An indicator to assess risks on water and air of pesticide spraying in crop fields

Frédéric Pierlot, Jonathan Marks-Perreau, Emma Soulé, Olivier Keichinger, Carole Bedos, Laetitia Prevost, Paul van Dijk, Christian Bockstaller

An indicator to assess risks on water and air of pesticide spraying in crop fields

Frédéric Pierlot [1,2] ; Jonathan Marks-Perreau [3] ; Emma Soulé [4] ; Olivier Keichinger [5] ; Carole Bedos [6] ; Laetitia Prevost [2] ; Paul Van Dijk[2] ; Christian Bockstaller [4]

[1] LAE, Université de Lorraine, INRAE, F54000 NANCY, France frederic.pierlot@univ-lorraine.fr

[2] Chambre Régionale d'Agriculture Grand Est, Laxou, France

[3] Arvalis, Institut du végétal, France

[4] LAE, Université de Lorraine, INRAE, F68000 COLMAR France

[5] Chercheur indépendant, France

[6] Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, 91120 Palaiseau, France

Highlights:

- Stakeholders need tools to assess environmental risks of pesticide
- I-Phy3 assesses the risk on 3 environmental compartments (air / surface water / ground water)
- Knowledge of recent studies and expertise on pesticides are combined in I-Phy3
- I-Phy3 yielded better validation results than previous versions of the indicator
- A good compromise between ease of use and predictive capability is offered by I-Phy3

Abstract

Stakeholders involved in actions to reduce the use and the impacts on the environment or human health of pesticides need operational tools to assess crop protection strategies in regard to these impacts. I-Phy3 brings together all improvements introduced since the first version of the indicator to better meet user's needs and requirements of integrating processes. I-Phy3 was deeply modified to ensure its predictive quality. I-Phy 3 is structured in three levels of aggregation in form of hierarchical fuzzy decision trees designed with the CONTRA method. At the 1st level, five basic subindicators assess the risk of contamination (RC) for the different transfer pathways involved in surface water, ground water and atmosphere contamination: leaching, runoff, drainage, drift, volatilization. At the 2nd level, RC subindicators are aggregated with a toxicity variable (human or aquatic) in a risk indicator. At the 3rd level, the global indicator I-Phy3 results from the aggregation of three risk indicators for groundwater, surface waters and air. I-Phy3 yielded better validation results than its previous versions. This effort to assess the predictive quality of the indicator should be pursued and completed by a feasibility and utility test by end-users. A subindicator on risk of soil contamination is a gap which remains to fill.

Keywords: Indigo method; risk assessment; I-Phy3; pesticide transfer pathways; water quality; air quality

1. Introduction

The incredible and securing rise of agriculture production since the end of the Second World War, was favored partially by the widespread use of pesticides. Nevertheless, side effects on the environment (Richardson, 1998) and in particular on water quality (Flury et al., 1995; Grung et al., 2015; Lopez et al., 2015; Real et al., 2005), air (Hulin et al., 2021; Lichiheb et al., 2015) and accumulation in soil (Silva et al., 2019; Tang and Maggi, 2021) have been observed. These contaminations lead to exposure of non-target ecosystems as well as populations that may have ecotoxicological and/or toxicological impacts. Consequently, regulations on pesticide use have been continuously reinforced since the Council Directive 79/117/EEC with the Council Directive EC 1107/2009 es and the Pesticide Package 2009/128/CE on sustainable use of pesticides. This was amended by the Commission Directive (EU) 2019/782 introducing a monitoring of risk by the HRI indicator (Methodology for calculating harmonized risk indicators for pesticides under Directive 2009/128/EC, 2021) (European Commission. Statistical Office of the European Union., 2021; (Lykogianni et al., 2021). In all cases, stakeholders involved in actions to reduce the use and impacts of pesticides need operational tools to assess crop protection strategies. The aim of such assessment may be to monitor and to report on the current status of environmental compartments quality, to produce references for the good management of crop protection and to work on innovative systems (Bockstaller et al., 2015).

The need for assessment tools dealing with pesticide issues has led to the development of numerous indicators. The simplest ones rely on the supplied quantities, the Quantity of Active Ingredients (QAI) or the Treatment Frequency Index (TFI) calculating the ratio of applied pesticide to the registered rate (Hossard et al., 2017; Uthes et al., 2019). Although those indicators have been developed to describe the evolution of pesticide use intensity, they are often used as main indicators to address the environmental effects due to pesticide spraying in environmental assessment method (Eckert et al., 2000; Vilain et al., 2008). Pesticide risk indicators (Levitan, 2000) requiring complementary variables such as active ingredient properties, crop management data and pedoclimatic variables are more elaborate and were reviewed by several authors (Feola et al., 2011; Keichinger et al., 2013; Maud et al., 2001; Reus et al., 2002). Some deserve more attention. While the Danish Pestice Load (Kudsk et al., 2018) is based on a quite simple scoring approach of pesticide properties,, other EPRIP2 (Trevisan et al., 2009),, POCER (Vercruysse and Steurbaut, 2002) or SYNOPS (Strassemeyer and Gutsche, 2010), the indicators of the HAIR project (Kruijne et al., 2011)) rest on quantitative equations addressing several factors (pesticide properties, soil climate, etc.). This array of indicators may be explained by the context of use (purpose, environmental covered addressed, scales, means, etc.) and the research of compromise between integration of process and feasibility (Bockstaller et al., 2015) Among those indicators, the I-Phy indicator, in its two versions, assessing risk for ground water, surface water and air, distinguished by its original construction of a fuzzy decision tree combining qualitative and quantitative data. It was developed for an assessment at the field level, first for arable cropping systems and later its was adapted to viticulture (Thiollet-Scholtus and Bockstaller, 2015). Its aim was to help advisers guide farmers in their choice of active ingredients and application methods at limit pressure on the environment. In the validation study of Pierlot et al. (2017), I-Phy1 (first version, van der Werf et Zimmer 1998) came out as one of the “best” indicators for transfer by leaching. Lindahl et Bockstaller (2012) proposed an upgraded version assessing the pesticide transfer by leaching to better take into account preferential transfer (I-Phy2), that was overlooked by I-Phy1. However, its predictive quality was not improved.

Furthermore, assessments of transfer and toxicity of pesticides in I-Phy1 were combined in a calculation of pesticide risk. Yet, stakeholders and advisers working directly with farmers require a separate assessment of transfer to gain insight of water bodies contamination with respect to the European rules for water bodies quality (DCE). This demand was motivated by their need to gain knowledge on water quality in terms of pesticide contamination. Ecotoxicological risk integrating transfer and toxicity remains a matter of interest when stakeholders consider the impact of pesticide use on biocenosis. In both cases, the implementation of assessment tools is aimed at helping them prioritize their actions.

Here is presented I-Phy3 which brings together all improvements introduced since the first version of the indicator to better meet users' needs and requirements of integrating processes to ensure its predictive quality. Based on the knowledge gained on pesticide transfers and risks for the last 20 years, the indicator was restructured to separate transfer assessment from risks on living organisms assessment, new input variables were introduced and most algorithms deeply revised. This article presents the new calculation algorithms for arable crops as well as results of the indicator evaluation for its predictive quality. Besides the design of the indicator and validation results, possibilities for implementing the indicator are discussed.

2. Material and methods

The design of I-Phy3 started with the selection of input variables (section 2.1). Those variables were aggregated into several subindicators and then in a global indicator I-Phy3. The methodology is described in section 2.2 and the structure of the indicator in section 2.3. The predictive quality of the indicator was then assessed by comparison of indicator outputs calculated on sites with measured values of water or air contamination. Section 2.4 presents the study sites used to collect measurements of water (EQUIPE project) or air (Repp'Air project) contaminations while section 2.5 presents the methodology to assess predictive quality of the indicator.

2.1. Selection of input variables

Input variables were selected from a literature review, starting from the work of van der Werf (1996) and on expertise based on experimental data (e.g. see section 2.3.2 for contamination of groundwater).

For pesticides properties, we used the revised Pesticide Properties Database PPDB which is updated since more than 20 years with more reliable value available for each active ingredient (a.i.) and used in many risk assessment studies (Lewis et al., 2016). PPDB can be browsed on a website (PPDB : Pesticide Properties Database, 2020) or purchased in form of an Access database. Values of pesticide properties were extracted from the latter (version 20/05/11).(Tomlin et al., 1995) For the DT50, the value proposed for laboratory rather than the "typical value" proposed by the database resulting from an expert work was selected. This choice resulted also from the calibration work with monitoring data in the Rhine plain (ERMES 2017, Koller et al. 2015) for four active ingredients (see section 2.3.3). For some few active ingredients, values as listed in Supplementary Materials S1 were modified. For some few other a.i., values were taken from other sources because of missing value in PPDB. Overall, 498 active ingredients were integrated in the database.

2.2. Design of indicators

The indicator I-Phy3 as well as all subindicators are calculated for a single active ingredient and then aggregated for a spraying program. They are expressed on a continuous performance scale between 0 (highest risk) and 10 (lowest risk), which allows an operational user of the indicators by non-scientific users. (Bockstaller et al., 2008). (Craheix et al., This scale can be easily inverted between 0 (lowest risk) and 10 (highest risk) for a use of I-Phy3 for specific risk assessment for which high values of indicators are usually associated to high risk".

Fuzzy decision trees were implemented to design the indicator I-Phy and its different subindicators like in the first version of the indicator (Roussel et al., 2000; van der Werf and Zimmer, 1998). Fuzzy decision trees present several advantages (Bockstaller et al., 2017): they rely on linguistic “if then” rules with a transparency, at least semantic through linguistic rules that are easy to understand for non-specialists. They cope with qualitative as well as quantitative heterogeneous information. Furthermore, combining decision trees with fuzzy logic makes it possible to mitigate threshold effects linked to the linguistic “if then” rules when they are Boolean, i.e., consisting of two alternatives yes/no. Fuzzy logic introduces fuzzy subsets to deal with the whole set of intermediate cases.

As shown in the simplified example , (Supplementary Materials S2), I-Phy3 and each subindicator require i) fuzzy subsets for each input variable with threshold values, ii) membership functions to calculate membership degree of each variable to the favourable (low risk) or unfavourable (high risk) subset, iii) decision trees. More details for fuzzy subsets and membership functions are given in Supplementary Materials S3, while decision trees are presented in section 2.3. Most decision trees were designed with the CONTRA method which supports the design of fuzzy decision tree in a transparent way(Bockstaller et al., 2017) were calibrated with a model (see section 2.3.3) or based on one variable (see sections 2.3.5 and 2.3.6). In this latter case, the membership degree to the favourable subset (expressed between 0 and 1) is transformed into an indicator score by multiplying the former by 10.

2.3. Description of the indicator

2.3.1. Overview of the indicator structure and calculation

As shown on Figure 1, I-Phy3 is structured in three levels of aggregation. At the 3^{rd} level, the global indicator I-Phy3 results from the aggregation of three risk indicators tackling three environmental compartments like in version 1 (van der Werf and Zimmer, 1998): 1) groundwater, 2) surface water and 3) air. At the 2^{nd} level, each of this indicator of risk consists in the aggregation of one or two sub indicators assessing the risk of contamination (RC) via specific transfer pathways, with a toxicity variable, Admissible Daily Intake (ADI) or Aquatox (highest toxicity level between fish, daphnia or alqua, for the human health and ecotoxicological impacts respectively, see van der Werf and Zimmer (1998)). At the 1^{st} level, five basic subindicators assess the risk of contamination (RC) namely, a) the groundwater contamination by pesticide leaching ($RCgw_{lea}$), b) surface water contamination by pesticide runoff or drainage transfer ($RCsw_{ru/d}$). Drainage transfer is characterized by an initial vertical transfer, *i.e.* leaching, to an impermeable layer where till drains have been set to evacuate rapidly water excess. This requires a specific calculation (see 3.1.4). c) surface water contamination by pesticide drift ($RCsw_{dr,}$); d) air contamination by pesticide volatilization ($RCair_{vol}$) and e)) air contamination by pesticide drift ($RCair_{dr}$).

