Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

SPRINT
SUSTAINABLE PLANT PROTECTION TRANSITION

Generalized PPPs' Application Patterns from regional to EU level

Chaideftou et al.

Deliverable number: 5.2
Due date: 28.02.2023 (month 30)
Dissemination level:
Lead beneficiary: MU
Deliverable type:
Version:

Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

PROJECT INFORMATION	
Project Title	Sustainable Plant Protection Transition; A global health approach
Project Acronym	SPRINT
Call Identifier	H2020-SFS-2019-2; Sustainable Food Security
Grant Agreement no.	862568
Starting Date	01-09-2020
End Date	31-08-2025
Project duration	60 months
Website address	www.sprint-h2020.eu
Project coordinator	Wageningen University
Overall Project coordinator	Prof. Dr. V. Geissen +31317485144 (violette.geissen@wur.nl)
Scientific Project Manager	V. Felix da Garca Silva, MSc (vera.felixdagracasilva@wur.nl)
REPORT INFORMATION	
Report Title	Generalized PPPs' Application Patterns from regional to EU level.
Principle Author(s)	Evgenia Chaideftou (MU), Shiva Sabzevari (MU), Mazen Ismael (MU), Václav Novotný (MU), Zdenka Bednářová (MU), Richard Hůlek (MU), Veronika Hofmanová, Jakub Hofman (MU)
Principle Author e-mail	jakub.hofman@recetox.muni.cz
Deliverable Number	5.2
Work Package	5
Editor(s)	
E--Mail(s)	
Report Due Date	28.02.2023 (month 30)
Report publish date	28.02.2023 (month 30)

Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

Table of Contents

Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

List of abbreviations and acronyms

Abbreviation / acronym	Definition (explanation)
a.s.	**a**ctive **s**ubstance
AGES	Österreichische **A**gentur für **G**esundheit und **E**rnährungs**s**icherheit GmbH (Austrian Agency for Health and Food Safety)
AT	**Aus**t**r**ia
Avoindata	OpenData in Finnish
BBCH	Abbreviation of crop-plant growth-stage scale, derived from the names of the originally participating stakeholders: "Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie". (https://en.wikipedia.org/wiki/BBCH-scale)
BVL	**B**undesamt für **V**erbraucherschutz und **L**ebensmittelsicherheit (Federal Office of Consumer Protection and Food Safety of Germany)
CSV	**C**omma **S**eparated **V**alues
Ctgb	Het **C**ollege voor de **t**oelating van **g**ewasbeschermingsmiddelen en **b**iociden (The Board for the Authorisation of Plant Protection Products and Biocides of the Netherlands)
CZ	**C**zech **R**epublic
D 5.2	**D**eliverable **5.2**
DBT	**D**ata **B**uild **T**ool
DGAV	**D**ireção-**G**eral da **A**limentação e **V**eterinária (Directorate for Health Protection Means of Portuguese Ministry of Agriculture)
DWH	**d**ata **w**are**h**ouse
EAGRI	EAGRI (Ministry of Agriculture of the Czech Republic)
EC	**E**uropean **C**ommission
EE	**E**stonia
EPA DK	The **E**nvironmental **P**rotection **A**gency of **D**enmar**k**
EPPO	**E**uropean and Mediterranean **P**lant **P**rotection **O**rganization
EU	**E**uropean **U**nion
EU MS(s)	**E**uropean **U**nion **M**ember **S**tate(s)
EU-Cropmap	**E**uropean **U**nion **C**rops **m**ap
GenAP	**Gen**eralised **A**pplication **P**attern
ICPS	**C**entro **I**nternationale per gli Antiparassitari e la **P**reventione **S**anitaria (International Center for Pesticides and Health Risk Prevention, Italy)
JSON	**J**ava**S**cript **O**bject **N**otation
Kemi	**Kemi**kalieinspektionen (Swedish Chemicals Agency)

kemidigi	**kemi**kaalitieto (chemical) **digi**tal data in one service (Chemical-products Finnish register)
M7	**M**ilestone **7** of the Sprint project
M8	**M**ilestone **8** of the Sprint project
NUTS2	NUTS - **N**omenclature of **T**erritorial **U**nits for **S**tatistics; NUTS **2**: basic regions for the application of regional policies
PDF	**P**ortable **D**ocument **F**ormat
PESTIDOC	PESTIDOC web platform of ICPS
PMAIS	sisene **p**ortaali **Ma**aelumini**s**teerium (Official portal of the Estonian Ministry of Rural Affairs)
PPP(s)	**P**lant **P**rotection **P**roduct(s)
PT	**P**or**t**ugal
PTA	Põllumajandus- ja Toiduamet (Estonian Agriculture and Food Board)
SIFITO	**Si**stema de Gestão das Autorizações de Produtos **Fito**farmacêuticos (Management System for Plant Protection Products Authorizations of DGAV Portugal)
SPF/FOD	**S**ervice **p**ublic **f**ederal/**F**ederale **o**verheids**d**ienst (Federal Public Service of Belgium)
SPRINT	Sustainable Plant Protection Transition - A global health approach: European H2020 project
Tukes	**Tu**rvallisuus- ja **ke**mikaalivira**s**to (Finnish Safety and Chemicals Agency)
UKSUP	**Ú**stredný **k**ontrolný a **s**kúšobný **ú**stav **p**oľnohospodársky (Central Agricultural Inspection and Testing Institute of Slovakia)
UKZUZ	**Ú**střední **k**ontrolní a **z**kušební **ú**stav **z**emědělský (Central Institute of Supervising and Testing in Agriculture)
VAT	**V**alstybiné **A**ugalikinkystés **T**arnyba (The State Plant Service under the Ministry of Agriculture, Lithuania)
WP5	**W**ork**p**ackage 5 of the SPRINT project
XML	e**X**tensible **M**arkup **L**anguage

Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

1. Introduction

Within the goals of the SPRINT Toolbox (Milestone 7 – M7)[1] and of WP5 of the SPRINT project, it was planned to estimate the inputs (emissions) into the European environment for the active substances (a.s.) of Plant Protection Products (PPPs). The first appraisal related to the specific PPP data needed for these estimations was presented in Milestone 8 (M8)[2]. As presented in M8, among the major data that are used for the SPRINT Toolbox estimations and predictions are:

- Data inputs from the national PPP-registers – national data on registered PPPs (i.e., authorized according to EC Regulation 1107/2009) – defining application patterns for each a.s.: crop, pest, application rate, frequency, and time.
- The general EU list of the approved and non-approved active substances (a.s.) as the most comprehensive database of pesticides' properties.
- The annual usages of the a.s. for specific crops for the countries of smaller regions (e.g., NUTS3).
- Geographically referenced data on the specific crop occurrences at sufficient resolution (field) for each year.
- Pesticide properties and other necessary data.

The state of the national PPP-registers was described in Chapter 2 of M8.

In this Report, the **Generalised Application Patterns (GenAP)** of authorized PPPs were estimated following a thorough search and analysis of authorization data from the PPP-registers of European countries. **Generalized Application Patterns are the tables/databases containing all possible authorized uses (crop specific application rates, time, and number of applications per year) of active substances in the country**.

Results are exemplified for four Member States of the European Union (EU MSs) selected based on specific criteria and approach. Because the selected countries represent all three PPP authorization zones and within the zones the portfolio of PPPs is usually very similar, the GenAP might represent the typical situation on PPPs in EU.

We describe the followed strategy on analyzing the quality and types of data, its retrieval, and processing, explaining also filtering and selections, that were applied during the process. Finally, the potential and power, the faced challenges and limitations, and implications for a short- and long-term utilization of data and tools are concisely discussed in a constructive manner.

[1] Hofman et al. 2021. Structure, functionality, data inputs and outputs of the Toolbox. Report of Milestone 7, Work Package 5, SPRINT Project.
[2] Hofman et al. 2022. Database of relevant European data for upscaling – collected and harmonized database. Report of Milestone 8, Work Package 5, SPRINT Project.

2. Country selection for PPPs' Generalized Application Patterns in Europe

2.1 Selection of member state databases

The method applied for the selection of Member State Databases for deriving Generalized Application Patterns (GenAP) of Plant Protection Products (PPPs) in Europe was stepwise using two main combined modes: (i) utilizing the first-phase appraisal of the European PPP-registers status (Chapter 2 and Tables 2.2-2.4 of M8) and (ii) a decision-tree approach (**Fig. 1a**).

2.1.1 Utilizing the first-phase appraisal of the European PPP-registers status

In M8[3], and thus, in the first-phase appraisal of the European PPP-registers status, the main search criteria for the PPP-registers of the European countries were **availability, accessibility, and data resolution**. It is noted that accessibility was added as a criterion, distinguished from availability because the fact that there is data deposited in a country is not a prerequisite for easy or adequate access to it. Therefore, within the criterion of accessibility lies a different question: how easy, and if the documents can be retrieved from an online link. Data resolution is a criterion that does not define the quality of data, standalone, but also the detail of data that can be retrieved. The specific properties assigned to each of the three criteria that explicitly describe them, are explained precisely in Chapter 2 of M8.

In Table 2.2 (summary) of M8, the overall scores per search criterion (availability, accessibility, resolution of related data) were displayed. This appraisal was structured upon a color scale indicating the state of the PPP-register of each country searched. Namely, each search criterion was assigned a score indicated with a scale of four distinct colors, that is described in the below key (see also Chapter 2 of M8).

- Grey: answers difficult to identify through source/link (for instance when no answer can be provided, or no information could be found at the period of searching).
- Green: answers "yes" or "sufficient/high" to the question of the property related to the status of the criterion
- Yellow: answers "possible" or "partly" or "in progress" to the question of the property related to the status of the criterion
- Red: answers "no" or "lacking/low" to the question of the property related to the status of the criterion

It is noted that high variability was observed among different quality/relevance criteria, and across countries.

[3] Hofman et al. 2022. Database of relevant European data for upscaling – collected and harmonized database. Report of Milestone 8, Work Package 5, SPRINT Project.

2.1.2 Using a decision-tree approach

Overall, 3 main steps were employed in this combined-mode scheme:

1. Second phase of exploring the sources of PPP-registers (Step 1, questions 1-4 in Fig. 1a; Chapter 2.2)
2. Evaluation of the resolute data quality of PPP-registers (Step 2, question 5 in Fig. 1a; Chapter 2.3)
3. Selection of PPP-register databases of European Member States (EU MSs) (Step 3, questions 6-8 in Fig. 1a; Chapter 2.1.3)

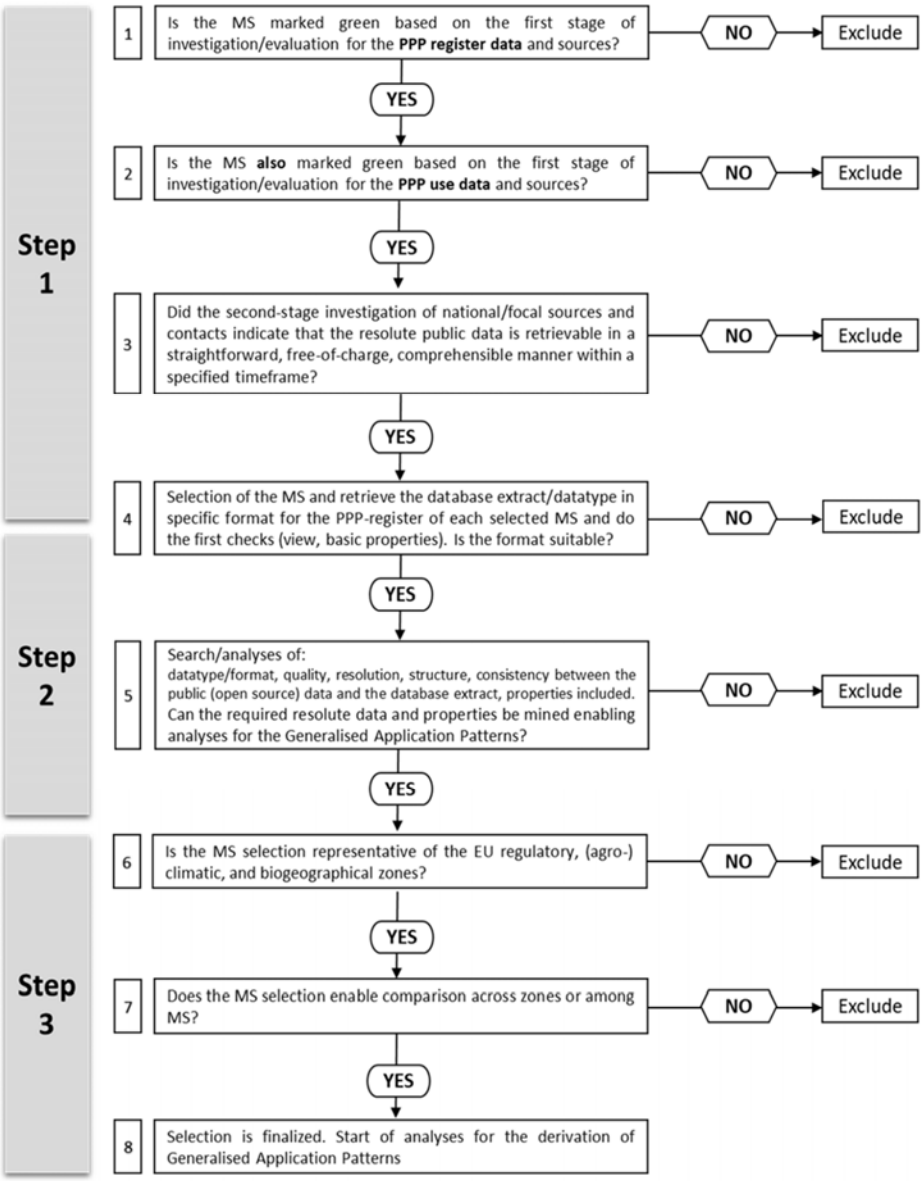


Figure 1a. Decision tree towards the final selection of MS(s) databases of national PPP-registers.

