Topsoil-River water Pesticide Fate Model Documentation

Author: Artur Radomyski[1]

Reviewers: Reynold Chow[2], , Daniel D.M. Figueiredo[4], Lingtong Gai[2], João P. Nunes[2], Joost Vlaming[3], Anke Huss[4], Jakub Hofman[1]

Version: 0.2

[1] RECETOX, Faculty of Science, Masaryk University, Brno, Czechia
[2] Soil Physics and Land Management Group, Wageningen University & Research, P.O. Box 47, 6700, AA, Wageningen, the Netherlands
[3] Envista Consultancy, Arnhem, the Netherlands
[4] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands

E-mail contact: artur.radomyski@recetox.muni.cz

Table of Contents

Abbreviations:

AS: Active Substance

PPP: Plant Protection Product

PEC: Predicted Environmental Concentration

SPM: Suspended Particulate Matter

RQ: Risk Quotient

MS: Member State

CZ: Czech Republic

NL: Netherlands

DK: Denmark

PT: Portugal

LEVEL 1. Basic knowledge

Purpose

Goal and potential regulatory frameworks

Mixture of agrochemicals released to environment during farming activities may potentially reach surrounding environmental media such as topsoil and adjacent flowing water bodies. Overland processes including surface runoff and soil erosion by wind and water may play important role in distribution of agrochemicals between environmental compartments. A multitude of parameters can influence occurrence and magnitude of these processes, thus affecting number, amount and residency of agrochemicals in a given environmental medium and as a result contribute to temporal and spatial variability in levels of exposure and potential risks to biota.

Fate model of organic agrochemicals addressing spatial and temporal variability of farming and environmental processes is described. Relevant processes and parameters are addressed and their role in calculating of Predicted Environmental Concentrations (PECs) is presented. The model provides time-dependent and geo-referenced deterministic PEC estimation in topsoil and river water at individual field and river segment level on the one hand and frequency distribution or lumped PEC values at coarser scales on the other hand. Temporal profile of calculated PECs includes initial values right after application of active substance (AS) and time averaged values over duration of the simulation scenario. Subsequently, geo-PECs can be used to derive time-dependent and spatially distributed RQs for relevant terrestrial and aquatic endpoints.

In order to allow transparency and reproducibility of formulated fate and exposure models the documentation is structured according to the “Standard documentation of chemical exposure models” (CEN Workshop Agreement (CWA), 2015). Presentation of exposure models following the standard documentation protocol is believed to facilitate a more rigorous formulation of models’ description and the understanding by users (Ciffroy et al., 2016).

Applicability domain

The model estimations are primarily driven by chemical application patterns, chemical properties, pedo-climatic data, and river hydrology. Therefore, the model can reflect various spatial and temporal characteristics of the considered parameters, and application patterns when building site specific environmental scenarios. Consequently, scenarios are scale dependent, what in turn is reflected by applying weighting factors over the parameters within the scenario’s target area.

Temporal

The model has been designed to operate on a daily resolution to reflect AS application cycles and climate data. The model calculates time weighted average concentration in soil and water over simulation period, that is starting from the calendar day and month of PPP application until last day specified in the simulation scenario.

Spatial

Spatial scope of the model concerns small agricultural catchments, and in particular individual farms, and edge-of-field surface water bodies. Fine scale simulation results are then aggregated at coarser scales to produce geo-referenced single or distribution of PEC values for administrative or geographical units such as country districts, river buffers and small catchments. The model reflects geographic, pedo-climatic and farming characteristics of four EU countries: Czech Republic, Denmark, Netherlands and Portugal.

Chemicals

The model describes environmental fate of organic agrochemicals released to environment and transported to agricultural topsoil and edge-of-field surface water bodies.

Model mode and type

Predictive vs diagnostic

It is a predictive model.

Prescriptive vs screening

The model is intended for screening.

Steady state vs dynamic

The steady-state model simulates agrichemical mass in topsoil and river water on a daily or time-weighted basis. The exchanges of contaminants between Pore water and Soil Particles are assumed to be at equilibrium (i.e. represented by a partition coefficient at equilibrium). The application of the model just after AS release must then be considered with cautious if equilibrium condition between pore water and soil particles is not respected.

Analytical vs numerical

The model applies analytical solutions.

Empirical vs mechanistic

Several sub modules of the model related to simulating chemical losses via surface runoff that is effect of crop interception, terrain slope and precipitation on surface runoff occurrence and depth originate from empirical relationships. On the other hand, exchanges between media in topsoil are based on mechanistic assumptions.

Lumped vs distributed

At field level the model provides distributed output, and lumped results are delivered for the target spatial unit, river buffers and basins.

Deterministic vs probabilistic

The model provides deterministic output at field level and at coarser scale frequency distribution of PECs.

Model components

Model constituents can be grouped into the following categories:

Parameters are defined as a term in the model for a desired exposure scenario. They remain fixed during a model run or simulation but can be changed in different runs as a method for conducting sensitivity analysis or to achieve calibration goals.

Forcing variables are defined as time-dependent input variables influencing state variables. They can include e.g. meteorological, agronomical, biological factors that show a temporal variability (e.g. rain, temperature, seasonal land use coverage, and seasonal activity pattern of animals).

Intermediate variables are defined as a dependent variable calculated within the model, either fixed during model run if calculated from parameters only, or time-dependent when derived from forcing variables. They can be used to monitor model performance, and usually not available to end users for decision-making.

Regulatory variables can be both fixed and time-dependent, lumped or distributed. They are available to end-users and serve a decision-making purpose.

Medium is defined as an environmental or biota compartment assumed to contain a given quantity of the chemical. The quantity of the chemical in the media is governed by loadings/losses.

Loading is defined as the rate of release/input of the chemical of interest to the receiving system.

Loss is defined as the rate of output of the chemical of interest from the receiving system.

Exchanges are defined as the transfer of the chemical of interest between two media of the system.

Parameters, forcing and state variables are additionally grouped into several groups based on their purpose in the model: *climate*, *farming*, *chemical*, *topsoil*, *crop*, and *hydrology*.

Compartments and processes

Media considered

Currently model describes temporal and spatial concentration profile of agrichemicals in the top 20 cm layer of soil, the topsoil compartment, and surface river water compartment. Both media are depicted as colour-filled rectangles on ***Figure 1***.

Loadings and losses

Loadings and losses are depicted on ***Figure 1*** as one-headed arrows coming from or in the considered medium.

Loadings to topsoil may include:

1. Direct application on topsoil from chemical usage, interceptor factor is then 0
2. Weathering from plant leaves of contaminants previously intercepted from the atmosphere, interceptor factor is then > 0. Alternatively, it can be modelled as a delayed leaves-to-soil transfer.

Loading to topsoil includes:

1. Load of chemicals from topsoil surface runoff to surface water

Losses from topsoil:

1. Degradation of contaminant in the topsoil medium,
2. Chemical transported from topsoil in surface runoff towards the river.

Losses from surface water:

1. Degradation of contaminant in the river medium.

Exchanges

Transfer between media are depicted as two-headed arrows on ***Figure 1***.

1. Sorption/desorption. The equilibrium partitioning between concentrations in soil pore water and in soil particles is used to calculate the fraction of a contaminant dissolved in water contained in the soil compartment and bound to soil particles.

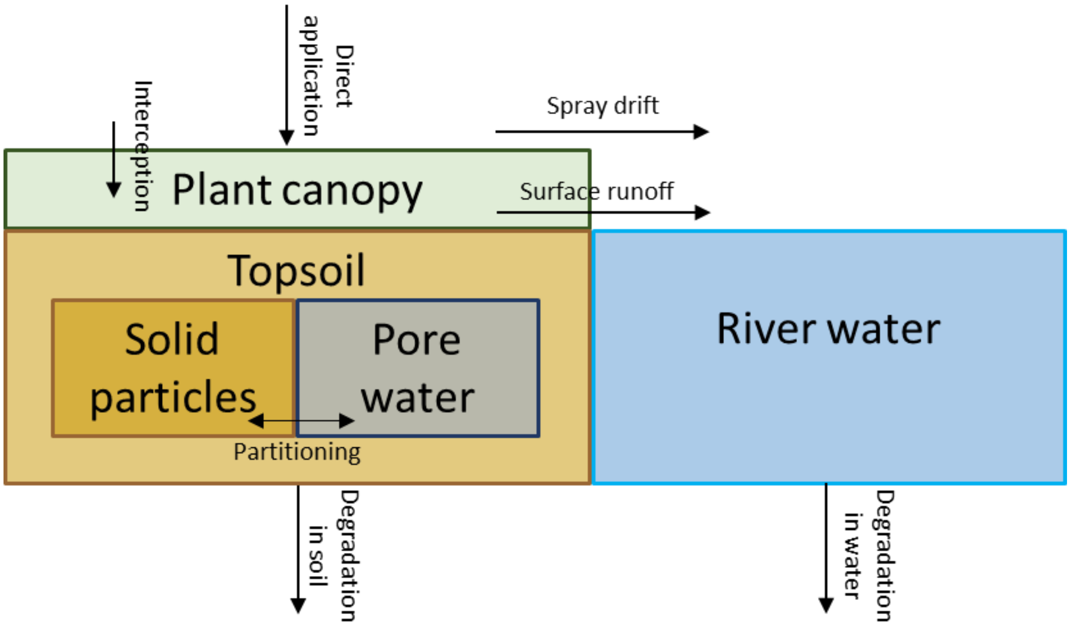


Figure 1 Compartments and processes included in the model.

Forcing variables

Application scheme

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly in calculating
AS application start date	t_app_startday	date (calendar day)	N	Decides which day of the month the AS application takes place and therefore how much rain is considered in calculating surface runoff	Depth of surface runoff
AS application month	t_app_month	date (calendar month)	N	Decides which month the AS application takes place and therefore how much rain is considered in calculating surface runoff	Depth of surface runoff
AS applications number	n_app	(-)	N	Number of AS applications within time averaging period. It is used to correct application rate in case there are multiple applications during time averaging window.	Corrected application rate
Averaging window	Δt_app_ave	days	N	Time period since the application start date over which concentrations are averaged.	Time-weighed concentration in topsoil
AS mass applied	M_as_crop	kg/year	N	Annual mass of AS applied in a given area. Usually available at a coarser scale and requires scaling down to individual fields.	1. AS mass applied on individual patches 2. Application rate
AS hectarage treated	A_as_crop	ha/year	N	If available, it is a treated area with AS	1. Fraction of treated area

				per annum. Usually available at a coarser scale and requires scaling down to individual fields.	2. Corrected application rate of field-crop-AS

Table 1 Forcing variables related to agrichemical application scheme

Climate

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly in calculating
Precipitation	P_hour (if available) P_day	mm/hr mm/day	Y	Forcing variable affecting occurrence and magnitude of surface runoff.	1. Depth of surface runoff 2. Time of surface runoff occurrence

Table 2 Forcing variables related to climate

Hydrology

River discharge	Q_river	m^3/day	Y	It is used for calculating flow rate-dependent AS concentration in river water.	PEC surface water

Table 3 Forcing variables related to river system

Parameters

Chemical properties

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly in calculating
Degradation half-life in soil	DT50_soil	days	N	It describes how much time is needed to decrease initial amount of substance by half	AS degradation rate in river water
Degradation half-life in water	DT50_water	days	N		PEC river water
Water-Organic carbon partitioning coefficient	K_oc	ml/g	N	Sorption coefficient data is a measure of the tendency of a chemical to bind to soils, corrected for soil organic carbon content.	1. Fraction of AS solved in water phase 2. Fraction of AS bound to solid phase

Table 4 Parameters related to physico-chemical properties of agrichemicals

Topsoil

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly in calculating
Field-crop area ID	A_field_crop_id	ha	Y	Geo-referenced area of individual field patches growing crops in a given season. Crops unique ID can be an individual crop name and/or code, crop group name and/or code.	1. Volume of topsoil 2. Agricultural area weighting factor 3. Geo-referencing AS application data
Ecotoxicological averaging depth	H_ top_soil	m	N	Eco-toxicological evaluation depth for concentration in total soil	Volume of topsoil
Sand content	Sand_ top_soil	%	Y	Sand content of individual patches	Depth of surface runoff
Clay content	Clay_ top_soil	%	Y	Clay content of individual patches	Depth of surface runoff
Soil organic carbon content	OC_top_soil	%	Y	Decides about mobility of a substance in pore water-soil particle system.	1. Fraction of AS solved in water phase 2. Fraction of AS bound to solid phase
Soil bulk density	rho_ top_soil	g/cm^3	Y	It is the dry bulk density of the soil.	Mass of topsoil

Table 5 Parameters related to topsoil properties

Crop

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly for
Crop-AS ID	Crop_AS_id	(-)	Y	Crop identification, that is names and codes, are used as	Geo-referencing AS application data

				matching keys for mapping non-geo-referenced PPP usage data. Codes and groupings are specific to crop coding system of a country.	
EPPO ID	EPPO_id	(-)	N	It is an additional crop taxonomical identification key to facilitate translation between crop naming systems.	Geo-referencing crop interception factor
Crop interception factor	IF_crop	(-)	N	Provides information what fraction of applied AS doesn't reach soil.	AS fraction intercepted

Table 6 Parameters related crop names and coding

Hydrology

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly for
River basin area	A_basin	ha	Y	It is defined here as a target area for calculating scaled PEC values in topsoil and river water.	1. Fraction of intersected agricultural patches (Total field-crop area target) 2. Fraction of intersected river segments
River width	W_river	m	Y	Used for calculating variable buffer area around river segments.	Buffer width corrected
River length	L_river	m	Y	Used mainly as a scaling parameter for calculating concentration in surface water for target area.	River segment weighting factor
River segment buffer width	W_buffer	m	Y	These are constants that determine area around river segments from which total flux of AS potentially entering given river segment is calculated from.	1. River segment weighting factor 2. Buffer width corrected 3. Buffer area

Table 7 Parameters related to river network and channel geometry

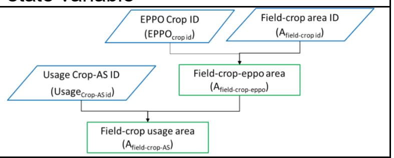
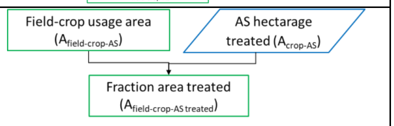
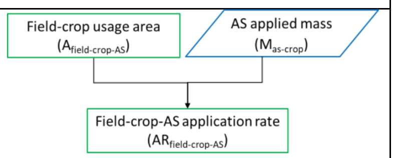
Topography

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Used directly in calculating
Terrain slope	Slope_terrain	%	Y	It is a primary topographical factor affecting if any and/or how much of surface runoff can be generated from individual patches on a given rainy day.	Depth of surface runoff

Table 8 Parameters related to terrain topography

Intermediate variables

Site-specific application rate at field level

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the state variable
Field-crop usage area	A_field_crop_AS	ha	Y	When geo-referenced, crop types are used as a matching and then aggregation key to assign individual patches to a crop type.	
Fraction field-crop area treated	F_field_crop_AS_treated	(-)	Y	It represents a **unique** fraction of AS treated crop area by considering reported treated areas from a given region.	
Application rate of field-crop-AS	A_field_as_crop	kg/ha	Y	When geo-referenced, AS is used as a matching and aggregation key to assign individual AS mass to a crop specific area to calculate application rates.	

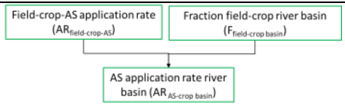
Corrected application rate of field-crop-AS	A_field_as_crop_corr	kg/ha	Y	It represents a **unique** application rate corrected for reported treated area from a given region.	

Table 9 State variables related to mapping individual fields to specific crop types and AS usage

Site-specific application rate in river buffer

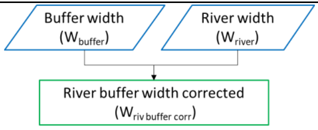
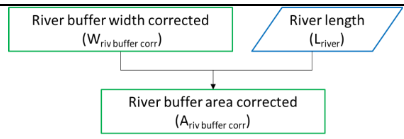
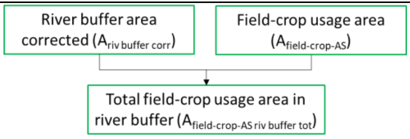
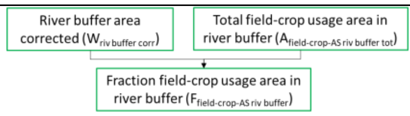
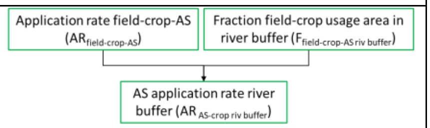
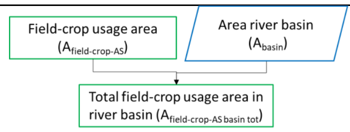
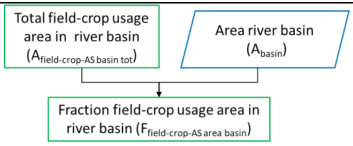
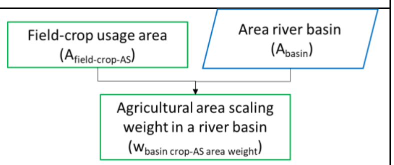
Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the state variable
Buffer width corrected	W_buffer_corr	m	Y	Buffer around river segment corrected of actual river segment width, it assures that it starts from riverbanks.	
Buffer area corrected	A_buffer_corr	ha	Y	Buffer around river segment corrected of actual river segment width, it assures that it starts from riverbanks.	
Total field area in a buffer	A_field_riv_buffer	ha	Y	It is agricultural area intersected with river buffer.	
Agricultural area weighting factor for a river buffer	F_field_crop_riv_buffer	(-)	Y	This factor is a unitless fraction of agricultural patches area occupying a river buffer. If patches are grouped by ASs and crop types, then more specific aerial fractions are obtained.	
Application rate field-crop-AS target spatial unit	AR_as_crop_riv_buffer	kg/ha	Y	It is corrected application rate for the fraction of total individual patches present in a target spatial unit.	