The five risks of contamination are calculated for 1 kg of active ingredient (a.i.) and modulated by the actual application rate which results from the calculation of a non-intercepted rate of a.i. by the treated crop (see section 2.3.8): each of these five subindicators is transformed in a risk between 0 (no) and 1 (high) and weighted by the percentage of sprayed area within the field. It is then retransformed in a value of performance according to Equation 1. It was assumed that there is a proportional relation between sprayed area and level of contamination (Melland et al., 2016):

$$RC_{final} = 10\left(1-\left(1-\frac{RC_{100\%}}{10}\right)\right)\frac{\%Area}{100} \qquad \text{(Equation 1)}$$

Where:

RC_{ffinal}: final result for a risk of contamination (see Figure 1)

$RC_{100\%}$: result for a risk of contamination calculated for 100% of sprayed area

%area: percentage of sprayed area within a field.

I-Phy and each subindicator result from the aggregation based on a fuzzy decision tree.

2.3.2. Groundwater contamination ($RCgw_{lea}$)

Two variables of previous versions were kept: the GUS and the leaching potential for which a new calculation method was proposed (Supplementary Materials S4.1.). A variable identified as playing an important role was added: the water status of the soil when pesticides are applied (Pierlot et al., 2017) depending on climate and soil conditions. This variable was assessed through the application period (Supplementary Materials S4.2). Finally, solubility as a fourth variable was added. This property was identified as explaining of the discrepancy between the low transfer risk calculated by previous indicator version and the alarming contamination level of groundwater in Rhine Plain for the nicosulfuron active ingredient (ERMES, 2017). Nicosulfuron is characterized by a very high solubility in water (7500 mg.L-1). The role of solubility in pesticide leaching is confirmed by literature (Elliott et al., 2000). While pesticide properties are directly retrieved from data bases, the two other variables have to be calculated (Supplementary Materials 4.1 and 4.2).

No prominent weight was assigned to the GUS, like the version 1, regarding the aggregation of the four input variables in a fuzzy decision tree (Table 1). This is justified by the fact that in situations of preferential transfer, even active ingredient with favourable GUS like glyphosate can be leached (Vereecken, 2005). We selected weights and modified some decision rules proposed automatically by CONTRA as made possible by this method, to adjust indicators outputs to results of groundwater monitoring of the Rhine Plain for four active ingredients among the most used pesticides for the main crop of the region, maize' (dmat-p, mesotrion, nicolsulfuron, s-metolachlor), (ERMES, 2017; Koller et al., 2015). Table 1 depicts the structure of the indicator.

2.3.3. Surface water contamination by drainage/runoff ($RCsw_{d/r}$)

The calculation of this subindicator, coming from the work of Wohlfart (2008) is based on the runoff potential and a availability variable that is more elaborated than the use of the single DT50 in the first version (van der Werf and Zimmer 1998). The runoff potential depends on slope, on soil properties, texture, crusting sensitivity, hydromorphy, and also on management, tillage and implantation of buffers strip (Supplementary Materials S5). The availability variable is inspired from quantitative indicators like Synops (Strassemeyer and Gutsche, 2010) and Eprip2 (Trevisan et al., 2009). Equation 2 shows the calculation of the availability variable:

$$\mathbf{Avai = r \,.\, e^{-Ln(2)/(DT50.t)}} \qquad \text{(Equation 2)}$$

Where:

Avai : availability of active ingredient

DT50 : soil half-life of active ingredient (days)

t : time in days between date of spraying and date of next runoff event with a default value of 3 for a worst-case situation

r: A reduction coefficient assessing the reduction pesticide amount for runoff when the active ingredient is incorporated (Mickelson et al., 2001). Considering the results of these authors, we proposed a default value of 0.5 when incorporation and 1 when no incorporation.

Table 2 depicts the fuzzy decision tree aggregating those two variables. Outputs of the decision trees were calibrated with help of the PRZMv3.12 model (Carsel et al., 1986) for two rates (Wohlfahrt, 2008) and interpolated for 1 kg. Input variables belong either to the favourable set (F) or to unfavourable set (U).

2.3.4. Particular case of drained plot

It has been clearly demonstrated that drained fields show a high risk of rapid transfer of pesticide to adjacent surface water bodies through subsurface pipes (Brown and van Beinum, 2009). Although the initial transfer process consists in a preferential vertical transfer, final impacted compartment is surface water and not groundwater. This is supported by the variables playing a role in the determinism of pesticide loss in drained field (Brown and van Beinum, 2009), Kd (non-normalized KOC) and DT50 which are both aggregated in the GUS variable main variable for the groundwater subindicator (see section 2.3.1.). To cope with this specificity, was assumed that in drained field, pesticides are mainly transferable to surface water and not groundwater although the transfer process is based initially on vertical leaching like for $RCgw_{lea}$ (see section 2.3.1.) when there is no higher risk by runoff.

Therefore, the following specific decision rules was introduced:

If the field is drained then $RCsw_{d/r}$ = MIN($RCsw_{d/r}$, $RCgw_{lea}$)

If the field is drained and $RCsw_{d/r}$ = $RCgw_{lea}$ then $RCgw_{lea}$ =10 (to avoid double counting of risk)

Where:

$RCsw_{d/r}$: contamination risk for surface water through drainage or runoff (see section 2.3.3)

$RCgw_{lea}$: contamination risk for groundwater through leaching (see section 2.3.2)

2.3.5. Surface water contamination by drift ($RCsw_{dr}$)

While in the previous version of the indicator, coefficients were taken from a table resulting from a collection of expert judgment and measured values, the equations of Trevisan et al. (2009) were used here to calculate spray drift. This spray drift potential is divided by 5 when anti-drift nozzles are used. These spray drift values are transformed in a score between 0 and 10 as described in Supplementary Materials S6.

2.3.6. Air contamination by volatilization ($RCair_{vol}$)

Like for the subindicator *spray drift to surface water* (see previous section), the decision tree used in previous version of the indicator was replaced by quantitative equations calculating pesticide volatilization (expressed in µg/m²/hr) in function of pesticide properties (Woodrow et al., 1997). We added some abatement factors taking into account effect of soil components (mulch, etc.), pesticide properties (penetration ability), and field edge (presence of hedges or trees reducing transfer to outside of field) which is expected to contribute to decrease the transfer risk. The calculation is given by Equation 3 and more details are given in Supplementary Materials S7.

$$Vol_{tot} = (1 - c_{edge}).((1 - c_{form}).(1 - ic).Vol_{sol} + (1 - c_{prod}).ic.Vol_{plant})$$

(Equation 3)

Where:

Vol_{tot}: total volatilization (expressed in µg/m²/hr)

Vol_{sol}: volatilization from soil (expressed in µg/m²/hr), calculated according to Woodrow et al. (1997): $LnVol_{sol}$= 28.335+1.6158.Ln(Pv/(KOC.Sol) with: Pv: pressure vapor (Pa), KOC: soil adsorption coefficient (mg.L^{-1}), Sol: water solubility (mg.L^{-1})

Vol_{plant}: volatilization from plant (expressed in µg/m²/hr), calculated according to (Woodrow et al., 1997): $LnVol_{plant}$= 11.779+0.85543.Ln(Pv) with: Pv; pressure vapor (Pa)

ic: interception of pesticide by crop (see 3.1.8 and Supplementary Materials S8.1)

c_{edge}: abatement coefficient (between 0 and 1) due to field edge, more precisely to the presence of a hedge reducing pesticide transfer to outside of field. Four variables are used to assess it: type of plant (persistent or deciduous), hedge density (number of field sides with a hedge and porosity of hedge), hedge height, spraying month (to assess the presence of leaves or not)

C_{form}: abatement coefficient (between 0 and 1) due to the product formulation

c_{prod}: abatement coefficient (between 0 and 1) due to penetration ability of the product in plant which limits pesticide volatilization and is assessed with pesticide mechanism of action, the octanol water coefficient (LogKow) and the use of an adjuvant to facilitate penetration.

2.3.7. Air contamination by drift ($RCair_{dr}$)

This risk of air contamination by spray drift was not covered by the original version of I-Phy (van der Werf and Zimmer, 1998) but was added to the version for wine growing activity (with a simplified assessment including only the type of sprayer. A more elaborated decision tree taking into account additional relevant variables like speed sprayer, the sprayer height, the use of antidrift nozzle and the air pressure (Pressure) were included in this new version of the indicator according to the study of Bahrouni, Sinfort, et Hamza (2010) (Table 3). The CONTRA method (was used to aggregate the five input variables in a fuzzy decision tree. This indicator is weighted by an abatement coefficient, c_{edge} like for $RCair_{vol}$ (see Equation 3 and Supplementary Materials 7.2).

2.3.8. Integration of the a.i. application rate effect

In I-Phy3, the method proposed by Lindahl and Bockstaller (2012) was used to integrate the effect of the a.i. application rate in the calculation of each contamination subindicator. They proposed to calculate an effective dose of a.i. available for transfer by weighting the initial dose by an interception rate as shown in equation 4:

$Dose_{eff}$ = (1-ic).$Dose_{ini}$ (Equation 4)

$Dose_{eff}$: effective rate of pesticide

ic: interception coefficient (see Supplementary Materials S98.1)

$Dose_{ini}$: initial rate of pesticide

They designed an algorithm which makes possible to reduce (if the dose is higher than 1 kg.ha^{-1}) or increase (if the rate is lower than 1 kg.ha^{-1}) the indicator value obtained for 1 kg/ha^{-1} a.i., in function of this effective dose. More details on the determination of the interception rate and the calculation of the indicator in function of the effective dose are given in Supplementary Materials S 8.

2.3.9 . Risk by compartment

Three subindicators of environmental risk, for ground water (I-Phy_{gw}), surface water (I-Phy_{sw}), and air (I-Phy_{air}) result from the aggregation of the five subindicators of contamination risk (RC) with toxicity variables (see Figure 1). For groundwater, $RCgw_{lea}$ is aggregated with the daily admissible intake (ADI) because this source is for human water supply. For surface water, two subindicators, $RCsw_{r/d}$ and $RCsw_{dr}$ are aggregated with a toxicity variable based on the highest toxicity between Aquatox tackling toxicity for aquatic organisms and ADI for human toxicity (Roussel et al., 2000). Aquatox results from the highest toxicity for fish, daphnia and algae (van der Werf and Zimmer, 1998). For air, two subindicators, $RCair_{vol}$ and $RCair_{dr}$ are aggregated with ADI for human toxicity.

Regarding aggregation rules, a weight of 60 % is given to the contamination risk and 40 % to toxicity for groundwater (Supplementary Materials S9). For surface water and air, 30 % is given to each contamination risk and again 40 % to the toxicity with a small modification in the second case (Table 4). We considered that a situation with a high risk of contamination is more undesirable than a situation of toxic a.i. without any risk of contamination. This may be justified by water potability threshold of 0.1 μg^{-1} applied to all a.i. which leads water manager to focus on water contamination. Furthermore, uncertainty exists on actual toxicity of a.i. which is not well assessed by regulation tests (Centner, 2021) so that a situation with apparently no toxicity may present a risk for human health. Similar decision rules to those for I-Phy_{sw} are set to I-Phy_{air}.

2.3.10. Final aggregated indicator (I-Phy)

The three subindicators of pesticide risk per compartment, ground (I-Phy_{gw})- and surface water (I-Phy_{sw}), air (I-Phy_{air}) are aggregated with the same weight of 33% given to each compartment . However, to limit compensation, the score was reduced to 6 when there was a risk totally unfavourable for one compartment (I-Phy_{gw} or I-Phy_{sw} or I-Phy_{air}=0) and to 2 when two compartments were concerned by a totally unfavourable risk (Table 5). We consider that if there is a risk maximal for one compartment, the value should be clearly under 7. This value 7 is a reference value expressing an acceptable risk for the environment used for the set of indicators of the INDIGO method to which belongs I-Phy (Bockstaller et al., 1997).