In the next lines (Chapters 2.2-2.4) the combined modes' scheme (Fig. 1a) is described thoroughly.

2.2 Second phase of exploring the sources of PPP-registers

The **27 EU MSs** were selected as the set for investigation in the second phase, as none of the EU MSs was scored with grey for any of the three criteria in the M8 analysis, indicating that the difficulty to identify specific properties was decreased and the possibility to further clarify the status of the PPP-register was increased.

Firstly, the overall scores for the 27 EU MSs of the first appraisal (set derived from Table 2.2 of M8) were considered and are shown in **Table 1** below.

Table 1. Overall scoring of the state of national PPP-registers across the 27 EU MSs based on a color scale in the first phase of investigation (Chapter 2 M8[4]). Color scores: green: high score; yellow: indicates ongoing process or partly, and red indicates no or a low score up to the date of exploration.

nr	country	country code	Availability SCORE	Accessibility SCORE	Resolution SCORE
1	Austria	AT			
2	Belgium	BE			
3	Bulgaria	BG			
4	Croatia	HR			
5	Cyprus	CY			
6	Czech Republic (Czechia)	CZ			
7	Denmark	DK			
8	Estonia	EE			
9	Finland	FI			
10	France	FR			
11	Germany	DE			
12	Greece (Hellas)	GR			
13	Hungary	HU			
14	Ireland	IE			
15	Italy	IT			
16	Latvia	LV			
17	Lithuania	LT			
18	Luxembourg	LU			
19	Malta	MT			
20	Netherlands	NL			
21	Poland	PL			
22	Portugal	PT			
23	Romania	RO			
24	Slovakia	SK			
25	Slovenia	SI			
26	Spain	ES			
27	Sweden	SE			

From Table 1, the "greener" EU MSs were selected in a priority-mode (the MSs that were greened for all 3 criteria were selected first, and then the selection continued to the less green

[4] Hofman et al. 2022. Database of relevant European data for upscaling – collected and harmonized database. Report of Milestone 8, Work Package 5, SPRINT Project.

and yellow MSs), to search the state of the national PPP-registers of the EU MSs in the second phase (see Step 1, question 1 in Fig. 1a).

The color scores in each "priority set" of EU MSs was then compared with the respective color scores for the use-data of PPPs (Tables 2.5-2.7, M8), to allow for building a robust data-frame for optimum utilization towards the inputs/emissions estimates of PPP a.s. (see Step 1, question 2 in Fig. 1a).

Then, the color-scoring for each EU MS was verified through contacting the national sources/contacts again with a set of more resolute and specific queries. Based on the responses, it was clarified in more depth (compared to the first appraisal) whether the resolute public data can be provided in a straightforward manner (directly or in brief time, free of charge and comprehensive enough). Following this, the publicly available data were retrieved or shared/provided by the respective national authorities/agencies, and the data types/formats and main content were checked (see Step1, questions 3 and 4 in Fig. 1a). An updated table (**Table 2**) of overall scores of the PPP-register state resulting from these main checks (Step 1, questions 1-4 in Fig. 1a) was created and is shown below.

Table 2. Overall scoring of the state of national PPP-registers across the 27 EU MSs based on a color scale in the second phase of investigation. For EU MSs, except for the four selected ones, the second-phase investigation is still ongoing as a more long-term procedure and thus the color scoring presented is subject to updates (yellow and red color scores).

nr	country	country code	Availability	Accessibility	Resolution
			SCORE	SCORE	SCORE
1	Austria	AT			
2	Belgium	BE			
3	Bulgaria	BG			
4	Croatia	HR			
5	Cyprus	CY			
6	Czech Republic (Czechia)	CZ			
7	Denmark	DK			
8	Estonia	EE			
9	Finland	FI			
10	France	FR			
11	Germany	DE			
12	Greece (Hellas)	GR			
13	Hungary	HU			
14	Ireland	IE			
15	Italy	IT			
16	Latvia	LV			
17	Lithuania	LT			
18	Luxembourg	LU			
19	Malta	MT			
20	Netherlands	NL			
21	Poland	PL			
22	Portugal	PT			
23	Romania	RO			
24	Slovakia	SK			
25	Slovenia	SI			
26	Spain	ES			
27	Sweden	SE			

2.3 Evaluation of the resolute data quality of PPP-registers

Then, the data were cross-checked and compared with the data displayed openly at the List of databases on registered plant protection products in the EPPO region[5], and in the links/sources indicated by regular contacts with the respective national authorities/agencies. It was searched if the preferred properties and resolute data are available or can be retrieved and mined adequately for estimating the GenAP (see 5 in Fig. 1a). This process resulted in drawing a more definitive conclusion on the PPP-register state and datatypes for 15 out of the total 27 EU MSs (see **Table 3**, and **Fig. 1b**).

Table 3. Overview of the current state of PPP-register database formats, status, and availability for 15 EU MSs for which the second-phase investigation evolved to a more thorough level.

Nr.	EU MSs	Regulatory zone	Data sources	Data type	Data format
1	Czech Republic CZ	Central	Central Institute for Supervising and Testing in Agriculture (UKZUZ)[6] and electronic data exchange service of Czech Ministry of Agriculture (EAGRI)[7]	Full database than can be loaded to MU DWH	XML, CSV
2	Austria AT	Central	Austrian Agency for Health and Food Safety (AGES)[8]	Database extract that can be loaded to MU DWH	XML, JSON
3	Netherlands NL	Central	The Board for the Authorisation of Plant Protection Products and Biocides (Ctgb)[9]	overview download	Excel
4	Germany DE	Central	Federal Office of Consumer Protection and Food Safety (BVL)[10]	Database extract that can be ordered	Microsoft Access
5	Slovakia SK	Central	Central Agricultural Inspection and Testing Institute (UKSUP)[11]	Part of database extract	XML
6	Slovenia SI	Central	Administration of the Republic of Slovenia Food safety, Veterinary sector, and Plant protection[12]	Data repository that can be extracted only as files at ministerial authority (only link to search available)	/
7	Belgium BE	Central	Federal Public Service (SPF/FOD)[13]	National repository under upgrading process (new platform); currently not available as database extract	/

[5] The EPPO databases of registered PPPs: https://www.eppo.int/ACTIVITIES/plant_protection_products/registered_products
[6] https://eagri.cz/public/web/en/ukzuz/portal
[7] https://eagri.cz/public/web/mze
[8] www.ages.at
[9] www.ctgb.nl
[10] www.bvl.bund.de/DE/Home/home_node.htm
[11] www.uksup.sk
[12] www.gov.si
[13] www.health.belgium.be

Nr.	EU MSs	Regulatory zone	Data sources	Data type	Data format
8	Estonia EE	North	Agriculture and Food Board (PTA)[14], Official portal of the Estonian Ministry of Rural Affairs (PMAIS)[15] & Estonian Open Data	Database extract that can be loaded to MU DWH	CSV
9	Finland FI	North	Chemical digital data in one service of the Finnish register (kemidigi)[16], Finnish Safety and Chemicals Agency (Tukes)[17], & other	Part of database extract available; data to be retrieved otherwise	(Avoindata service[18]) JSON; CSV & PDF
10	Denmark DK	North	The Environmental Protection Agency of Denmark (EPA DK)[19], and the private organisation SEGES Innovation (List of approved PPPs in DK[20])	overview list download	Excel
11	Sweden SE	North	Swedish Chemicals Agency (kemi)[21]	currently cannot be extracted	/
12	Lithuania LT	North	The State Plant Service under the Ministry of Agriculture (VAT)[22]	repository not disposable as part of confidential data	PDF-files at link
13	Portugal PT	Southern	Directorate for Health Protection Means of Portuguese Ministry of Agriculture (DGAV)[23] and Management System for Plant Protection Products Authorizations of DGAV (SIFITO)[24]	Database extract that can be loaded to MU DWH	CSV
14	Italy IT	Southern	International Center for Pesticides and Health Risk Prevention (ICPS)[25] and Ministry of Health & other	PESTIDOC[26]: part of database extract; overview download; data to be retrieved otherwise	Microsoft access; CSV
15	Croatia HR	Southern	Croatian Agency of Agriculture and Food (HAPIH)[27], and Directorate General for Agricultural Land, Plant Production and Market (Ministry of Agriculture)	Repository under constructions; only link to search available	Database extract currently cannot be provided

[14] www.pta.agri.ee
[15] https://portaal.agri.ee
[16] https://www.kemidigi.fi/kasvinsuojeluainerekisteri/haku
[17] https://tukes.fi/etusivu
[18] Plant protection products - Dataset avoindata.fi
[19] https://eng.mst.dk
[20] https://middeldatabasen.dk
[21] https://www.kemi.se
[22] https://vatis.vatzum.lt/aapSarasas
[23] www.dgav.pt
[24] https://sifito.dgav.pt/Account/Login?ReturnUrl=%2F
[25] www.icps.it
[26] ICPS Web application: PESTIDOC Detailed pesticide technical-scientific documentation (https://www.icps.it)
[27] www.hapih.hr

Figure 1b. The EU Member States for which the investigation of PPP-register data evolved during the second phase are marked with yellow triangles. The second phase allowed for drawing a more resolute conclusion towards the final country selection. For the unmarked 12 EU MSs the second-phase investigation is still ongoing as a more long-term procedure.

2.4 Selection of PPP-register databases of EU member states

It was further checked whether the EU regulatory, (agro)climatic and biogeographical zones are represented in the selected candidates (EU MSs for estimating GenAP) and if comparison between them was allowed, and when these conditions were also met the EU MSs selection

was finalized (Step 3, questions 6-8 of Fig. 1a). Apart from the **Czech Republic** (prototype), the selected EU MSs are **Austria**, **Estonia**, and **Portugal** (**Fig. 1c**, **Table 4**). It should be clarified that Austria was selected as a second MSs of the central zones not only to allow for current and future comparisons between MSs within the same zone, but also because the central EU regulatory zone includes more MSs (11) than the North (6) or the South (9) zones (see also **Fig. 1d**).

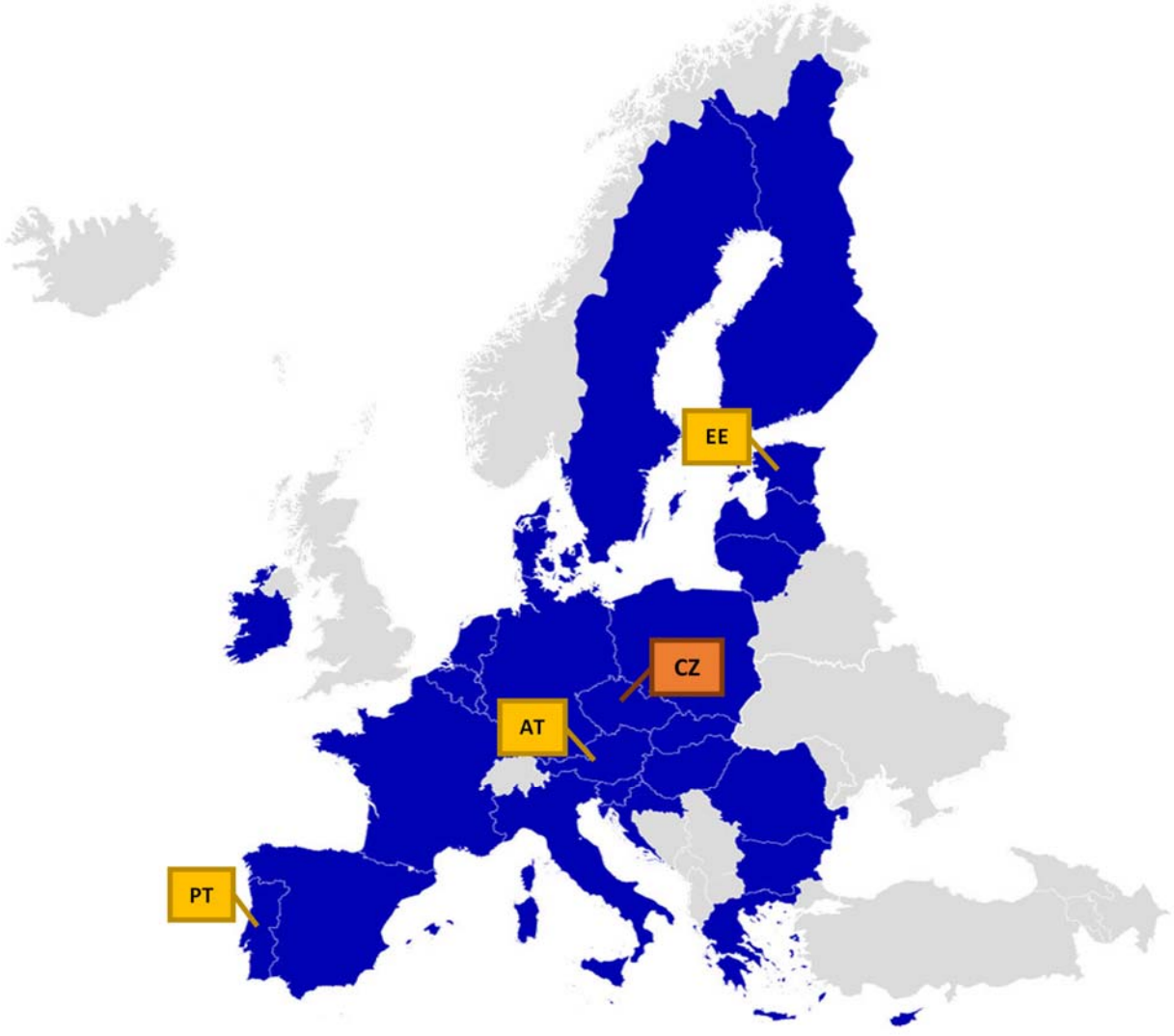


Figure 1c. Selected EU Member States for PPPs' Generalized Application Patterns. Czech Republic (CZ) is marked with red colors to indicate that CZ is the reference model (**prototype**) in the study design of Generalized Application Patterns of PPPs.