TABLE 10 STATE VARIABLES RELATED TO OBTAINING CROP AND CHEMICAL SPECIFC APPLICATION RATES FOR A TARGET SPATIAL UNIT – RIVER BUFFER

Site-specific application rate at river basin level

Total field area in target spatial unit	A_field_crop_area_tot_target	ha	Y	Total area of individual patches intersected by target spatial unit (e.g., a river catchment). It can be additionally grouped into crop type and applied AS.	
Agricultural area fraction in a river basin	F_field_crop_AS_basin	(-)	Y	This factor is a unitless fraction of agricultural patches area occupying a target spatial unit. If patches are grouped by ASs and crop types, then more specific aerial fractions are obtained.	
Agricultural area scaling weight in a river basin	w_field_area_eight	(-)	Y	It is a weight based on contribution of individual patches to the total agricultural area in a target spatial unit.	

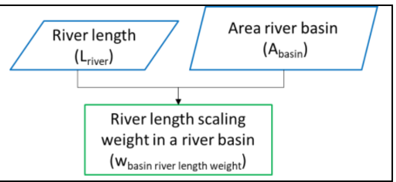
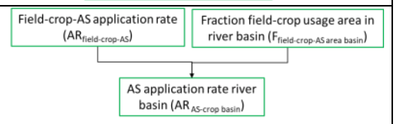
River length scaling weight in a river basin	w_field_area_eight	(-)	Y	It is a weight based on contribution of individual river segments to the total river network length in a target spatial unit.	
Application rate field-crop-AS target spatial unit	AR_as_crop_target	kg/ha	Y	It is corrected application rate for the fraction of total individual patches present in a target spatial unit.	

Table 11 State variables related to obtaining crop and chemical specifc application rates for a target spatial unit – river basin

Surface runoff

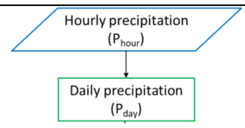
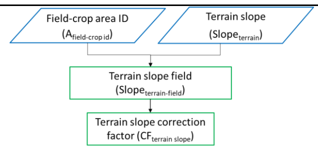
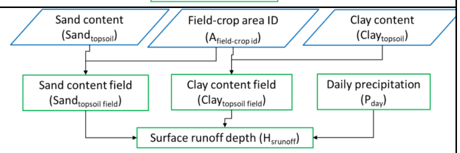
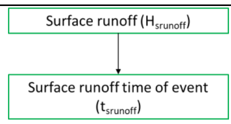
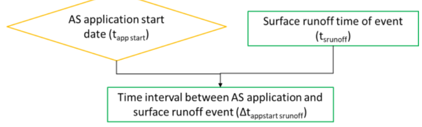
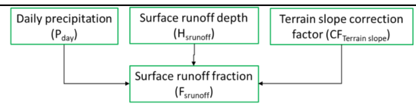
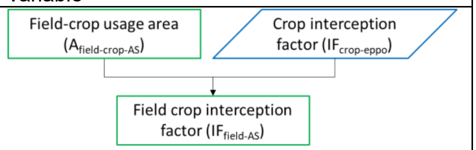
Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the state variable
Mean daily precipitation	P_day	mm/day	Y	If applicable, it is calculated from hourly and/or daily precipitation data for the target area.	
Terrain slope correction factor	CF_terrain_slope	(-)	Y		
Surface runoff depth	H_srunoff	mm/day	Y	Describes how much surface runoff is generated from individual patches on a given rainy day. It considers site specific soil texture.	
Surface runoff time of event	t_srunoff	day	Y	These are days when surface runoff is generated from individual patches.	
Time interval between AS application and surface runoff event	Δt_days_to_srunoff	days	Y	It describes time passed between consecutive surface runoff generating events starting from first day of the simulation	
Fraction of surface runoff	F_srunoff	(-)	Y	It is a ratio of surface runoff to precipitation. It expresses fraction of daily precipitation turned into runoff corrected by terrain slope of individual patches.	

Table 12 State variables related to calculating effect of surface runoff on chemical load

Effect of crop interception on chemical load

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the state variable
Field interception factor	IF_field	(-)	Y	It is geo-referenced crop specific interception factor.	

AS fraction intercepted	F_field_if	(-)	Y	AS fraction reaching soil surface after crop interception	Field crop interception factor ($IF_{field\text{-}AS}$) → AS fraction intercepted ($F_{AS\ IF}$)

Table 13 State variables related to calculating effect of crop interception factor on chemical load

Degradation rates

AS degradation rate in soil	k_DT50_soil	1/day	N	Degradation of the chemical in soil and river water is assumed to follow a linear first-order kinetics.	AS degradation half-life in soil ($DT50_{soil}$) → AS degradation rate in soil ($k_{DT50soil}$)
AS degradation rate in river water	k_DT50_riverwater	1/day	N		AS degradation half-life in water ($DT50_{water}$) → AS degradation rate in water ($k_{DT50water}$)

Table 14 State variables related to calculating degradation rates in topsoil and water

Partitioning between topsoil pore water and soil particles

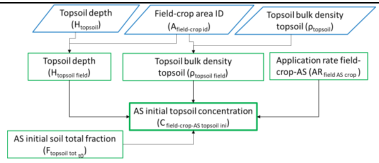
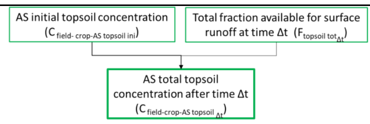
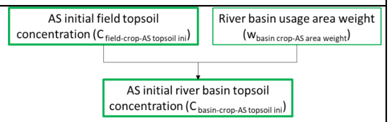
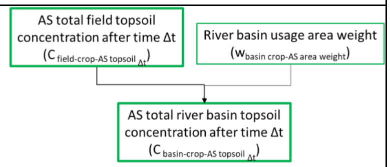
Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the state variable
Distribution coefficient at the interface pore Water-soil particles	K_d_soil_water	ml/g	Y	Describes the exchanges of contaminants between Pore water and soil particles. It is assumed to be at equilibrium. It can geo-referenced when corrected for local organic carbon content.	OC topsoil content ($OC_{topsoil}$) + Field-crop area ID ($A_{field\text{-}crop\ id}$) → Soil organic carbon content field ($OC_{topsoil\ field}$); Soil organic carbon content field ($OC_{topsoil\ field}$) + Water-Organic carbon partitioning coefficient (K_{OC}) → Distribution coefficient at the interface pore Water-soil particles ($K_{d\ soil\ water}$)
AS initial fraction dissolved in soil pore water	F_AS_soil_water_ini	(-)	Y	Describes initial fraction of AS solved in water phase	Distribution coefficient at the interface pore Water-soil particles ($K_{d\ soil\ water}$) → AS initial fraction dissolved in soil pore water ($F_{topsoil\ water_{t0}}$)
AS initial fraction bound to soil particles	F_AS_soil_solid_ini	(-)	Y	Describes initial fraction of AS bound to solid phase	Distribution coefficient at the interface pore Water-soil particles ($K_{d\ soil\ water}$) → AS initial fraction bound to soil particles ($F_{topsoil\ solid_{t0}}$)
AS initial soil total fraction	F_AS_top_soil_tot	(-)	Y	Initial total fraction of AS in topsoil	AS initial fraction dissolved in soil pore water ($F_{topsoil\ water_{t0}}$) + AS initial fraction bound to soil particles ($F_{topsoil\ solid_{t0}}$) → AS initial soil total fraction ($F_{topsoil\ tot\ t0}$)
AS fraction dissolved in soil pore water after time lag	Δt_F_AS_pore_water	(-)	Y	It is remaining AS fraction dissolved in soil pore water after time from AS application	AS initial fraction dissolved in soil pore water ($F_{topsoilwater_{t0}}$) + Time interval between AS application and surface runoff ($\Delta t_{appstart\ to\ runoff}$) + AS degradation rate in soil ($k_{DT50soil}$) + Fraction AS intercepted ($F_{field\ IF}$) → AS fraction dissolved in soil pore water after time Δt ($F_{topsoil\ water_{\Delta t}}$)
AS fraction bound to soil solids after time lag	Δt_F_AS_soil_solid	(-)	Y	It is remaining AS fraction bound to soil solids after time from AS application	AS initial fraction bound to soil particles ($F_{topsoil\ solid_{t0}}$) + Time lag between AS application and surface runoff ($\Delta t_{appstart\ to\ runoff}$) + AS degradation rate in soil ($k_{DT50soil}$) + AS fraction intercepted ($F_{AS\ IF}$) → AS fraction bound to soil particles after time Δt ($F_{topsoil\ solid_{\Delta t}}$)
AS soil total fraction after time lag	Δt_F_AS_top_soil_tot		(-)	It is the AS fraction in total top soil at each day of simulation	AS fraction dissolved in soil pore water after time Δt ($F_{topsoilwater_{\Delta t}}$) + AS fraction bound to soil particles after time Δt ($F_{topsoil\ solid_{\Delta t}}$) → Total fraction available for surface runoff at time Δt ($F_{topsoil\ tot_{\Delta t}}$)

Table 15 State variables related to calculating fraction of chemicals partitioned between topsoil pore water and soil particles

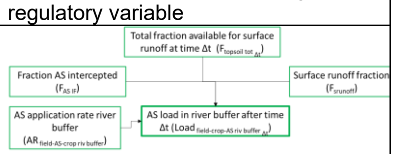
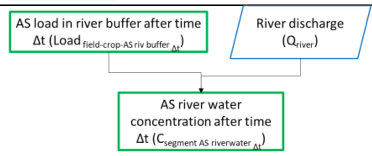
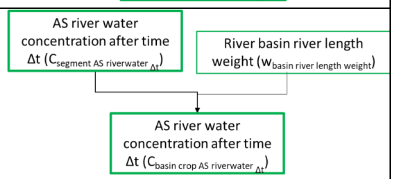
Regulatory variables

Topsoil

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the regulatory variable

AS initial field total topsoil concentration	C_field_crop_AS_topsoil_ini	(-)	Y	Georeferenced total initial amount of AS in the topsoil at equilibrium immediately after application. It applies to any target spatial unit where AS is applied such as river basin or river buffer	
AS total field topsoil concentration after time lag	Δt_C_field_AS_topsoil	(-)	Y	Georeferenced total AS amount in topsoil after time interval and corrected of effect of degradation processes. It applies to any target spatial unit where AS is applied such as river basin or river buffer	
AS initial river basin total topsoil concentration	C_basin_crop_AS_topsoil_ini	(-)	Y	It is a georeferenced initial concentration in topsoil weighted by size of individual patches. Represents upscaled field concentration in a target spatial unit.	
AS total river basin topsoil concentration after time lag	Δt_C_basin_crop_AS_topsoil	(-)	Y	It is a georeferenced lagged concentration in topsoil weighted by size of individual patches. Represents upscaled field concentration in a target spatial unit.	

River water

Full Name	Abbreviated name	Unit	Spatial (Y/N)	Purpose/explanation	Process followed for calculating the regulatory variable
AS load individual field in river buffer	Δt_load_topsoil_rivbuff	mg/day	Y	Georeferenced total mass of AS entering river segment from the buffer area	
Concentration in surface river water after time lag	Δt_C_segment_riverwater	mg/m^3	Y	Georeferenced total concentration of AS in a river segment	
Concentration in surface river water in a basin after time lag	Δt_C_basin_riverwater			Georeferenced total concentration of AS in all river segments in a river basins, weighted by individual segments lengths as fraction of total river network length.	

LEVEL 2. Process knowledge

Theoretical background

In order to estimate the amount of pesticide that is transported into the surface water by the overland surface runoff it is necessary to determine the proportion of AS in the application dose (AD), which will reach the ground and will be dissolved in the water phase of the soil. Similarly, it is important to know how much pesticide will be bound to the solid phase of the soil in order to estimate the amount of pesticide transported into the surface water by erosion. The proportion of AD available for runoff and erosion will be influenced by such factors as soil temperature, soil organic carbon content and pH values of the soil.

For consistency the overall structure of the model follows the OECD model (‘‘simplified formula for indirect loadings caused by run-off’’) developed to estimate the percentage of the applied amount of a pesticide that is dissolved in runoff water and, thus, can reach a surface water body (Gutsche & Strassemeyer, 2007). The OECD model calculates the dissolved percentage as a product of functions each of which describes the reduction of the maximum available percentage due to one of the following parameters and processes respectively: plant interception, degradation, sorption, depth of surface water runoff, slope and width of functioning buffer strips (Schriever et al., 2007).

PEC topsoil and PEC surface water are deterministically calculated as point estimates at individual field and river segment respectively. Then, weighted mean PEC topsoil is derived from individual concentrations at field level topsoils by weighting of aggregated by crop and AS areas of fields intersected with the river basins. On the other hand, weighted mean of PEC surface water is derived by weighting individual stretch lengths (Boeije et al., 2000).

Calculated mass of AS in a volume unit of the topsoil on individual farm expressed as $PEC_{topsoil}$, can be taken up by soil species via respiration of soil pore water and ingested soil particles. Plant canopy may potentially accumulate AS via root uptake. $PEC_{surface\ water}$ generally reflects the cumulative inputs of a substance from the area of application in a river stretch by diffuse sources, and total concentration of ASs to which aquatic species might be exposed to via respiration of water and/or ingestion of suspended particulate matter. $PEC_{surface\ water}$ at the outlet of a river catchment represent the sum of all non-point source inputs estimated for a given catchment (Röpke et al., 2004). Due to heterogeneity of different catchments in terms of their environmental parameters (soil, precipitation, land use, etc.) controlling pesticide , $PEC_{surface\ water}$ is expected to show significant variation at large spatial scales such as country or a region. The underlying quantity for $PEC_{surface\ water}$ calculations are the expected mean daily load from the overland surface runoff as described above. Then the $PEC_{surface\ water}$ is yielded by the ratio of the mean daily input into various flowing surface water bodies characterised by their daily discharge.

Application scheme

Motivation: Application scheme includes information about when, how much and what surface area is covered by application of AS. Application timing is dependent on application zone; therefore, the model should be run for areas where uniform application timing can be assumed.

Selected approach(es): The major source of uncertainty in AS application scheme is lack of information on how much chemical is actually used on individual parcels, and what area is actually treated with a given substance. Before running the model, crops grown on individual fields are matched with reported chemical usage data (M_as) for individual crops. Reported chemical volume sales data can be used as an alternative (Eurostat, 2022). The usage or sales data for individual chemicals are commonly reported for districts, region or a country as a whole. In order to represent such lumped usage data at individual field level the ratio of actual field areas to the treated area (A_as) in a region for which A_as is reported can be calculated and used as a correction factor when calculating application doses (F_field_crop_area_target). The "treated area fractions" are then specific to a given area allowing for downscaling of reported AS usage quantities.

Alternative approach(es) (if known): Often times single presupposed application rates are used as inputs in modelling exercises (Schriever et al., 2007a). On the other hand, PPP registration databases report maximum/minimum and average application rates and application frequencies, which are specific to a country and a registered product/active substance, what can be used as a prior to inform the model.

Canopy interception

Motivation: In higher tiers of the exposure assessment, crop interception and subsequent dissipation at the crop canopy may be included. Since the introduction of the FOCUS groundwater scenarios, it has been common practice to reduce the application rate by the fraction that is intercepted by the crop canopy and to apply this reduced fraction to the soil

Selected approaches: The implemented model doesn't simulate explicitly canopy processes instead it uses crop interception factors defined for generic FOCUS crop groups. Original crop interception tables currently used for PEC soil and groundwater calculations were compiled by members of the FOCUS groundwater group (FOCUS, 2000). These tables show the fraction of the dose reaching the soil as the sum of wash-off from the canopy and the fraction of the dose that reaches the soil directly. For the purpose of the model the FOCUS crop groups were "disentangled" prior to running simulation using EPPO codes to match specific crops grown on individual parcels. Additional references complementing crop interceptor factors compiled by FOCUS is provided in the literature review by van Beinum & Beulke (2010).

Alternative approach(es) (if known): More detailed approach can be applied using most important properties affecting canopy processes such as the half-life for the decline of the dislodgeable residue on plants and the wash-off factor. EFSA PPR Panel (2012) proposed to use as default values in the exposure assessment a wash-off factor of 0.1 1/mm (100 1/m) and a half-life for the dislodgeable foliar residue on plants of 10 days. The effect of dissipation at the crop canopy and foliar wash-off should be also included when the substance is applied to the crop canopy (*Figure 2*) (European Food Safety Authority, 2017).

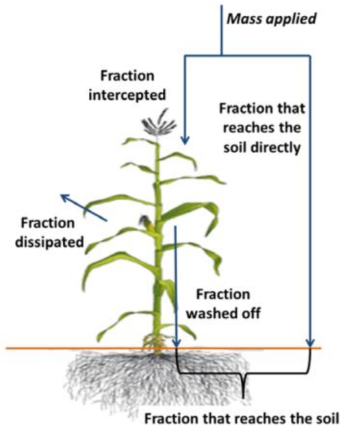


Figure 2 Canopy related processes reducing amount of applied chemical reaching topsoil

Surface runoff

Motivation: Surface runoff is one of the most important pathways for pesticides to enter surface waters from watersheds. It is a loss process from soils and can be a significant secondary input into freshwaters because they serve as a "sink" collecting water and/or particle fluxes from potentially wide areas, especially during rainfall. Mathematical models are employed to characterize spatio-temporal variability of surface runoff within landscapes, but they are often data intensive and therefore difficult to parameterise. Simpler spatially-explicit models were thus taken into consideration here.