2.3.11. Implementation of the indicator

Calculations of the indicator are run on an Excel Sheet with one sheet in which all data on input variables are entered. Each line corresponds to a calculation for one active ingredient. Stable data (e.g. field characteristics) have to be copied from line to line. Users have access to all the detail of calculations with results expressed with a color code (see Figure 3).

2.4. Study sites

Measured data of environmental compartment contamination (groundwater, surface water, air) from several study sites (Figure 2) were used and compared with outputs of I-Phy3 for validation. The sites of the EQUIPE project (see 2.3.1) provided data on the transfer by drainage or runoff to surface water and vertical transfer by leaching. For the transfer to air, the data of the project Repp'Air was used. For each treatment, the I-Phy 3 indicator was calculated with the help of an Excel sheet calculator with the aim of comparing the results with measurement data.

2.4.1. Sites of the Equipe project

The EQUIPE (2014-2017) project aimed to assess the predictive quality of pesticide indicators addressing transfers to surface and ground water. To do so, outputs of 26 indicators (among them Synops, Eprip, I-Phy1, I-Phy2...) and a mechanistic model (MACRO) were compared to measure pesticide transfers at plot levels at four sites with different climate and soil conditions, and transfer pathways (Pierlot et al., 2017). The complete description of the 3 sites and the indicators and model tested is detailed in Supplementary Materials S10) The Jaillière experimental station, located in the Pays de la Loire region (France), is under the influence of an oceanic climate, with a brown hydromorphic clay-textured soil, resulting from alterite shale. This experimentation site consists of 10 agricultural plots of 0.5 to 1 ha each, where water from drainage and runoff (saturation overland flow) are collected separately; ii) The experimental station of the Magneraud, located in the Nouvelle-Aquitaine region (France), is also under the influence of oceanic climate and is composed mainly of clayed and silty limestone soils developed on sand-stone strata characterized by alternating layers of hard limestone and marl. This site is made up of 14 lysimetric plots of 1 m2 surface, with no vertical walls and no soil shuffle; and iii) the Geispitzen experimental station is located in the hills of the lower Sundgau district (Alsace region, France) and has an attenuated oceanic climate. The hills are covered with loess-derived soils of silt loam texture overlying Oligocene molasses and marls. A sloping field (5%) of about 9 ha was divided into 3 bordered fields with measuring flumes and automatic water samplers at the down slope borders just upslope of a ditch drainage catchment runoff.

2.342. Sites of the Repp'Air project

The 7 measurement sites selected for the Repp'Air project came from historical sites monitored by regional Association of Air Quality Survey in association with Chambers of Agriculture in order to have different agricultural systems: arable crops, vineyards, arboriculture, mixed cropping-livestock, and "mixed" sites with different types of crops. Farm practice surveys were conducted during the 3 monitoring campaigns (2017, 2018, 2019) and for each site, to help in the interpretation of local air contamination data. These investigations were conducted over a radius of 1 km around the air sampler installed at each site. The choice of the 1 km radius was a compromise between technical feasibility (particularly in the wine-growing zone, where the number of plots, often smaller than in field crops, is greater in a given area, implying a greater number of farmers) and a surface area in agreement with atmospheric dispersion patterns at the local scale . Such radius value was sufficient to find a correlation between pesticide in precipitation and land use (Grynkiewicz et al., 2001)). Atmospheric samples were collected for a whole week during the spraying period (in average 27 weeks), this for 3 years on the 7 sites concerned by the project, I.e. a total of 567 samples. Pesticide contents of each sample were analyzed in an external accredited laboratory and allowed to quantify a.i. concentration in the atmosphere.

2.5. Evaluation of the predictive quality of I-Phy3

Following Pierlot et al. (2017), two tests were carried out to compare outputs of I-Phy3 subindicators assessing the pesticide transfer to environmental compartments with measured data (RC_{gw} ; RC_{sw}, RC_{air}, see section 2.3.). First, a classical correlation test between indicator outputs and measurements was carried out to calculate a correlation coefficient r, and not the determination coefficient r^2. The significance of the results by calculating the p-value was also tested. Then, we ran a probability test consisting in comparing the rank of indicator outputs and measurements through a contingency table. A similar ranking means that the result of the indicator' calculation appears to be correct while when the indicator rank is lower than the rank of the measured value, it is considered as an underestimation and when it is higher than the rank of the measured value, it is considered as an overestimation. The probability considered in the test is the sum of correct and overestimation. Indeed, I-Phy assesses a potential risk (which can occur or not depending on climate events for example), so the positive result considered in this test are the sum of well-predicted events of transfer and overestimation. (see a theoretical example in Table 6). Pierlot et al. (2017) set the rule that an indicator is considered as acceptable when the probability is higher than 60% and the correct estimation is higher than 40% to avoid considering an indicator whose results would systematically predict a high risk regardless of the context. This general analysis was completed by detailing the proportion of values in each class to assess the distribution of values and to check that results are not only due to one class (*e.g.* the class no risk, no pesticide in water).

For these analyses, regarding water contamination, the following measured variables available in the EQUIPE project were used:

- frequency of exceedance of the threshold of the water quality standard of drinking water: 0.1 µg.L^{-1} (fd0.1)

fd0.1= $\frac{n1_{ijk}}{n_{ijk}}$ (Equation 5)

with $n1_{ijk}$: number of measurements with concentration > 0.1µg.L^{-1} for active ingredient i on plot j at sampling time k; n_{ijk}: total number of measurements for active ingredient i on plot j and sampling time k. The sampling was stopped when no a.i. was detected 3 consecutive weeks and lasted one year maximum after the spraying date (Pierlot et al., 2017)

- cumulated flux of active ingredient in mg/ha (ftotal) during the measurement period

ftotal = $\Sigma(f_{ijk})$ (Equation 6)

with f_{ijk} : flux of active ingredient i on plot j and sampling time k; $f_{ijk} = c_{ijk} . w_{jk}$ with c_{ijk}: concentration of active ingredient i on plot j and sampling time k (µg.L^{-1}) w_{jk}: water flux (drainage or runoff) from plot j during sampling time k (L)

- weighted average concentration on the period in µg/L (CMP)

CMP = $\frac{\Sigma c_{ijk} \times w_{jk}}{\Sigma w_{jk}}$ (Equation 7)

with c_{ijk} : concentration of active ingredient i on plot j and sampling time k (µg.L-^{1})

with c_{ijk} : concentration of active ingredient i on plot j and sampling time k (µg.L-^{1})

with w_{jk} : water flux (drainage or runoff) from plot j during sampling time k (L)

Regarding the assessment of predictive quality for the atmospheric compartment, the correlation between indicator outputs and a value calculated from measurements was analyzed. Those from the REPP'AIR project were atmospheric concentrations in each site. For a given site and year, the pesticide concentration in the sample was considered as resulting from the volatilization and drift after spraying on fields from a buffer of 1km radius around the sampler. Since temporal scales differed between spraying date (day), pesticide concentration (week) and the indicator (year), it was not possible to compare the assessment of volatilization of transfer by I-Phy calculated at field level and raw weekly concentrations. For each a.i. of one site and one year, the weekly concentrations were plotted against the area sprayed with a.i. during the week. The slope of the linear regression ion between concentration and area was derived as a proxy of the measured volatilization risk. This means that for a given sprayed area, the higher the slope, the higher the concentration expressing a higher volatilization. This slope was compared to the mean of the indicator weighted by sprayed area for each a.i. and for the sampling period. In this case, we worked on a limited number of points, so that it was not possible to run a probability test.

3. Results

(Elliott et al., 2000; ERMES, 2017; Lindahl and Bockstaller, 2012; Melland et al., 2016; Roussel et al., 2000; Vereecken, 2005; van der Werf and Zimmer, 1998)(ERMES, 2017; Koller et al., 2015)(Brown and van Beinum, 2009; Buczko and Kuchenbuch, 2007; Carsel et al., 1986; Mickelson et al., 2001; Pierlot et al., 2017; Strassemeyer and Gutsche, 2010; Trevisan et al., 2009; Wohlfart, 2008)
(Bahrouni et al., 2010; Bockstaller et al., 2017; Roussel et al., 2000; Thiollet-Scholtus and Bockstaller, 2015; van der Werf and Zimmer, 1998; Woodrow et al., 1997)

(Centner, 2021; van der Werf and Zimmer, 1998)

3.1. Examples of calculation

I-Phy3 was calculated for 4 to 7 fields of 33 arable farms from the Champagne Crayeuse (East of France) presenting diversified rotations with winter wheat, winter and spring barley, sugar beet, potatoes, winter rapeseed, etc. for the harvest year of 2020. Figure 3 presents three levels of results. From the top, at a first level, results for different level of intensity are shown in for winter barley. They vary between 10and 2.5 for an intensive program with the herbicide chlortoluron. This active ingredient presents a high risk for groundwater as shown at the second level. Finally, explanation can be found at the third level. Chortoluron has very unfavorable property regarding the GUS, the soil is sensitive to leaching and spraying period in autumn are unfavourable because the soil becomes wet.

Tables at second and third levels on Figure 3 were directly taken from the Excel calculator presenting the results of the final indicator and its sub indicators (see Figure 1). A continuous colour code is used to provide information on the level of risk. Results of the sub indicators (e.g. $RCeso_{lea}$) are completed with the membership degree of each variable to the favourable set (F), (see Supplementary 2). When this value was equal to 1, the variable is totally favourable, i.e. it does not present a risk for the environment. This helps identify the ones which influence the calculated risk and by this way can help users to identify levers to improve the indicator and to reduce risk on the environment.
The example shows the ability of the indicator to differentiate crop management with different level of intensity as well as risk level between active ingredients, and the possibility to explain the results.

3.2. Predictive quality of contamination risk subindicators

3.2.1. Predictive quality assessment for the water compartment

Table 7 shows that the highest correlation between measured data in the fourth studied sites (one for each transfer pathway) and the outputs of the risk of contamination indicator for ground or surface water were obtained for the frequency of exceeding the threshold of the water quality standard of drinking water (fd0.1), with value of correlation coefficient close to 0.50 (see Supplementary Materials S12) and even more for $RC_{swr/d}$ at Geispitzen. (r=0.66). Such values of r are close to those found by Pierlot et al; (2017) for indicator with the same degree of complexity. These results are better for three sites or equal for one site (Jaillière runoff) than those of the previous versions of I-Phy. The comparison between either the cumulated flux of active ingredient in $mg.ha^{-1}$ (ftotal) during the measurement period or the weighted average concentration on the period in $\mu g.L^{-1}$ (CMP), and outputs of the indicator yielded lower value of coefficient between 0.08 and 0.35. For the site of Magneraud (leaching) and Geispitzen (hortonian runoff), the new version yielded better value of correlation coefficients than the previous one, which however remain at a lower level than for fd0.1, and much lower than 0.50.

The probability test reveals that the risk of contamination indicator for ground or surface water yielded results meeting the criteria set for the test for fd0.1 and CMP for the fourth studied sites (except for the site of Geispitzen for fd0.1). (Table 8). For ftotal, the only test meeting the criteria set was for the site of Le Magneraud with a probability of 68% and correct estimation of 49%. In comparison with the previous version of I-Phy, the new one obtains better results than I-Phy2 for the three sites of La Jailliere runoff, Le Magneraud and Geispitzen for all the tests, especially for the correct estimations. Compared to I-Phy1, I-Phy3 obtained better results for the sites of Le Magneraud and Geispitzen for all the measured data whereas it surpassed only for CMP for the site of La Jailliere runoff. From the analysis of the cases showing the highest discrepancy (Table 9), it came out that 3 a.i. play a major role: epoxiconazole, diflufenican and isoproturon explain 35 cases out of 50 .

3.2.2. Predictive quality assessment for the air compartment

Figure 4 shows a relatively clear correlation with a r of 0.73 between the slope coefficient of the cumulative fluxes of a.i. and the means of risk of air contamination by volatilization ($RCair_{vol}$), this for only 12 a.i. for which there was enough data on the studied sites. Such a value is satisfying regarding the elaboration degree of the indicator and when they are compared to the coefficient for transfer to water. Nevertheless, unlike the risk of ground or surface water contamination, the outputs of risk of air contamination were not compared directly to contamination measurements at plot scale but to a calculated value derived from measurements at a larger scale.