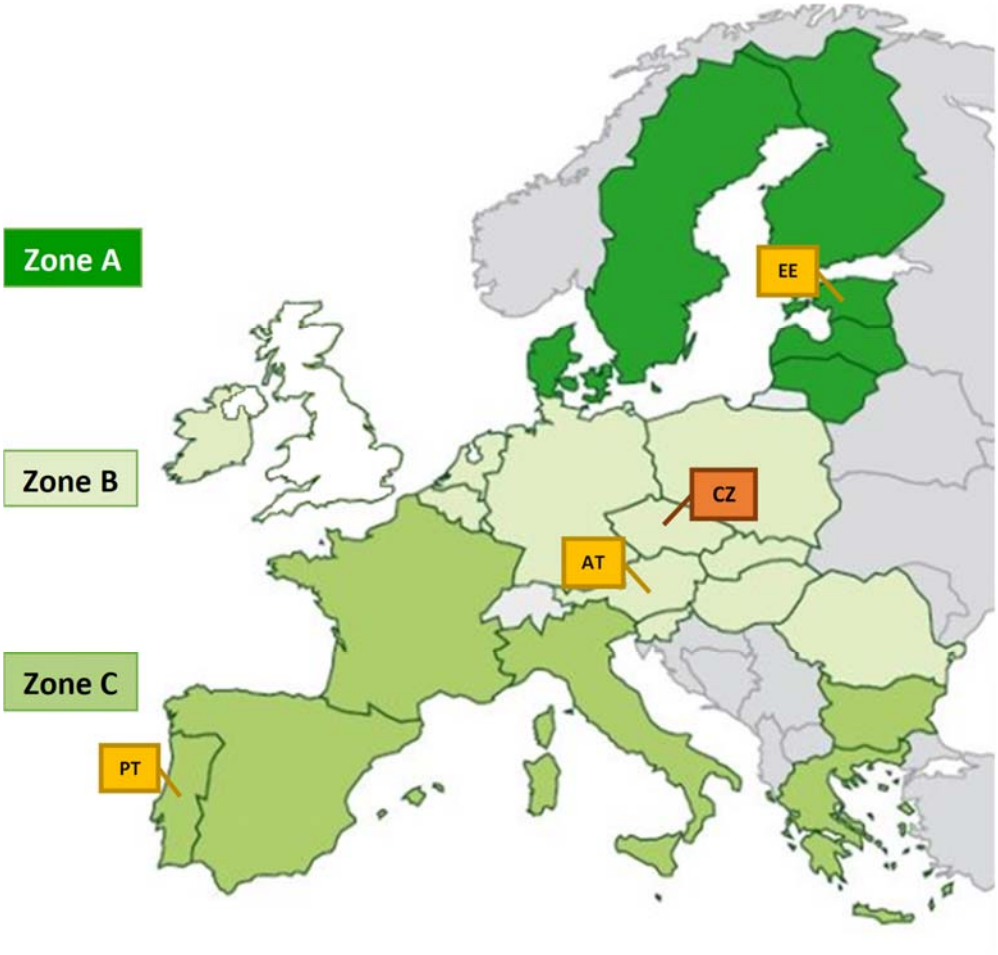


Figure 1d. Selected EU Member States for PPPs' Generalized Application Patterns across the EU regulatory zones of PPP authorization according to Annex I of (EC) Regulation 1107/2009 [28]. Zone A is the North zone and indicated with dark green color on the map, zone B is the Central (lightest green color) and zone C is the Southern zone (light green). It is noted that in this map Croatia was not marked yet as an EU Member State, while currently comprises an EU MS of the southern regulatory zone (Zone C).

[28] Annex I – Definition of zones for the authorisation of plant protection products as referred to in Article 3(17) of Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market. EUR-Lex - 02009R1107-20221121 - EN - EUR-Lex (europa.eu)

Table 4. Table of the selected EU Member States (MSs) of the study for the Generalized Application Patterns (GenAP) of PPPs. The respective EU regulatory, agro-climatic, and biogeographical zones are shown for each EU MS.

EU Member State (acronym)[29] [30]	EU Regulatory Zone[31] [32]	European (agro-)climatic Zone[33] [34]	European Biogeographical Region[35]	EPPO climatic zones for PPP efficacy evaluation[36]
Estonia (EE)	North (A)	Nemoral	Boreal	North-East
Czech Republic (CZ)	Central (B)	Continental/Pannonian	Continental/Pannonian	Maritime
Austria (AT)	Central (B)	Continental/Pannonian/Boreal South	Continental/Alpine	Maritime
Portugal (PT)	Southern (C)	Mediterranean/Maritime south	Mediterranean/Atlantic	Mediterranean

[29] European Union (EU) in EPPO Global Database

[30] Council of the European Union, *The Member States of the European Union*, Publications Office, 2020, https://data.europa.eu/doi/10.2860/082123

[31] The (inter) zonal system for plant protection products | Plant Protection Products | Board for the Authorisation of Plant Protection Products and Biocides (ctgb.nl)

[32] Annex I – Definition of zones for the authorisation of plant protection products as referred to in Article 3(17) of Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market. EUR-Lex - 02009R1107-20221121 - EN - EUR-Lex (europa.eu)

[33] Observed climate zones in the period 1975-1995 (left) and 1996-2016 (right) — European Environment Agency (europa.eu)

[34] Ceglar, A., Zampieri, M., Toreti, A., Dentener, F., Observed northward migration of agro-climate zones in Europe will further accelerate under climate change. Submitted to Earth's Future.

[35] Biogeographical regions in Europe — European Environment Agency (europa.eu)

[36] https://www.eppo.int/media/uploaded_images/ACTIVITIES/plant_protect_products/zonal_assessment/EPPO_zonal_maps.pdf

3. Raw data from national PPP-registers and their retrieval

The main properties/parameters that are displayed online and are part of the content of the database extracts in the studied PPP-registers of the four selected EU MSs are shown in **Table 5**.

Differences, limitations, and challenges of raw data are in these PPP-register databases are explained in the below lines.

Open links (websites) versus database extracts

The information displayed on a website is not identical to the information included in a database extract or in a set of database extracts that is provided/retrieved. This stands especially for the EU MSs that reposit PPP-register data at multiple sources.

Language

The national PPP-registers of EU MSs are including information in the national languages. This is true for the links of authorities that entail the search options for PPPs, but the English version can be also available (e.g., AT, Estonian). However, the database extracts and downloads include information in the national language at a resolute level.

Structure

The data are organized in sets of tables that are based on a different approach for each EU MS. Even when similarities are identified, the structure across EU MS is variable. This approach can be linked with the way the data are displayed publicly on the authorities' websites.

For instance,

- CZ has a very detailed multi-layer organized database with many separated code lists (for crops, pests, a.s., uses, applications ...) that are interlinked in the very complicated nested structure.
- PT distinguishes the cancelled from approved authorizations at different CSV exports, while also separates into different CSV exports the information on the PPP registration (approval identification numbers and status), and the information on the PPP application that is listed in different columns (crops, pests, application rates and intervals).
- AT follows a structural approach of layering from general data to thorough data that are specified to applications in each crop-type.
- EE distinguishes related data in a comparable way like PT.

Content organization (columns and properties)

Apart from variability in language, there is variability in reporting specific data under each property or parameter. For example, the Portuguese database includes in the same column the application-rate value and its unit. In the AT database these are reported into distinguished columns entitled as "application rate" and "unit" respectively. Likewise, the

active ingredients entailed in each formulated product are reported by an EU MS in distinguished columns while in other cases into one single column.

Terminology & Regulations

Specific types of authorizations or uses (sometimes called applications / indications etc.) are grouped or indicated in diverse ways. In the AT schema, for instance, specific types are precisely categorized in accordance with the Article number in a respective Regulation[37]. While in other EU MSs, acronyms, letters, or words are used to specify such categories. This categorization is not always distinguished under the same parameter or property of the PPP-register across the four EU MSs.

[37] Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market.

Table 5. Main properties found in the 4 studied national PPP-registers. x indicates the property is the content of the database extract.

Nr.	Parameter/Property	CZ	AT	EE	PT
1	Formulated products PPPs (commercial names) authorized in the EU	x	x	x	x
2	Authorisation identification code of approval-decision related to formulated PPP	x	x	x	x
3	Authorisation holder/decision owner and Marketing holder- Organisation	x	x	x	x
4	The date the decision of authorisation approval starts (is valid)	x	x	x	x
5	The date the decision of marketing approval starts (is valid)	x			x
6	The date the decision of authorisation approval expires	x	x	x	x
7	The date the decision of marketing approval expires	x	x	x	x
8	The date of end of use	x	x		x
9	Specific authorisation types related to formulated PPP use (e.g., professional)	x	x	x	x
10	Formulated PPP (preparation) type	x	x	x	x
11	General biological function of PPP related to mode of action	x	x	x	x
12	Specific mode of action description/notes for each PPP active substance	x	x		x
13	The pure content of the active substances in the formulated product	x	x	x	x
14	The way of formulated PPP application	x	x	x	x
15	The growing field of PPP application (e.g. outdoor)	x	x	x	x
16	Crop the formulated PPP is applied (common crop-names in national language)	x	x	x	x
17	Crop the formulated PPP is applied (Scientific name of specific crop-plant taxon)	x	x		x
18	Crop the formulated PPP is applied (EPPO code of specific crop clas)	x	x		x
19	Pest targeted by the formulated PPP application (common pest-names in national language)	x	x	x	x
20	Pest targeted by the formulated PPP application (Scientific name of specific pest taxon)	x	x		x
21	Pest targeted by the formulated PPP application (EPPO code of specific pest clas)	x	x		x
22	The (min-max) application rates the formulated PPP is applied in each crop	x	x	x	x
23	The intervals between (multiple) applications	x	x	x	x
24	The number (max times) of applications per season/year	x	x	x	x
25	The BBCH stage as upper- and lower- bound numbers	x	x	x	x
26	Phenology, growth, application and use information and details	x	x	x	x
27	Pre-harvest intervals/ safety intervals	x	x		x
28	Units to application rates and pure contents (e.g., kg/ha)	x	x	x	x
29	Mitigation required: buffer zones and/or drift-reduction	x	x		x
30	Classification and Labeling information	x	x	x	x
31	Restrictions	x	x	x	x

4. Data processing, analyses and extraction

4.1 Data import, control, analysis, and processing

As shown in the general scheme in **Fig. 2**, from the selected MSs (see above) the data were **retrieved** (downloaded) in appropriate format: XML for CZ, XML and JSON for AT, CSV for EE, and PT. The data were imported into the **MU data warehouse (MU DWH)**, the database running on **PostgreSQL**[38]. MU DWH integrates data and information collected from various sources into one comprehensive database, the data from sources is processed in the staging to be finalized at the mart level where the data changes to another shape useful for reporting or analyzing. The GitLab[39] has been used to share the current state of the code in the development branch.

Then, the first **control and exploration** were performed. If needed and possible, the **corrections** of the source data were managed with cooperation of the national PPP authorities (source of the data).

After the needed data format and content was established, data were **analyzed and engineered** by multiple ways (re-organized, corrected, grouped, filtered, aggregated, concatenated, translated etc.) as needed to process towards **extraction of the desired data for the GenAP**.

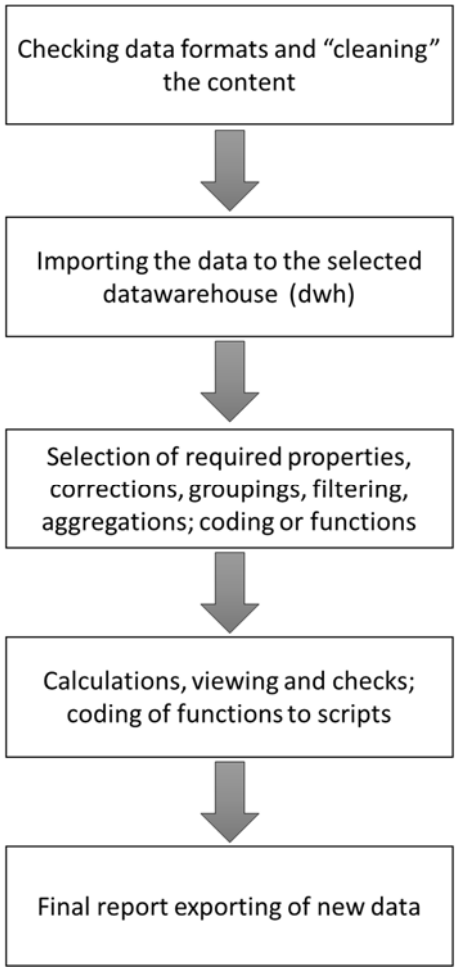


Figure 2. Simplified flowchart of data processing towards deriving the output of Generalized Application Patterns of PPPs.

