grids 250 m: Global gridded soil information based on machine learning. *PLoS One*, **12**, 1–40.
- Kašuba, V., Milić, M., Rozgaj, R., Kopjar, N., Mladinić, M., Žunec, S., Vrdoljak, A., Pavičić, I., Čermak, A., Pizent, A., Lovaković, B., & Želježić, D. 2017. Effects of low doses of glyphosate on DNA damage, cell proliferation and oxidative stress in the HepG2 cell line. *Environmental Science and Pollution Research*, **24**, 19267–19281.
- la Cecilia, D., & Maggi, F. 2018. Analysis of glyphosate degradation in a soil microcosm. *Environmental Pollution*, **233**, 201 – 207.
- la Cecilia, D., Porta, G.M., Tang, F.H.M., Riva, M., & Maggi, F. 2020. Probabilistic indicators for soil and groundwater contamination risk assessment. *Under Review in Ecological Indicators*.
- la Cecilia, Daniele, Tang, Fiona HM, Coleman, Nicholas V, Conoley, Chris, Vervoort, R Willem,

- & Maggi, Federico. 2018. Glyphosate dispersion, degradation, and aquifer contamination in vineyards and wheat fields in the Po Valley, Italy. *Water research*, **146**, 37–54.
- Laitinen, Pirkko, Rämö, Sari, Nikunen, Unto, Jauhiainen, Lauri, Siimes, Katri, & Turtola, Eila. 2009. Glyphosate and phosphorus leaching and residues in boreal sandy soil. *Plant and soil*, **323**(1-2), 267–283.
- Landrigan, P. J., Fuller, R., Acosta, N.J.R, Adeyi, O., Arnold, R., Basu, N., BaldÃ©, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breyse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A. Murray, C.J.L., Ndahimananjara, J.D., Perera, F., PotoÄ•nik, J., Preker, A.S., Ramesh, J., Rockstrm, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., van Schayck, O.C.P., Yadama, G.N., Yumkella, K., & Zhong, M. 2018. The Lancet Commission on Pollution and Health. *The Lancet*, **391**, 462–512.
- Lewis, Kathleen A, Tzilivakis, John, Warner, Douglas J, & Green, Andrew. 2016. An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment: An International Journal*, **22**(4), 1050–1064.
- Li, H., Joshi, S. R., & Jaisi, D. P. 2016. Degradation and isotope source tracking of glyphosate and aminomethylphosphonic acid. *J. Agric. Food Chem.*, **64**, 529–538.
- Maggi, F., Gu, C., Riley, W. J., Hornberger, G. M., Venterea, R. T., Xu, T., Spycher, N., Steefel, C., Miller, N. L., & Oldenburg, C. M. 2008. A mechanistic treatment of the dominant soil nitrogen cycling processes: Model development, testing, and application. *J Geophys Res Biogeosci*, **113**(G2), 1–13.
- Maggi, Federico. 2019a. BRTSim, a general-purpose computational solver for hydrological, biogeochemical, and ecosystem dynamics. *arXiv, preprint arXiv:1903.07015*.
- Maggi, Federico. 2019b. BRTSim v4.0 Release a, Bio-Reactive transport Simulator; A general-purpose multiphase and multispecies computational solver for biogeochemical reaction-advection-dispersion processes in porous and non-porous media. *User Manual and Technical Guide*, pp. 79.
- Maggi, Federico, Tang, Fiona HM, la Cecilia, Daniele, & McBratney, Alexander. 2019. PEST-CHEMGRIDS, global gridded maps of the top 20 crop-specific pesticide