4. Discussion

4.1 Originality of the I-Phy3 indicator

I-Phy3 can be classified in the same class as I-Phy1 and I-Phy2 in the typology proposed by Pierlot et al. (2017) classifying pesticide indicators assessing transfer risk to water in function of their design. They are calculated with pesticide properties and use data, crop management and field data (soil, slope, etc.). It does not only consist in a simple scoring of variables according to expert opinion or an aggregation separating risk linked to pesticide properties, and risk linked to soil and climate. Variables are integrated according to knowledge on processes, with some calibration procedure for some subindicators. Furthermore, through the decision rules, calculations seem to be more easy to grasp than indicators based on quantitative equations like EPRIP2 (Trevisan et al., 2009), POCER (Vercruysse and Steurbaut, 2002) or SYNOPS (Strassemeyer and Gutsche, 2010). Regarding integration of toxicity variables, those are aggregated in a qualitative way with transparent assumption in I-Phy3, while in EPRIP2, SYNOPS or POCER, a risk ratio (concentration in the environment compartment/concentration threshold for toxicological effect in this compartment) is used resulting in a quantitative assessment. Exposure of living beings to pesticide is assessed with more precision in POCER or in models used in Life Cycle Analysis (Gentil et al., 2020) taking into account behavior of target living beings (*e.g.* for the ingestion exposure pathway) than in the other indicators. Regarding I-Phy3, a sub indicator assessing exposure and effect may be developed in the future by a separated decision tree and aggregated at the second level with risk of contamination (RC) replacing the aggregation of RC with a toxicity variable (Figure 1). A work is ongoing on pesticide effects on human health taking into account variables and knowledge inputs from the POCER indicator (Vercruysse and Steurbaut, 2002) and the more elaborated Browse model (Butler Ellis et al., 2017).

Aggregation procedure using fuzzy decision tree is also very original in comparison with other indicators as pointed out by several authors (Feola et al., 2011; Keichinger et al., 2013; Maud et al., 2001; Reus et al., 2002). This aggregation method presents different advantages like the readability through linguistic rules, the possibility to cope with qualitative and quantitative variables, the mitigation of threshold effect, section 2.2. For I-Phy3, the CONTRA method (Bockstaller et al., 2017) was implemented to design fuzzy decision tree in order to enhance transparency of the aggregation procedure. This confers a supplementary advantage to the aggregation method, while aggregation is often criticized for a lack of transparency. However, calculation of final result for a given decision tree may remain a “black box” without additional information of intermediate calculation (Bockstaller et al., 2017). This problem was partially solved as discussed further in section 4.4.

4.2. Novelties of I-Phy3 compared to the previous versions

The structure of I-Phy3 was totally changed compared to the initial version (van der Werf and Zimmer, 1998) with the addition of a third level making it possible to deliver an assessment of the contamination risk disconnected from the toxicity of the pesticide. Contamination of environmental compartments is of major concern for many stakeholders working on water quality management due to the current drinking water standards based on a concentration threshold of 0.1 $\mu g.L^{-1}$ independently from toxicity. This supplementary level may confer more complexity to the indicator but this might not be a problem (see section 4.4.).

We tried to integrate more processes into the design of I-Phy3 to consolidate the scientific basis. Nevertheless, the metamodelization approach derived from the mechanistic MACRO model and implemented for the groundwater sub indicator in I-Phy2 was left. I-Phy2 version did not yield satisfying results regarding its predictive quality for pesticide leaching (see Table 7). This discrepancy may be due to a parametrization of MACRO which did not deliver better results for 6 out 7 parameter sets in the study of Pierlot et al. (2017). The new ground water subindicator based on a simpler structure than this of I-Phy2 yielded better validation results than I-Phy2 and slightly better results than I-Phy1.
The runoff surface water subindicator had already been improved in the second version (Wohlfahrt, 2008). Besides the interception coefficient used for all indicators (ic, see Equation 4), it is the only subindicator that entails the temporal dimension in an explicit way in the availability variable (see Equation 2). Nevertheless, for most usage, this temporal variable giving the time between spraying and variable is set to 3 days, which is a worst-case value like in other indicators like Synops (Strassemeyer et al., 2017) and Eprip2 (Trevisan et al., 2009). After all, it is still possible in the calculator to change the value and to make the indicator more sensitive to the spraying date and the delay with the transfer event (*i.e.* significant rain). The other subindicators were totally changed with additional information required, especially on spraying conditions for spray drift to air and physical conditions of field margins. It was

assumed that information on spraying conditions and field margins characteristics are easily accessible too.

Effect of tillage and pesticide incorporation were better integrated in the new version of I-Phy by means of a much broader knowledge basis than for the previous version of I-Phy. Now the effect was quantified more precisely than with a rough “expert value”, especially for effect of tillage on runoff (meta-analysis of Elias, Wang, et Jacinthe (2018)). Such meta-analyses would be useful to parametrize the effect of tillage on vertical transfer and the effect of pesticide incorporation, for which some experimental results exist but remain fragmentary.

In I-Phy1, the effect of pesticide dose was assessed separately from the risk on environmental compartment, the latter including the effect of crop interception. Like in I-Phy2, pesticide dose and crop interception were combined since the amount that can be transferred from soil surface to water is not the sprayed dose but depends on the interception by crop canopy. Although a part of this amount intercepted by crop canopy can be washed off, we considered like Rosenbaum et al. (2015) that in good practices conditions, pesticides are not sprayed just before an important rainfall so that this fraction can be neglected in this approach. This integration makes it possible to avoid giving systematically a favorable value to pesticide rate with low application rate like it was the case in I-Phy1. As shown in Lopez et al. (2015) for metsulfuron-methyl as well as in the ERMES monitoring program for nicosulfuron, sulfonyl-urea herbicide sprayed at low rate (less 50 g/ha) are detected in groundwater and sometimes at concentration exceeding quality water standards (Koller et al., 2015). With the new calculation method, even for low a.i. application rates, unfavorable values for risk of contamination subindicators may be found.

4.3 Design of the indicator

The indicator relies on an approach combining a qualitative (decision tree) with a quantitative approach (fuzzy subsets) which present several advantages as pointed out previously. But the outputs of the indicators are not expressed in quantitative physical or ecotoxicological values. In particular, the indicator does not deliver quantitative information on contamination levels in the environmental compartment, so that stakeholders have no information on the exceeding of given standards like this for drinking water. This would require quantitative models which are in most cases complicated to implement due to the type and amount of data required a calibration procedure to carry out carefully to avoid false prediction as pointed out by Pierlot et al. 2017 for MACRO. Another drawback of quantitative models is their reduced scope to one or two environmental compartments. PestLCI 2.0 is an exception by covering the same compartments as I-Phy3 (Dijkman et al., 2012) and providing percentages of emissions from the initial rate in each compartment. However, it does not calculate concentration in the environment and is only for about 100 active ingredients in comparison with about 500 for I-Phy3.

In the assessment of the pesticide transfer pathways by many indicators, climate variables are not directly included although variables like especially rainfall amount plays a significant role in pesticide transfer (Baran et al., 2021). This would require additional data and may complexify calculations. One way would be to integrate them in the leaching and runoff potential variable in function of location and even of the year. This is possible manually in the calculator for dry year; for example, it is easy to change the value into 0 (low potential). But adding an actual value of the year may hide the effect of change of practice which is not the objective of the indicator. Furthermore, intra annual effects of climate are taken into account by the period of application according to the recommendation of Pierlot et al. (2017) for the groundwater subindicator or by the availability variable (see Equation 5 for the surface subindicator). But in case of transfer to air, climate is not included at all, neither for spray drift or volatilization. Wind speed, a major driver for transfer to air (Lavin and Hageman, 2013), remains too difficult to get for each treatment. Introducing such a variable would be an avenue for progress.

I-Phy3 like PestLCI 2.0 did not address contamination of soil by pesticides although several recent studies revealed a “hidden reality” of pesticide in soils (Riedo et al., 2021; Silva et al., 2019), especially for glyphosate (Silva et al., 2018), even in organic farming (Riedo et al., 2021). Furthermore, these last authors found a negative relation between the amount of pesticide residue in soil and microbial biomass and specifically the abundance of arbuscular mycorrhizal fungi, a widespread group of beneficial plant symbionts likewise other parameters of soil biological activities (Wołejko et al., 2020). The lack of a soil subindicator was due to knowledge gaps which these recent studies tend to bridge. Besides the issue of soil pesticide residues in terms of amount and concentration, another aspect concerns the temporal dimension which should be included in future. Although a part of an active ingredient is adsorbed in an irreversible way in the form of non-extractable bound residues (bounsten) another non-negligible part may be released after several months (Suddaby et al., 2016). Furthermore, the cumulative effect of repeated treatments should also be addressed.

Last, the contamination by metabolites released by the degradation of active ingredients should also be assessed (Baran et al., 2021) as it has been pointed out by pioneer studies (Dana W. Kolpin et al., 2000; D.W. Kolpin et al., 2000). But the integration of such an assessment would probably complexify the indicator, exceeding an acceptable level. This would require quantitative knowledge on the nature of the metabolite formed, the percentage of a.i. transformed in this product, properties (KOC, DT50, etc.), etc., data that does not currently exist in databases. except for recently marketed a.i. (Lopez et al., 2015). In any case, some available information on metabolites was added indirectly in the database. For example, in case when an active ingredient is rapidly degraded in its metabolites, (*e.g.* different form in glyphosate acid, iodosulfuron-methyl in metsulfuron-metyl), properties of the metabolite are attributed to a.i. For few pesticides like metazachlore, dimetolachlor (see Supplementary Materials S1), selected values for DT50 appear to be too favourable while they present metabolites susceptible to be transferred to water bodies (Reemtsma et al., 2013). In this case we decided to attribute a more unfavorable value to the DT50 of this active ingredient.

4.4 Predictive quality

Datasets used to assess the predictive quality of the water contamination risk were consequent. This was also the case for the air component in this study with several sites over the country and a 3 years campaign. Nevertheless, as pointed out by Pierlot et al. (2017), the effort should be pursued. Ideally, this work would require a broader combination of soil type, slopes and climatic conditions, and much more new active ingredients.

The simplification performed in the design of I-Phy3 may explain the mixed results in the validation test. Indeed, the compromise between accessibility of data and explanation of mechanical processes leads us to ignore some variables like soil moisture, which is important to explain transfers in pathway like agricultural drainage (Guimont et al., 2005) or wind speed as explained above. It was considered that the additional cost for collecting usable data does not compensate by loss of information for I-Phy3. Furthermore, these data are useful on a time step of a few days when I-Phy3 estimates transfer risk at the scale of the growing season.

The results obtained for I-Phy3 were overall better than those obtained with the previous versions. If the values of correlation coefficient are not always good, the results in the probability test are satisfying, particularly for the frequency of exceeding the threshold of the drinking water quality (fd0.1), particularly important for the stakeholders. The result of the test combines acceptable and overestimation of risk. Thus, it shows the ability of the indicator to assess a potential contamination which could or not be observed, depending on other factors like climate. Unacceptable underestimations were observed for one site at la Jaillière runoff dataset. It was showed that 3 a.i. (isoproturon, diflufenican and epoxiconazole) play a major role in this underestimation. On this site, runoff was mostly to overland flow due to saturation of the soil profile, a process which may be poorly covered by the indicator.

4.5 Utilization

The design of the I-Phy 3 indicator was not intended to be used directly by farmers themselves, although this may be possible if farmers have time and support to interpret the results. It rather targets advisers trained by scientists who may deliver interpreted results to feed the recommendations to farmers in addition to technical advice. The statement of Box (1976): “all models are wrong but some are useful”, may be applied to I-Phy3. In spite of its mixed results regarding its predictive quality, it may be used beyond the simple results obtained by the indicator. Indeed, it makes possible initiation of discussion on pesticide use strategies with farmers integrating environmental aspects provided by the indicator and other aspects (economic, management of weed and pest resistance, …) not considered by the indicators. Furthermore, if it does not give a precise value , it provides a positioning to some threshold, like for instance to classify results in three classes (acceptable/mixed/unacceptable). But in this case, we reintroduce a threshold effect that was avoided by the design method (see section 2.2.). This is also the case with quantitative mechanistic models as PEARL (Tiktak et al., 2012) which deliver continuous values but can be presented in class when they are mapped for stakeholders.