[38] https://www.postgresql.org
[39] https://about.gitlab.com

These processes with the data were done in PostgreSQL using **library of SQL scripts that were developed specifically for the SPRINT project** and that enable not only to build GenAP but also detailed, repeatable, recordable, and trackable analysis, exploration, and extraction of the national PPP-register data (see **Annex D** as an example). The database was created using the DBT[40] framework. Database schema and ETL (Extract, Transform, Load) processes are stored in source code (primarily SQL and Python[41]). The whole data processing pipeline can be executed anytime to process new and updated data from relevant sources. Source code can be executed by individual researchers in their own isolated environment as well to enable parallel development without affecting each other. For the daily and user-friendly connection to MU DWH, work with the datasets, and for the building of the scripts, DBeaver universal database tool[42] was used.

The Czech PPP-register is presented for the public using the web search engine interface[43]. For the SPRINT project, the data are retrieved from electronic data exchange service of Czech Ministry of Agriculture (EAGRI)[44] as a large XML file. From this file, a set of interlinked tables are sourced to form a relational database.

The Austrian PPP-register is publicly available on the search engine of AGES[45] in English and German version. There, detailed information can be found comparable to the CZ PPP-register database. Also, Excel files can be exported with the lists of chemical PPP products, microbiological PPP products, a.s. information and some statistics. All information of the PPP-register at the search engine of AGES is reposited and was provided to MU by AGES as a database extract in the format of XML together with two schemas (JSON swagger format; the properties and structure of the database, and German and corresponding English terms of properties included). The AT PPP-register is not listing adjuvants, as there is no national regulation/approval system in place in AT, only data on PPPs.

The Portuguese PPP-register, SIFITO, the Management System of Plant Protection Product Authorizations of PT, contains all PPP authorizations in force, with all amendments made included. At the SIFITO website[46], four CSV files can be exported, and these comprise the database extracts of the Portuguese PPP-register.

The Estonian PPP-register contains a set of data that are openly displayed using the search engine of the Estonian Ministry of Agriculture[47]. The information is displayed in Estonian and English language. Information can be directly retrieved in CSV file extension format as open data at the website of the Estonian Agriculture and Food Board (Ministry of Agriculture)[48]. All enclosed data is in Estonian and are organized in three different CSV exports. The Estonian PPP-register is updated regularly.

[40] https://www.getdbt.com
[41] https://www.python.org
[42] https://dbeaver.io
[43] https://eagri.cz/public/app/eagriapp/POR/Vyhledavani.aspx?type=0&vyhledat=A&stamp=1673591483569
[44] https://eagri.cz/public/web/mze/farmar/elektronicka-vymena-dat/prehled-vystavenych-sluzeb
[45] https://psmregister.baes.gv.at/psmregister/faces/main;jsessionid=-2-aOdyzIW7YN2dZa0khkejaJkPLJ0n8NOlY-A1fmz1BLViYPpQe!1367250002?_adf.ctrl-state=7zff4eltt_4
[46] https://sifito.dgav.pt/Account/Login?ReturnUrl=%2F
[47] https://portaal.agri.ee/avalik/#/taimekaitse/taimekaitsevahendid-otsing
[48] https://avaandmed.eesti.ee/datasets/taimekaitsevahendite-register

4.2 Building the structure and content in the data warehouse

Czech PPP-register data are retrieved from electronic data exchange service of Czech Ministry of Agriculture (EAGRI)[49] as a large XML file. The XML has been imported to the database as an object and then, the PHP Symfony framework has been used to manage and run the SQL code through PostgreSQL to parse the XML. Symfony manages running SQL codes for creating the source and temp tables respecting the relationships including creating the constraints (primary and foreign keys) to ensure the accuracy of the parsing. Following that, Symfony runs SQL codes for inserting the data parsed from XML in the tables. Most of the entities were without identifiers (which are used in the database as primary keys) that was mandatory for creation of the hash values from the content of the entity as an identifier. As the result, a set of interlinked tables formed a relational database with the following entity relationship diagram (**Fig. 3**).

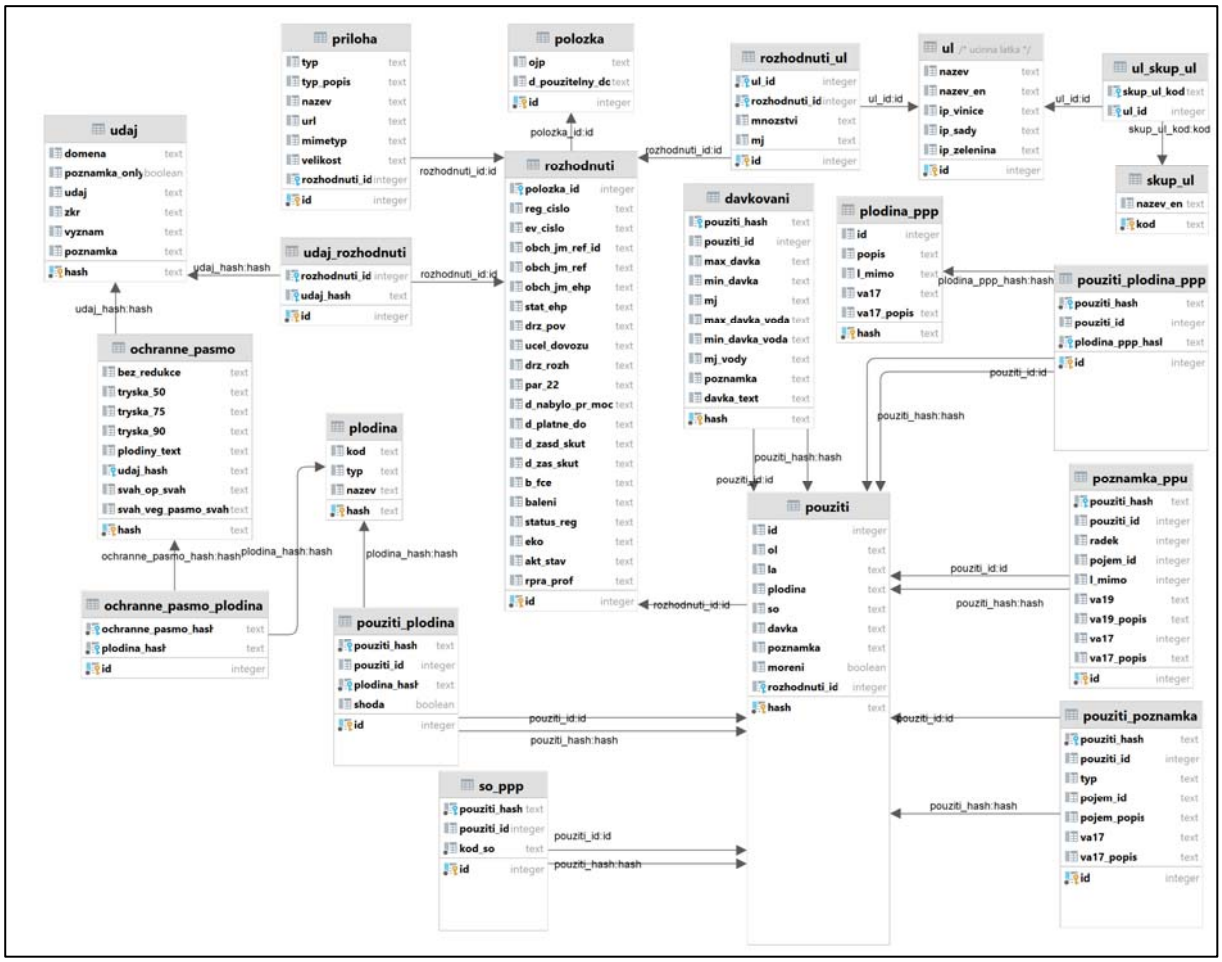


Figure 3. Entity Relationship diagram (ERD) following import of the CR PPP-register data from EAGRI XML to the MU DWH. Terms used in each box represent quotations that correspond to specific parameters and properties of the CZ PPP-register. The actual parameters and properties included in each PPP-register and used for the GenAP output of CZ are described in Tables 5 and 6.

[49] https://eagri.cz/public/web/mze/farmar/elektronicka-vymena-dat/prehled-vystavenych-sluzeb

Regarding **the Austrian PPP-register data**, similar to the XML data of the EAGRI of CZ, SQL has been used through PostgreSQL for parsing the XML of AT and has been managed by PHP Symfony. The problem of missing the identifiers has been solved by hashing the content of the entities. The generated outcome of this process is demonstrated in **Fig. 4**.

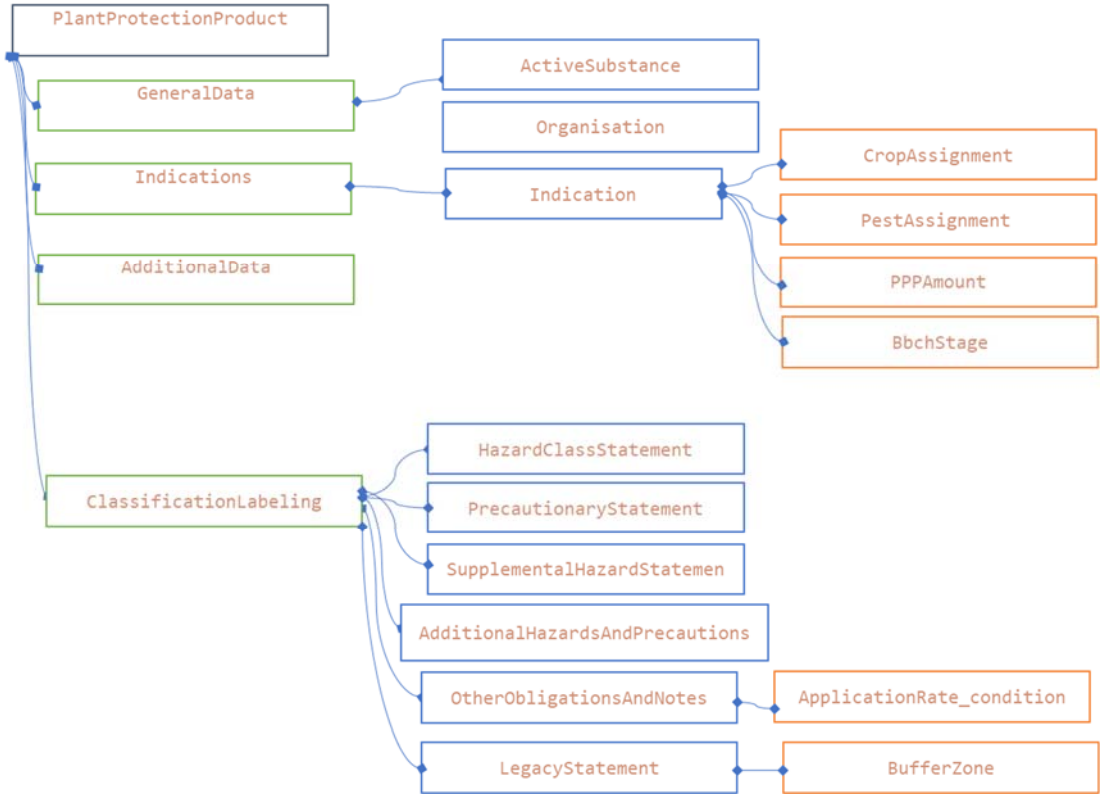


Figure 4. This Entity Relationship Diagram (ERD) was generated from XML database extract and related schema (*.xsd), provided by AGES, using PostgreSQL, PHP Symfony framework to export the data to the DWH. Terms used in each box represent quotations corresponding to specific parameters and properties of the AT PPP-register. The actual parameters and properties included in each PPP-register and used for the GenAP output of AT are described in Tables 5 and 7."

Data from **the Portuguese PPP-register** has been obtained as Excel files which could be imported to the database using DBT (Data Build Tool) in case we convert the files to CSV, but the Excel from Microsoft was not efficient in converting those files because of the encoding of the Portuguese. The best way was to convert the content to CSV files using Python code.

Data from **the Estonian PPP-register** were downloaded as Excel tables and processed with Excel Editor Query (see detailed description in Chapter 4.3.4).

4.3 Data processing to extract the Generalized Application Patterns

In Figure 2 (Chapter 2.2 of this Report), the general overview of data processing and analyses is displayed and can be accessed.

4.3.1 Czech PPP-register data processing and GenAP extraction

As described above, data in MU DWH were organized in the complicated relational database containing multiple interlinked tables (see **Fig. 3**). Using SQL scripts, the data were analyzed, transformed, and processed by number of ways not only to extract GenAP at the end (see below a brief description), but also to deeply explore and deliver valuable information about PPPs authorized in CR (see **Annex B**).

The major and central entity in the database is a **decision** (= rozhodnuti). This is a unique authorization record with the clearly set the dates (start of validity, end of validity, end of use, end of marketing) and also other important parameters of the given authorization, e.g., authorization mode (inc. minority uses) and status, biological function, authorized user (e.g., professionals), parallel trade data etc. The decision is related to many other entities that further specify the unique authorization record for PPP. These specify **PPP** (name, validity), **a.s.** (name, concentration in PPP), **use** and **application** (crop, pest, pre-harvest interval, application rate, time, number per year, seed treatment and other conditions), and statements (labelling, safety sentences, buffer zones).

To extract GenAP from the data, information from the various tables needed to be connected and extracted. Thus, many tables were **joined** using their constraints, primary and foreign keys: decisions (rozhodnuti), PPP (polozka), a.s. (ul), use (pouziti), application (davkovani), crop (plodina), using the interlinking helping tables. Also at this step, the data were joined with several **newly created tables** (see below more details) containing application time (BBCH) data, maximum number of applications per year, and other application notes. Tables containing crop EPPO codes, and EPPO foreign language name database were joined to enable crops translation from Czech language.