Selected approach(es): The depth of surface runoff (mm) is specified according to results of Lutz and Maniak (OECD, 1998) (Berenzen et al., 2005; Probst et al., 2005) for vegetated dry soils, which represent conditions in middle and late vegetation period. It is a piecewise function that depends on which soil texture dominates the desired spatial unit. The run-off model estimates the volume of water generated daily from the desired spatial unit after precipitation events which is then used to estimate AS application dose available in the run-off water. This approach was selected due to its broad range of applications as demonstrated by numerous published case studies (Ippolito et al., 2015; Kattwinkel et al., 2011; Schriever, Von Der Ohe, et al., 2007b; Verro et al., 2002).

Alternative approach(es) (if known): Curve-Number method (CN) is commonly used in estimating surface runoff from various types of surfaces (Hawkins et al., 2009; Steenhuis et al., 1996). It is based on the principle that the ratio of run-off and precipitation is proportional to the ratio of actual and potential infiltration. The potential infiltration depends on local parameters describing land usage and soil properties, which are expressed in curve numbers. Another approach may consider global wash-off rate constant. It describes the transport of contaminants in water flowing over the soil surface and finally reaching freshwater systems (rivers and/or lakes).

AS partitioning between soil particles and pore water

Motivation: Exchanges of contaminants between the dissolved (i.e. porewater) and the particulate phases of soil govern how much chemical is available for surface runoff and what fraction of the applied amount will enter surface water.

Selected approaches: Exchanges of contaminants between porewater and soil particles are assumed to be equilibrated and thus described by a distribution (or partition) coefficient K_d_soil_water. For organic pollutants, exchanges are governed by a hydrophobic sorption mechanism and the distribution coefficient K_d_soil_water is related to the octanol-water partition coefficient Koc, and the concentration of organic carbon in the soil particles OC_top_soil

Alternative approach(es) (if known): When equilibrium condition between porewater and soil particles is not respected, e.g. just after a direct application, the model must then be considered with cautious. In such a case, exchanges between water and SPM should be described by non-equilibrium kinetics, using sorption and desorption kinetic rate constants.

AS degradation in soil

Motivation: Degradation can be a significant loss term in soils, because the pesticide concentration in runoff water at the beginning of a runoff event highly depends on the substance decay as well as the retention capacity of the crop and soil.

Selected approaches: The individual processes that are responsible for degradation (e.g. biodegradation, photolysis, hydrolysis) are not distinguished here but they are added into an aggregated loss rate. Degradation is assumed to follow linear first-order kinetics.

Alternative approach(es) (if known): Transformation processes can distinguish the following mechanistic processes: (i) hydrolysis (acid- and base-catalyzed hydrolysis can be modelled according a thermodynamic principle); (ii) photolysis (that should involve light intensity); (iii) microbial degradation (that can be calculated by a maximum microbial degradation rate, eventually attenuated by anaerobic conditions, suboptimal temperature and suboptimal pH). Such processes are considered individually in some specific models. However, such an approach requires a lot of parameters that are poorly available for most of the chemicals and a global loss rate was preferred.

AS degradation in river water

Motivation: Degradation can be a significant loss term in freshwater column, in particular for chemicals stored in sediments.

Selected approaches: The individual processes that are responsible for degradation (e.g. biodegradation, photolysis, hydrolysis) are not distinguished here but they are added into an aggregated loss rate. Degradation is assumed to to follow linear first-order kinetics.

Alternative approach(es) (if known): Degradation processes are generally simulated by pseudo first order degradation rates (i.e. degradation proportional to the concentration of contaminants in the media). However, transformation processes can distinguish the following mechanistic processes: (i) hydrolysis (acid- and base-catalyzed hydrolysis can be modelled according to a thermodynamic principle); (ii) photolysis (that should involve light intensity); (iii) microbial degradation (that can be calculated by a maximum microbial degradation rate, eventually attenuated by anaerobic conditions, suboptimal temperature and suboptimal pH). Such processes are considered individually in some specific models dedicated to freshwaters (e.g. Aquatox). However, such an approach requires a lot of parameters that are poorly available for most of the chemicals.

LEVEL 3. Input data

Model input data consist of forcing variables and parameters needed to be specified a priori to initialise the model run.

Data processing and spatial operations

This section describes what sources and data transformation procedures were used to obtain parameter values. At individual field and river segment scale finer datasets were used, when available, and aggregated at coarser scales. However, if not available, less resolute datasets were used especially for coarser target areas, that is river basin or a country district. All spatial input data were converted to spatial objects, *SpatVector* and *SpatRast,* using *vect()* and *rast()* function respectively, from the ***terra*** package. Spatial objects were subsequently transformed using *project()* function, to the World Geodetic System (WGS84; https://epsg.io/3857), which is a commonly used metric projection system in meters.

All site-specific parameter values representing desired spatial target unit are derived through spatial operations. The latter involve two steps: 1) preparatory spatial data processing at country level, “outside” of the model. Initial preparatory steps are taken to compile datasets which are then used as inputs for creating simulation exposure scenarios. *Figure 3* shows example of all attributes defined for all parcels in CZ within the parameter category “Application scheme” and “Crop”.

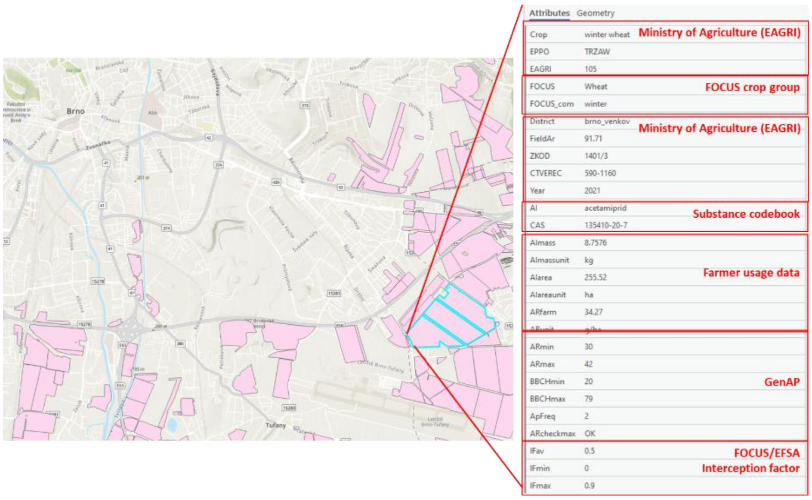


Figure 3 Example of elements of application scheme data set compiled for Czech Republic. Same attributes constitute application schemes of other countries.

More detailed steps included in preparation of “Application scheme” data are shown for the Czech Republic in APPENDIX 1. Flowchart of AS usage and crop data preparation using Czech data. and APPENDIX 2.

Flowchart of AS usage data spatialisation using Czech data.

Major part of the elements of the work flow are transferrable to other regions and countries.

2) Spatial operations integral to the model at fine and target spatial unit level. These include spatial operations at field and river segment level, bounded by user defined spatial target unit that is country district, or NUTS region, and the number of intersected river basins. *SpatRast* objects are cropped to match the extent of the desired target area given by *SpatVector* polygons for which further calculations are performed. All initial processing of raster and vector data takes place at individual field level using country specific parcel and crop vector maps. Subsequently, template polygons defined as “target spatial units” for which spatial statistics from the intersected raster and vector input data objects are extracted. The “target spatial units” are river basins polygons intersected with the country district(s) or NUTS region borders. Two functions with additional arguments, *zonal()* and *aggregate()* are applied. Depending on the map’s communication objective and in order to summarise raster and vector data at the desired target area either function is used. More detailed breakdown of possible levels of map visualisation with respect to communication objectives is given in Table 16.

Technical features					Communication objectives				
Mapping type	Granularity/target area for modelling	Recommended temporal resolution	Spatial object type	Input dataset origin	Target area for visualisation	AS*	Crops	Regulatory variable	RQ-relevant temporal window (days)
Distributed	Field	daily	Vector	Local	District	3	All	PEC topsoil	Initial/56/365
Distributed	River segment	daily	Vector	Local	District	3	All	PEC surface water	
Distributed	Field	daily	Raster	Local	Whole country	3	All	PEC topsoil	
Distributed	River segments	daily	Raster	Local	Whole country	3	All	PEC surface water	
Lumped	River basin	daily/monthly	Vector	Local/Global	Whole country	3	All	PEC topsoil	
Lumped	River basin	daily/monthly	Vector	Local/Global	Whole country	3	All	PEC surface water	

Table 16 Summary table of levels of visualisation according to communication objective of maps. * Tebuconazole, glyphosate and acetamiprid for demonstration.

Local, high-resolution datasets were preferred for obtaining parameter and forcing variables values. Sources for these datasets are diversified coming from governmental bodies like cadastre, statistical and meteorological authorities. Such data are oftentimes publicly available at high temporal and spatial resolutions. However, if not available, global digital datasets were selected. For consistency Hydrosheds database version 1.0 was preferred since it contains actively maintained large number of well documented datasets and secondary products (Lehner et al., 2008). Products such as the HydroATLAS Version 1.0 (Linke et al., 2019) offers data for 56 variables, partitioned into 281 individual attributes and organized in six categories: hydrology; physiography; climate; land cover & use; soils & geology; and anthropogenic influences.

Forcing variables

Application scheme

These time dependent variables are to be specified prior to running simulation in order to define set of input values required to run the model.

AS application start date (t_app_startday)

Physical/chemical/biological/empirical meaning:

It decides about calendar day when AS application takes place.

Factors influencing parameter value:

Time of first AS application is site specific, meaning that all crops in the investigated zone are treated at the same time.

Role in the model:

It is used to initiate simulation run. From this day duration of simulation in days is specified

Database used for parameter estimation:

Its value is specified by end user

Parameter default value:

Its value is specified by end user

AS application month (t_app_month)

Physical/chemical/biological/empirical meaning:

It decides about calendar month when AS application takes place.

Factors influencing parameter value:

Time of first AS application is site specific, meaning that all crops in the investigated zone are treated at the same time.

Role in the model:

It is used to initiate simulation run. It decides about average daily and/or monthly precipitation used in the simulation.

Database used for parameter estimation:

Its value is specified by end user.

Parameter default value:

Its value is specified by end user.

AS applications number (n_app)

Physical/chemical/biological/empirical meaning:

Crop and AS dependent number of PPP applications during growing season.

Factors influencing parameter value:

Depends on type of crop, time of application and season.

Role in the model:

Decide if the application dose is taken multiple times during simulation.

Database used for parameter estimation:

It is user defined value. Reference application frequencies and number of annual applications can be found in national PPP register databases.

Parameter default value:

It should be defined case by case.

Averaging window (Δt_app_ave)

Physical/chemical/biological/empirical meaning:

It is a user defined parameter describing duration of the simulation scenario.

Factors influencing parameter value:

Value of averaging window should be based on expert judgement and consider actual application scheme.

Role in the model:

The TWA concentrations are calculated for periods over a use-defined duration, e.g., 365 days after the occurrence of the initial concentration in topsoil.

Database used for parameter estimation:

Defined by end-user.

Parameter default value:

Its parameter value should reflect the desired analysis objective and be grounded in local application scheme specific to the region, chemicals, and crops.

AS mass applied (M_as_crop)

&

AS hectarage treated (A_as_crop)

Physical/chemical/biological/empirical meaning:

M_as_crop signifies amount in number of kilograms of the active substances used on a certain crop. A_as_crop is the physical area treated with plant protection products in hectares. It can be a whole crop parcel or only parts of a plot. This is a unique parameter in the sense that values on the area treated are only indicated if country confirmed that several treatments on the same physical area are reported separately, so that double counting can be avoided.

Factors influencing parameter value:

The accuracy of reported AS usage areas and masses by quality of data in the national reports. In case of A_as_crop, the same physical area can also be treated several times, which can lead to double counting.

Country	Country AS use metadata sources
Czech Republic	Czech Statistical Office
Netherlands	Statistics Netherlands
Denmark	Ministry of Environment and Food of Denmark
Portugal	Statistics Portugal

Table 17 Compiled AS usage data sources for individual countries

Role in the model:

Once defined for each parcel it describes initial conditions of the simulation scenario.

Database used for parameter estimation:

Country authorities develop sampling and statistical procedures for collecting and processing usage data following overarching mandatory pesticide usage/sales data collection scheme according to (European Parliament, 2009). Generally active substance's usage and sale data should be collected by the Member States according to the European Statistics Code of Practice (European Commission, 2018), using reference metadata template for the chemicals listed by (European Commission, 2017). It describes statistical concepts and methodologies used for the collection and generation of data. They provide information on data quality and, since they are strongly content-oriented, assist users in interpreting the data. Reference metadata, unlike structural metadata, can be decoupled from the data. Reference metadata for pesticide usage and sales can be consulted via EUROSTAT following link to Pesticide use in agriculture (aei_pestuse) and Pesticide sales (aei_fm_salpest09) respectively.

Country	Country sources for AS use and application data
Czech Republic	Czech Ministry of Agriculture (EAGRI)
Netherlands	Statistics Netherlands' database (CBS)
Denmark	Ministry of Environment of Denmark (MST)
Portugal	Statistics Portugal

Table 18 Main sources for AS usage data in four selected countries

Parameter default value:

M_as_crop and A_as_crop values are defined by, first country, then each chemical and crop combination.

Climate

Precipitation requires specification of a dataset from which either 1) average daily rainfall is calculated, or 2) monthly rainfall is extracted without further processing.

Precipitation (P_day)

Physical/chemical/biological/empirical meaning:

Amount of rainfall falling on a single day is one of the most important parameters which governs occurrence and depth of generated surface runoff.

Factors influencing parameter value:

Date (t_app_startday) and Averaging window (Δt_app_ave) of the simulation affect amount of precipitation considered in the simulated scenario.

Role in the model:

Depth (mm) of daily rainfall is a forcing variable directly used in calculating surface runoff. It may change on a daily basis, providing that, such data are available then the calculated depth of surface runoff will vary daily.

Database used for parameter estimation:

If hourly/daily precipitation is not available from country specific datasets shared via local cadastre or meteorological authorities, average precipitation values is used from WorldClim v1.4 (Hijmans et al., 2005). WorldClim is a database of interpolated global climate surfaces (excluding Antarctica) at a spatial resolution of 30 arc-seconds. Input data for the generation of WorldClim were gathered from a variety of sources (~70,000 stations) and, where possible, were restricted to records from 1950 to 2000. WorldClim applied the thin-plate smoothing spline algorithm implemented in the ANUSPLIN package for interpolation, using latitude, longitude, and elevation as independent variables. It is an element included in HydroATLAS.

Parameter default value:

Depends on the location. Precipitation values are extracted for individual river basins. If hourly precipitation data are available daily averages are derived for meteorological stations within the borders of a country districts or a region.

Parameters

Chemical properties

Degradation half-life in soil (DT50_soil)

Physical/chemical/biological/empirical meaning:

Several processes can contribute to the degradation of a chemical in soils: biodegradation, photolysis, hydrolysis. A major degradation pathway is aerobic or anaerobic biodegradation. For some compound classes (e.g. alkyl halides, carboxylic acids, and organophosphate esters), hydrolysis is an important abiotic degradation process. Indirect photolysis may take place at the soil surface layer via sunlight-excited dissolved organic matter. All these processes are added in the model into an aggregated loss rate. Degradation is assumed to follow linear first-order kinetics.

Factors influencing parameter value:

Microbial degradation can be attenuated by anaerobic conditions, suboptimal temperature and suboptimal pH. Photolysis involves light intensity and is then dependent on light penetration.

Role in the model:

Degradation processes are simulated by pseudo first order degradation rates (i.e. degradation proportional to the concentration of contaminants in the media). They are used in the model to simulate the loss term associated to biodegradation, photolysis and hydrolysis.

Database used for parameter estimation:

Chemical related parameters are extracted from the PPDB database (Lewis et al., 2016). Alternatively, QSAR from the VEGA tool (Benfenati et al., 2013) are used in cases when PPDB doesn't contain required information.

Parameter default value:

DT50_soil parameter value is specified for chemicals selected to be included in the simulation and extracted from the PPDB on a chemical-by-chemical basis.

Degradation half-life in water (DT50_water)

Physical/chemical/biological/empirical meaning:

Several processes can contribute to the degradation of a chemical in water: biodegradation, photolysis, hydrolysis. In water, a major degradation pathway is aerobic biodegradation. For some compound classes (e.g. alkyl halides, carboxylic acids, and organophosphate esters), hydrolysis is an important abiotic degradation process. Indirect photolysis may take place in the aqueous surface layer via sunlight-excited dissolved organic matter. All these processes are added in the model into aggregated loss rates for water. Degradation is assumed to follow linear first-order kinetics.

Factors influencing parameter value:

Microbial degradation can be attenuated by anaerobic conditions, suboptimal temperature and suboptimal pH. Photolysis involves light intensity and is then dependent on light penetration.

Role in the model:

Database used for parameter estimation:

Chemical related parameters are extracted from the PPDB database (Lewis et al., 2016). Alternatively, QSAR from the VEGA tool (Benfenati et al., 2013) are used in cases when PPDB doesn't contain required information.