For the council advice, as stated by Bockstaller et al. (2008), both aggregated indicators and non-aggregated sub-indicators have to be used in parallel. The aggregated I-Phy indicator may be associated with other sustainability indicators in a global assessment when there is a need to reduce the number of indicators. In this case, non-aggregated subindicators should complete and explain the global value in the analysis. The aggregated I-Phy indicator may also serve to rank pesticides when all the environmental issues are considered by stakeholders while a non-aggregated subindicators, especially risk of contamination indicators may be implemented by stakeholders dealing with one environmental compartment (e.g. groundwater).

Currently the indicator is calculated with the help of an Excel sheet calculator, which facilitates its implementation because the software is a basic software on most of the users' offices. The tool is transparent for the user, who can see all the decision rules. The problem is that it may also lead to misuse because the user can potentially and accidentally change the decisions rules. This can be solved by protection of some cells. However, misuse also occurs in copying and pasting lines, deleting formulas in the cell. To mitigate the problem, we can imagine a further online-version of the tool with data fields to be completed and calculations of the different sub-indicators provided. In any case, the tool should remain transparent and not only provide calculation results but also the intermediate values to interpret them and identify the variable(s) which plays the major role in the determination of the risk.

Last but not least, the I-Phy indicator is designed to predict the risk of pesticide transfer at field level. Some of the stakeholders need to assess the risk of transfer at a higher level like a watershed or an administrative region (i.e. the managers of the Water Agencies in France or the advisers in charge of water catchments). Some previous research work ((Wohlfahrt et al., 2010) shows that the contribution of the different plots to the watershed depends on the size of the plot in comparison to the size of the watershed and the distance and the connectivity between the plot and the hydrological network. To simplify the use of the indicator at watershed scale and because they have no precise information about the real water flows in the watershed, advisers consider only the size of the plot in comparison to the size of the watershed. Another approach is to map the distribution of the indicator values set in small classes like a 3 levels scale as traffic lights without any spatial aggregation, as discussed at the beginning of this section.

5. Conclusion

This new version of the I-Phy indicator and its subindicators provides major changes in comparison with previous versions, to better integrate processes of transfer of the pesticides to the environmental compartment. The separation between contamination and toxicity, as well as the transfer pathways are some examples to meet requirements of potential users like adviser or stakeholders. For the air compartment, a spray drift subindicator was added but like previous versions, whereas soil is still not addressed by this indicator, partially due to knowledge gaps. With the increasing focus on this compartment in publications, this gap is about to be filled and an addition of a new subindicator will be possible. While the effect of degradation products of the a.i. is covered by I-Phy3 in an indirect way, their risks should be covered in an explicit way to meet social concern for this issue. Last, the effort to assess the predictive quality of the indicator should be pursued and should be completed by a feasibility and utility test among end-users. This new version should still be confronted with other datasets including recent active ingredients with the aim of improving the predictive quality.

Acknowledgment

Thanks: The authors wish to express their thanks to Cyril Pallarès and Eve Chetien from ATMO Grand Est and to Arnaud Jouart, Natacha Kozack and Johanna Zaffagni from the Chambre Régionale D'Agriculture Grand Est.

Funding: the authors will thank the funders of the different project which were used to elaborate I-Phy3: Equipe (ONEMA, Ecophyto) Repp'Air (ADEME, CasDAR),)

References:

Bahrouni H, Sinfort C, Hamza E. An approach for pesticide loss estimation adapted to field crops in Mediterranean conditions. vol. Sustainable biosystems for bio-engineering. Savoie,P., Villeneuve,J.,Morisette,R.(Eds), Quebec: 2010, p. 1.

Baran N, Surdyk N, Auterives C. Pesticides in groundwater at a national scale (France): Impact of regulations, molecular properties, uses, hydrogeology and climatic conditions. Sci Total Environ 2021;791:148137. https://doi.org/10.1016/j.scitotenv.2021.148137.

Bockstaller C, Beauchet S, Manneville V, Amiaud B, Botreau R. A tool to design fuzzy decision trees for sustainability assessment. Environ Model Softw 2017;97:130–44. https://doi.org/10.1016/j.envsoft.2017.07.011.

Bockstaller C, Feschet P, Angevin F. Issues in evaluating sustainability of farming systems with indicators. OCL 2015;22:D102. https://doi.org/10.1051/ocl/2014052.

Bockstaller C, Guichard L, Makowski D, Aveline A, Girardin P, Plantureux S. Agri-environmental indicators to assess cropping and farming systems. A review. Agron Sustain Dev 2008;28:139–49. https://doi.org/10.1051/agro:2007052.

Box GEP. Science and Statistics. J Am Stat Assoc 1976;71:791–9. https://doi.org/10.1080/01621459.1976.10480949.

Brown CD, van Beinum W. Pesticide transport via sub-surface drains in Europe. Environ Pollut 2009;157:3314–24. https://doi.org/10.1016/j.envpol.2009.06.029.

Buczko U, Kuchenbuch RO. Phosphorus indices as risk-assessment tools in the U.S.A. and Europe—a review. J Plant Nutr Soil Sci 2007;170:445–60. https://doi.org/10.1002/jpln.200725134.

Butler Ellis MC, van de Zande JC, van den Berg F, Kennedy MC, O'Sullivan CM, Jacobs CM, et al. The BROWSE model for predicting exposures of residents and bystanders to agricultural use of plant protection products: An overview. Biosyst Eng 2017;154:92–104. https://doi.org/10.1016/j.biosystemseng.2016.08.017.

Carsel RF, Nixon WB, Ballantlne LG. Comparison of pesticide root zone model predictions with observed concentrations for the tobacco pesticide metalaxyl in unsaturated zone soils. Environ Toxicol Chem 1986;5:345–53. https://doi.org/10.1002/etc.5620050403.

Centner TJ. Pesticide Usage Is Compromising People's Health in the United States: Ideas for Reducing Damages. Agriculture 2021;11:486. https://doi.org/10.3390/agriculture11060486.

Craheix D, Bergez J-E, Angevin F, Bockstaller C, Bohanec M, Colomb B, et al. Guidelines to design models assessing agricultural sustainability, based upon feedbacks from the DEXi decision support system. Agron Sustain Dev 2015;35:1431–47. https://doi.org/10.1007/s13593-015-0315-0.

Dijkman TJ, Birkved M, Hauschild MZ. PestLCI 2.0: a second generation model for estimating emissions of pesticides from arable land in LCA. Int J Life Cycle Assess 2012;17:973–86. https://doi.org/10.1007/s11367-012-0439-2.

Eckert H, Breitschuh G, Sauerbeck DR. Criteria and standards for sustainable agriculture. J Plant Nutr Soil Sci 2000;163:337–51. https://doi.org/10.1002/1522-2624(200008)163:4<337::AID-JPLN337>3.0.CO;2-3.

Elias D, Wang L, Jacinthe P-A. A meta-analysis of pesticide loss in runoff under conventional tillage and no-till management. Environ Monit Assess 2018;190:79. https://doi.org/10.1007/s10661-017-6441-1.

Elliott JA, Cessna AJ, Nicholaichuk W, Tollefson LC. Leaching Rates and Preferential Flow of Selected Herbicides through Tilled and Untilled Soil. J Environ Qual 2000;29:1650–6. https://doi.org/10.2134/jeq2000.00472425002900050036x.

ERMES. Etat de la nappe phréatique d'Alsace et des aquifères du Sundgau. Premiers résultats sur les nitrates et les pesticides 2016. ERMES; 2017.

European Commission. Statistical Office of the European Union. Methodology for calculating harmonised risk indicators for pesticides under Directive 2009/128/EC: 2021 edition. LU: Publications Office; 2021.

Feola G, Rahn E, Binder CR. Suitability of pesticide risk indicators for Less Developed Countries: A comparison. Agric Ecosyst Environ 2011;142:238–45. https://doi.org/10.1016/j.agee.2011.05.014.

Flury M, Leuenberger J, Studer B, Flühler H. Transport of Anions and Herbicides in a Loamy and a Sandy Field Soil. Water Resour Res 1995;31:823–35. https://doi.org/10.1029/94WR02852.

Gentil C, Basset-Mens C, Manteaux S, Mottes C, Maillard E, Biard Y, et al. Coupling pesticide emission and toxicity characterization models for LCA: Application to open-field tomato production in Martinique. J Clean Prod 2020;277:124099. https://doi.org/10.1016/j.jclepro.2020.124099.

Grung M, Lin Y, Zhang H, Steen AO, Huang J, Zhang G, et al. Pesticide levels and environmental risk in aquatic environments in China — A review. Environ Int 2015;81:87–97. https://doi.org/10.1016/j.envint.2015.04.013.

Grynkiewicz M, Polkowska Ż, Górecki T, Namieśnik J. Pesticides in precipitation in the Gdańsk region (Poland). Chemosphere 2001;43:303–12. https://doi.org/10.1016/S0045-6535(00)00130-2.

Guimont S, Perrin-ganier C, Real B, Schiavon M. Effects of soil moisture and treatment volume on bentazon mobility in soil. Agron Sustain Dev 2005;25:323–9. https://doi.org/10.1051/agro:2005012.

Hossard L, Guichard L, Pelosi C, Makowski D. Lack of evidence for a decrease in synthetic pesticide use on the main arable crops in France. Sci Total Environ 2017;575:152–61. https://doi.org/10.1016/j.scitotenv.2016.10.008.

Hulin M, Leroux C, Mathieu A, Gouzy A, Berthet A, Boivin A, et al. Monitoring of pesticides in ambient air: Prioritization of substances. Sci Total Environ 2021;753:141722. https://doi.org/10.1016/j.scitotenv.2020.141722.

Keichinger O, Benoit P, Boivin A, Bourrain X, Briand O, Chabert A, et al. GUIDE : développement d'un outil d'aide à la sélection d'indicateurs de risques liés à la présence des produits