The resulting data were filtered to **valid and use-of-stocks decisions** for GenAP presented in this deliverable. However, using decision validity data, the GenAP can be also generated backwards for the history, for instance for year **2021**, where SPRINT sampling campaign was performed. As described in the previous chapters, garden, and home use (non-professional users) and parallel trade authorizations were also filtered out. The data were also filtered to keep only authorizations where a.s. concentration in PPP and application rate were expressed in **units** that enable recalculation of the PPP based application rate to a.s. based application rate. That means that only units for a.s. concentration in PPP like g/kg or g/l were kept and units for application rate like g/ha or l/ha were kept. Other units were filtered out; however, these accounted for only about 5% of all authorizations. Last, **prioritized crops and a.s. were filtered** (see the following chapter) were filtered to reduce the size of GenAP.

In the resulting data, **a.s. based application rates were calculated** from the PPP based application rate and a.s. concentration in PPP. **BBCH values** were extracted from the relevant part of the database, where they were present as numbers (sometimes multiple alternative values for the same PPP x crop authorization) or even in the text application notes (from these they were extracted using SQL string functions). **The maximum number of applications per year** were extracted from another type of application note by SQL string functions. The **database of EPPO codes** and names for crops in other languages were downloaded from EPPO website[50], converted to database that was joined to the data and enabling each crop name in Czech language to be **correctly and automatically translated** to English or Latin language (as set preferred in the EPPO database).

The last step to GenAP was **grouping of the data** resulting from the previous steps to get **distinct a.s. x crop combinations** and for them, i) **typical application rates** calculated as the minimum and maximum of all application rates authorized for the given a.s. x crop combination, ii) **typical application time** calculated as the minimum of all lower BBCH bound values and maximum of all upper BBCH bound values authorized for the given a.s. x crop combination, iii) **maximum number of applications per year** authorized for the given a.s. x crop combination. The resulting GenAP table was exported to MS Excel table for more comfortable viewing. This is presented in **Annex A1**.

4.3.2 Austrian PPP-register data processing and GenAP extraction

The XML content was imported to a source schema in the MU DWH. The data were organized according to original tagged and nested structure in XML and unique hashes were generated (as primary keys), resulting in the data relations displayed in **Figure 4**. Then, the database structure and data content were viewed and checked using Visual Studio Code[51]. Clarifications on corresponding fields and differences in terms between the online search and the database extract were provided by the chief specialists of AGES.

Data was processed in DBeaver using PostgreSQL. The different tables were **fully joined**, and **filtering** was applied. Using SQL scripts, the data were analyzed, transformed, and processed by number of ways to extract GenAP.

The major tables that were joined via their relationships were **crop_assignement, pest_assignement, active_substance, general_data, and indications**. The main central identity in the database is the registry number and the product name. In a comparable way to the CZ PPP-register, these major tables specify all needed information about the PPP application, crop, pest, and authorization decision. The AT schema includes information on crop and pest by common name in German, scientific name of the taxa, and corresponding EPPO codes. In the "indications" table the maximum application rate of the formulated PPP is included, therefore only this was used in the below described estimations.

Filtering out data was applied in a harmonized mode with the CZ database processing. As described above for CZ, garden and home use, non-professional uses, as well as seed

[50] https://data.eppo.int
[51] https://code.visualstudio.com

treatment were filtered out. The data were also filtered out according to units like CZ to enable estimations. Last, **prioritized crops and a.s. were filtered** (see the following chapter) were filtered.

After filtering out undesirable data, the units for a.s. concentration and application rate have been checked. For the a.s. concentration, the units, g/kg or g/l were the only two units in this case after all the data filtrations. For the application rate the unit conversion was carried out, when necessary, into Kg/ha or L/ha. Then the calculation was carried out to obtain the application rate for each active substance. In the resulting data, **a.s. based application rates were calculated** from the PPP based application rate and a.s. concentration in PPP.

The last step to GenAP was **grouping of the data** resulting from the previous steps to get **distinct a.s. x crop combinations** and for them, i) **typical application rates** calculated as the minimum and maximum of all application rates authorized for the given a.s. x crop combination, ii) **typical application time** calculated as the minimum of all lower BBCH bound values and maximum of all upper BBCH bound values authorized for the given a.s. x crop combination, iii) **maximum number of applications per year** authorized for the given a.s. x crop combination. The resulting GenAP table was exported to MS Excel table for more comfortable viewing. This is presented in **Annex A2**.

4.3.3 Portuguese PPP-register data processing and GenAP extraction

From 4 files retrieved from the SIFITO website, two CSV files entail information on the **PPP approval decisions, applications and uses** (crops, pests, product names and registration codes, active substances, and content, etc.). The other two entail respective information for the cancelled ones. This register is detailed like the AT and the CZ registers of PPPs. However, the information related to specific properties is included in one column (e.g., the minimum with the maximum application rate values with the unit of the product). Like the AT, includes information on crop and pest by common name in Portuguese, scientific name of the taxa, and corresponding EPPO codes. Like CZ, apart from indicating the lower and upper BBCH bound with numbers, it entails descriptive information written as text, or combination of BBCH values and description (as phenological stage and season).

The Portuguese database is organized in a way that includes one register number for each product (like the AT and the Estonian registers) and contains the first (parent authorization) date and the expiration date (this is updated based on the latest authorization decision). However, the same preparation may have more than one authorization in force.

Only the authorized (and not the cancelled) PPP data were utilized to derive GenAP for PT. The two CSV file extensions on authorized products and uses were imported and "cleaned" in the DWH (staging). The data were **connected via full join** and the "product name" and "registry number". SQL scripts were used to clean, correct, and filter the data. For each PPP, the presented a.s. in the formulation were reported in the same column separated by "plus" sign and in the next column their related concentration in the formulation was reported respectively in the same format. All the crop types for which each PPP is supposed to be used were also reported in one column in front of each PPP. First, the information related to each a.s. was separated (including name of the active substance, concentration in the formulation, etc.) and inserted in the table as different rows. Each row was designated to one a.s. in one PPP formulation. The same procedure was done for the crop types related to each. Afterward, the cleaning of the data was done by separation of

the digits from the texts in the necessary columns, mostly related to the values like concentration, application rate, etc. from their units. At the end units were checked for a.s. concentration and application rate and were harmonized in such a way that the final application rate was **g of a.s. per ha**. The information about **BBCH values** was extracted from descriptions reported in phenology column. It was seen that in some cases more than one BBCH stage (2 to 4 BBCH stages) for PPP application was reported; however, it was merely 10% of the cases so we ignored them and considered only the first BBCH lower and upper limit. Filtering out was harmonized with the strategy described in previous sections. Finally, the application rate was calculated, and the data was aggregated for each a.s. and crop type into GenAP table as described for AT and CZ databases.

The resulting GenAP table was exported to MS Excel table for more comfortable viewing. This is presented in **Annex A3**.

4.3.4 Estonian PPP-register data processing and GenAP extraction

The CSV files retrieved from the website of the Estonian Agriculture and Food Board include:

1. product names, registry codes, type, method of application, name of the authorization holder and their address.
2. crops and the pests based on the registry code and product name.
3. the active ingredients and their amounts based on the registry (product) code.

The PPP-register properties and content were explored, and requested clarifications were provided by the Estonian Agriculture and Food Board. Properties and terms were translated and determined, and then were verified by chief specialists of the Estonian Agriculture and Food Board.

From these CSV exports data on BBCH per application in each crop was missing. Data on crop conditions was provided directly to MU by the Plant Protection and Fertilisers Department of the Estonian Agriculture and Food Board, as an Excel export.

The data were checked according to product name and registry code, and the data were inserted on BBCH per application rate, and maximum number of applications in each crop, using the search engine of the Estonian Ministry of Agriculture[52]. Therefore, a fourth file was created including the missing information. The Excel format was then saved as CSV extension. It is noted all four CSV extensions have as a common property the "registry code". It is also noted that in the case of Estonian PPP-register, interdependencies of authorization approvals are not reposited as CSV or displayed online, and each registry code comprises a unique code that corresponds to a product commercial name.

At the time of data processing, the Estonian database consisted of 4021 entries of formulated PPP applications in crops. This is a relatively small database compared with other registers that contain over 60.000 entries on PPP authorizations. Such datasets can be mined and processed in Microsoft Excel using its Power Query Editor.

[52] https://portaal.agri.ee/avalik/#/taimekaitse/taimekaitsevahendid-otsing/

First, all files were transformed into Excel worksheet format, and then table connections were created. All table connections appeared in the "Queries and connections" pane and comprise queries that can be combined (either merged or summed).

All the connection tables were then merged gradually, by merging two connection tables each time via their common properties (matching columns), into one joint connection table that entailed all data (from the former four Excel worksheets) merged (over 7023 loaded row counts; Excel uses fuzzy-logic algorithm to combine the connection tables). By this process identical (for the same a.s.) duplicate rows are created that were deleted. The query dependencies were viewed and checked. Potential syntax errors in codes were checked in the Advanced Editor. The joined tables were then expanded and the columns that were needed for the estimation of GenAP were selected. The rest of the columns that were not needed were deleted.

The Estonian database has the advantage that it is specified whether the PPP is chemical or other, along with the growing field of PPP application, as well as the crop categories (major or rare) of Estonia, in distinguished columns. The filters applied were selection of outdoor application of chemical only PPP in the major crop classes of EE. Second filtering excluded other types, like adjuvants/safeners, auxiliary products, inorganic substances, etc., as described in Chapter 3 of D5.2. The major crops were, in addition, filtered as described in Chapter 3 of D5.2. Following the above, a dataset of 1800 rows, 30 crop classes and 97 (PPP chemical) a.s. was created and saved as Excel worksheet to enable error corrections for specific Estonian characters.

Where needed, the application rate of the formulated PPP (product) was converted to g/ha. The GenAP was then estimated for each a.s. and each crop class using the functions MINIFS, MAXIFS for minimum and maximum application rates within each crop class and across all formulated PPP including the same a.s.. In cases where only the maximum application rates were reported, the estimations for minimum of all application rate of a.s. were not performed due to missing information (blank cells in the GenAP output of EE). Likewise, the widest BBCH range (estimating the minimum lower bound and the maximum upper bound for each crop class and a.s.) were estimated. Data was further explored using the functions UNIQUE and COUNT. The output is presented in Annex **A4.**

4.4 Selections and data filtering

Selections (or data filtering) during their analysis and GenAP extraction were made according to:

- Active substance
- Crop
- Specific authorization types

Active substances

PPP-registers of EU MSs are characterized by high variability in the a.s. they list. Here the term means the chemical active substance, either organic or inorganic, and the microorganism or bioagent with pesticidal mode of action authorized for use in crop plant protection, as well as adjuvants or auxiliary products. For estimation of PPPs' GenAP in the

current study, the organic chemical a.s. of PPPs were selected. The selection was conducted in accordance with the types listed in the EU Pesticides Database[53], by excluding:

- basic risk substances
- low-risk active substances
- microorganisms

Combined with the above, the selection was conducted based on expert knowledge (scientific expertise and open-literature data[54] [55]) in the field of PPP assessments and related properties:

- Adjuvants and auxiliary products were also excluded, as, although foreseen in the Reg. (EC) 2009/1107, there is no EU-agreed risk assessment process for them comparable to organic chemical PPPs.
- Likewise, inorganic substances (like metals) known to hold specific properties that drive specific conditions in a risk assessment process (such as copper) or fall under rare particular application-patterns, were excluded.
- Microorganismal or bioagent (e.g., nematode) PPPs that might not be listed in the EU Pesticides Database were excluded based on expert knowledge.

This was also true for inorganic substances that also fall under the category of food and feed additives, and their properties are well-known and/or are similar with listed substances in the EU Pesticides Database.

Crops

Different crop types hold varied importance in each of the four EU MSs of the current study, in terms of polygons' frequency (nr), and area (in hectares (ha)). For the estimation of the PPPs' GenAP in each EU MS of the four, these criteria (frequency, and area) were taken into consideration and crop selection was conducted in a stepwise mode:

1. The frequency of polygons for all the crop types included in the first European crop Type Map (EU-Cropmap)[56] (d'Andrimont et al. 2021)[57] were retrieved after depicting the spatial data of EU-Cropmap using the ESRI ArcGIS Pro 3.0.0 software. The crop classes of higher frequencies were considered for each MSs (CZ, AT, EE, PT).
2. The Eurostat database of Agriculture[58] was used. The dataset of Eurostat Statistics for the "Crops by classes of utilized agricultural area in number of farms and

[53] EU Pesticides Database - Active substances (europa.eu)
[54] Data Sources - PubChem (nih.gov)
[55] EFSA Journal - Wiley Online Library
[56] Joint Research Centre Data Catalogue - European Union Crop Type Map - European Commission (europa.eu)
[57] d'Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M., & van der Velde, M. (2021). From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. Remote Sensing of Environment, 266, 112708. https://doi.org/10.1016/j.rse.2021.112708
[58] The Eurostat database of Agriculture: https://ec.europa.eu/eurostat/web/agriculture/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_ff6jlD0oti4U&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view

hectare by NUTS 2 regions"[59] were used to derive the data on utilized agricultural area and crop classes for each of the four EU MSs with two units of measure: polygon (seeded field) and hectare.