Parameter default value:

DT50_water parameter value is specified for chemicals selected to be included in the simulation and extracted from the PPDB on a chemical-by-chemical basis.

Water-Organic carbon partitioning coefficient (K_oc)

Physical/chemical/biological/empirical meaning:

Organic carbon is assumed to be the main particulate media interacting with hydrophobic chemicals potentially present in soil. The Water-Organic Carbon partition coefficient represents the ratio at equilibrium of the chemical associated to particulate organic matter and present in water respectively.

Factors influencing parameter value:

The main limitation of the approach described above is that the variability in the soil composition is only described by the content of organic carbon. The validity of this assumption can be disputable especially for ionizable compounds (ter Laak et al., 2006). In particular, the organic carbon content alone as the descriptor of soil composition is not sufficient to predict the soil–water distribution of chemicals that do not exclusively sorb to organic matter. Furthermore, the composition of organic carbon itself can vary substantially and influence sorption. Complexation is another process neglected by the model, although it may have significant impact for some compounds.

Role in the model:

The exchanges of contaminants between Pore Water and Soil particles are assumed to be at equilibrium and represented by a Partition coefficient at equilibrium K_d_soil_water. Interactions of chemicals with particles are assumed to be governed by lipophilic sorption onto organic matter. The K_oc parameter, together with the organic fraction in soil, allows calculating the partition coefficient at equilibrium K_d_soil_water.

Database used for parameter estimation:

Chemical related parameters are extracted from the PPDB database (Lewis et al., 2016). Alternatively, QSAR from the VEGA tool (Benfenati et al., 2013) are used in cases when PPDB doesn't contain required information.

Parameter default value:

K_oc parameter value is specified for chemicals selected to be included in the simulation and extracted from the PPDB on a chemical-by-chemical basis.

Topsoil

Parameters related to calculation of variables required to obtain AS concentration in the topsoil are presented in more details. Description, parameter values with their transformation if applicable and relevant references are provided.

Field area (A_field)

Physical/chemical/biological/empirical meaning:

Area (parcel/plot) on which one crop is cultivated and treated with at least one active substance, and which can be identified by a unique ID.

Factors influencing parameter value:

Area of an individual parcel may undergo changes from year to year.

Role in the model:

It is used to calculated volume of soil when soil organic carbon content is given in stock units, tonnes/ha, for instance in the HYDROSHEDS database. Then, it is used as an upscaling factor to derive fraction of treated area in any target spatial unit.

Database used for parameter estimation:

Country	Sources for parcel maps
Czech Republic 🇨🇿	Czech Ministry of Agriculture (EAGRI) / EuroCrops
Netherlands 🇳🇱	Netherlands Enterprise Agency (RVO) / EuroCrops
Denmark 🇩🇰	EuroCrops
Portugal 🇵🇹	EuroCrops

Table 19 Main sources for AS parcel area data in four selected countries

Parameter default value:

It is country and year specific spatial information for each individual field.

Ecotoxicological averaging depth (H_ top_soil)

Physical/chemical/biological/empirical meaning:

Much of the chemical transformation in general takes place in the top layer of the soil. The top layer is important for its functions it provides to crops; it is also a part of upper soil which is directly affected by agricultural practices and exposed to immediate effects of weather. For the simulated concentration in total soil and the concentration in pore water and particulate matter ecotoxicologically relevant soil depth must be taken into account.

Factors influencing parameter value:

The parameter value is in fact affected by range of different cropping and application systems due to e.g., varying distribution of organic matter with depth in permanent crops and in annual crops.

Role in the model:

It is used to calculate mass of chemical in a top volume, expressed as mass using soil bulk density, of the topsoil.

Database used for parameter estimation:

A set of possible ecotoxicologically relevant depths are considered following the description of the default “Ecotoxicological evaluation depth” values for the topsoil particulate matter and pore water in the PERSAM tool.

Parameter default value:

The EFSA soil exposure guidance (European Food Safety Authority, 2017) proposes to use the assessment depths of 5 cm and 20 cm (the latter in case of soil incorporation) in line with the currently applied procedure (FOCUS, 1997). In fact, default values for ecotoxicologically relevant depth in the topsoil pore and particulate matter are separately defined in the PERSAM tool, having 1cm and 5cm, respectively.

Soil organic carbon content (OC_top_soil)

&

Clay content (Clay_ top_soil)

&

Sand content (Sand_ top_soil)

&

Soil bulk density (rho_ top_soil)

Physical/chemical/biological/empirical meaning:

OC_top_soil describes soil organic carbon (SOC) stored in soil organic matter (SOM) is divided between living soil biota and dead biotic material derived from biomass. Detritus resulting from plant senescence is the major source of soil organic carbon. Plant materials, with cell walls high in cellulose and lignin, are decomposed and retained as humus. Cellulose and starches readily degrade, resulting in short residence times. More persistent forms of organic C include lignin, humus, organic matter encapsulated in soil aggregates, and charcoal. These resist alteration and have long residence times. Soil organic carbon tends to be concentrated in the topsoil. Topsoil ranges from 0.5% to 3.0% organic carbon for most upland soils. Soils with less than 0.5% organic C are mostly limited to desert areas. Soils containing greater than 12–18% organic carbon are generally classified as organic soils. Soil organic carbon (SOC) concentrations in sandy soils influence soil bulk density which decreases with an increase in SOC (Périé & Ouimet, 2008).

SOC influences soil structure by binding soil particles together, improving aggregation, and increasing porosity. This leads to better water infiltration and aeration. SOC affects soil color, often darkening it. SOC is the primary food source for soil microorganisms. It drives microbial activity, which is essential for nutrient cycling and decomposition. Higher SOC supports a diverse and abundant soil food web. SOC acts as a reservoir of essential plant nutrients like nitrogen, phosphorus, and sulfur. It increases the soil's cation exchange capacity (CEC), enhancing its ability to retain nutrients. It buffers soil pH, reducing the impact of acidity or alkalinity.

Clay_ top_soil refers clay that is the smallest soil particles (<0.002 mm). It contributes to soil texture, making soils feel sticky and plastic when wet. Clay has a high surface area, which influences water holding capacity and nutrient retention. Clay can reduce permeability, causing waterlogging. Clay can protect organic matter, affecting microbial activity. It can provide habitats for some soil organisms. Clay minerals have a high CEC, meaning they can hold onto positively charged ions (cations) like calcium, magnesium, and potassium. It influences soil pH buffering capacity.

Sand_ top_soil is the sand content and refers to the largest soil particles (0.05-2 mm). It creates a coarse texture, making soils feel gritty. Sand has low water holding capacity and high permeability. Sandy soils have high aeration. Sandy soils tend to have lower microbial activity due to low water and nutrient retention. They can provide habitats for specific soil organisms adapted to dry conditions. Sand has a low CEC, meaning it has limited ability to hold onto nutrients. It has minimal influence on soil pH buffering.

rho_ top_soil that is soil bulk density is the mass of dry soil per unit volume, including pore space. It indicates soil compaction. High bulk density reduces pore space, limiting water infiltration, aeration, and root growth. High bulk density restricts root growth and reduces soil organism activity. Lower bulk density is essential for healthy plant growth. Bulk density indirectly affects nutrient availability by influencing water and air movement.

Factors influencing parameter value:

With respect to interplay with other parameters SOC is strongly influenced by clay content. Clay can protect organic matter from decomposition. Sand content has an inverse relationship with SOC; sandy soils tend to have lower SOC due to faster decomposition. Bulk density is inversely related to SOC; higher SOC typically leads to lower bulk density. Other Factors having influence on SOC include climate (temperature, rainfall), vegetation type and management, tillage practices, fertilization and erosion.

Regarding the interactions with other parameters clay content directly affects bulk density, generally leading to higher bulk density. It influences SOC storage, as clay can protect organic matter and determines soil texture, which interacts with sand and silt content. Othe factors affecting this parameter parent material (the rock from which the soil formed), weathering processes and time.

Interactions of Sand content with other parameter include inverse relatetionship with clay content and bulk density. It influences SOC storage, generally leading to lower SOC and it greatly impacts soil texture, and therefore water infiltration. Other factors having an impact on sand content include parent material, weathering processes and erosion.

Bulk density is inversely related to SOC and porosity. It is directly related to clay content, as clay tends to increase bulk density. Sand tends to increase bulk density also, but not as much as clay. Other factors affecting soil bulk density are compaction from machinery or livestock, tillage practices, soil texture, organic matter and water content.

Role in the model:

Sand_top_soil, Clay_top_soil, OC_top_soil and rho_top_soil are used to calculate PEC of individual parcels and PECs distribution for the intended use area in each target spatial unit crops are present. More specifically Sand_top_soil and Clay_top_soil are directly used to calculate daily surface runoff depth generated on individual parcels. OC_top_soil is directly used to calculate fraction of active substance bound to soil particulates and dissolved in soil pore water. rho_top_soil together with ecotoxicologically relevant depth (H_top_soil) are directly used in calculating initial PEC.

Database used for parameter estimation:

Topsoil physical properties for Europe (based on LUCAS topsoil data) is a dataset (GIS maps) (2016) consisting of 7 soil property maps that have been derived using soil point data from the LUCAS 2009 soil survey (around 20,000 points) for EU-25, using hybrid approaches like regression kriging (Ballabio et al., 2016). Properties: clay, silt and sand content, coarse fragments, and derived products bulk density, USDA soil textural classes and water capacity, all are available at 500m resolution. These physical properties are generally stable in a relatively long timespan (decades). rho_top_soil was originally obtained by Ballabio et al. (2016), from the packing density and mapped clay content following equation by Jones et al. (2003):

$$\rho_{bulk\ density} = \rho_{packing\ density} - 0.009 \times Clay_{content}$$

The LUCAS data set was created in 2009 in response to the scarcity of harmonised up-to-date OC data. The data set contains some 20,000 soil samples taken in 23 MSs. These spatial layers are highly requested for modelling activities in erosion by water and wind, biodiversity modelling, water capacity, crop growth, vegetation, soil conservation, moisture, land use, ecological analysis, groundwater vulnerability and hydrology.

Parameter default value:

Parameter values of the soil specific parameters are defined for each individual fields by extracting statistics, an average of all pixels cropped with individual parcel polygons and using *zonal()* function.

Crop

Crop name/code (Crop_code)

&

EPPO code (eppo_code)

Physical/chemical/biological/empirical meaning:

Crop names and codes represent national systems for naming and coding of individual crops and crop groups.

EPPO codes (Ayllón-Benitez et al., 2023).

Factors influencing parameter value:

Role in the model:

Crop names and codes are used to assign crop specific variables such as interception factors and plant protection product usage data. EPPO codes are mostly used as a key for joining various datasets such as national crop names and codes with the Focus crops. National crop codes are used predominantly for mapping. Whenever possible they were attributed with respective EPPO codes. Additionally, EuroCrop codes were used as crop attribute to further harmonise crop coding across countries.

Database used for parameter estimation:

Since Focus Crops cover only 31 generic crop groups the goal of assigning EPPO codes to national crops was to roughly match them by plant family or genus, and species, taxon where possible with the focus group to obtain interception factor. The EPPO codes were attributed to specific plants in order to find a match with FOCUS crop groups according to van de Zande & ter Horst (2019). The matching system developed therein was then used as an indication for finding connections between various coding and naming sets.

Crop_name_pestdose_en	Crop_name_pestdose_nl	Crop_pestdose_code_eppo_focus	Crop_name_rvo_nl	Crop_code_rvo	EC_hcat_n	EC_trans_n	Focus_crop	Crop_code_pestdose_eppo_focus
Carrots	Bos- en waspeen	Carrot production	Waspeen, productie	2783	carrots_daucus	Carrot, production	Carrots	DAUCS
Carrots	Bos- en waspeen	Carrot, seeds and propagating material	Waspeen, zaden en opkweekmateriaal	2784	carrots_daucus	Carrot, seeds and propagating material	Carrots	DAUCS

Table 20 Extract from the compiled dataset of the Dutch crop names, groups, codes matched with the EPPO codes, FOCUS crops and EuroCrops. Column explanation: 1) <u>Crop_name_pestdose_en</u>: names of crops in English with calculated pesticide application dose, 2) <u>Crop_name_pestdose_nl:</u> names of crops in Dutch with calculated pesticide application dose, 3) <u>Crop_code_rvo:</u> crop codes according to the RVO, 4) Crop_name_rvo_nl names of crops in Dutch according to the RVO, 5) EC_hcat_n: EuroCrop names, 6) <u>EC_trans_n:</u> EuroCrop translations, correspond directly with english names according to the RVO,7) <u>Crop_code_pestdose_eppo_focus</u>: EPPO codes attributed to individual or group of crops (obtained from the Statistic Netherlands (CBS)).

An example of data processing involving Dutch crops. In cases of plant groups without an EPPO code but matched to generic Focus groups, the EPPO code (Crop_code_ pestdose _eppo_focus) is assigned higher taxa to reflect position of a respective Focus crop e.g., 1) Shrubs and hedges (1067/1081) were assigned to Beans (Focus_group) by code 1LEGF (family *Fabaceae*), 2) Ornamental crops such as flower bulb and flower tuber cultivation or cut flowers (1010/1013/1016…) were oftentimes assigned, genus e.g., 1ALLG, 1NARG, 1AMYG, otherwise species level EPPO code was used were possible. Another problematic instance where given EPPO code isn't assigned to Focus crop. e.g., Cucumber, CUMSA. It belongs to fruiting vegetables (4.3 DTG), together with zucchini and Gherkin assigned with CUUPG/CUUPM and CUMSG EPPO codes and matched with Potato Focus group. For simplicity cucumber was assigned together with zucchini and Gherkin to Potatoes. This reflects the characteristic of the CBS crop grouping which may represent several individual crops and/or cultivation methods. Matching multiple plants from the national crop map to generic plant groups from CBS, e.g., summer/winter oilseed rape (1923/1922) and total rapeseed. Another similar example involved total pepper, total tomato, total seed potato. Some were deleted e.g., Avenue and park trees deemed as irrelevant.

The EuroCrops project developed a Hierarchical Crop and Agricultural Taxonomy (HCAT) following EAGLE principles (Arnold et al., 2013) to harmonize all crop classifications. The HCAT classifies each crop into a 10-digit code and a related readable name. It consists of 6 nested levels, with higher levels containing broader information of land use and lower levels containing very detailed levels of crops.

EPPO Global Database (GD). It provides basic information for species of interest to agriculture, forestry and plant protection: plants (cultivated and wild) and pests (including pathogens and invasive alien plants). For each species: scientific names, synonyms, common names in different languages, taxonomic position, and EPPO Codes are given.

Parameter default value:

EPPO, national and EuroCrop crop codes and names are extracted on a crop-by-crop and country-by-country basis for individual crops and crop groups.

Crop interception factor (IF_crop)

Physical/chemical/biological/empirical meaning:

Interception factor IF_crop is an important determinant influencing exposure of terrestrial and aquatic organism to applied PPPs via spraying. Interception by crops reduces the amount of the plant protection product that reaches the ground underneath the crop. Only the plant protection products that reach the ground are taken into account in regulatory calculations of predicted environmental concentrations (PECs) in soil, surface water and groundwater.

Factors influencing parameter value:

Crop interception is related to crop growth development. The seasonal trend of the crop development stage depends to a large extent on climatological conditions. Crop interception should not be included in calculations for ridge–furrow systems unless the spray is targeted on just the crop canopy or the crop canopy has closed between the rows (European Food Safety Authority, 2017).

Role in the model:

IF_crop is used to calculate fraction of application dose reaching soil after interception by the plant canopy.

Database used for parameter estimation:

There are two major sources for developing the lookup table of interception factors implemented in the model. First one is the original updated FOCUS crops (Olesen & Jensen, 2013) and the second one matched with EPPO codes according to van de Zande & ter Horst (2019).