phytopharmaceutiques dans les milieux aquatiques - Mise au point, applications et perspectives. Innov Agron 2013:1–13.
Koller R, Martin L, Rapp O. Observatoire des fuites de Phytosanitaires sous parcelles agricoles en Alsace: Synthèse pluriannuelle 2006-2014 et confrontation aux résultats d'ERMES Alsace. ARAA; 2015.
Kolpin Dana W., Barbash JE, Gilliom RJ. Pesticides in Ground Water of the United States, 1992-1996. Ground Water 2000;38:858–63. https://doi.org/10.1111/j.1745-6584.2000.tb00684.x.
Kolpin D.W., Thurman EM, Linhart SM. Finding minimal herbicide concentrations in ground water? Try looking for their degradates. Sci Total Environ 2000;248:115–22. https://doi.org/10.1016/S0048-9697(99)00535-5.
Kruijne R, Deneer J, Lahr J, Vlaming J. Hair 2010 Documentation: Calculating risk indicators related to agricultural use of pesticides within the European Union. 2011.
Kudsk P, Jørgensen LN, Ørum JE. Pesticide Load—A new Danish pesticide risk indicator with multiple applications. Land Use Policy 2018;70:384–93. https://doi.org/10.1016/j.landusepol.2017.11.010.
Lavin KS, Hageman KJ. Contributions of Long-Range and Regional Atmospheric Transport on Pesticide Concentrations along a Transect Crossing a Mountain Divide. Environ Sci Technol 2013:130109121249000. https://doi.org/10.1021/es304497e.
Levitan L. "How to" and "why": assessing the enviro-social impacts of pesticides. Crop Prot 2000;19:629–36. https://doi.org/10.1016/S0261-2194(00)00083-1.
Lewis KA, Tzilivakis J, Warner DJ, Green A. An international database for pesticide risk assessments and management. Hum Ecol Risk Assess Int J 2016;22:1050–64. https://doi.org/10.1080/10807039.2015.1133242.
Lichiheb N, Bedos C, Personne E, Barriuso E. Synthèse des connaissances sur le transfert des pesticides vers l'atmosphère par volatilisation depuis les plantes. Pollut Atmos Clim Santé Société 2015.
Lindahl AML, Bockstaller C. An indicator of pesticide leaching risk to groundwater. Ecol Indic 2012;23:95–108. https://doi.org/10.1016/j.ecolind.2012.03.014.
Lopez B, Ollivier P, Togola A, Baran N, Ghestem J-P. Screening of French groundwater for regulated and emerging contaminants. Sci Total Environ 2015;518–519:562–73. https://doi.org/10.1016/j.scitotenv.2015.01.110.
Lykogianni M, Bempelou E, Karamaouna F, Aliferis KA. Do pesticides promote or hinder sustainability in agriculture? The challenge of sustainable use of pesticides in modern agriculture. Sci Total Environ 2021;795:148625. https://doi.org/10.1016/j.scitotenv.2021.148625.
Maud J, Edwards-Jones G, Quin F. Comparative evaluation of pesticide risk indices for policy development and assessment in the United Kingdom. Agric Ecosyst Environ 2001;86:59–73. https://doi.org/10.1016/S0167-8809(00)00258-9.
Melland AR, Silburn DM, McHugh AD, Fillols E, Rojas-Ponce S, Baillie C, et al. Spot Spraying Reduces Herbicide Concentrations in Runoff. J Agric Food Chem 2016;64:4009–20. https://doi.org/10.1021/acs.jafc.5b03688.
Methodology for calculating harmonised risk indicators for pesticides under Directive 2009/128/EC: 2021 edition. 2021 edition. Luxembourg: Publications Office of the European Union; 2021.
Mickelson SK, Boyd P, Baker JL, Ahmed SI. Tillage and herbicide incorporation effects on residue cover, runoff, erosion, and herbicide loss. Soil Tillage Res 2001;60:55–66. https://doi.org/10.1016/S0167-1987(01)00170-2.
Pierlot F. Evaluation des impacts des produits phytosanitaires sur la qualité de l'eau à l'aide d'indicateurs : amélioration et validation d'un indicateur de risque de contamination, application de la parcelle au bassin versant. 2017.
Pierlot F, Marks-Perreau J, Real B, Carluer N, Constant T, Lioeddine A, et al. Predictive quality of 26 pesticide risk indicators and one flow model: A multisite assessment for water contamination. Sci Total Environ 2017;605–606:655–65. https://doi.org/10.1016/j.scitotenv.2017.06.112.
PPDB : Pesticide Properties Database. 2020 : https://sitem.herts.ac.uk/aeru/iupac/

Real B, Dutertre A, Eschenbrenner G, Bonnifet J-P, Lasserre D. Résultats de 10 campagnes d'expérimentation : les transferts de produits phytosanitaires vers les eaux varient selon les types de sol. Perspect Agric 2005:20–4.
Reemtsma T, Alder L, Banasiak U. Emerging pesticide metabolites in groundwater and surface water as determined by the application of a multimethod for 150 pesticide metabolites. Water Res 2013;47:5535–45. https://doi.org/10.1016/j.watres.2013.06.031.
Reus J, Leendertse P, Bockstaller C, Fomsgaard I, Gutsche V, Lewis K, et al. Comparison and evaluation of eight pesticide environmental risk indicators developed in Europe and recommendations for future use. Agric Ecosyst Environ 2002;90:177–87. https://doi.org/10.1016/S0167-8809(01)00197-9.
Richardson M. Pesticides - friend or foe? Water Sci Technol 1998;37:19–25. https://doi.org/10.1016/S0273-1223(98)00257-1.
Riedo J, Wettstein FE, Rösch A, Herzog C, Banerjee S, Büchi L, et al. Widespread Occurrence of Pesticides in Organically Managed Agricultural Soils—the Ghost of a Conventional Agricultural Past? Environ Sci Technol 2021;55:2919–28. https://doi.org/10.1021/acs.est.0c06405.
Rosenbaum RK, Anton A, Bengoa X, Bjørn A, Brain R, Bulle C, et al. The Glasgow consensus on the delineation between pesticide emission inventory and impact assessment for LCA. Int J Life Cycle Assess 2015;20:765–76. https://doi.org/10.1007/s11367-015-0871-1.
Roussel O, Cavelier A, Werf HMG van der. Adaptation and use of a fuzzy expert system to assess the environmental effect of pesticides applied to field crops. Agric Ecosyst Environ 2000;80:143–58. https://doi.org/10.1016/S0167-8809(00)00142-0.
Sadok W, Angevin F, Bergez J-E, Bockstaller C, Colomb B, Guichard L, et al. MASC, a qualitative multi-attribute decision model for ex ante assessment of the sustainability of cropping systems. Agron Sustain Dev 2009;29:447–61. https://doi.org/10.1051/agro/2009006.
Silva V, Mol HGJ, Zomer P, Tienstra M, Ritsema CJ, Geissen V. Pesticide residues in European agricultural soils – A hidden reality unfolded. Sci Total Environ 2019;653:1532–45. https://doi.org/10.1016/j.scitotenv.2018.10.441.
Silva V, Montanarella L, Jones A, Fernández-Ugalde O, Mol HGJ, Ritsema CJ, et al. Distribution of glyphosate and aminomethylphosphonic acid (AMPA) in agricultural topsoils of the European Union. Sci Total Environ 2018;621:1352–9. https://doi.org/10.1016/j.scitotenv.2017.10.093.
Strassemeyer J, Daehmlow D, Dominic AR, Lorenz S, Golla B. SYNOPS-WEB, an online tool for environmental risk assessment to evaluate pesticide strategies on field level. Crop Prot 2017;97:28–44. https://doi.org/10.1016/j.cropro.2016.11.036.
Strassemeyer J, Gutsche V. The approach of the German pesticide risk indicator SYNOPS in frame of the National Action Plan for Sustainable Use of Pesticides. OECD; 2010.
Suddaby LA, Beulke S, van Beinum W, Oliver RG, Kuet S, Brown CD. Long-term experiments to investigate irreversibility in sorption of pesticides to soil. Chemosphere 2016;162:40–7. https://doi.org/10.1016/j.chemosphere.2016.07.062.
Tang FHM, Maggi F. Pesticide mixtures in soil: a global outlook. Environ Res Lett 2021. https://doi.org/10.1088/1748-9326/abe5d6.
Thiollet-Scholtus M, Bockstaller C. Using indicators to assess the environmental impacts of wine growing activity: The INDIGO® method. Eur J Agron 2015;62:13–25. https://doi.org/10.1016/j.eja.2014.09.001.
Tiktak A, Hendriks RF, Boesten JJ. Simulation of movement of pesticides towards drains with a preferential flow version of PEARL. Pest Manag Sci 2012;68:290–302. https://doi.org/10.1002/ps.2262.
Tomlin CDS, British Crop Protection Council, Royal Society of Chemistry, editors. The pesticide manual: incorporating The Agrochemicals handbook ; a world compendium. 10. ed., reprinted. Farnham: British Crop Protection Council [u.a.]; 1995.
Trevisan M, Di Guardo A, Balderacchi M. An environmental indicator to drive sustainable pest management practices. Environ Model Softw 2009;24:994–1002. https://doi.org/10.1016/j.envsoft.2008.12.008.

Uthes S, Heyer I, Kaiser A, Zander P, Bockstaller C, Desjeux Y, et al. Costs, quantity and toxicity: Comparison of pesticide indicators collected from FADN farms in four EU-countries. Ecol Indic 2019;104:695–703. https://doi.org/10.1016/j.ecolind.2019.05.028.
Vercruysse F, Steurbaut W. POCER, the pesticide occupational and environmental risk indicator. Crop Prot 2002;21:307–15. https://doi.org/10.1016/S0261-2194(01)00102-8.
Vereecken H. Mobility and leaching of glyphosate: a review. Pest Manag Sci 2005;61:1139–51. https://doi.org/10.1002/ps.1122.
Vilain L, Boisset K, Girardin P, Guillaumin A, Mouchet C, Viaux P, et al. La méthode IDEA: indicateurs de durabilité des exploitations agricoles : guide d'utilisation. Dijon: Éducagri éd.; 2008.
van der Werf HMG. Assessing the impact of pesticides on the environment. Agric Ecosyst Environ 1996;60:81–96. https://doi.org/10.1016/S0167-8809(96)01096-1.
van der Werf HMG, Zimmer C. An indicator of pesticide environmental impact based on a fuzzy expert system. Chemosphere 1998;36:2225–49. https://doi.org/10.1016/S0045-6535(97)10194-1.
Wohlfahrt J, Colin F, Assaghir Z, Bockstaller C. Assessing the impact of the spatial arrangement of agricultural practices on pesticide runoff in small catchments: Combining hydrological modeling and supervised learning. Ecol Indic 2010;10:826–39. https://doi.org/10.1016/j.ecolind.2010.01.004.
Wohlfart J. Développement d'un indicateur d'exposition des eaux de surface aux pertes de pesticides à l'échelle du bassin versant. Thèse. INPL, 2008.
Wołejko E, Jabłońska-Trypuć A, Wydro U, Butarewicz A, Łozowicka B. Soil biological activity as an indicator of soil pollution with pesticides – A review. Appl Soil Ecol 2020;147:103356. https://doi.org/10.1016/j.apsoil.2019.09.006.
Woodrow JE, Seiber JN, Baker LW. Correlation Techniques for Estimating Pesticide Volatilization Flux and Downwind Concentrations. Environ Sci Technol 1997;31:523–9. https://doi.org/10.1021/es960357w.

Table 1: Snapshot of CONTRA "decision tree" tab (Bockstaller et al., 2017), showing the decision tree design of the subindicator, pesticide leaching risk of contamination ($RCgw_{lea}$). The groundwater ubiquity index (GUS), the application period (Appli), the leaching potential of soil (LixPot) and the solubility of the active ingredient in water (Solu) were aggregated with respectively a weight of 40% for GUS and 20% for the three other variables, giving the calibrated $RCgw_{lea}$ expressed between 0 (high risk) and 10 (low risk). Some modifications were given to the decision rules yielding the final $RCgw_{lea}$ according to the explanations given in the last column. Input variables belong either to the favourable set (F) or to unfavourable set (U), (see Table S3).

GUS	Appli	LeaPot	Solu	Calibrated RCgwlea	Correction (absolute value)	Final $RCgw_{lea}$	Explanation on correction
F	F	F	F	10,0		10,0	
F	F	F	D	8,0	1	9,0	if GUS is favourable, effect of solubility is reduced
F	F	D	F	8,0		8,0	
F	F	D	D	6,0	1	7,0	if GUS is favourable, effect of solubility is reduced
F	D	F	F	8,0		8,0	
F	D	F	D	6,0	0,5	6,5	if GUS is favourable, effect of solubility is reduced
F	D	D	F	6,0		6,0	
F	D	D	D	4,0		4,0	
D	F	F	F	6,0	-0,5	5,5	If GUS is unfavourable, risk is increased to avoid a similar results with previous lines
D	F	F	D	4,0	0,5	4,5	If praying period is favourable (dry soil), risk is reduced
D	F	D	F	4,0	0,5	4,5	If praying period is favourable (dry soil), risk is reduced
D	F	D	D	2,0	0,5	2,5	If praying period is favourable (dry soil), risk is reduced
D	D	F	F	4,0		4,0	
D	D	F	D	2,0		2,0	
D	D	D	F	2,0		2,0	
D	D	D	D	0,0		0,0	

Table 2: Decision tree of the subindicator, pesticide runoff risk of contamination ($RCsw_{r/d}$) in case for runoff. RuPot is the runoff potential and Avai is the availability of the pesticide (see Equation 5). Input variables belong either to the favourable set (F) or to unfavourable set (U), (see Table S3).