3. The Eurostat sources on pesticide use in agriculture (reference metadata)[60] were used to identify the sources of each EU MS (links to National Statistics Agencies' data[61] [62] [63] (except for AT that we directly used the information manager of Austrian Statistics Agency[64]). This data was also checked in comparison with 1 and 2. It is noted that in some cases data on crop production was also available on the Statistics Agencies' websites, and these, where possible, were also considered.
4. The Eurostat Classification Detail List[65] [66] was used to crosscheck the crop classes.
5. The crop categories of high importance consistent with the above cross-checks were selected.

It is noted that for Portugal, the EU-Cropmap data do not include crop categories of importance for this EU MS, and therefore the rest criteria and sources listed above were considered. It is also noted that the Estonian database enabled selection of main crop categories therefore the stepwise mode approach was applied following this.

Specific authorization types

Authorizations for Professional uses were only included. Garden or home use were excluded. The parallel trade was also excluded from the next analyses because they "copy" the data from standard registrations (there must be always a reference decision with the completely same use for the parallel trade). Seed treatment uses were also excluded because they cannot be clearly recalculated to field application rates.

4.5 Data limitations and challenges

Other challenges and issues treated were: i) Retrieving the BBCH information to specify the widest BBCH range for each crop and a.s.; ii) Identification of identical terms and properties in the national language and assignment to unique parameter in English in a harmonized way; iii) Manipulation of different units (e.g., dispensers, or tabs or pipettes); iv) Manipulation of dictation and typo- errors and of particular symbols during data processing; v) Achieving harmonized processes in differently organized and structured databases.

[59] The Eurostat dataset of "Crops by classes of utilised agricultural area in number of farms and hectare by NUTS 2 regions": https://ec.europa.eu/eurostat/databrowser/view/ef_lus_allcrops/default/map?lang=en
[60] The Eurostat sources: https://ec.europa.eu/eurostat/cache/metadata/en/aei_pestuse_esms.htm
[61] Czech Statistics: Spotřeba v jednotlivých letech (ÚKZÚZ) (eagri.cz)
[62] Estonian Statistics: https://andmed.stat.ee/en/stat/keskkond__pollumajanduskeskkond
[63] Portuguese Statistics: Statistics Portugal - Web Portal (ine.pt)
[64] Austrian Statistics: Crop production and farming - STATISTICS AUSTRIA - The Information Manager (statistik.at)
[65] The Classification Detail List of Agricultural Products of Eurostat: Europa - RAMON - Classification Detail List
[66] The Eurostat RAMON-Reference And Management Of Nomenclatures: Europa - RAMON - Classifications Download List

5 Output – Generalized Application Patterns

DISCLAIMER:

The data extracted from the national PPP-registers in this report are for the internal use within the SPRINT project. They cannot be used without the permission of the authors of this deliverable. The links and password to the tables cannot be shared with the third parties. The tables cannot be edited, copied or screenshots taken.

Also, the colleagues within the project SPRINT are kindly asked to contact firstly the authors before they use the data contained in the Generalized Application Patterns.

The content of the GenAP tables as well as the coding scripts are subjects of ongoing research and development (are updated and changed) in contrast with the GenAP tables presented in this report, that will not be updated beyond the delivery date.

The output of PPPs' GenAP explored for the four selected Eu MSs (CZ, AT, PT, EE) is presented in **Annexes A1 - A4**. The links and passwords to access the documents (Excel outputs) are given in **Table 6**. The tables cannot be edited or copied. The **extracted parameters** displayed in Annexes A1-A4 are listed and explained in **Table 7**.

Table 6. The links and password of access to the GenAP of CZ, AT, PT, EE.

Annex Nr	Annex name	link	password
A1	GEN_APP _PATT CZ	https://ucnmuni-my.sharepoint.com/:x:/g/personal/22745_muni_cz/EebJUjGTroNMkwkCKqQhQoYBLiFCAZFrBGTKH6TK4xGNgA	Sprint-GenAP(2023)
A2	GEN_APP _PATT AT	https://ucnmuni-my.sharepoint.com/:x:/g/personal/22745_muni_cz/EXOACsswBOBHmjGIWekhfhABKgGCz7ZWTXwW-CEclkcTIg	Sprint-GenAP(2023)
A3	GEN_APP _PATT PT	https://ucnmuni-my.sharepoint.com/:x:/g/personal/22745_muni_cz/EZdHry9ZH45KiIOB9EM9QIgB3TUOZ8A1Eivk9HOAlHjiug?e=D6A0AB	Sprint-GenAP(2023)
A4	GEN_AP_ PATT EE	https://ucnmuni-my.sharepoint.com/:x:/g/personal/22745_muni_cz/Ea9g1B1KGStAlYeyO7nYDdcBobzYDyZiWTu1gK_hiLcfWA?e=mnk3wd	Sprint-GenAP(2023)

Table 7. List of columns used and included in the Output of estimations of PPPs' GenAP in the four selected EU MSs. See also **Annexes A** and **B**.

Nr.	Column	Property description
1	as_name	the name of the active substance in English
2	crop_name	the common name of the crop in English for each country
6	min_as_appl_rate_g_ha	the minimum of all application rates authorized for the given a.s. x crop combination in g/ha
7	max_as_appl_rate_g_ha	the maximum of all application rates authorized for the given a.s. x crop combination in g/ha
8	min_from_bbch	the minimum of the lower BBCH bound of all authorized uses for the given a.s. x crop combination
9	max_till_bbch	the maximum of the upper BBCH bound of all authorized uses for the given a.s. x crop combination
10	max_number_appl	the maximum number of applications of all authorized uses for the given a.s. x crop combination

6 Conclusions

PPP data (national PPP-registers) are the most challenging data to collect and harmonize. So far, we have done an overview of these data in all European countries and described their parameters (metadata) in the project SPRINT Milestone 8. In this deliverable, we presented how we proceeded in a more detailed exploration of the resolute data of the national PPP-registers of the 27 EU MSs. For four data-rich EU MSs (CZ, AT, PT, EE), we have explored in detail the structure, content, and properties of these data to represent the central, northern, and southern zones of Europe. The CZ PPP-register was considered as a reference model (prototype), and we have translated it to the data warehouse (DWH) enabling flexible analysis by SQL scripts. The same strategy was followed for the other three data-rich EU MSs as was described in the chapters of this report.

The compilation of these four databases was achieved through long-term, systematic, and continuous contacts with the national authorities on data-structure reorganization, clarifications on methods of data collection or estimates, provision of information not available online, and unfolding data aggregation into comprehensive datasets. **Generalized most probable steps needed** to follow in each country considering the difficulties and time constraints expected, based on the experience with the four datasets of the current report are:

- A common harmonized platform with sources to PPP-registers
- More resolute data related to PPP application regimes, crops and pests targeted, available for each PPP in a format that allows for easier data analysis
- Common methodology and terminology in data analyzation and estimations
- English version available
- Clarification of the precision and frequency needed for data collection and updates (a more coordinated and systematic scheme is necessary for comparability across countries)
- To overcome confidentiality issues

It was challenging to compile and analyze all the data needed as explained in the previous chapters. However, **it is clearly demonstrated that the current database tools (e.g., SQL) pose very high predictive value regarding the PPP strategies and decisions**.

7 Acknowledgements

Table of Acknowledgements per EU MS listing the institutions/organizations that retrieved/shared open data and provided valuable clarifications to the repositories of the national PPP-registers for the selected four EU MSs of the current research presented in this report.

Nr.	EU MS	EU regulatory zone	Institution / Organization	Contribution type
1	Estonia	North	PTA Plant Protection and Fertilisers Department Agriculture and Food Board www.pta.agri.ee	Data provision and links; clarifications
2	Czech Republic	Central	EAGRI Ministry of Agriculture of the Czech Republic https://eagri.cz/public/web/mze UKZUZ Central Institute of Supervising and Testing in Agriculture https://eagri.cz/public/web/en/ukzuz/portal	Database provision and links; clarifications; changes in the XML export
3	Austria	Central	Department for Plant Protection Products Authorization Institute for Plant Protection Products Division for Food Security AGES - Austrian Agency for Health and Food Safety www.ages.at	Database provision; clarifications
4	Portugal	South	DGAV Direção-Geral de Alimentação e Veterinária Plant Protection Products Management and Authorization Unit Directorate for Health Protection Means https://www.dgav.pt	Data links and clarifications

8 Overview of attached ANNEXES

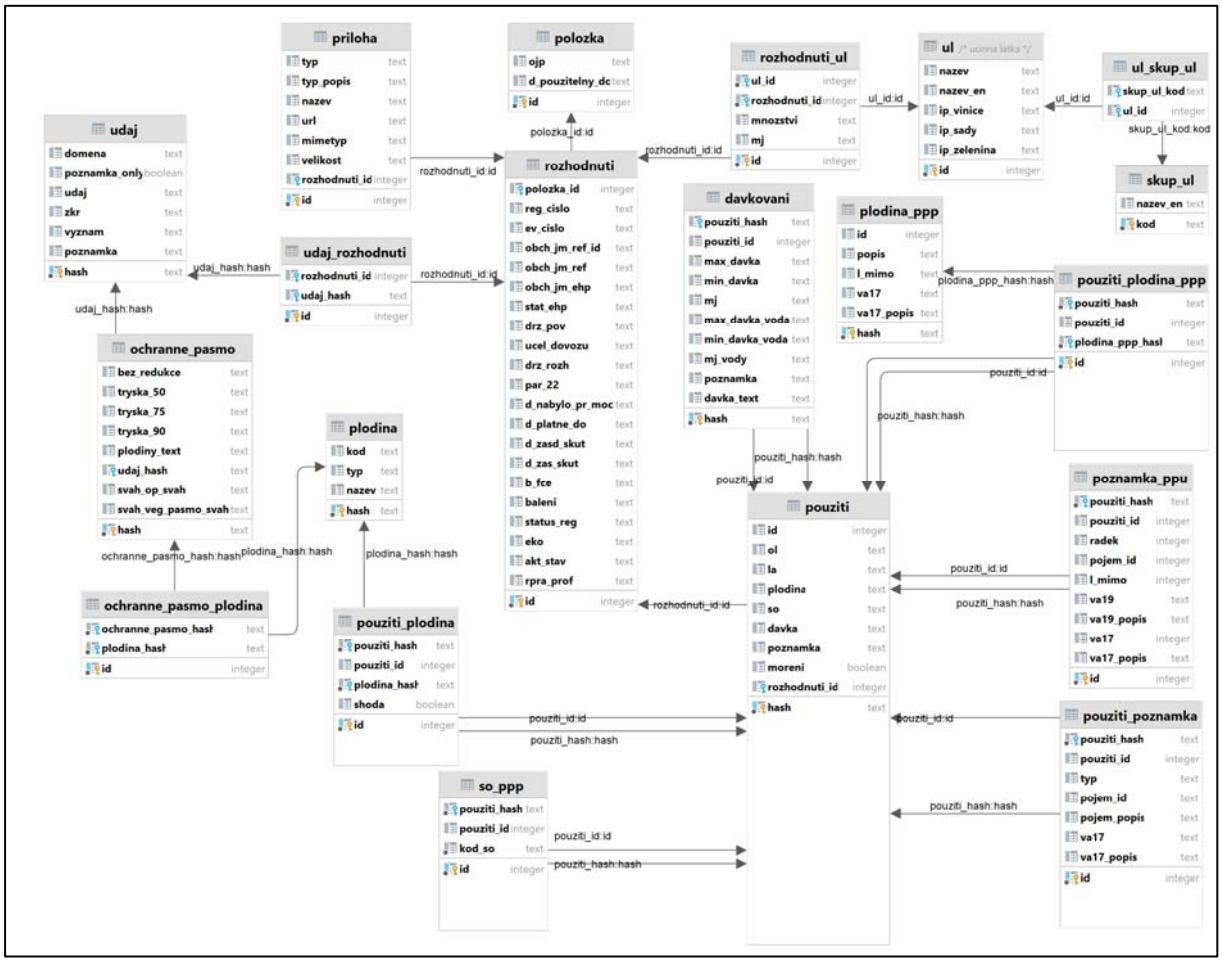
Annex	Title	Description
A1	Gen_APP_PATT CZ (Output of GenAP for Czech Republic (CZ))	The results of the extracted Generalised Application Patterns for the Czech Republic (CZ) as document of Excel format
A2	Gen_APP_PATT AT (Output of GenAP for Austria (AT))	The results of the extracted Generalised Application Patterns for Austria (AT) as document of Excel format
A3	Gen_APP_PATT PT (Output of GenAP for Portugal (PT))	The results of the extracted Generalised Application Patterns for Portugal (PT) as document of Excel format
A4	Gen_APP_PATT EE (Output of GenAP for Estonia (EE))	The results of the extracted Generalised Application Patterns for Estonia (EE) as document of Excel format
B	Czech PPP-register exploration and analysis	Czech PPP-register exploration and analysis showing the power of DWH to describe national PPP authorization data from multiple perspectives

Disclaimer: This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation program under grant agreement number 862568.

Annex B - Czech PPP register exploration and analysis

Czech register of Plant Protection Products (PPP) is presented for the public using the web search engine interface. For the SPRINT project, it is retrieved from electronic data exchange service of Czech Ministry of Agriculture (EAGRI) as a large XML file. From this file, as set of interlinked tables are sourced to form a **relational database** with the following entity relationship diagram (**Fig. 1**).

Fig. 1: *Entity relationship diagram of the Czech register of plant protection products*



Apparently, the register contains a plenty of details about current and historic (almost 25 years back) PPP authorizations in the Czech Republic. Beyond the goals of the SPRINT project WP5 and the deliverable D 5.2 (to extract the Generalized Application Patterns), these data have **enormous potential** to reveal interesting numbers, figures and statistics about PPP registered in the Czech Republic, their active substances (AS), target crops, use and application details. Therefore, in this Annex, several **examples of interesting outcomes** of such exploration and data-mining are presented. These following analyses are done on the data from 21st February 2023.