Parameter default value:

crop_name_focus	crop_code_eppo	crop_focus_comment	inter_frac_bbch.ave	inter_frac_bbch.min	inter_frac_bbch.max
Apples	1MSPG	application on crop in row	0.6	0.5	0.65
Apples	1MSPG	between crop in row, grass	0.9	0.9	0.9
Apples	1ATIG	application on crop in row	0.6	0.5	0.65
Apples	1ATIG	between crop in row, grass	0.9	0.9	0.9
Apples	MABSD	application on crop in row	0.6	0.5	0.65
Apples	MABSD	between crop in row, grass	0.9	0.9	0.9
Apples	CRWMS	application on crop in row	0.6	0.5	0.65
Apples	CRWMS	between crop in row, grass	0.9	0.9	0.9
Apples	1MORG	application on crop in row	0.6	0.5	0.65
Apples	1MORG	between crop in row, grass	0.9	0.9	0.9
Apples	ROSVL	application on crop in row	0.6	0.5	0.65
Apples	ROSVL	between crop in row, grass	0.9	0.9	0.9
Apples	MAHAQ	application on crop in row	0.6	0.5	0.65
Apples	MAHAQ	between crop in row, grass	0.9	0.9	0.9
Apples	CYLAV	application on crop in row	0.6	0.5	0.65
Apples	CYLAV	between crop in row, grass	0.9	0.9	0.9
Apples	CSNSS	application on crop in row	0.6	0.5	0.65
Apples	CSNSS	between crop in row, grass	0.9	0.9	0.9
Apples	IUGRE	application on crop in row	0.6	0.5	0.65
Apples	IUGRE	between crop in row, grass	0.9	0.9	0.9
Apples	FIUCA	application on crop in row	0.6	0.5	0.65
Apples	FIUCA	between crop in row, grass	0.9	0.9	0.9
Apples	ATIDE	application on crop in row	0.6	0.5	0.65
Apples	ATIDE	between crop in row, grass	0.9	0.9	0.9
Apples	PYUCO	application on crop in row	0.6	0.5	0.65
Apples	PYUCO	between crop in row, grass	0.9	0.9	0.9
Apples	PYUPC	application on crop in row	0.6	0.5	0.65
Apples	PYUPC	between crop in row, grass	0.9	0.9	0.9
Apples	CYDOB	application on crop in row	0.6	0.5	0.65
Apples	CYDOB	between crop in row, grass	0.9	0.9	0.9
Apples	MSPGE	application on crop in row	0.6	0.5	0.65
Apples	MSPGE	between crop in row, grass	0.9	0.9	0.9
Apples	PRNAV	application on crop in row	0.6	0.5	0.65
Apples	PRNAV	between crop in row, grass	0.9	0.9	0.9
Apples	PRNCE	application on crop in row	0.6	0.5	0.65
Apples	PRNCE	between crop in row, grass	0.9	0.9	0.9
Apples	PRNDO	application on crop in row	0.6	0.5	0.65
Apples	PRNDO	between crop in row, grass	0.9	0.9	0.9
Apples	PRNDT	application on crop in row	0.6	0.5	0.65
Apples	PRNDT	between crop in row, grass	0.9	0.9	0.9
Apples	PRNAR	application on crop in row	0.6	0.5	0.65
Apples	PRNAR	between crop in row, grass	0.9	0.9	0.9
Apples	PRNPS	application on crop in row	0.6	0.5	0.65
Apples	PRNPS	between crop in row, grass	0.9	0.9	0.9
Apples	PRNPN	application on crop in row	0.6	0.5	0.65
Apples	PRNPN	between crop in row, grass	0.9	0.9	0.9
Beans	LACSA	field, veg.	0.43	0	0.8
Beans	LACSC	field, veg.	0.43	0	0.8
Beans	LACSP	field, veg.	0.43	0	0.8
Beans	LACSO	field, veg.	0.43	0	0.8

Beans	1LEGF	field, veg.	0.43	0	0.8
Beans	1VICG	field, veg.	0.43	0	0.8
Beans	PHSVX	field, veg.	0.43	0	0.8
Beans	VICFX	field, veg.	0.43	0	0.8
Beans	LACSS	field, veg.	0.43	0	0.8
Beans	CICEN	field, veg.	0.43	0	0.8
Beans	SPQOL	field, veg.	0.43	0	0.8
Beans	TEATE	field, veg.	0.43	0	0.8
Beans	BRSRE	field, veg.	0.43	0	0.8
Beans	AMADU	field, veg.	0.43	0	0.8
Beans	ATXHO	field, veg.	0.43	0	0.8
Beans	POROS	field, veg.	0.43	0	0.8
Beans	CLAPE	field, veg.	0.43	0	0.8
Beans	NAAOF	field, veg.	0.43	0	0.8
Beans	VLLLO	field, veg.	0.43	0	0.8
Beans	ERUVE	field, veg.	0.43	0	0.8
Beans	ASTTR	field, veg.	0.43	0	0.8
Beans	LEPSA	field, veg.	0.43	0	0.8
Beans	PHSVN	field, veg.	0.43	0	0.8
Beans	PHSCO	field, veg.	0.43	0	0.8
Beans	VIGSC	field, veg.	0.43	0	0.8
Beans	VIGSI	field, veg.	0.43	0	0.8
Beans	PHSLU	field, veg.	0.43	0	0.8
Beans	ASPOF	field, veg.	0.43	0	0.8
Beans	APUGD	field, veg.	0.43	0	0.8
Beans	CYUCA	field, veg.	0.43	0	0.8
Beans	RHERH	field, veg.	0.43	0	0.8
Beans	ALLPO	field, veg.	0.43	0	0.8
Beans	CRMMA	field, veg.	0.43	0	0.8
Beans	OCIBA	field, veg.	0.43	0	0.8
Beans	ALLSC	field, veg.	0.43	0	0.8
Beans	ALLTU	field, veg.	0.43	0	0.8
Beans	STIHO	field, veg.	0.43	0	0.8
Beans	STIMO	field, veg.	0.43	0	0.8
Beans	MLSOF	field, veg.	0.43	0	0.8
Beans	AFEGR	field, veg.	0.43	0	0.8
Beans	ARTDR	field, veg.	0.43	0	0.8
Beans	HYSOF	field, veg.	0.43	0	0.8
Beans	ARNCE	field, veg.	0.43	0	0.8
Beans	CORSA	field, veg.	0.43	0	0.8
Beans	PARCR	field, veg.	0.43	0	0.8
Beans	LEWOF	field, veg.	0.43	0	0.8
Beans	MAHJO	field, veg.	0.43	0	0.8
Beans	ORIVU	field, veg.	0.43	0	0.8
Beans	MENSS	field, veg.	0.43	0	0.8
Beans	SANMI	field, veg.	0.43	0	0.8
Beans	RMSOF	field, veg.	0.43	0	0.8
Beans	SALOF	field, veg.	0.43	0	0.8
Beans	THYVU	field, veg.	0.43	0	0.8
Beans	FOEVD	field, veg.	0.43	0	0.8
Beans	APUGS	field, veg.	0.43	0	0.8
Beans	RUMAC	field, veg.	0.43	0	0.8
Beans	CUUPG	field, veg.	0.43	0	0.8
Beans	TAGER	field, veg.	0.43	0	0.8
Beans	TOPMA	field, veg.	0.43	0	0.8
Beans	CLDOF	field, veg.	0.43	0	0.8
Beans	ANKAR	field, veg.	0.43	0	0.8
Beans	PIMSA	field, veg.	0.43	0	0.8
Beans	PARCT	field, veg.	0.43	0	0.8
Beans	LOBIN	field, veg.	0.43	0	0.8
Beans	DIKLA	field, veg.	0.43	0	0.8
Beans	VIOTR	field, veg.	0.43	0	0.8
Beans	MATCH	field, veg.	0.43	0	0.8
Beans	RUDPU	field, veg.	0.43	0	0.8
Beans	VALOF	field, veg.	0.43	0	0.8
Beans	PNXGI	field, veg.	0.43	0	0.8
Bushberries	1RUBG	application on crop in row	0.583333	0.4	0.75
Bushberries	1RUBG	between crop in row, grass	0.9	0.9	0.9
Bushberries	1SOUG	application on crop in row	0.583333	0.4	0.75
Bushberries	1SOUG	between crop in row, grass	0.9	0.9	0.9
Bushberries	1LONG	application on crop in row	0.583333	0.4	0.75
Bushberries	1LONG	between crop in row, grass	0.9	0.9	0.9
Bushberries	AMEAL	application on crop in row	0.583333	0.4	0.75
Bushberries	AMEAL	between crop in row, grass	0.9	0.9	0.9
Bushberries	RIBNG	application on crop in row	0.583333	0.4	0.75
Bushberries	RIBNG	between crop in row, grass	0.9	0.9	0.9
Bushberries	SOUDO	application on crop in row	0.583333	0.4	0.75
Bushberries	SOUDO	between crop in row, grass	0.9	0.9	0.9
Bushberries	SOUAU	application on crop in row	0.583333	0.4	0.75
Bushberries	SOUAU	between crop in row, grass	0.9	0.9	0.9
Bushberries	ABOME	application on crop in row	0.583333	0.4	0.75
Bushberries	ABOME	between crop in row, grass	0.9	0.9	0.9
Bushberries	RIBAU	application on crop in row	0.583333	0.4	0.75
Bushberries	RIBAU	between crop in row, grass	0.9	0.9	0.9
Bushberries	SAMNI	application on crop in row	0.583333	0.4	0.75
Bushberries	SAMNI	between crop in row, grass	0.9	0.9	0.9
Bushberries	RIBRU	application on crop in row	0.583333	0.4	0.75
Bushberries	RIBRU	between crop in row, grass	0.9	0.9	0.9
Bushberries	RIBNI	application on crop in row	0.583333	0.4	0.75
Bushberries	RIBNI	between crop in row, grass	0.9	0.9	0.9
Bushberries	RIBUC	application on crop in row	0.583333	0.4	0.75
Bushberries	RIBUC	between crop in row, grass	0.9	0.9	0.9
Bushberries	VACCO	application on crop in row	0.583333	0.4	0.75
Bushberries	VACCO	between crop in row, grass	0.9	0.9	0.9

Bushberries	VACMY	application on crop in row	0.583333	0.4	0.75
Bushberries	VACMY	between crop in row, grass	0.9	0.9	0.9
Bushberries	VACVI	application on crop in row	0.583333	0.4	0.75
Bushberries	VACVI	between crop in row, grass	0.9	0.9	0.9
Bushberries	VACOX	application on crop in row	0.583333	0.4	0.75
Bushberries	VACOX	between crop in row, grass	0.9	0.9	0.9
Bushberries	VACMA	application on crop in row	0.583333	0.4	0.75
Bushberries	VACMA	between crop in row, grass	0.9	0.9	0.9
Bushberries	MORSS	application on crop in row	0.583333	0.4	0.75
Bushberries	MORSS	between crop in row, grass	0.9	0.9	0.9
Bushberries	ROSSS	application on crop in row	0.583333	0.4	0.75
Bushberries	ROSSS	between crop in row, grass	0.9	0.9	0.9
Bushberries	ATIAR	application on crop in row	0.583333	0.4	0.75
Bushberries	ATIAR	between crop in row, grass	0.9	0.9	0.9
Bushberries	SAMSS	application on crop in row	0.583333	0.4	0.75
Bushberries	SAMSS	between crop in row, grass	0.9	0.9	0.9
Bushberries	ABOSS	application on crop in row	0.583333	0.4	0.75
Bushberries	ABOSS	between crop in row, grass	0.9	0.9	0.9
Bushberries	HIORH	application on crop in row	0.583333	0.4	0.75
Bushberries	HIORH	between crop in row, grass	0.9	0.9	0.9
Bushberries	LONCO	application on crop in row	0.583333	0.4	0.75
Bushberries	LONCO	between crop in row, grass	0.9	0.9	0.9
Bushberries	VITVI	application on crop in row	0.583333	0.4	0.75
Bushberries	VITVI	between crop in row, grass	0.9	0.9	0.9
Bushberries	RUBFR	application on crop in row	0.583333	0.4	0.75
Bushberries	RUBFR	between crop in row, grass	0.9	0.9	0.9
Bushberries	RUBID	application on crop in row	0.583333	0.4	0.75
Bushberries	RUBID	between crop in row, grass	0.9	0.9	0.9
Bushberries	RUBPH	application on crop in row	0.583333	0.4	0.75
Bushberries	RUBPH	between crop in row, grass	0.9	0.9	0.9
Bushberries	RUBCA	application on crop in row	0.583333	0.4	0.75
Bushberries	RUBCA	between crop in row, grass	0.9	0.9	0.9
Bushberries	RUBLO	application on crop in row	0.583333	0.4	0.75
Bushberries	RUBLO	between crop in row, grass	0.9	0.9	0.9
Bushberries	CAHSI	application on crop in row	0.583333	0.4	0.75
Bushberries	CAHSI	between crop in row, grass	0.9	0.9	0.9
Cabbage	BRSOL	NA	0.491667	0	0.9
Cabbage	BRSOF	NA	0.491667	0	0.9
Cabbage	BRSOB	NA	0.491667	0	0.9
Cabbage	BRSOK	NA	0.491667	0	0.9
Cabbage	BRSPK	NA	0.491667	0	0.9
Cabbage	BRSOC	NA	0.491667	0	0.9
Cabbage	BRSOG	NA	0.491667	0	0.9
Carrots	DAUCS	NA	0.541667	0	0.8
Carrots	SIUSI	NA	0.541667	0	0.8
Carrots	PARCT	NA	0.541667	0	0.8
Carrots	PAVSA	NA	0.541667	0	0.8
Carrots	TARKS	NA	0.541667	0	0.8
Citrus	NA	application on crop in row	0.8	0.8	0.8
Cotton	1MALG	NA	0.51	0	0.9
Grass	1GRAF	pasture	0.9	0.9	0.9
Grass	LOLPE	pasture	0.9	0.9	0.9
Grass	GGGPE	pasture	0.9	0.9	0.9
Grass	LOLMU	pasture	0.9	0.9	0.9
Grass	ARREL	pasture	0.9	0.9	0.9
Grass	LOLMG	pasture	0.9	0.9	0.9
Grass	LOLBO	pasture	0.9	0.9	0.9
Grass	FESRU	pasture	0.9	0.9	0.9
Grass	FESOV	pasture	0.9	0.9	0.9
Grass	FESAR	pasture	0.9	0.9	0.9
Grass	POARR	pasture	0.9	0.9	0.9
Grass	POAPA	pasture	0.9	0.9	0.9
Grass	POANE	pasture	0.9	0.9	0.9
Grass	FESPR	pasture	0.9	0.9	0.9
Grass	PHLPR	pasture	0.9	0.9	0.9
Grass	DACGL	pasture	0.9	0.9	0.9
Grass	AGSSS	pasture	0.9	0.9	0.9
Grass	CYXCR	pasture	0.9	0.9	0.9
Grass	DECCA	pasture	0.9	0.9	0.9
Grass	KOLSS	pasture	0.9	0.9	0.9
Grass	LOLSS	pasture	0.9	0.9	0.9
Grass	AVESG	pasture	0.9	0.9	0.9
Grass	PHCTA	pasture	0.9	0.9	0.9
Grass	SPRAR	pasture	0.9	0.9	0.9
Grass	TAGER	pasture	0.9	0.9	0.9
Grass	SOLSI	pasture	0.9	0.9	0.9
Grass	SORSU	pasture	0.9	0.9	0.9
Grass	CMASA	pasture	0.9	0.9	0.9
Grass	BRSRR	pasture	0.9	0.9	0.9
Grass	ERUVE	pasture	0.9	0.9	0.9
Grass	GUIAB	pasture	0.9	0.9	0.9
Hops	HUMLU	application on crop in row	0.575	0.4	0.75
Hops	HUMLU	between crop in row, bare soil	0	0	0
Lineseed	PAPSO	NA	0.5	0	0.9
Lineseed	CRYCA	NA	0.5	0	0.9
Lineseed	LIUUT	NA	0.5	0	0.9
Lineseed	OEOSS	NA	0.5	0	0.9
Lineseed	CMASA	NA	0.5	0	0.9
Lineseed	CRMAB	NA	0.5	0	0.9
Lineseed	TRFPR	NA	0.5	0	0.9
Lineseed	TRFRE	NA	0.5	0	0.9
Lineseed	TRFHY	NA	0.5	0	0.9
Lineseed	TRFIN	NA	0.5	0	0.9
Lineseed	TRFRS	NA	0.5	0	0.9