RuPot	Avai	$RCsw_{r/d}$
F	F	10
F	U	5.5
U	F	6.7
U	U	0

Table 3: Snapshot of CONTRA "decision tree" tab (Bockstaller et al., 2017), showing the decision tree design of the subindicator, risk of air contamination by drift ($RCair_{dr}$). The type of sprayer (Sprayer), the speed sprayer (Speed), the sprayer height (Height), the use of antidrift

nuzzle (Nuzzle) and the air pressure (Pressure) were aggregated with the same weight of 20% for all variables, giving the calibrated $RCair_{dr}$,expressed between 0 (high risk) and 10 (low risk). Some modifications were given to the decision rules yielding the final $RCair_{dr}$ according to the explanations given in the last column. Input variables belong either to the favourable set (F) or to unfavourable set (U), (see Table S3).

Sprayer	Speed	Height	Nuzzle	Pressure	Calibrated $RCair_{dr}$	Correction (absolute value)	Final $Rcair_{dr}$	Explanation on correction
F	F	F	F	F	10	-1	9,0	Drift is possible even if all variables are favourable
F	F	F	F	U	8		8,0	
F	F	F	U	F	8		8,0	
F	F	F	U	U	6		6,0	
F	F	U	F	F	8		8,0	
F	F	U	F	D	6		6,0	
F	F	U	U	F	6		6,0	
F	F	U	U	U	4		4,0	
F	U	F	F	F	8	-0.5	7.5	If sprayer speed is unfavourable then an additional negative effect
F	U	F	F	U	6	-0.5	5.5	If sprayer speed is unfavourable then an additional negative effect
F	U	F	U	F	6	-0.5	5.5	If sprayer speed is unfavourable then an additional negative effect
F	U	F	U	U	4	-0.5	3.5	If sprayer speed is unfavourable then an additional negative effect
F	U	U	F	F	6	-0.5	5.5	If sprayer speed is unfavourable then an additional negative effect
F	U	U	F	U	4	-0.5	3.5	If sprayer speed is unfavourable then an additional negative effect
F	U	U	U	F	4	-0.5	3.5	If sprayer speed is unfavourable then an additional negative effect
F	U	U	U	U	2	-0.5	1.5	If sprayer speed is unfavourable then an additional negative effect
U	F	F	F	F	8	-8	0	If sprayer type defavourable, no possibility to reduce drift
U	F	F	F	U	6	-6	0	If sprayer type defavourable, no possibility to reduce drift
U	F	F	U	F	6	-6	0	If sprayer type defavourable, no possibility to reduce drift
U	F	F	U	U	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	F	U	F	F	6	-6	0	If sprayer type defavourable, no possibility to reduce drift
U	F	U	F	U	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	F	U	U	F	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	F	U	U	U	2	-2	0	If sprayer type defavourable, no possibility to reduce drift
U	U	F	F	F	6	-6	0	If sprayer type defavourable, no possibility to reduce drift
U	U	F	F	U	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	U	F	U	F	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	U	F	U	U	2	-2	0	If sprayer type defavourable, no possibility to reduce drift
U	U	U	F	F	4	-4	0	If sprayer type defavourable, no possibility to reduce drift
U	U	U	F	U	2	-2	0	If sprayer type defavourable, no possibility to reduce drift
U	U	U	U	F	2	-2	0	If sprayer type defavourable, no possibility to reduce drift
U	U	U	U	U	0		0	If sprayer type defavourable, no possibility to reduce drift

Table 4: Snapshot of CONTRA "decision tree" tab (Bockstaller et al., 2017), showing the decision tree design of the subindicator, pesticide risk for surface water ($I\text{-}Phy_{sw}$). The contamination risk for surface water through drainage or runoff ($RCsw_{d/r}$:), the contamination risk for surface water through drift ($RCsw_{dr}$), and the toxicity variable for aquatic organisms (TOX) were aggregated with respectively a weight of 33%, giving the calibrated $I\text{-}Phy_{sw}$ expressed between 0 (high risk) and 10 (low risk). Some modifications are given to the decision rules yielding the final $I\text{-}Phy_{sw}$ according to the explanations given in the last column. Input variables belong either to the favourable set (F) or to unfavourable set (U), (see Table S3).

$RCsw_{r/d}$	$RCsw_{dr}$	Tox	Calibrated I-Phy_{sw}	Correction (absolute value)	Final I-Phy_{sw}	Explanation on correction
F	F	F	10		10	
F	F	U	6		6	
F	U	F	7		7	
F	U	U	3	-1	2	Transfer for one pathway and toxicity unfavourable: decrease of 1 point to highlight the unacceptable situation for stakeholder
U	F	F	7		7	
U	F	U	3	-1	2	Transfer for one pathway and toxicity unfavourable: decrease of 1 point to highlight the unacceptable situation for stakeholder
U	U	F	4		4	
U	U	U	0		0	

Table 5: Snapshot of CONTRA "decision tree" tab (Bockstaller et al., 2017), showing the decision tree design calculating the final I-Phy indicator. The subindicators assessing the risk for three environmental compartments, pesticide risk respectively for groundwater (I-Phy_{gw}), for surface water (I-Phy_{sw}) and for air (I-Phy_{air}) were aggregated with the same weight of 33%, giving the calibrated I-Phy expressed between 0 (high risk) and 10 (low risk). Some modifications were given to the decision rules yielding the final I-Phy according to the explanations given in the last column. Input variables belong either to the favourable set (F) or to unfavourable set (U), (see Table S3). Membership functions are sinusoidal.

I-Phy_{gw}	I-Phy_{sw}	I-Phy_{air}	Calibrated I-Phy	Correction (absolute value)	Final I-Phy	Explanation on correction
F	F	F	10		10	
F	F	U	6.7	-0.7	6	To set at 6
F	U	F	6.7	-0;7	6	To set at 6
F	U	U	3.3	-1.3	2	To set at 2
U	F	F	6.7	-0.7	6	To set at 6
U	F	U	3.3	-1.3	2	To set at 2
U	U	F	3.3	-1.3	2	To set at 2
U	U	U	0		0	

Table 6: theoretical example of probability test comparing classes of an indicator (in colum, class 5 shows a higher value effect than 1) with a measurement (increasing values show a higher effect). The probability is the sum of correct estimation (cases colored in green) and overestimation (cases colored in blue). The cases colored in brown are considered as underestimation.

	Result of measured data (in class)				
Result of Indicator (in class)	**From 0 to 20**	**From 20 to 40**	**From 40 to 60**	**From 60 to 80**	**Fomr 80 to 100**
From 8 to 0	**20**	**8**	**8**	**9**	**2**
From 6 to 8	**16**	**5**	**11**	**11**	**35**

From 4 to 6	3	1	12	22	3
From 2 to 4	23	0	0	1	5
From 0 to 2	4	0	1	0	2

Table 7: correlation test between the measured data of the 3 sites and 4 transfer pathway and the concerned subindicator RC for the 3 versions of I-Phy. The results are in bold when subindicators of I-Phy3 performs better than in the previous versions.

Component of I-Phy	Site	Transfer pathway	fd0.1			ftotal			CMP		
			I-Phy3	I-Phy 1	I-Phy 2	I-Phy3	I-Phy 1	I-Phy 2	I-Phy 3	I-Phy 1	I-Phy 2
RC_{gw-lea}	La Jaillière	Drainage	**0.48**	0.36	0.24	0.35	0.49	0.19	0.35	0.41	0.13
$RC_{sw-r/d}$	La Jaillière	Saturation runoff	**0.49**	0.44	0.49	0.27	0.26	0.28	0.32	0.38	0.31
$RC_{sw-r/d}$	Geispitzen	Hortonien runoff	**0.66**	0.31	0.32	**0.15**	-0.15	-0.04	**0.35**	0.21	0.34
RC_{gw-lea}	Le Magneraud	Leaching	**0.46**	0.3	-0.17	**0.17**	0.05	-0.11	**0.08**	-0.01	-0.09

Table 8: probability test (see Figure 2) for the 3 sites and 4 transfer pathways for each concerned subindicator RC (see Figure 3), comparing to the measured transfer of active ingredient In blue the probability is over 70%, in yellow, the probability is between 50% and 70% and in red, between 40% et 50%

Component of I-Phy 3	site	fd 0.1				ftotal				CMP			
		probablity	correct	overestimation	underestimation	probablity	correct	overestimation	underestimation	probablity	correct	overestimation	underestimation
Jaillère ruissellement	I-Phy 3 esu (note)	53%	43%	10%	47%	52%	39%	13%	48%	53%	41%	12%	47%
Jaillère drainage	I-Phy 3 Reso (note)	74%	44%	30%	26%	57%	38%	19%	43%	68%	45%	23%	32%
Magneraud	I-Phy 3 Reso (note)	81%	56%	26%	19%	68%	49%	18%	32%	79%	55%	24%	21%
Geispitzen	I-Phy 3 Resu (note)	73%	28%	45%	28%	80%	28%	53%	20%	73%	40%	33%	28%

Table 9: detailed probability test (see Figure 2) for the site of La Jaillière runoff for the frequency of exceedance of the threshold of 0,1µg.L^{-1}).

	Fd 0.1 (%)				
Class	20	40	60	80	100
1	80	8	8	9	14
2	16	8	11	11	27
3	3	1	2	1	3
4	0	0	0	0	5
5	0	0	0	0	0

	n	%
Correct	90	43%
overestimation	20	10%
underestimation	97	47%
Total	207	

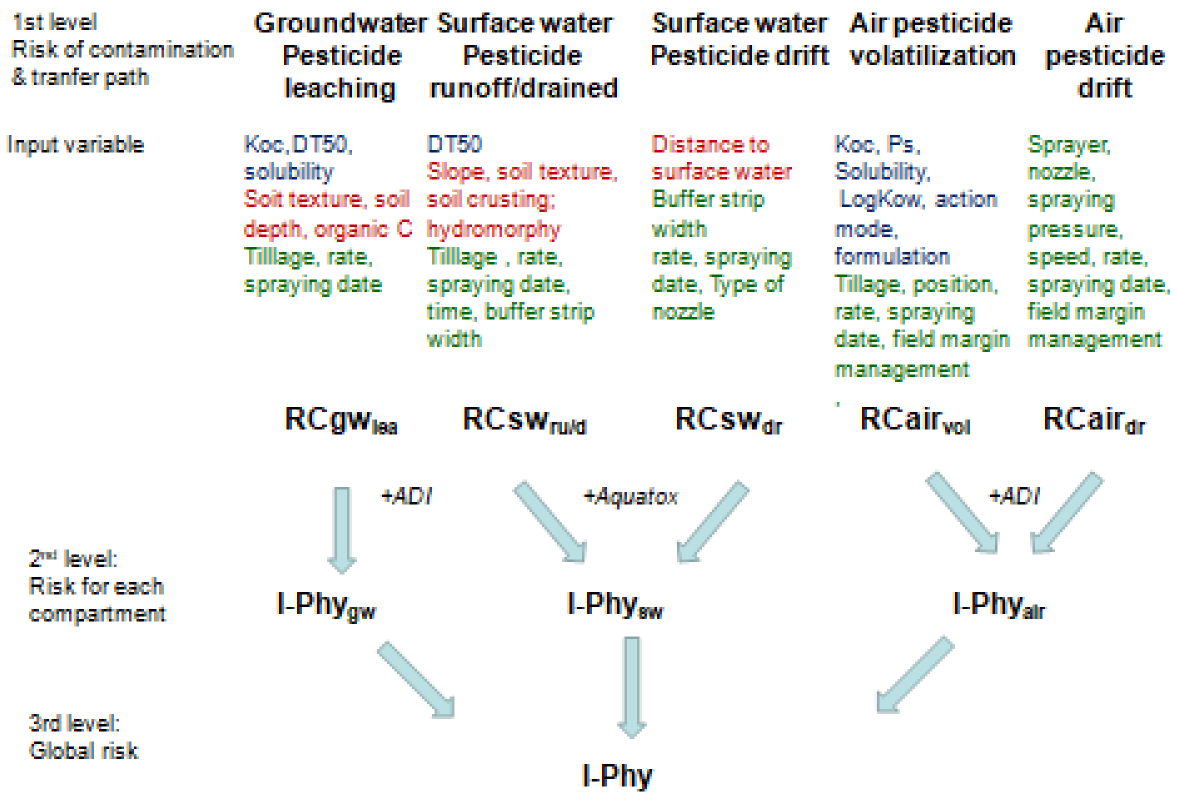

Figure 1: Overview of the calculation of I-Phy3 with three levels of aggregation and main input variables, in blue pesticide properties, in red soil and topographic variables, in green: management variable (KOC: adsorption coefficient, DT50: half-life, Solubility: solubility in water, Ps: vapor pressure, LogKow: logarithm of the octanoal-water coefficient, time: time between spraying and runoff event). For each aggregation, a fuzzy decision tree (see Figure S2) was implemented with information on fuzzy subsets linked to each variable given in Table S2.