- application rates from 2015 to 2025. *Scientific data*, **6**(1), 1–20.
- McBratney, Alex, Field, Damien J, & Koch, Andrea. 2014. The dimensions of soil security. *Geoderma*, **213**, 203–213.
- Menne, Matthew J, Durre, Imke, Vose, Russell S, Gleason, Byron E, & Houston, Tamara G. 2012. An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, **29**(7), 897–910.
- Moechnig, Mike, & Deneke, Darrell. 2011. *Harvest Aid Weed Control in Small Grain. Fact Sheets. Paper 153*. [http://openprairie.sdstate.edu/extension\\_fact/153](http://openprairie.sdstate.edu/extension_fact/153).
- Monfreda, Chad, Ramankutty, Navin, & Foley, Jonathan A. 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global biogeochemical cycles*, **22**(1).
- Napoli, Marco, Marta, Anna Dalla, Zanchi, Camillo A, & Orlandini, Simone. 2016. Transport of glyphosate and aminomethylphosphonic acid under two soil management practices in an Italian vineyard. *Journal of environmental quality*, **45**(5), 1713–1721.
- NOAA/OAR/ESRL PSD, Boulder, Colorado USA. 2019. *CPC Global Unified Precipitation dataset*. <https://www.esrl.noaa.gov/psd/>.
- PAN. 2019. *Pesticide Action Network International Consolidated List of Banned Pesticides. Database accessed on 26.07.2019 at "http://pan-international.org/pan-international-consolidated-list-of-banned-pesticides/"*.
- Paris, P., Pace, E., Presicce, D.P., Maschio, G., Ursino, S., Bisceglie, S., Cornetta, T., Citro, L., Pacifico, R., Giardina, S., Esposito, D., Romoli, D., Floridi, E., & Tornato, A. 2016. *Rapporto nazionale pesticidi nelle acque dati 2013-2014, Edizione 2016, Rapporti 244/2016, ISPRA, Italia, pp 121. ISBN: 9788844807702*.
- Paudel, P., Negusse, A., & Jaisi, D. P. 2015. Birnessite-catalyzed degradation of glyphosate: a mechanistic study aided by kinetics batch studies and NMR spectroscopy. *Soil Sci. Soc. Am. J.*, **79**, 815–825.
- Pelletier, JD, Broxton, PD, Hazenberg, P, Zeng, X, Troch, PA, Niu, G, Williams, ZC, Brunke, MA, & Gochis, D. 2016. Global 1-km gridded thickness of soil, regolith, and sedimentary deposit layers. *ORNL DAAC*.
- Peruzzo, Pablo J, Porta, Atilio A, & Ronco, Alicia E. 2008. Levels of glyphosate in surface waters, sediments and soils associated with direct sowing soybean cultivation in north pampasic

- region of Argentina. *Environmental Pollution*, **156**(1), 61–66.
- Potter, P, Ramankutty, N, Bennett, EM, & Donner, SD. 2011a. Global fertilizer and manure, version 1: nitrogen fertilizer application. *NASA Socioeconomic Data and Applications Center*.
- Potter, P, Ramankutty, N, Bennett, EM, & Donner, SD. 2011b. Global fertilizer and manure, version 1: phosphorus fertilizer application. *Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)*. doi, **10**, H4FQ9TJR.
- Ramankutty, Navin, Evan, Amato T, Monfreda, Chad, & Foley, Jonathan A. 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**(1).
- Sacks, William J, Deryng, Delphine, Foley, Jonathan A, & Ramankutty, Navin. 2010. Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, **19**(5), 607–620.
- Sandvik, B. 2009. World Borders Dataset. *Thematic Mapping*.
- Sidoli, P., Baran, N., & Angulo-Jaramillo, R. 2016. Glyphosate and AMPA adsorption in soils: laboratory experiments and pedotransfer rules. *Environ Sci Pollut Res*, **23**, 5733–5742.
- Silva, Vera, Montanarella, Luca, Jones, Arwyn, Fernández-Ugalde, Oihane, Mol, Hans GJ, Ritsema, Coen J, & Geissen, Violette. 2018. Distribution of glyphosate and aminomethylphosphonic acid (AMPA) in agricultural topsoils of the European Union. *Science of the total environment*, **621**, 1352–1359.
- Simonsen, L., Fomsgaard, I.S., Svensmark, B., & Spliid, N.H. 2008. Fate and availability of glyphosate and AMPA in agricultural soil. *J Environ Sci Heal B*, **43**, 365–375.
- Székács, András, & Darvas, Béla. 2018. Re-registration challenges of glyphosate in the European Union. *Frontiers in Environmental Science*, **6**, 78.
- Tang, Fiona HM, Jeffries, Thomas C, Vervoort, R Willem, Conoley, Chris, Coleman, Nicholas V, & Maggi, Federico. 2019. Microcosm experiments and kinetic modeling of glyphosate biodegradation in soils and sediments. *Science of the Total Environment*, **658**, 105–115.
- Thenkabail, PS, Knox, J, Ozdogan, M, Gumma, MK, Congalton, R, Wu, Z, Milesi, C, Finkral, A, Marshall, M, Mariotto, I, *et al.* 2016. NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security Support Analysis Data (GFSAD) Crop Dominance 2010 Global 1 km V001.