Lineseed	TRFAL	NA	0.5	0	0.9
Lineseed	MEDLU	NA	0.5	0	0.9
Lineseed	LOTCO	NA	0.5	0	0.9
Lineseed	MEUAL	NA	0.5	0	0.9
Lineseed	LUPSS	NA	0.5	0	0.9
Lineseed	VICSA	NA	0.5	0	0.9
Lineseed	OROSA	NA	0.5	0	0.9
Lineseed	CEOAR	NA	0.5	0	0.9
Lineseed	ONBVI	NA	0.5	0	0.9
Maize	ZEAMX	NA	0.48	0	0.9
Maize	ZEAMS	NA	0.48	0	0.9
Oilseed rape	BRSNS	NA	0.58	0	0.9
Oilseed rape	BRSNW	NA	0.58	0	0.9
Oilseed rape	SINAL	NA	0.58	0	0.9
Oilseed rape	BRSNI	NA	0.58	0	0.9
Oilseed rape	RAPSA	NA	0.58	0	0.9
Oilseed rape	BRSOM	NA	0.58	0	0.9
Oilseed rape	CRYCA	NA	0.58	0	0.9
Oilseed rape	PAPSO	NA	0.58	0	0.9
Olives	NA	application on crop in row	0.7	0.7	0.7
Olives	NA	between crop in row, bare soil	0	0	0
Onions	1ALLG	NA	0.291667	0	0.6
Onions	ALLCE	NA	0.291667	0	0.6
Onions	ALLAS	NA	0.291667	0	0.6
Onions	ALLSA	NA	0.291667	0	0.6
Onions	1PAEF	NA	0.291667	0	0.6
Onions	1GLAG	NA	0.291667	0	0.6
Onions	1NARG	NA	0.291667	0	0.6
Onions	1TULG	NA	0.291667	0	0.6
Onions	1DAHG	NA	0.291667	0	0.6
Onions	1IRIG	NA	0.291667	0	0.6
Onions	1ALLG	NA	0.291667	0	0.6
Onions	1CHYG	NA	0.291667	0	0.6
Onions	MUSBO	NA	0.291667	0	0.6
Onions	1HKYG	NA	0.291667	0	0.6
Onions	1CVOG	NA	0.291667	0	0.6
Onions	1ZNTG	NA	0.291667	0	0.6
Onions	1AMYG	NA	0.291667	0	0.6
Onions	1LILF	NA	0.291667	0	0.6
Onions	MUSCO	NA	0.291667	0	0.6
Peas	PIBSA	NA	0.52	0	0.85
Peas	PIBSZ	NA	0.52	0	0.85
Peas	LENCU	NA	0.52	0	0.85
Peas	CIAER	NA	0.52	0	0.85
Peas	TTGPU	NA	0.52	0	0.85
Peas	PIBSX	NA	0.52	0	0.85
Potatoes	1SOLG	NA	0.42	0	0.85
Potatoes	SOLTU	NA	0.42	0	0.85
Potatoes	CUMSG	NA	0.42	0	0.85
Potatoes	CUUPG	NA	0.42	0	0.85
Potatoes	CUMSA	NA	0.42	0	0.85
Potatoes	CUUPM	NA	0.42	0	0.85
Potatoes	CUUPE	NA	0.42	0	0.85
Soybean	GLXMA	NA	0.48	0	0.85
Strawberry	FRAAN	NA	0.4	0	0.6
Sugar beets	BEAVV	NA	0.6	0	0.9
Sugar beets	BEAVA	NA	0.6	0	0.9
Sugar beets	BEAVV	NA	0.6	0	0.9
Sugar beets	CICIF	NA	0.6	0	0.9
Sugar beets	RAPSR	NA	0.6	0	0.9
Sugar beets	RAPSN	NA	0.6	0	0.9
Sugar beets	BRSOG	NA	0.6	0	0.9
Sugar beets	BRSNA	NA	0.6	0	0.9
Sugar beets	HELTU	NA	0.6	0	0.9
Sugar beets	STASB	NA	0.6	0	0.9
Sugar beets	IPOBA	NA	0.6	0	0.9
Sugar beets	BEAVD	NA	0.6	0	0.9
Sugar beets	APUGR	NA	0.6	0	0.9
Sugar beets	SCVHI	NA	0.6	0	0.9
Sugar beets	TROPS	NA	0.6	0	0.9
Sugar beets	ARWLA	NA	0.6	0	0.9
Sugar beets	SIUSS	NA	0.6	0	0.9
Sugar beets	FOEVA	NA	0.6	0	0.9
Sunflowers	HELAN	NA	0.47	0	0.9
Sunflowers	CNISA	NA	0.47	0	0.9
Sunflowers	URTSS	NA	0.47	0	0.9
Sunflowers	BRSJU	NA	0.47	0	0.9
Sunflowers	BRSRA	NA	0.47	0	0.9
Sunflowers	1PAPG	NA	0.47	0	0.9
Sunflowers	CAUTI	NA	0.47	0	0.9
Sunflowers	GUIAB	NA	0.47	0	0.9
Sunflowers	BRSNI	NA	0.47	0	0.9
Tobacco	LOBIN	NA	0.6	0	0.9
Tomato	LYPES	NA	0.5	0	0.8
Tomato	VANPL	NA	0.5	0	0.8
Vines	1VITG	application on crop in row	0.57	0.4	0.75
Vines	1VITG	between crop in row, bare soil	0	0	0
Vines	VITVI	application on crop in row	0.57	0.4	0.75
Vines	VITVI	between crop in row, bare soil	0	0	0
Wheat	TRZAW	winter	0.5	0	0.9
Oats	AVESA	NA	0.5	0	0.9
Rye	SECCW	winter	0.5	0	0.9
Triticale	TTLWI	winter	0.5	0	0.9
Wheat	TRZAS	spring	0.5	0	0.9

Rye	SECCS	spring	0.5	0	0.9
Barley	HORVW	winter	0.5	0	0.9
Spelt	TRZSP	winter	0.5	0	0.9
Canary grass	PHACA	winter	0.5	0	0.9
Barley	HORVS	spring	0.5	0	0.9

Table 21

Hydrology

River basin area (A_basin)

Physical/chemical/biological/empirical meaning:

The area of land from which all surface run-off flows through a sequence of streams, rivers and, possibly, lakes into the sea at a single river mouth, estuary or delta.

Factors influencing parameter value:

Sub basin breakdown of sub-watersheds depends on confluence where the inflowing branches (i.e., a tributary and its main stem) exceed a certain size threshold, typically measured as the number of upstream pixels or the upstream catchment area. HydroBASINS follows the same concept and divides a basin into two sub-basins at every location where two river branches meet which each have an individual upstream area of at least 100 km2. It should be noted that this concept still allows for smaller sub-basins to occur, namely the inter-basins between the tributaries (which can have any smaller size).

A second critical feature of sub-basin delineations is the way the sub-basins are grouped or coded to allow for the breakout of nested sub-basins at different scales, or to navigate within the sub-basin network from up- to downstream. One of the easiest methods for navigation is to provide the ID of the next downstream object, which allows for moving from object to object in order to traverse the network.

As for nesting and topological concepts, the 'Pfafstetter' coding system is frequently used due to its relative simplicity and ease of application. The basic principle of the Pfafstetter coding is that a larger basin is sequentially subdivided into 9 smaller units (the 4 largest tributaries, coded with even numbers, and the 5 inter-basins, coded with odd numbers) Figure 4. The HydroBASINS product follows the Pfafstetter concept and provides levels 1 to 12 globally.

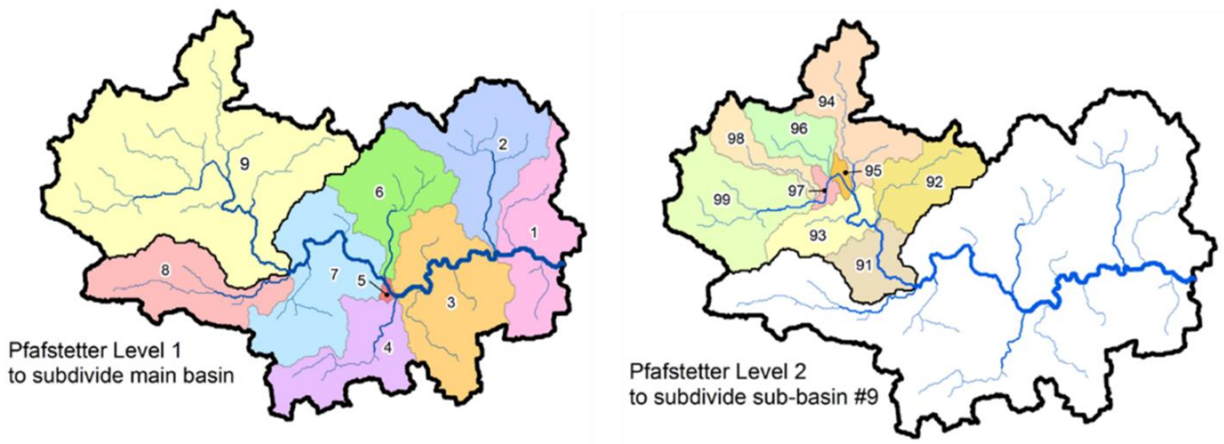


Figure 4 Nested grouping and coding applied to delineating watersheds into smaller units.

Role in the model:

River basin polygons are basic area units for any crop, AS, topsoil and river water related data aggregation and zonal processing.

Database used for parameter estimation:

The HYDROSHEDS database (version 1.0) and its product HYDROBASINS in particular have been used predominantly for obtaining river basins vector data. It consists of a series of polygon layers that depict watershed boundaries and sub-basin delineations at a global scale. The goal of this product is to provide a seamless global coverage of consistently sized and hierarchically nested sub-basins at different scales (from tens to millions of square kilometers), supported by a coding scheme that allows for analysis of watershed topology such as up- and downstream connectivity (Lehner & Grill, 2013).

The HydroSHEDS database provides hydrographic data layers that allow for the derivation of watershed boundaries for any given location based on the near-global, high-resolution SRTM digital elevation model. Using this hydrographic information, watersheds were delineated in a consistent manner at different scales, and a hierarchical sub-basin breakdown was created following the topological concept of the Pfafstetter coding system (Verdin & Verdin, 1999). The resulting polygon layers are termed HydroBASINS and represent a subset of the HydroSHEDS database.

Country	Sources for river basin spatial data
Czech Republic	The HydroSHEDS database
Netherlands	
Denmark	
Portugal	

Table 22 Data sources used for obtaining country level river basin polygon data.

Parameter default value:

Polygons of individual river basins are extracted from the HydroBASINS database for individual countries together with the required attributes such as river water volume, discharge, total length etc.

River discharge (Q_river)

&

River width (W_river)

&

River length (L_river)

Physical/chemical/biological/empirical meaning:

Watercourse network consists of a natural or man-made flowing watercourse or stream. W_river and L_river (and depth, not considered here) are two of the attributes used to describe river channel geometry.

Q_river is the volumetric flow rate (L T^{-1}). For the steady state flow rate, it is often calculated with the continuity equation. Water flow rate in rivers is governed by hydrological regime and is relevant for assessing ecological and chemical status and ecological potential of flowing surface water bodies.

Factors influencing parameter value:

Channel geometry is closely related to river discharge (Allen et al., 1994). A river's drainage area and discharge typically increase as tributaries join the main channel. This increase in discharge is accommodated by an increase in channel hydraulic geometry. Such close dependency of channel dimensions has been suggested also for other river characteristics such as sediment load, bed load, band and bed material composition, streamside vegetation, and channel alignment. In general, the relationship between channel morphology and transport of water and sediment is likely to be explained by (quasi)equilibrium that develops between channel's discharge (Q, m3/s), width (W, m), depth (D, m), and velocity. These variables can be interrelated in the form of three simple exponential equations of the general form $Y = a\mathrm{X}^b$, where X is river discharge (Q) and Y width, depth or velocity taking coefficients *a* and exponents *b* specific to each of the three relationships. The continuity equation of the form $Q = W \times D \times V$ must be satisfied.

Role in the model:

Length and width of the river segment are used to delineate polygons around river segments, which then are used as river buffer to detect and intersect all agricultural fields present within buffer. River length is also used as an upscaling factor for calculating weighted quantities, such $PEC_{\text{surface water}}$ for individual river basins.

Database used for parameter estimation:

Georeferenced river network, and hydrological elements metadata available from national cadastre authorities were preferred such as Czech and Dutch geoportals ČÚZK and PDOK, respectively. These data are characterised by fine resolution and respecting INSPIRE directive data quality requirements. If such data are not available appropriate datasets from the HYDROSHEDS database and its products were extracted. In particular discharge and runoff, which estimates for HydroATLAS are based on long-term (1971–2000) average 'naturalized' discharge and runoff values provided by the state-of-the-art global integrated water balance model WaterGAP (Döll et al., 2003, model version 2.2 as of 2014). The WaterGAP data were spatially downscaled from their original 0.5 degree pixel resolution (~50 km at the equator) to the 15 arc-second (~500 m) resolution of the HydroSHEDS river network using geo-statistical techniques (Lehner & Grill, 2013). Preliminary tests against approximately 3000 global gauging stations indicate a good overall correlation for the long-term averages, but also reveal larger uncertainties, in particular in the minimum and maximum statistics, for areas that are dominated by snow, glaciers, wetlands, and (semi-)arid conditions.

Country	Sources for parcel maps
Czech Republic 🇨🇿	Geoportál ČÚZK
Netherlands 🇳🇱	Publieke Dienstverlening Op de Kaart (PDOK)
Denmark 🇩🇰	The HydroSHEDS database
Portugal 🇵🇹	The HydroSHEDS database

Table 23

Parameter default value:

River segment buffer width (W_buffer)

Physical/chemical/biological/empirical meaning:

The width of individual rivers and river segments has to be specified in order to delineate polygons around them and calculate area for each element.

Factors influencing parameter value:

Field connectivity can be considered when deciding the width of the buffer.

Role in the model:

Width of the buffer affect if and how much of agricultural area surround given river segment. The total field area within the river buffer is likely to contribute under certain weather and seasonal conditions to AS inputs into surface river water.

Database used for parameter estimation:

It is user defined based on judgment.

Parameter default value:

It is user defined based on judgment.

Topography

Terrain slope (Slope_terrain)

Physical/chemical/biological/empirical meaning:

Terrain slope refers to the steepness or inclination of a land surface. It's the rate of change in elevation over a given horizontal distance. It's often expressed as a percentage, a ratio, or in degrees. A steeper slope means a more rapid change in elevation.

Factors influencing parameter value:

Role in the model:

Slope is a significant factor determining the speed and volume of overland surface runoff. Steeper slopes lead to faster runoff velocities. Water accelerates as it moves down a steeper incline due to gravity. This results in greater volumes of runoff, as water has less time to infiltrate the soil what may increase the erosive power of runoff, leading to soil erosion.

Gentler Slopes cause slower runoff velocities and allow more time for water to infiltrate the soil, reducing the volume of surface runoff and reducing the risk of soil erosion.

Database used for parameter estimation:

Terrain slope was extracted from The Harmonized World Soil Database. HWSD v2.0 is a 30 arc-second raster database (Fischer et al., 2021) containing a comprehensive global soil inventory that offers detailed insights into soil properties, including their morphology, chemistry, and physical characteristics, with a focus on a 1 km resolution. The HWSD v2.0 features a GIS raster image file that is connected to a soil attribute database. This database provides comprehensive information on the composition of soil units in nearly 30,000 soil mapping units.

The original Digital Elevation Model data (DEM), provided by the NASA Shuttle Radar Topographic Mission (SRTM), have been mosaiced into a seamless global coverage, and are available for download as 5˚ x 5˚ tiles, in geographic coordinate system - WGS84 datum. The available data cover a raster of 24 rows by 72 columns of 5˚ x 5˚ latitude/longitude tiles, from north 60 degree latitude to 56 degree south.

These processed SRTM data, with a resolution of 3 arc second (approximately 90m at the equator), i.e. 6000 rows by 6000 columns for each 5˚ x 5˚ tile, have been used for calculating: (i) terrain slope gradients for each 3 arc-sec grid cell; (ii) aspect of terrain slopes for each 3 arc-sec grid cell; (iii) terrain slope class by 3 arc-sec grid cell; and (iv) aspect class of terrain slope by 3 arc-sec grid cell. Products (iii) and (iv) were then aggregated to provide distributions of slope gradient and slope aspect classes by 30 arc-sec grid cell and for a 5'x5' latitude/longitude grid.

Parameter default value:

The complete dataset comprises one elevation map describing median elevation in each grid cell, eight slope and four aspect maps describing percentage distributions of the respective slope or aspect classes. The sum of all classes for slopes and aspects respectively is 100 percentages.

LEVEL 4. Equations

Intermediate variables

Site-specific application rate at field level

Calculation schematics related to obtaining field-level application rates are repeated here and complemented by formulas.

Field-crop usage area (A_field_crop_AS)

This process involves operations on relational data. It represents matching geo-referenced crops on individual fields, with not geo-referenced PPP usage data for each crop-AS unique combinations, and with EPPO crop ID for matching additional attributes. The key variable for joining data sets is "ID", being a unique crop code and/or name. The primary key "ID" variable are georeferenced crop names/codes. Dataset joining can be easily implemented by using *merge(x, y, all.x = TRUE),* where x is geo-referenced crops' IDs from the country parcel maps and y any other dataset to be joined, providing that it has correct crop ID variable. Conceptually this operation can be described as union of two Cartesian products (Equation 1).

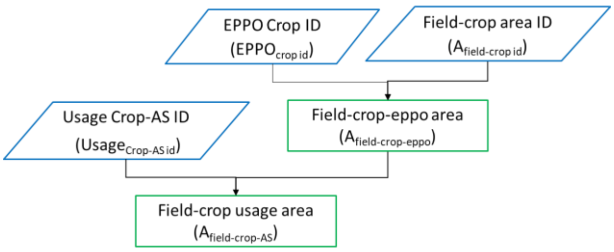


Figure 5

$$A_{field-crop-AS} = (A_{field-crop\ id} \times Usage_{crop-AS\ id}) \cup (A_{field-crop\ id} \times EPPO_{crop\ id})$$

Equation 1

Fraction field-crop area treated (F_field_crop)

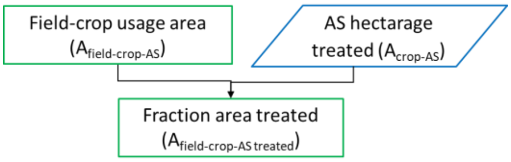


Figure 6

$$F_{field-AS-crop\ area\ treated} = \frac{\sum_{i=1}^{patch} A_{field-crop}}{A_{crop-AS}}$$

Equation 2

Application rate of field-crop-AS (A_field_as_crop)

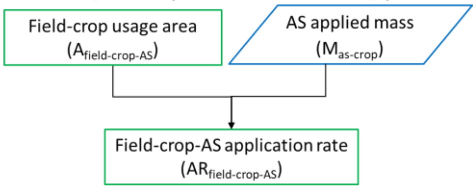


Figure 7

$$AR_{field-crop-AS} = \frac{M_{as-crop}}{A_{field-crop}}$$

Equation 3

Corrected application rate of field-crop-AS (A_field_as_crop_corr)

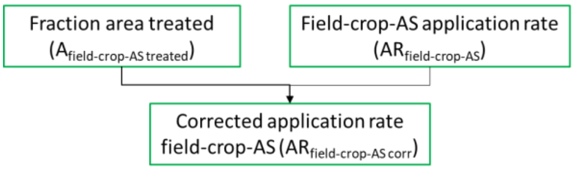


Figure 8

$$AR_{field-crop-AS\ corr} = AR_{field-crop-AS} \times F_{field-AS-crop\ area\ treated}$$

Equation 4

Site-specific application rate in river buffer

Buffer width corrected (W_buffer_corr)

&

Buffer area corrected (A_buffer_corr)

Buffers are polygons representing the area within a given distance of a geometric feature; regardless of whether the input is a point, line or polygon, the output is a polygon. More precisely it is unary transformation operation that works on a per-geometry basis and return for each geometry a new geometry.