01. Decisions, PPP and registration numbers

The major and central identity in the database is **a decision** (= rozhodnuti). This is a unique authorization record for PPP specifying the allowed uses (crop, pest, application rate, time, number per year, pre-harvest interval, and other conditions). Each decision has clearly set the dates: **start of validity**, end of validity, **end of use**, end of marketing. Currently (XML database retrieved from EAGRI from 2023-02-21), there are **21,193 decisions with 3121 valid decisions, 436 use-of-stocks (UOS) decisions, and 17,636 non-valid decisions**.

In each decision, along with specification of PPP and its use, also other specifications for the authorization are available: i) is PPP for professional-only use or also for non-professional?; ii) is PPP authorized within parallel trade?; iii) is the registration standard or minor use?; iv) who is registering PPP?

Out of 3557 valid and UOS decisions, **non-professional use** is in 354 decisions (10 %). These are dominantly for garden or home use and should be filtered out from the agriculture related analyses of the data.

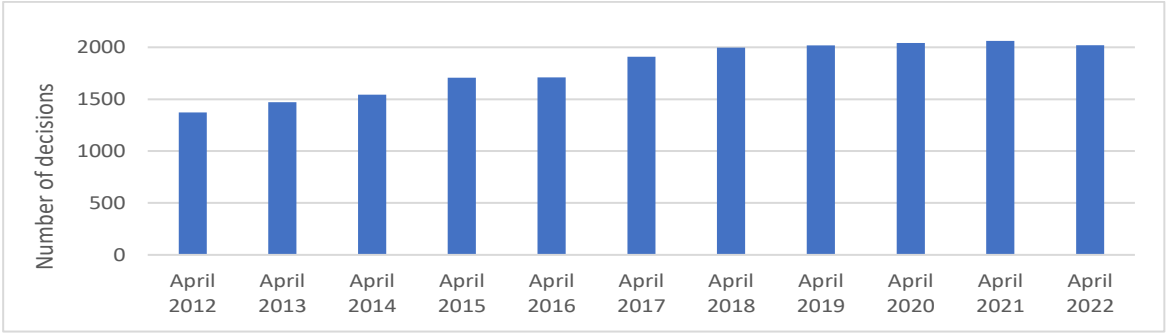
Out of 3557 valid and UOS decisions, **parallel trade** is in 1038 decisions (29 %). Out of all 21,193 decisions, there are 4257 decisions of parallel trade (20 %). The parallel trade decisions can be filtered out from the next analyses because they "copy" the data from standard registrations (there must be always a reference decision with the completely same use for the parallel trade).

Out of 3557 valid and UOS decisions, there are 63 decisions for **bioagens**, 25 decisions for **basic substances** and 194 decisions for **auxiliary products**.

After non-professional uses, parallel trade, bioagens, basic substances and auxiliary products are *filtered out*, there are **1961 valid and UOS decisions**. From these decisions, 232 (12 %) are minor use registrations. These must be kept for the next analyses, because minor use does not mean low amount used or small area of application.

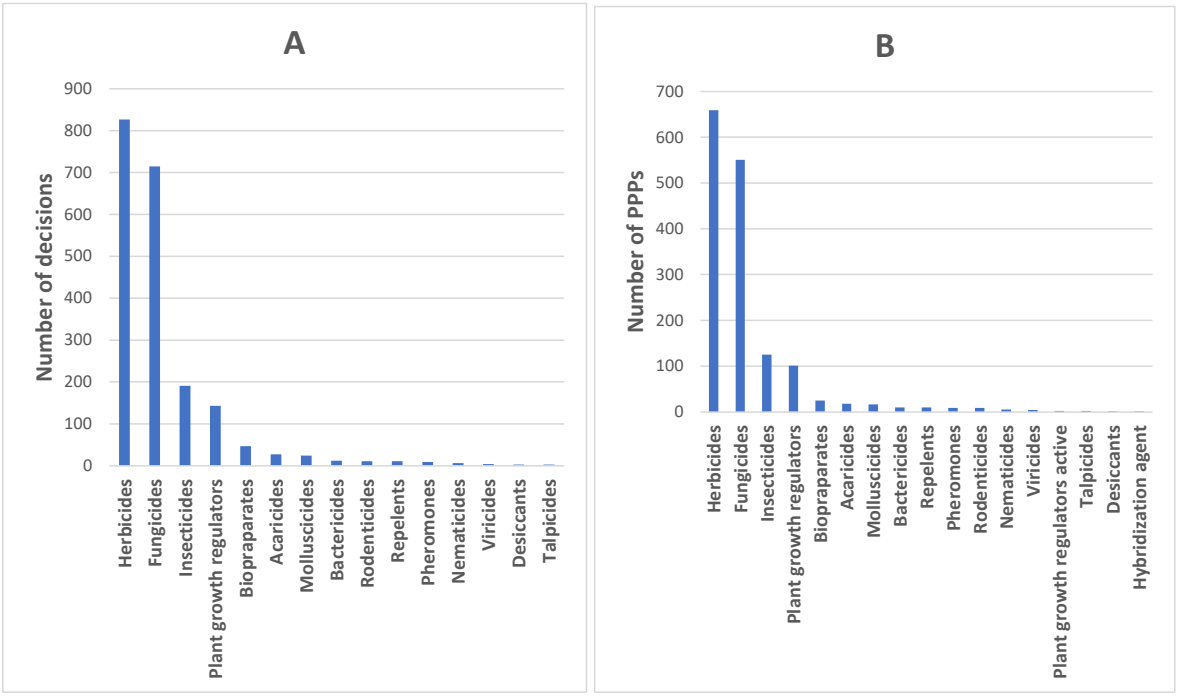
Using validity dates from the decisions, valid and UOS decisions for the given **time period** can be determined. For example, decisions valid for the *whole* April should have start of validity before 1st April and end of use after 30th April. For the historic years there can be artificially more decisions (plus about 10 %) retrieved, because the non-valid decisions do not contain information on professional/non-professional use and both are involved. The increase of the number of decisions is apparent from **Fig. 2**.

Fig. 2: *Number of decisions valid for April of each year 2012 to 2022. The decisions for parallel trade, bioagens, basic substances and auxiliary products are filtered out. Non-professional users are not filtered out.*



The decisions and PPP can be affiliated to **biological function** which is indicated in the register data. With that, numbers of decisions and PPP belonging to each biological function are shown in **Fig. 3**.

Fig. 3: *Distribution of the decisions (A) and PPP (B) among different biological functions for valid and use-of-stocks decisions (2023-02-21) with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out.*



Same PPP (defined by name and/or registration number) may have many decisions. There are 1961 valid and UOS decisions (non-professional uses, parallel trade, bioagens, basic substances and auxiliary products filtered out) for **different 1495 PPP**. Thus, one PPP may have up to 19 decisions because of different status of registration (valid or UOS) or mode of registration (standard, national, minority ...). When parallel trade is filtered out, **PPP name** has usually unique **registration number**. In rare cases (up to 1 %), same PPP name may have more registration numbers because of different mode of registration or different target use in the decision.

02. Active substances (AS) in PPP

For the 1961 valid and UOS decisions (non-professional uses, parallel trade, bioagens, basic substances and auxiliary products filtered out), there are **239 different AS** used. For each AS, it is defined its concentration in the PPP and also if it is allowed for the integrated production. The AS are linked to their **biological function** which is indicated in the decision. With that, list of AS available for each biological function in the Czech PPP register are shown in **Fig. 4** and **Table 1**.

Fig. 4: *Distribution of the AS among different biological functions for valid and use-of-stocks decisions (2023-02-21) with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out.*

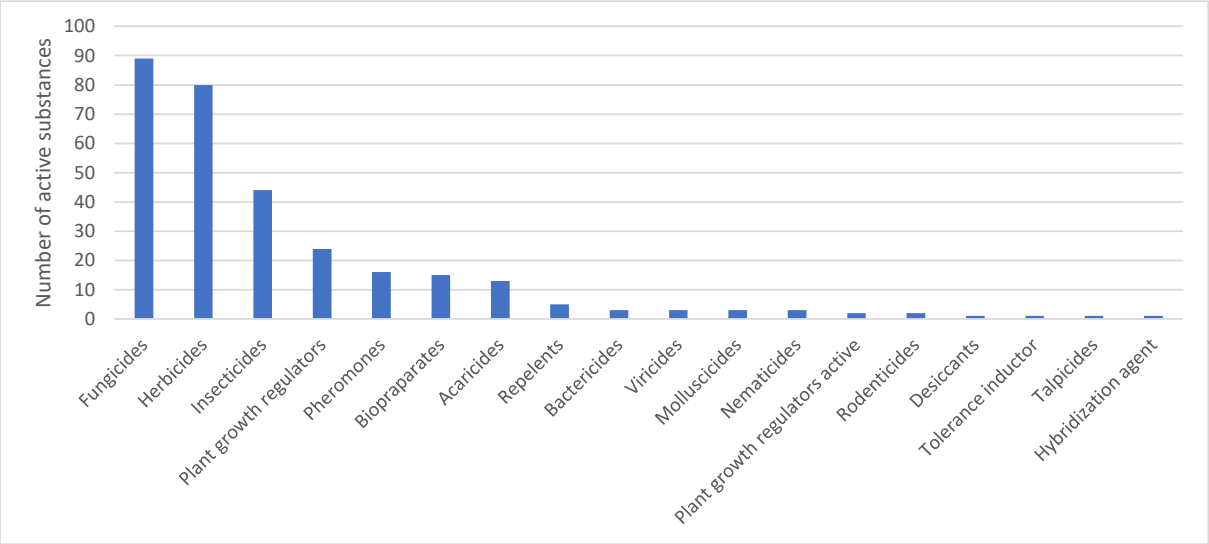


Table 1: *AS authorized under different biological functions for valid and use-of-stocks decisions (2023-02-21) with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out.*

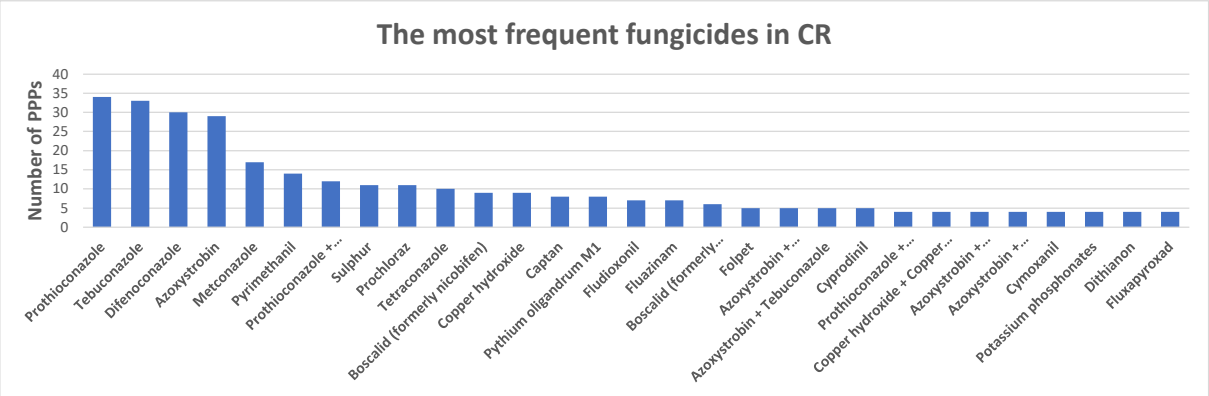
Biological function	AS names
Fungicides	Ametoctradin | Amisulbrom | Azoxystrobin | Bacillus amyloliquefaciens, strain FZB24 | Bacillus amyloliquefaciens strain MBI 600 | Bacillus amyloliquefaciens subsp. plantarum strain D747 | Bacillus pumilus QST 2808 | Bacillus subtilis str. QST 713 | Benalaxyl-M | Benthiavalicarb | benzoic acid | Benzovindiflupyr | Bixafen | Boscalid (formerly nicobifen) | Bromuconazole | Captan | Coniothyrium minitans Strain CON/M/91-08 (DSM 9660) | Copper hydroxide | Copper oxychloride | Cyazofamid | Cyflufenamid | Cymoxanil | Cyprodinil | Dazomet | Difenoconazole | Dimethomorph | Dimoxystrobin | disodium phosphonate | Dithianon | Fenhexamid | Fenpropidin | Fenpyrazamine | Fluazinam | Fludioxonil | Fluopicolide | Fluopyram | Fluoxastrobin | Flutolanil | Fluxapyroxad | Folpet | Fosetyl | Fosetyl-Al | Hymexazol | Imazalil | Ipconazole | Iprovalicarb | Isofetamid | Kresoxim-methyl | Mandestrobin | Mandipropamid | mefentrifluconazole | Mepiquat | Mepiquat-chloride | Meptyldinocap | Metalaxyl | Metalaxyl-M | Metconazole | Metiram | Metrafenone | Oxathiapiprolin | Paclobutrazol | Penconazole | Penthiopyrad | Potassium hydrogen carbonate | Potassium phosphonates | Prochloraz | Prohexadione | Propamocarb | Propamocarb hydrochloride | Proquinazid | Prothioconazole | Pseudomonas sp. strain DSZM 13134 | Pyraclostrobin | Pyrimethanil | Pythium oligandrum M1 | Sedaxane | silthiofam | Spiroxamine | Sulphur | Tebuconazole | Tetraconazole | Tribasic copper sulphate | Trichoderma asperellum strain T34 | Trichoderma atroviride, strain SC1 | Trichoderma harzianum strainT-22 | Trifloxystrobin | Triticonazole | Valifenalate | Zoxamide
Herbicides	2,4-D | Aclonifen | Amidosulfuron | Aminopyralid | Beflubutamid | Bentazone | Bifenox | carfentrazone ethyl | Chlorotoluron | Clethodim | Clomazone | Clopyralid | Cycloxydim | Dazomet | Dicamba | Dichlorprop-P | Diflufenican | Dimethachlor | Dimethenamid-P | Ethofumesate | Fenoxaprop-P-ethyl | Flazasulfuron | Florasulam | Fluazifop-P-butyl | Flufenacet | Flumioxazine | Flurochloridone | Fluroxypyr | Foramsulfuron | Glyphosate | Halauxifen-methyl | Imazamox | Iodosulfuron | Iodosulfuron-methyl-sodium | Isoxaflutole | Lenacil | MCPA | MCPB | Mecoprop-P | Mesosulfuron-methyl | Mesotrione | Metamitron | Metazachlor | Metobromuron | Metribuzin | Metsulfuron-methyl | Napropamide | Nicosulfuron | Pelargonic acid | Pendimethalin | Penoxsulam | Pethoxamid | Phenmedipham | Picloram | Picolinafen | Pinoxaden | Propaquizafop | Propoxycarbazone | Propyzamide | Prosulfocarb | Prosulfuron | Pyraflufen-ethyl | Pyridate | Pyroxsulam | Quinmerac | Quizalofop-P-ethyl | Quizalofop-P-tefuryl | Rimsulfuron | S-metolachlor | Sulcotrione | Sulfosulfuron | Tembotrione | Terbuthylazine | Thiencarbazone | Thiencarbazone-methyl | Thifensulfuron-methyl | Tribenuron-methyl | Triclopyr | Triflusulfuron | Tritosulfuron