- Trevisan, Marco, Di Guardo, Andrea, & Balderacchi, Matteo. 2009. An environmental indicator to drive sustainable pest management practices. *Environmental Modelling & Software*, **24**(8), 994–1002.
- UNEP. 2001. *Stockholm Convention on Persistent Organic Pollutants*. 2256 UNTS 119; 40 ILM 532.
- USDA, US. 2016. *Sprinkle Irrigation, Chapter 11, Part 623, National Engineering Handbook*. Washington, DC: USDA, Soil Conservation Service.
- Vonberg, D., Vanderborght, J., Cremer, N., Pätz, T., Herbst, M., & Vereecken, H. 2014. 20 years of long-term atrazine monitoring in a shallow aquifer in western Germany. *Water Research*, **50**, 294 – 306.
- WHO. 2010. The WHO recommended classification of pesticides by hazard and guidelines to classification 2009. *iris Institutional Repository for Information Sharing*, 1–78.
- Wilson, Clevo, & Tisdell, Clem. 2001. Why farmers continue to use pesticides despite environmental, health and sustainability costs. *Ecological Economics*, **39**, 449–462.
- Zhang, Y, Pena-Arancibia, J, McVicar, T, Chiew, F, Vaze, J, Zheng, H, & Wang, Y. 2016. *Monthly global observation-driven Penman-Monteith-Leuning (PML) evapotranspiration and components*. v2.
- Zhang, Yonggen, Schaap, Marcel G, & Zha, Yuanyuan. 2018. A High-Resolution Global Map of Soil Hydraulic Properties Produced by a Hierarchical Parameterization of a Physically Based Water Retention Model. *Water Resources Research*, **54**(12), 9774–9790.

**Highlights for**

**The global environmental hazard of glyphosate use**

Federico Maggi, Daniele la Cecilia, Fiona H.M. Tang, Alexander McBratney

1. Glyphosate is the most use herbicide worldwide but is controversial for its potential negative effects on human and environmental health.
2. Currently, there is no global-scale assessment of glyphosate use and environmental contamination risk.
3. We have deployed our newly-developed georeferenced PEST-CHEMGRIDS global scale pesticide use database and our BRTsim model to achieve a space- and time-resolved description of glyphosate dynamics over global croplands.
4. We have quantified natural biodegradation, soil residue accumulation, runoff potential, leaching to below the root zone, and persistence
5. We found low but pervasive contamination hazard in nearly all agroecosystems globally and a few geographic hotspots with mid to high contamination hazard