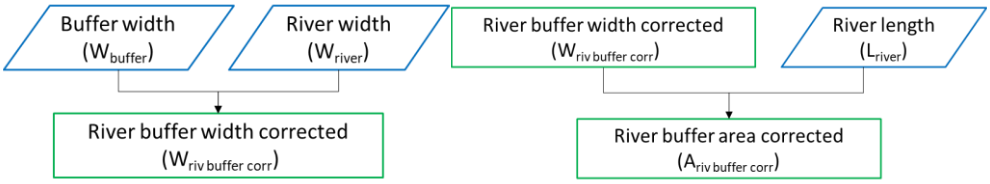


Figure 9

$$W_{riv\ buffer\ corr} = W_{river} + W_{buffer}$$

Equation 5

$$A_{riv\ buffer\ corr} = L_{river} \times W_{riv\ buffer\ corr}$$

Equation 6

Total field area in a buffer (A_field_riv_buffer)

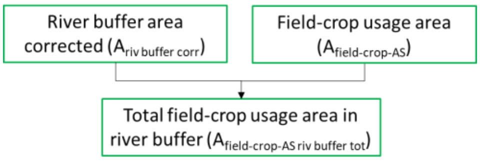


Figure 10

$$A_{field-crop-AS\ riv\ buff\ tot} = \left(\sum_{i=1}^{patches} \sum_{i=1}^{crop} \sum_{i=1}^{AS} A_{field-crop-AS} \right) \cap A_{riv\ buffer\ corr}$$

Equation 7

Agricultural area weighting factor for a river buffer (F_field_crop_riv_buffer)

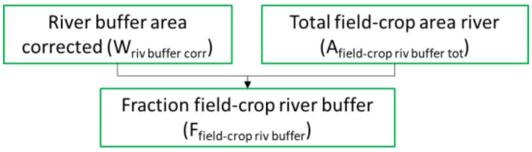


Figure 11

$$F_{field-crop-AS\ area\ riv\ buffer} = \frac{A_{field-crop-AS\ riv\ buff\ tot}}{A_{riv\ buffer\ corr}}$$

Equation 8

Application rate field-crop-AS target spatial unit (AR_as_crop_riv_buffer)

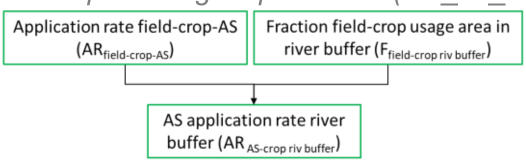


Figure 12

$$AR_{field-crop-AS\ riv\ buffer} = AR_{field-crop-AS} \times F_{field-crop-AS\ area\ riv\ buff}$$

Equation 9

Site-specific application rate at river basin level

Calculation schematics related to obtaining basin-level application rates are repeated here and complemented by formulas.

Total field area in target spatial unit (A_field_crop_area_tot_target)

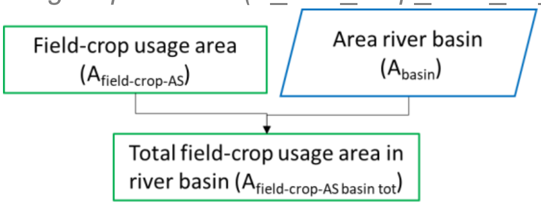


Figure 13

$$A_{field-crop-AS\ basin\ tot} = \left(\sum_{i=1}^{patches} \sum_{i=1}^{crop} \sum_{i=1}^{AS} A_{field-crop-AS} \right) \cap A_{basin}$$

Equation 10

Agricultural area fraction in a river basin (F_field_crop_AS_basin)

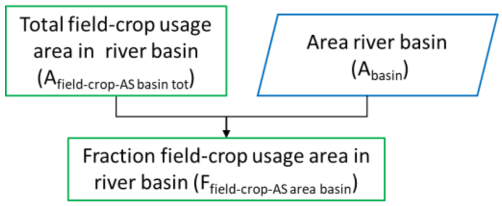


Figure 14

$$F_{field-crop-AS\ area\ basin} = \frac{A_{field-crop-AS\ basin\ tot}}{A_{basin}}$$

Equation 11

Agricultural area scaling weight in a river basin (w_field_area_weight)

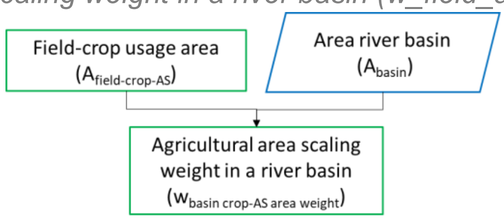


Figure 15

$$w_{field-basin\ area\ weight} = \frac{A_{field-crop-AS}}{\sum_{i=1}^{patches} A_{field-crop-AS}} \cap A_{basin}$$

Equation 12

River length scaling weight in a river basin (w_field_area_weight)

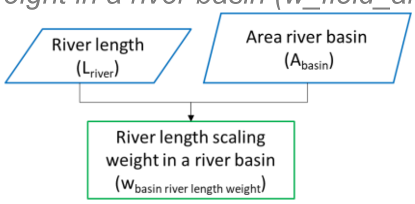


Figure 16

$$w_{river-basin\ length\ weight} = \frac{L_{river}}{\sum_{i=1}^{segments} L_{river}} \cap A_{basin}$$

Equation 13

Application rate field-crop-AS target spatial unit (AR_as_crop_target)

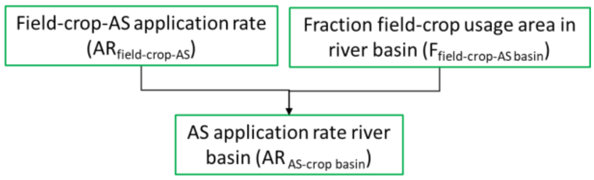


Figure 17

$$AR_{field-crop-AS\ basin} = AR_{field-crop-AS} \times F_{field-crop-AS\ area\ basin}$$

Equation 14

Surface runoff

Calculation schematics related to obtaining daily surface runoff depth and fraction of generated surface runoff are repeated here and complemented by formulas.

Mean daily precipitation (P_day)

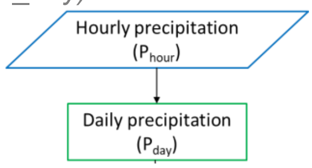


Figure 18

$$P_{day} = \frac{\sum_{i=1}^{P_{obs}} P_{hour}}{N_{P_{obs}}}$$

Equation 15

Terrain slope correction factor (CF_terrain_slope)

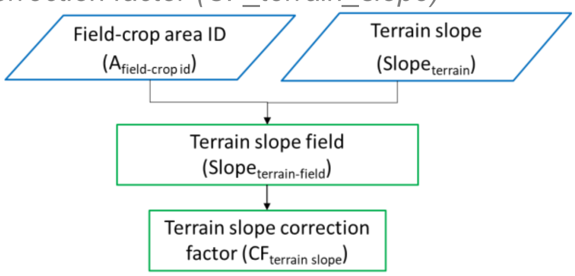


Figure 19

$$CF_{terrain\ slope} = \begin{cases} 0.00143 \times Slope_{terrain-field}^{\ 2} + 0.02153 \times Slope_{terrain-field}, & Slope_{terrain} \leq 20\% \\ 1, & Slope_{terrain} > 20\% \end{cases}$$

Equation 16

Surface runoff depth (H_srunoff)

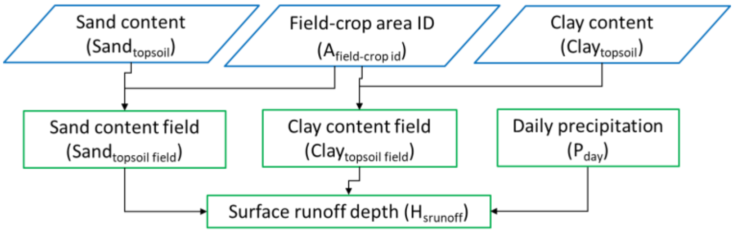


Figure 20

$$H_{srunoff} = \begin{cases} -5.86 \times 10^{-6} \times P_{day}{}^{3} + 2.63 \times 10^{-3} \times P_{day}{}^{2} - 1.14 \times 10^{-2} \times P_{day} - 1.164 \times 10^{-2}, & if\ Soil\ texture = Sand_{topsoil\ content} \\ -9.04 \times 10^{-6} \times P_{day}{}^{3} + 4.04 \times 10^{-3} \times P_{day}{}^{2} + 4.16 \times 10^{-3} \times P_{day} - 6.11 \times 10^{-2}, & if\ Soil\ texture = Clay_{topsoil\ content} \end{cases}$$

Equation 17

Surface runoff time of event (t_srunoff)

t_srunoff is a calendar day when H_srunoff is larger than 0.

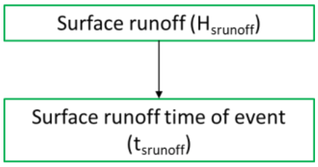


Figure 21

$$if\,(H_{srunoff} > 0,\ t_{srunoff} = day_{srunoff},\ 0)$$

Equation 18

Time interval between AS application and surface runoff event (Δt_days_to_srunoff)

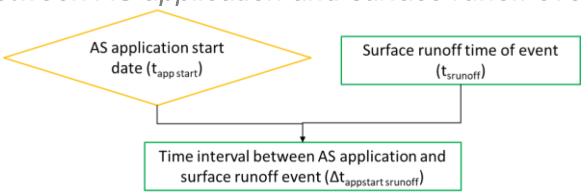


Figure 22

$$\Delta t_{days\ to\ srunoff} = t_{app\ start} + t_{srunoff}$$

Equation 19

Fraction of surface runoff (F_srunoff)

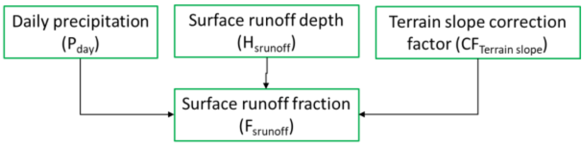


Figure 23

$$F_{srunoff} = \frac{H_{srunoff}}{P_{day}} \times CF_{terrain\ slope}$$

Equation 20

Effect of crop interception on chemical load

Calculation schematics related to obtaining effect of crop interception factors on chemical loss via surface runoff are repeated here and complemented by formulas.

Field interception factor (IF_field)

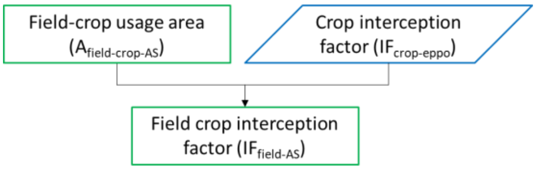


Figure 24

$$IF_{field-AS} = A_{field-crop-AS} \times IF_{crop-eppo}$$

Equation 21

AS fraction intercepted (F_field_IF)

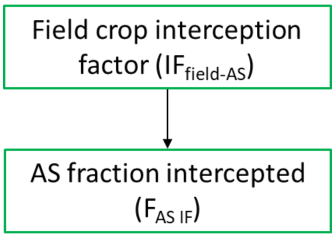


Figure 25

$$F_{AS\ IF} = 1 - \frac{IF_{field-AS}}{100}$$

Equation 22

Degradation rates

Calculation schematics related to obtaining individual chemical water and soil specific degradation rates are repeated here and complemented by formulas.

AS degradation rate in soil (k_DT50_soil)

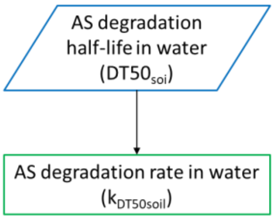


Figure 26

$$k_{DT50soil} = \frac{ln2}{DT50_{soil}}$$

Equation 23

AS degradation rate in river water (k_DT50_riverwater)

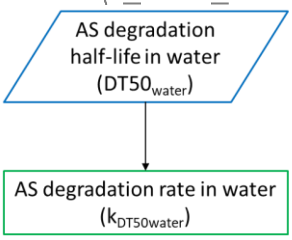


Figure 27

$$k_{DT50water} = \frac{ln2}{DT50_{water}}$$

Equation 24

Partitioning between topsoil pore water and soil particles

Calculation schematics related to obtaining initial and time-dependent total fractions of AS in topsoil are repeated here and complemented by formulas.

Distribution coefficient at the interface pore Water-soil particles (K_d_soil_water)

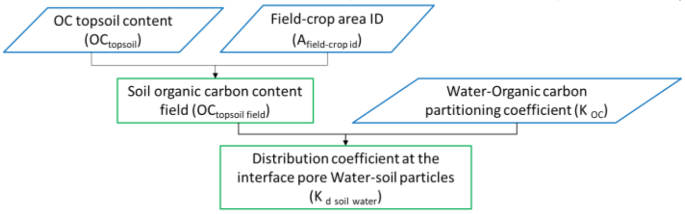


Figure 28

$$K_{d\ soil\ water} = \frac{OC_{topsoil\ field}}{100} \times K_{OC}$$

Equation 25

AS initial fraction dissolved in soil pore water (F_AS_soil_water_ini)

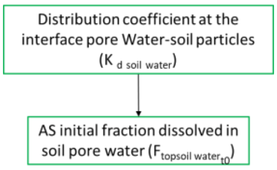


Figure 29

$$F_{topsoil\ water_{t0}} = \frac{1}{(1 + K_{d\ soil\ water})}$$

Equation 26

AS initial fraction bound to soil particles (F_AS_soil_solid_ini)

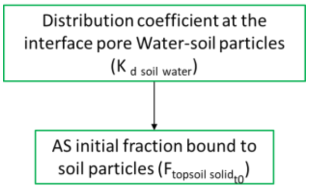


Figure 30

$$F_{topsoil\ solid_{t0}} = \frac{K_d}{(1 + K_d)}$$

Equation 27

AS initial soil total fraction (F_AS_top_soil_tot)

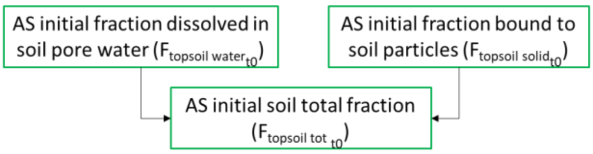


Figure 31

$$F_{topsoil\ tot\ t0} = F_{topsoil\ solid_{t0}} + F_{topsoil\ water_{t0}}$$

Equation 28

AS fraction dissolved in soil pore water after time lag (Δt_F_AS_pore_water)

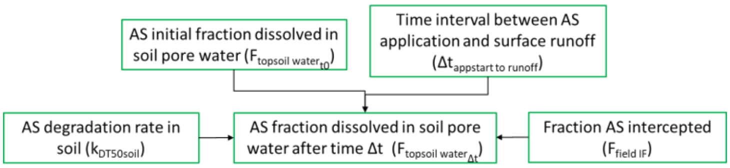


Figure 32

$$F_{topsoil\ water_{\Delta t}} = e^{-\Delta t_{days\ to\ runof} \times k_{DT50soil}} \times F_{field\ IF} \times F_{topsoil\ water_{t0}}$$

Equation 29

AS fraction bound to soil solids after time lag (Δt_F_AS_soil_solid)

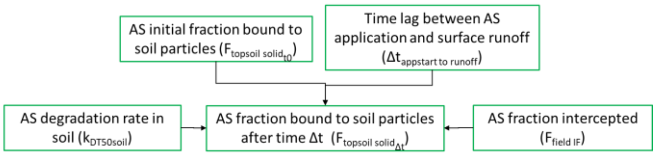


Figure 33

$$F_{topsoil\ solid_{\Delta t}} = e^{-\Delta t_{days\ to\ srunoff} \times k_{DT50soil}} \times F_{field\ IF} \times F_{topsoil\ solid_{t0}}$$

Equation 30

AS soil total fraction after time lag (Δt_F_AS_top_soil_tot)

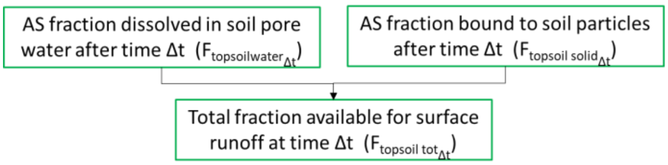


Figure 34

$$F_{topsoil\ tot\ \Delta t} = F_{topsoil\ solid_{\Delta t}} + F_{topsoil\ water_{\Delta t}}$$

Equation 31

Regulatory variables

Topsoil

Calculation schematics related to obtaining initial and time-dependent total concentrations of AS in topsoil are repeated here and complemented by formulas.