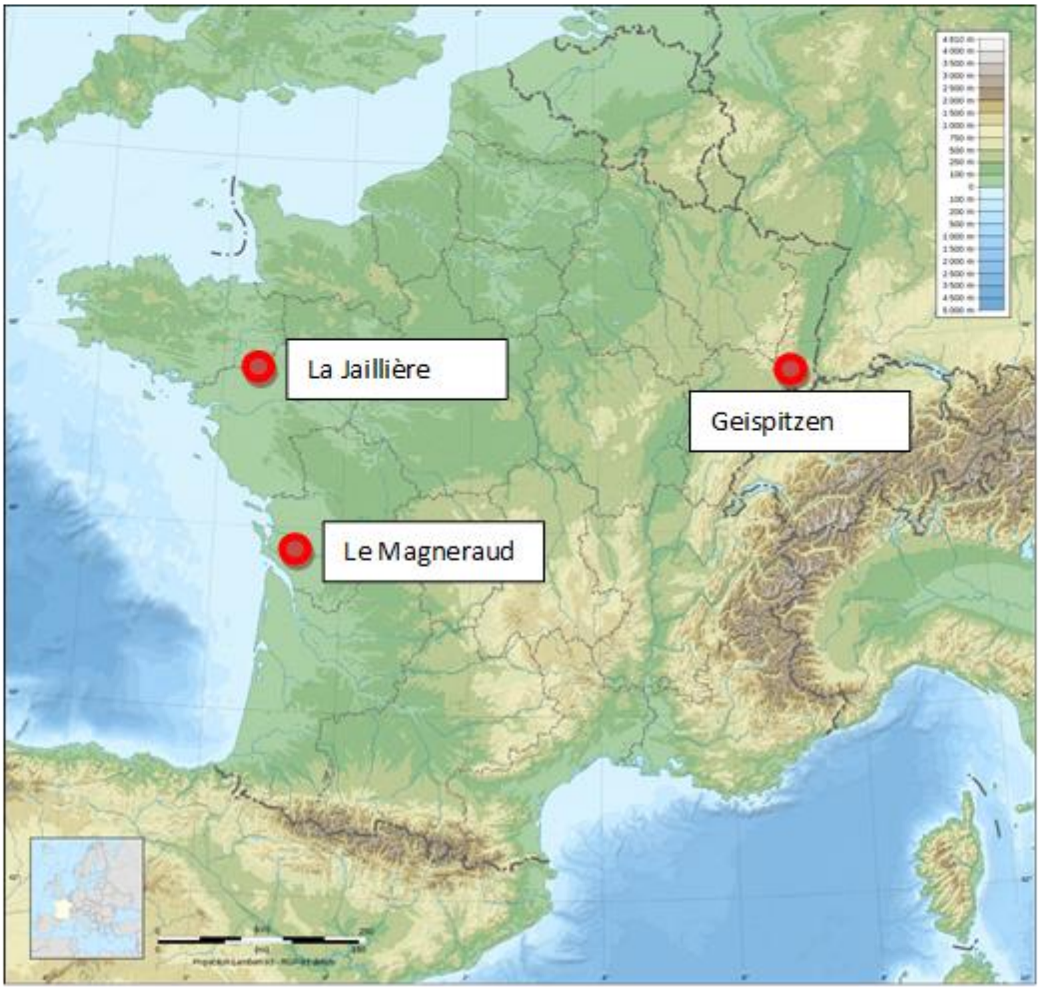

Figure 2: Repartitions of the sites used to assess predictive quality. With red bullet: sites for water quality and with blue bullets: sites for air quality

Year	Crop	Active ingredient	Corrected rate (g/ha)	% treated area	I-Phy	I-Phygw	RCgw	I-Physw	RCsw-rd	RCsw-dr	I-Phyair	RCair-vol	RCair-dr
2020	Spring barley	2,4-MCPA	150	100	4.6	4.3	3.6	6.9	6.6	10.0	4.7	6.1	4.9
2020	Spring barley	clopyralid	15	100	7.8	6.5	5.6	9.0	8.6	10.0	8.2	10.0	6.9
2020	Spring barley	fluroxypyr	30	100	9.0	8.8	7.0	9.1	8.5	10.0	8.8	10.0	5.8
2020	Spring barley	éthéphon	24	100	8.2	9.0	10.0	8.4	8.6	10.0	7.2	9.2	5.8
2020	Spring barley	prothioconazole	13	100	8.3	9.0	10.0	8.2	10.0	10.0	7.7	10.0	6.9
2020	Spring barley	fluoxastrobine	6	100	7.1	8.3	10.0	7.3	9.5	10.0	7.4	10.0	7.4

RCgw (weighted by treated area)	GUS	Application period (day number)	Potential of leaching	Solubility
3.6	0.00	0.56	0.10	0.00
5.6	0.00	0.56	0.10	0.00
7.0	0.31	0.56	0.10	0.00
10.0	1.00	1.00	0.10	0.00
10.0	1.00	1.00	0.10	1.00
10.0	0.98	1.00	0.10	1.00

Crop	Crop mnagement	Spraying programm*	I-Phy
Winter barley	Intensive (without chlortoluron)	3H, 2F, 1I, 3R	2.5
Winter barley	Intensive (without chlortoluron)	5H, 5F, 2I, 3R	4.1
Winter barley	Integrated	4H, 3F, 0I, 1R	6.5
Winter barley	Organic	0H, 0F, 0I, 0R	10.0

* Number of herbicdes (H) fungicides (F), insecticides (I), growth regulator (R)

Active ingredient	Rate (g/ha)	I-Phy	I-Phy $_{gw}$	RCgw $_{lea}$	I-Phy $_{sw}$	RCsw $_{r?d}$	RCsw $_{dr}$	I-Phy $_{air}$	RCair $_{vol}$	RCair $_{dr}$
flufenacet	200	4.2	5.1	5.2	3.7	4.6	10.0	6.2	10.0	7.4
diflufenican	100	6.5	8.9	8.2	3.8	4.7	10.0	8.5	10.0	9.4
chlortoluron	1250	3.5	3.5	2.1	2.9	2.7	9.7	5.7	9.9	8.5
esfenvalerate	6	6.8	8.5	10.0	5.4	7.6	10.0	8.2	10.0	7.9
cyprodinil	143	6.3	8.6	9.5	5.4	4.7	10.0	6.6	9.1	8.3
ethephon	143	6.3	8.0	8.1	6.8	5.1	10.0	6.3	8.2	9.3
chlormequat chlorure	285	5.3	5.5	4.7	6.9	4.5	10.0	6.9	10.0	9.3
prothioconazole	39	8.2	9.0	10.0	8.2	10.0	10.0	7.4	10.0	10.0
benzovindiflupyr	20	7.8	10.0	10.0	4.9	6.5	10.0	9.3	10.0	5.5
ethephon	54	7.4	9.0	9.9	7.1	6.2	10.0	6.9	8.9	5.8

Indicator	Membership value: 1 if favourable, 0 if unfavourable			
RCgw $_{lea}$	GUS	Spraying period	Leaching potential	Solubility
5.2	0.22	0.00	0.10	1.00
8.2	1.00	0.00	0.10	1.00
2.1	0.00	0.00	0.10	1.00
10.0	1.00	0.00	0.10	1.00
9.5	1.00	0.63	0.10	1.00
8.1	1.00	0.63	0.10	0.00
4.7	0.13	0.63	0.10	0.00
10.0	1.00	0.88	0.10	1.00
10.0	1.00	0.88	0.10	1.00
9.9	1.00	0.88	0.10	0.00

Figure 3: the detail for one subindicator RCeso. The membership degrees of the input variables show to which extent the variable is unfavourable (close to 0) or favourable (close to 1) and plays a role in the determination of the risk.

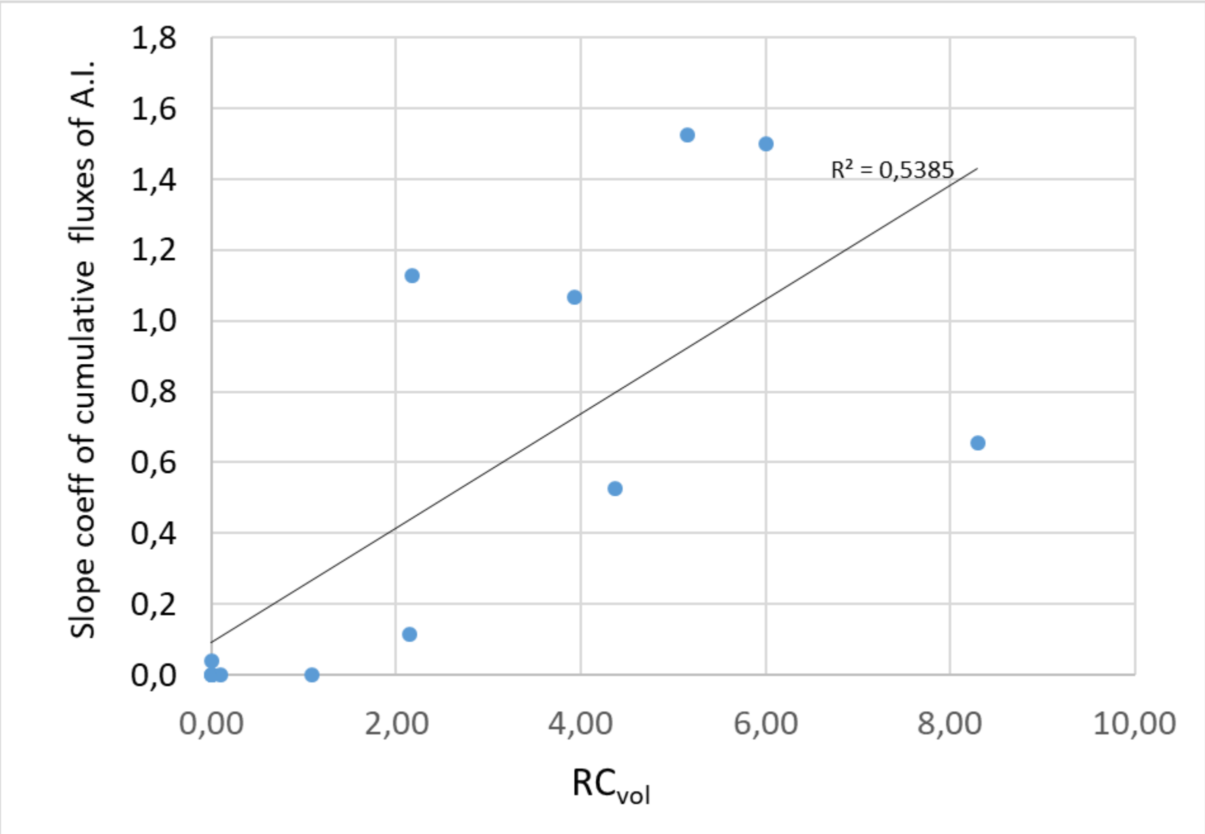

Figure 4: correlation between the slope coefficient of the cumulative fluxes of a.i. used as a proxy of volatilization risk (see SM XX) and the means of RC_{vol} for 12 a.i.