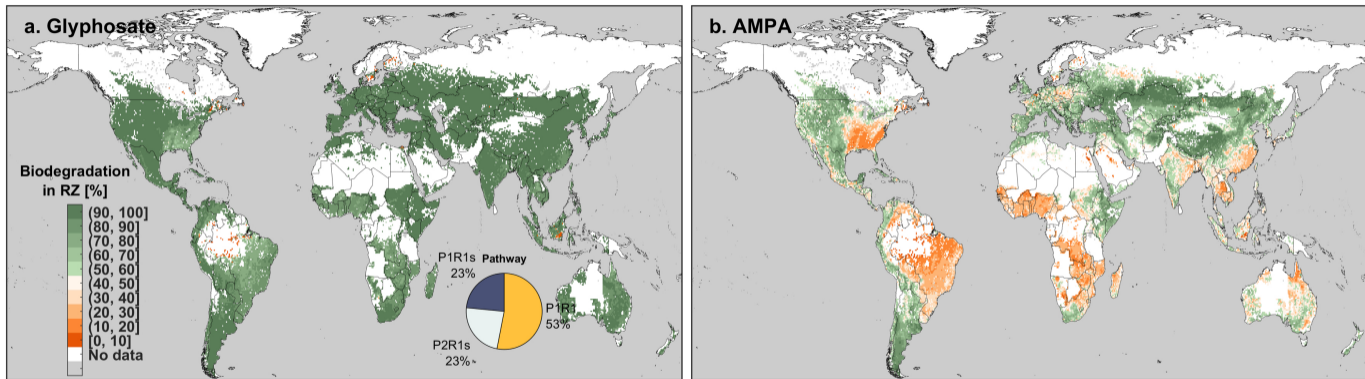


Figure 1

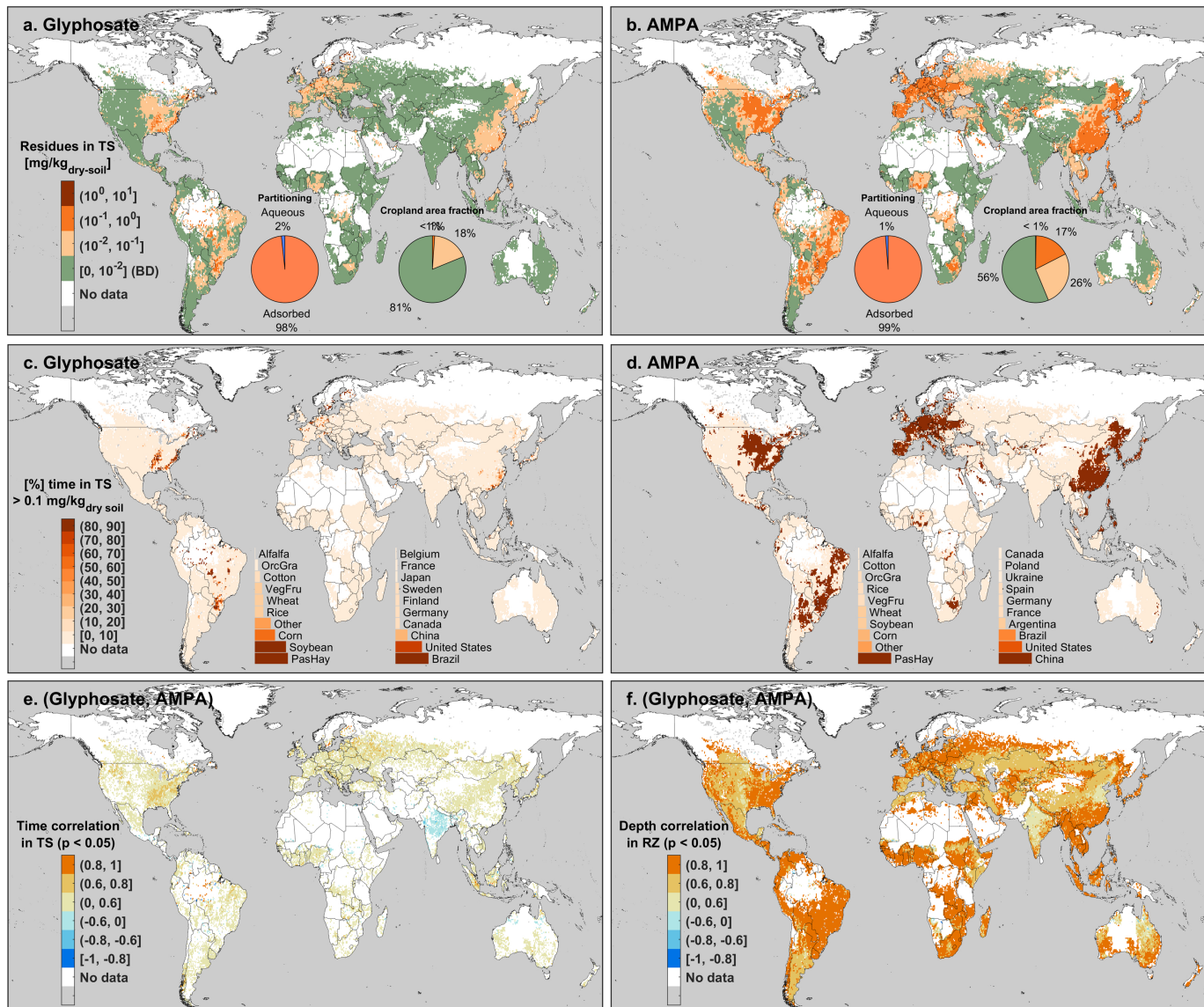


Figure 2

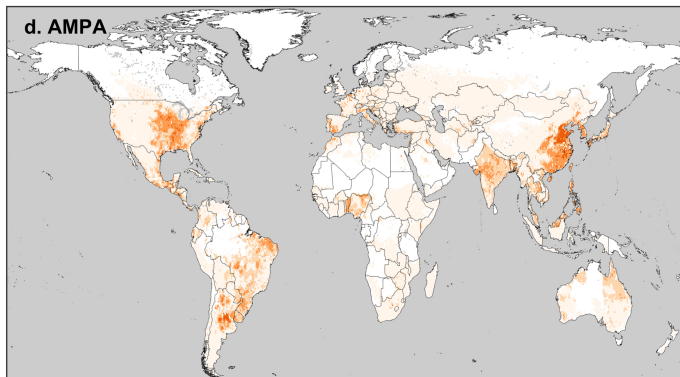
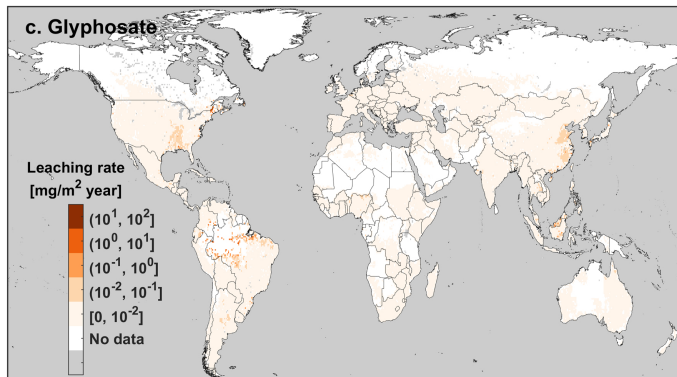