AS initial total topsoil concentration (C_field_crop_AS_topsoil_ini)

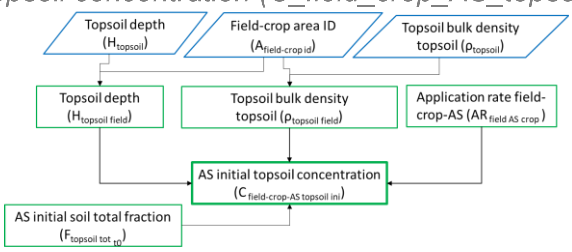


Figure 35

$$C_{field\ crop-AS\ topsoil\ ini} = \frac{AR_{field\ AS\ crop}}{H_{topsoil\ field} \times \rho_{topsoil\ field}} \times F_{topsoil\ tot_{t0}}$$

Equation 32

AS total topsoil concentration after time lag (Δt_C_field_AS_topsoil)

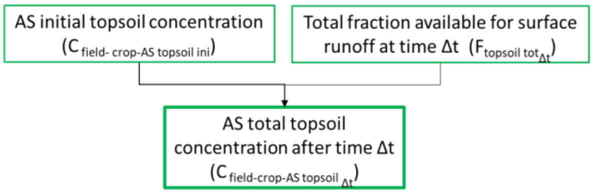


Figure 36

$$C_{field-crop-AS\ topsoil\ \Delta t} = C_{crop-AS\ topsoil\ target\ area\ ini} \times F_{topsoil\ tot\ \Delta t}$$

Equation 33

AS initial river basin total topsoil concentration (C_basin_crop_AS_topsoil_ini)

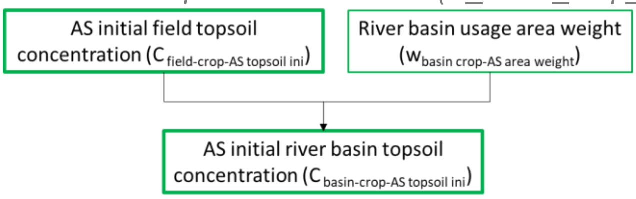


Figure 37

$$C_{basin-crop-AS\ topsoil\ ini} = \sum_{i=1}^{patches} (C_{field-crop-AS\ topsoil\ ini} \times w_{field-basin\ area\ weight})$$

Equation 34

AS total river basin topsoil concentration after time lag (Δt_C_basin_crop_AS_topsoil)

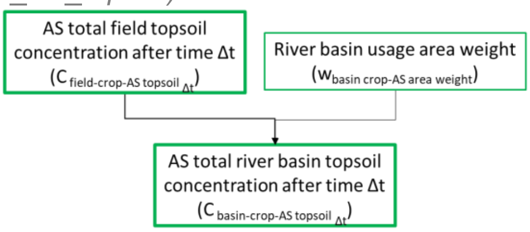


Figure 38

$$C_{basin-crop-AS\ topsoil\ \Delta t} = \sum_{i=1}^{patches} (C_{field-crop-AS\ topsoil\ \Delta t} \times w_{field-basin\ area\ weight})$$

Equation 35

River water

Calculation schematics related to obtaining chemical load from individual fields and time-dependent total concentrations of AS in river water are repeated here and complemented by formulas.

AS load individual field in river buffer (Δt_load_topsoil_rivbuff)

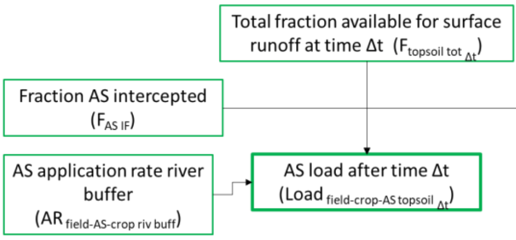


Figure 39

$$load_{field-crop-AS\,riv\,buffer}{}_{\Delta t} = AR_{field-AS-crop\,riv\,buffer} \times F_{AS\,IF} \times F_{topsoil\,tot_{\Delta t}} \times F_{srunoff}$$

Equation 36

Concentration in surface river water after time lag (Δt_C_riverwater)

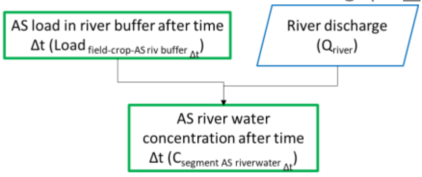


Figure 40

$$C_{segment\ AS\,riverwater_{\Delta t}} = \frac{load_{field-crop-AS\,riv\,buffer}{}_{\Delta t}}{Q_{river}}$$

Equation 37

Concentration in surface river water in a basin after time lag (Δt_C_basin_riverwater)

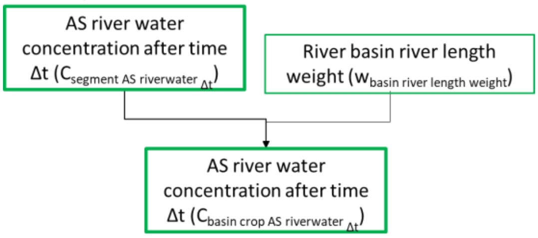


Figure 41

$$C_{basin-crop-AS-riverwater\ \Delta t} = \sum_{i=1}^{segments} \left(C_{segment-AS-riverwater\ \Delta t} \times w_{river-basin\ length\ weight}\right)$$

Equation 38

References

Allen, P. M., Arnold, J. C., & Byars, B. W. (1994). DOWNSTREAM CHANNEL GEOMETRY FOR USE IN PLANNING-LEVEL MODELS. *JAWRA Journal of the American Water Resources Association*, *30*(4), 663–671. https://doi.org/10.1111/j.1752-1688.1994.tb03321.x

Arnold, S., Kosztra, B., Banko, G., Smith, G., Hazeu, G. W., Bock, M., & Valcarcel, N. (2013). The EAGLE concept – A vision of a future European Land Monitoring Framework. In *EARSeL symposium proceedings "Towards Horizon 2020."*

Ayllón-Benitez, A., Bernabé-Diaz, J. A., Espinoza-Arias, P., Esnaola-Gonzalez, I., Beeckman, D. S. A., McCaig, B., Hanzlik, K., Cools, T., Castro Iragorri, C., & Palacios, N. (2023). EPPO ontology: A semantic-driven approach for plant and pest codes representation. *Frontiers in Artificial Intelligence*, *6*, 1131667. https://doi.org/10.3389/frai.2023.1131667

Ballabio, C., Panagos, P., & Monatanarella, L. (2016a). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, *261*, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006

Ballabio, C., Panagos, P., & Monatanarella, L. (2016b). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, *261*, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006

Benfenati, E., Manganaro, A., & Gini, G. (2013, December). VEGA-QSAR: AI inside a platform for predictive toxicology. *Proceedings of the Workshop on Popularize Artificial Intelligence (PAI 2013)*. http://ceur-ws.org/Vol-1107/paper8.pdf

Berenzen, N., Lentzen-Godding, A., Probst, M., Schulz, H., Schulz, R., & Liess, M. (2005). A comparison of predicted and measured levels of runoff-related pesticide concentrations in small lowland streams on a landscape level. *Chemosphere*, *58*(5), 683–691. https://doi.org/10.1016/j.chemosphere.2004.05.009

Boeije, G. M., Wagner, J.-O., Koormann, F., Vanrolleghem, P. A., Schowanek, D. R., & Feijtel, T. C. J. (2000). New PEC definitions for river basins applicable to GIS-based environmental exposure assessment. *Chemosphere*, *40*(3), 255–265. https://doi.org/10.1016/S0045-6535(99)00243-X

CEN Workshop Agreement (CWA). (2015). *Standard documentation of chemical exposure models* (No. 367). https://www.nen.nl/en/cwa-16938-2015-en-210819

Ciffroy, P., Altenpohl, A., Fait, G., Fransman, W., Paini, A., Radovnikovic, A., Simon-Cornu, M., Suciu, N., & Verdonck, F. (2016). Development of a standard documentation protocol for communicating exposure models. *Science of The Total Environment*, *568*, 557–565. https://doi.org/10.1016/j.scitotenv.2016.01.134

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *Journal of Hydrology*, *270*(1), 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

European Commission. (2017, February 16). *Commission Regulation (EU) 2017/269 of 16 February 2017 amending Regulation (EC) No 1185/2009 of the European Parliament and of the Council concerning statistics on pesticides, as regards the list of active substances*. https://eur-lex.europa.eu/eli/reg/2017/269/oj/eng

European Commission (Ed.). (2018). *European statistics code of practice: For the national statistical authorities and Eurostat (EU statistical authority)*. Publications Office. https://doi.org/10.2785/914491

European Food Safety Authority. (2017). EFSA Guidance Document for predicting environmental concentrations of active substances of plant protection products and transformation products of these active substances in soil. *EFS2*, *15*(10). https://doi.org/10.2903/j.efsa.2017.4982

European Parliament, Council of the European Union. (2009, November 25). *Regulation (EC) No 1185/2009 of the European Parliament and of the Council of 25 November 2009 concerning statistics on pesticides*. https://eur-lex.europa.eu/eli/reg/2009/1185/oj/eng

Eurostat. (2022). *Pesticide sales* [Dataset]. Eurostat. https://doi.org/10.2908/AEI_FM_SALPEST09

Fischer, G., Nachtergaele, F. O., Van Velthuizen, H., Chiozza, F., Franceschini, G., Henry, M., Muchoney, D., & Tramberend, S. (2021). *Global agro-ecological zones v4–model documentation*. Food & Agriculture Org.

Gutsche, V., & Strassemeyer, J. (2007). SYNOPS – ein Modell zur Bewertung des Umwelt-Risikopotentials von chemischen Pflanzenschutzmitteln. *Heft 9*, *59*(9), 197–210.

Hawkins, R., Ward, T. J., Woodward, D., & Mullem, J. A. (2009). *Curve number hydrology: State of the practice*. https://doi.org/10.1061/9780784410042

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Ippolito, A., Kattwinkel, M., Rasmussen, J. J., Schäfer, R. B., Fornaroli, R., & Liess, M. (2015). Modeling global distribution of agricultural insecticides in surface waters. *Environmental Pollution*, *198*, 54–60. https://doi.org/10.1016/j.envpol.2014.12.016

Jones, R. J. A., Spoor, G., & Thomasson, A. J. (2003). Vulnerability of subsoils in Europe to compaction: A preliminary analysis. *Soil and Tillage Research*, *73*(1), 131–143. https://doi.org/10.1016/S0167-1987(03)00106-5

Kattwinkel, M., Kühne, J.-V., Foit, K., & Liess, M. (2011). Climate change, agricultural insecticide exposure, and risk for freshwater communities. *Ecological Applications*, *21*(6), 2068–2081. https://doi.org/10.1890/10-1993.1

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world’s large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Lehner, B., Verdin, K., & Jarvis, A. (2008). New Global Hydrography Derived From Spaceborne Elevation Data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Lewis, K. A., Tzilivakis ,John, Warner ,Douglas J., & and Green, A. (2016). An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment: An International Journal*, *22*(4), 1050–1064. https://doi.org/10.1080/10807039.2015.1133242

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., & Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data*, *6*(1), 283. https://doi.org/10.1038/s41597-019-0300-6

Olesen, M. H., & Jensen, P. K. (2013). Collection and evaluation of relevant information on crop interception. *EFSA Supporting Publications*, *10*(6), 438E. https://doi.org/10.2903/sp.efsa.2013.EN-438

Périé, C., & Ouimet, R. (2008). Organic carbon, organic matter and bulk density relationships in boreal forest soils. *Canadian Journal of Soil Science*, *88*(3), 315–325. https://doi.org/10.4141/CJSS06008

Probst, M., Berenzen, N., Lentzen-Godding, A., & Schulz, R. (2005). Scenario-based simulation of runoff-related pesticide entries into small streams on a landscape level. *Ecotoxicology and Environmental Safety*, *62*(2), 145–159. https://doi.org/10.1016/j.ecoenv.2005.04.012

Röpke, B., Bach, M., & Frede, H. G. (2004). DRIPS - a decision support system estimating the quantity of diffuse pesticide pollution in German river basins. *Water Science and Technology*, *49*(3), 149–156. https://doi.org/10.2166/wst.2004.0184

Schriever, C. A., Ball, M. H., Holmes, C., Maund, S., & Liess, M. (2007). Agricultural intensity and landscape structure: Influences on the macroinvertebrate assemblages of small streams in northern Germany. *Environmental Toxicology and Chemistry*, *26*(2), 346–357. https://doi.org/10.1897/05-629R.1

Schriever, C. A., Von Der Ohe, P. C., & Liess, M. (2007a). Estimating pesticide runoff in small streams. *Chemosphere*, *68*(11), 2161–2171. https://doi.org/10.1016/j.chemosphere.2007.01.086

Schriever, C. A., Von Der Ohe, P. C., & Liess, M. (2007b). Estimating pesticide runoff in small streams. *Chemosphere*, *68*(11), 2161–2171. https://doi.org/10.1016/j.chemosphere.2007.01.086

Steenhuis, T. S., Winchell, M., Rossing, J., Zollweg, J. A., Walter, M. F., Hawkins, R. H., Steenhuis, T. S., Frankenberger, J. R., Winchell, M., & Zollweg, J. A. (1996). SCS Runoff Equation Revisited for Variable-Source Runoff Areas. *Journal of Irrigation and Drainage Engineering-American Society of Civil Engineers*, *122*(5), 319.

ter Laak, T. L., Gebbink, W. A., & Tolls†, J. (2006). Estimation of soil sorption coefficients of veterinary pharmaceuticals from soil properties. *Environmental Toxicology and Chemistry*, *25*(4), 933–941. https://doi.org/10.1897/05-229R.1

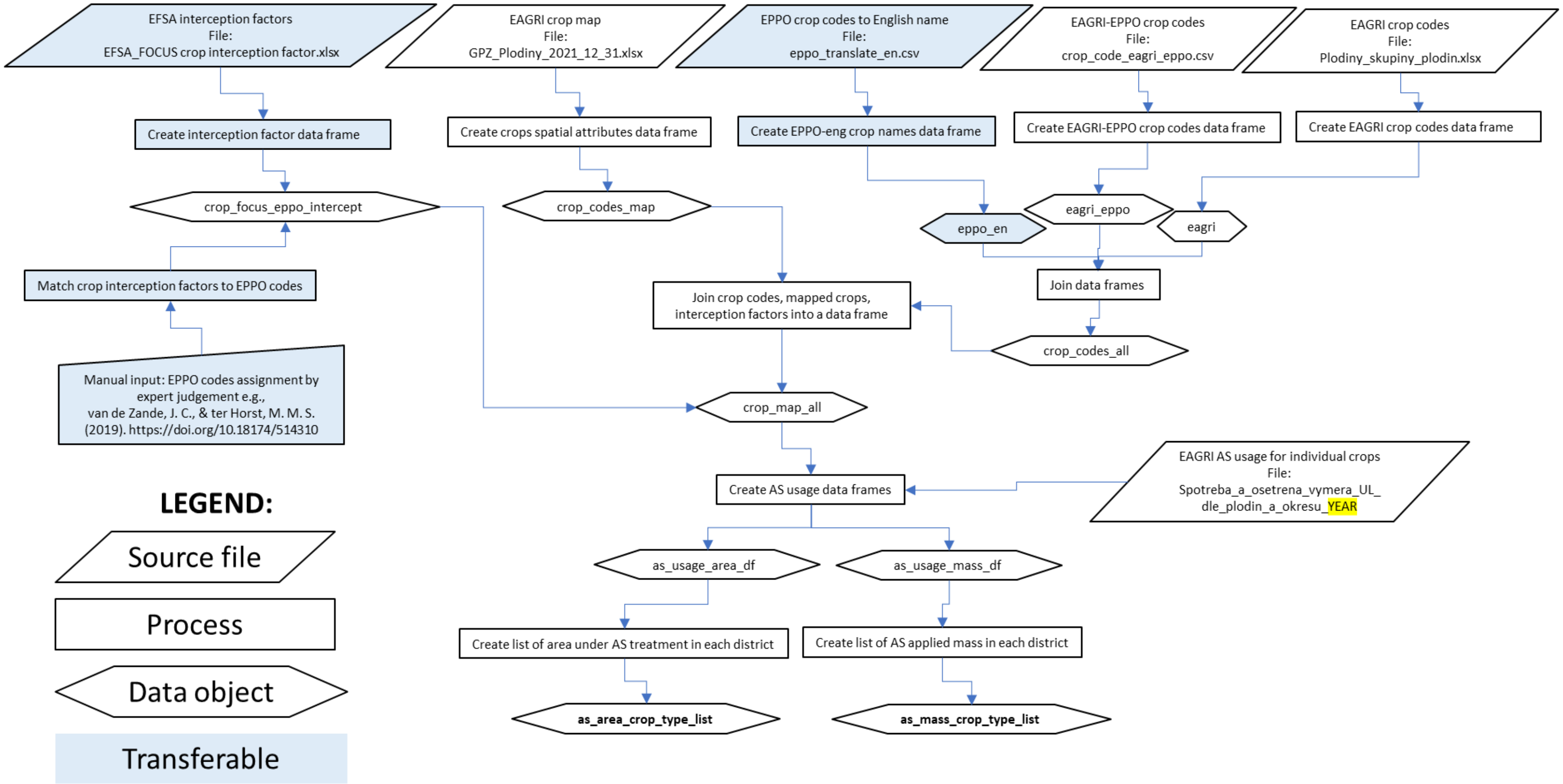
van Beinum, W., & Beulke, S. (2010). Collection and evaluation of relevant information on crop interception for the revision of the Guidance Document on Persistence in Soil. *EFSA Supporting Publications*, *7*(10). https://doi.org/10.2903/sp.efsa.2010.EN-73

van de Zande, J. C., & ter Horst, M. M. S. (2019). *Crop related aspects of crop canopy spray interception and spray drift from downward directed spray applications in field crops*. Wageningen Plant Research. https://doi.org/10.18174/514310

Verdin, K. L., & Verdin, J. P. (1999). A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology*, *218*(1), 1–12. https://doi.org/10.1016/S0022-1694(99)00011-6

Verro, R., Calliera, M., Maffioli, G., Auteri, D., Sala, S., Finizio, A., & Vighi, M. (2002). GIS-Based System for Surface Water Risk Assessment of Agricultural Chemicals. 1. Methodological Approach. *Environmental Science & Technology*, *36*(7), 1532–1538. https://doi.org/10.1021/es010089o

APPENDIX 1. Flowchart of AS usage and crop data preparation using Czech data.



APPENDIX 2.
Flowchart of AS usage data spatialisation using Czech data.