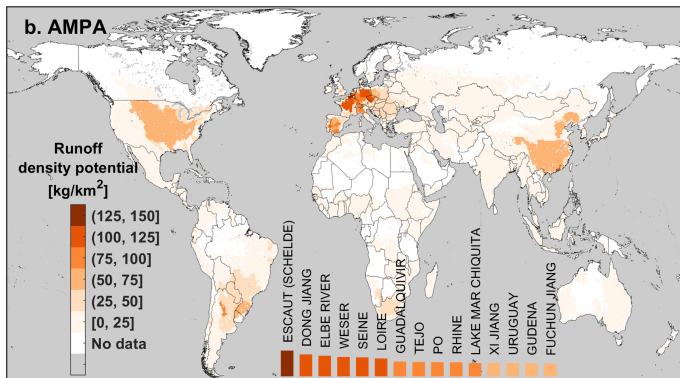
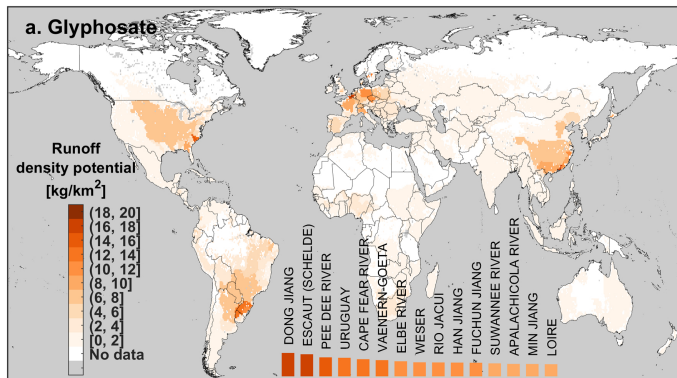


Figure 3

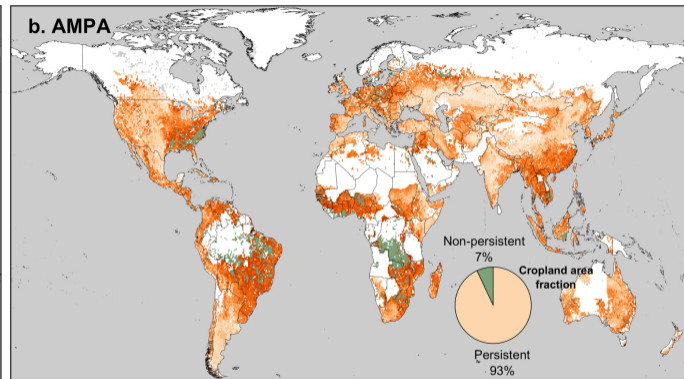
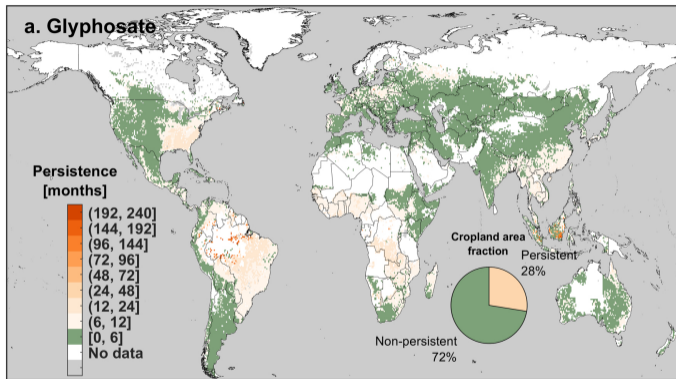


Figure 4

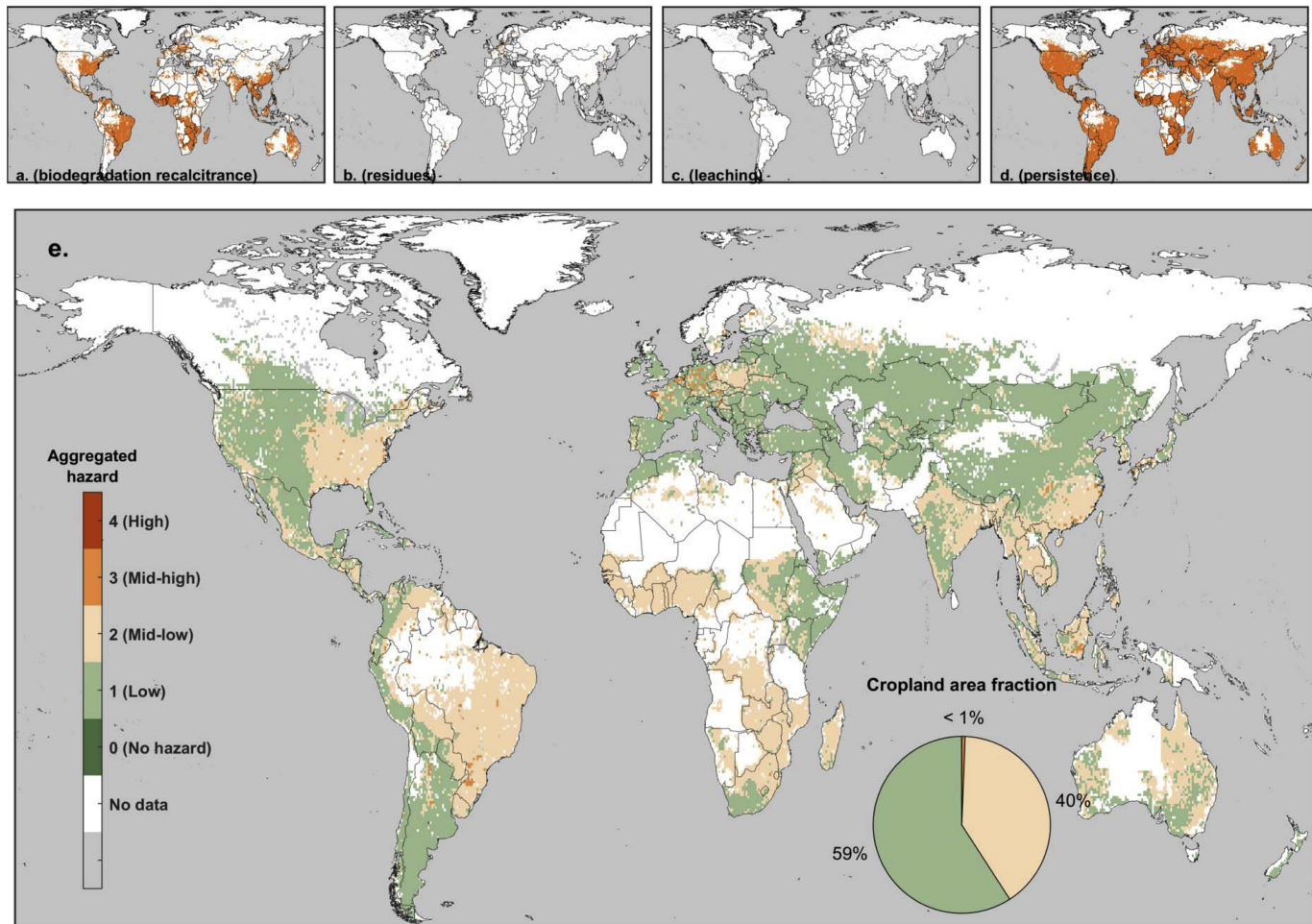


Figure 5

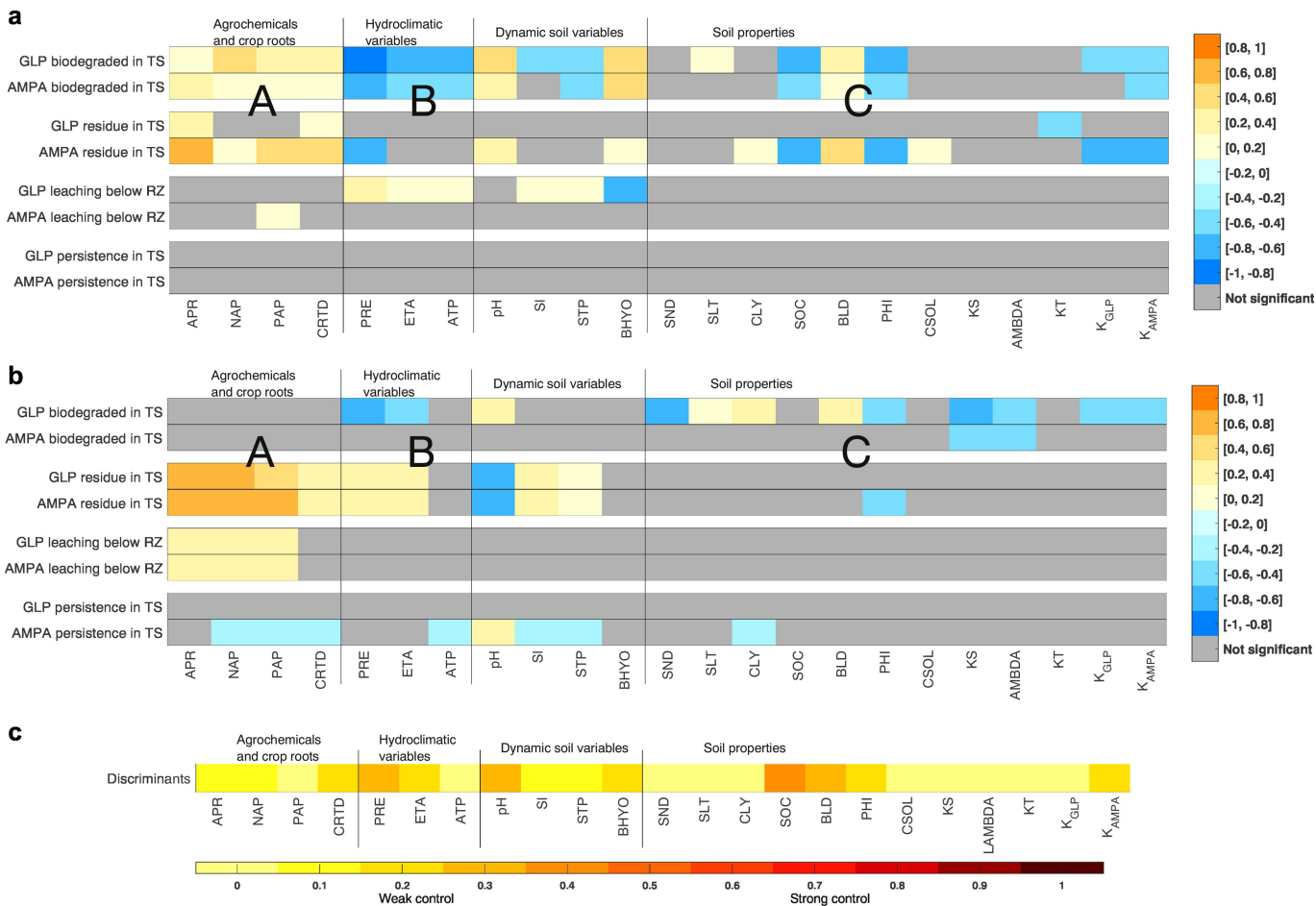


Figure 6