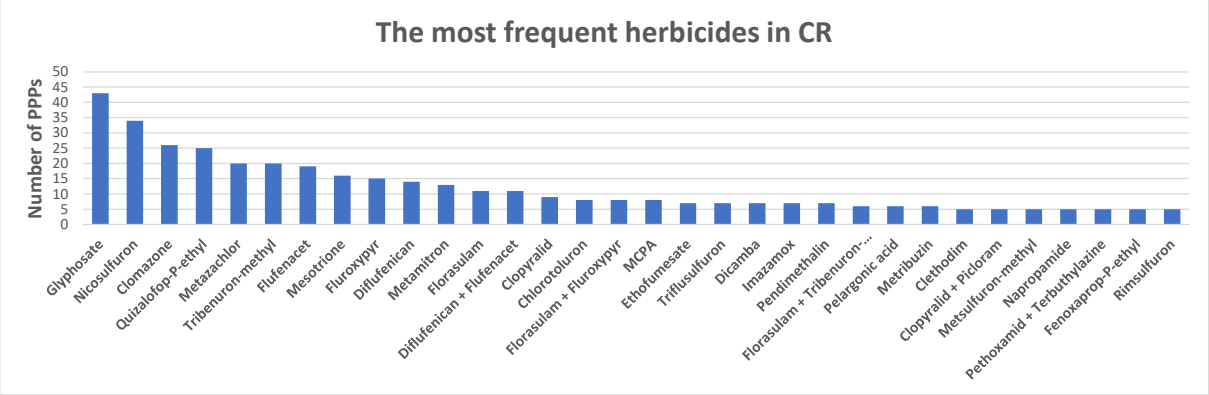
Insecticides	Abamectin | Acetamiprid | Alpha-cypermethrin | Aluminium phosphide | Azadirachtin | Bacillus thuringiensis ssp. aizawai strain GC-91 | Bacillus thuringiensis ssp. Kurstaki strain SA-11 | Bacillus thuringiensis subsp. kurstaki strain EG 2348 | Beauveria bassianastrain ATCC-74040 | Beauveria bassiana strain GHA | Buprofezin | Chlorantraniliprole | Cyantraniliprole | Cydia pomonella Granulovirus (CpGV) | Cypermethrin | Dazomet | Deltamethrin | Emamectin benzoate | Esfenvalerate | Etofenprox | (EZ)-7,9-dodecadien-1-yl- acetate | Fatty acid potassium salt | Flonicamid (IKI-220) | Fludioxonil | Flupyradifurone | Gamma-cyhalothrin | Lambda-cyhalothrin | Maltodextrin | Metalaxyl-M | milbemectin | Orange oil | Pirimicarb | Pirimiphos-methyl | Pyrethrins | Pyriproxyfen | Spinosad | Spirotetramat | Sulfoxaflor | Tau-fluvalinate | Tebufenozide | Tefluthrin | terpenoid blend QRD 460 | Thiamethoxam | Z-9-dodecenyl acetate
Plant growth regulators	1,4-dimethylnaphthalene | 1-methylcyklopropen | 1-naphthylacetic acid | 6-Benzyladenine | Chlormequat | Chlormequat chloride | Ethephon | Ethylene | Gibberellic acid | Gibberellins | maleic hydrazide | Mepiquat | Mepiquat-chloride | Metamitron | Metconazole | Orange oil | Paclobutrazol | Prohexadione | Prohexadione-calcium | Pyraclostrobin | Sodium 2-nitrophenolate | Sodium 4-nitrophenolate | Sodium 5-nitro-guaiacolate | Trinexapac-ethyl
Pheromones	Dodecan-1-ol | (E)-8-dodecen-1-yl acetate | (E,E)-8,10-dodecadien-1-ol | (E,E)-8,10-dodecadien-1-ol and n-tetradecyl-acetate | (E,E/Z)-7,9-dodecadienyl acetate | (E,Z)-3,8-tetradecadien-1-yl acetate | (EZ)-7,9-dodecadien-1-yl- acetate | (E/Z)-9-dodecenyl acetate | (E,Z,Z)-3,8,11-tetradecatrien1-yl acetate | Tetradecan-1-ol | (Z)-11-tetradecen-1-yl-acetate | Z-8-dodecenol | Z-8-dodecenyl acetate | Z-9-dodecenyl acetate | (Z)-9-tetradecen-1-yl-acetate | (Z)-tetradec-11-en-1-yl-acetate+ n-tetradecyl-acetate
Biopraparates	Bacillus amyloliquefaciens, strain FZB24 | Bacillus amyloliquefaciens strain MBI 600 | Bacillus subtilis str. QST 713 | Bacillus thuringiensis ssp. aizawai strain GC-91 | Bacillus thuringiensis ssp. Kurstaki strain SA-11 | Bacillus thuringiensis subsp. kurstaki strain EG 2348 | Coniothyrium minitans Strain CON/M/91-08 (DSM 9660) | Cydia pomonella Granulovirus (CpGV) | Mild Pepino mosaic virus isolate VC1 | Pepino mosaic virus strain CH2 isolate 1906 | Pseudomonas sp. strain DSZM 13134 | Pythium oligandrum M1 | Trichoderma asperellum strain T34 | Trichoderma atroviride, strain SC1 | Trichoderma harzianum strainT-22
Acaricides	Acequinocyl | Bifenazate | Cyflumetofen | Fatty acid potassium salt | Fenazaquin | Fenpyroximate | Hexythiazox | milbemectin | Orange oil | Pyridaben | Sulphur | Tebufenpyrad | terpenoid blend QRD 460
Repelents	Fat distilation residues | Fish oil | Quartz sand | Sheep fat | Ziram
Bactericides	benzoic acid | Copper hydroxide | Copper oxychloride
Viricides	benzoic acid | Mild Pepino mosaic virus isolate VC1 | Pepino mosaic virus strain CH2 isolate 1906
Molluscicides	Ferric phosphate | Ferric pyrophosphate | Metaldehyde
Nematicides	Bacillus firmus, strain I-1582 | Dazomet | Oxamyl
Plant growth regulators active	Chlormequat | Daminozide
Rodenticides	Aluminium phosphide | Zinc phosphide
Desiccants	Pyraflufen-ethyl
Tolerance inductor	COS-OGA
Talpicides	Aluminium phosphide
Hybridization agent	Sintofen

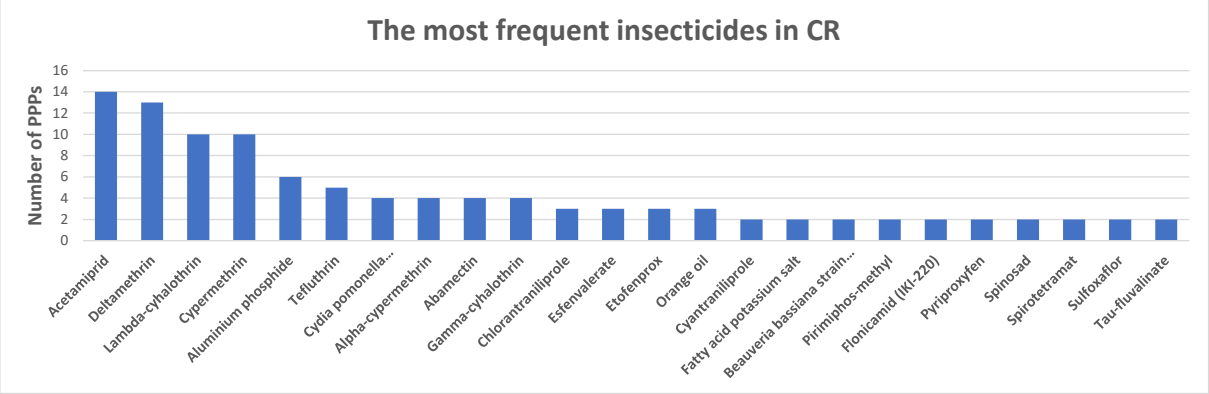
The **number of AS** in the same PPP can be up to 5 (1 PPP named Isomate C LR). Two have 4 AS (Vibrance TZ, Dicotex). There are 87 PPP with 3 AS, 422 PPP with 2 AS. However, 1035 out of 1447 PPP (72 %) have only one AS.

For valid and UOS decisions, same PPP name has always same composition given by AS contained and their concentration. In old, non-valid, decisions, rarely happened Regardless the concentration of the AS in different PPP, the most typical (by number of PPP) **composition of PPP** that have valid and UOS decisions are shown in **Fig. 5**.

Fig. 5: *The most typical (by number of PPP) composition of PPP in CR. Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out. Only compositions used in at least 4, 5 or 3 PPP were considered for fungicides, herbicides or insecticides, respectively.*







From all valid and UOS decisions (with filtering as described above) – 2601 records (because 1961 decisions are combined by multiple AS), the **units of AS in PPP** are in 1987 decisions "g/L" and in 530 decisions "g/kg". This is together 97 % of decisions. However, also another 14 types of units are used for AS in PPP, dominantly for microbial active substances, pheromones, plant growth regulators etc. Some of them hardly useful for the field application rate calculation (e.g. pieces/ml, mg per piece ...).

03. Use and application data for PPP

In Czech PPP register, there are often multiple uses for same decision. The **use is defined** by target crop, pest, application rate, time, number of applications, pre-harvest interval and other application conditions. For all 1961 valid and UOS decisions (with filtering as described above), there are 10,113 use records. These can be very similar or even same by means of crop, pest and application details. One decision may have up to 43 uses (given by unique combination of crop x pest x application details).

Interesting subgroup of use is **seed treatment**, that is 96 out of 1961 valid and UOS decisions (4.9 %).

Within the same use record in the database, there are only rarely multiple alternative application scenarios. Usually (in 97%), one use record corresponds to **one defined application**. For all 1961 valid and UOS decisions (with filtering as described above), there are 10,737 records when use data and application data are incorporated.

For all 1961 valid and UOS decisions (with filtering as described above), the **application units** are dominantly (96 %) l/ha, kg/ha, g/ha or ml/ha. However, there are also another 35 types of units, some hardly useful for the field application rate calculation (e.g. number of baits, g or ml per tonne, number of tablets or pieces ...).

04. Crops

The **crops** for PPP use are specified in the Czech register database by multiple ways. They are defined either as species names or general categories. Often, the target application is not defined as crop name but as various terms like "arable soil", "non-agricultural soil", "tree nursery". The application of PPP is always relevant for the given use which is authorized only for the given crop(s). Also many other details linked to the crop specification in the database. The most important is time of PPP application – usually defined by **plant growth stage (BBCH)**.

In the database, there are crops taken from the Czech EAGRI codelist and crops taken from the codelist of the national authority for PPP registration (Central Institute for Supervising and Testing in Agriculture; Ústřední kontrolní a zkušební ústav zemědělský - UKZUZ). There are 666 EAGRI crops distributed to 29 crop groups. However, they are less precisely connected to use/application records via crop name similarity. There are **514 distinct UKZUZ crops** (403 for the valid and UOS decisions) and they are not affiliated to crop groups. However, they are precisely connected to the use/application records. Therefore, these are preferred for the next data analyses.

There are often **multiple crops** (up to 24) for the same application scenario as defined by use/application record. However, 77 % of use/application records have only one crop. One decision can be related to multiple crops (up to 58) as is shown in **Fig. 6**. Same PPP can be authorized for multiple crops (up to 62) and also same AS is authorized for multiple crops (up to 105), as is shown in **Fig. 7**.

One specific crop can be, of course, involved in high number (hundreds) of decisions and affiliated use/application records. Of course, the given crop, is involved in many different PPP and can be treated by multiple AS from various purposes. The number of distinct PPP and AS that are related to specific crops are shown in **Fig. 8** (this is done only for valid and UOS decisions with filtering as described above).

Fig. 6: *The distribution of the number of crops that are involved in same decision. Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out.*

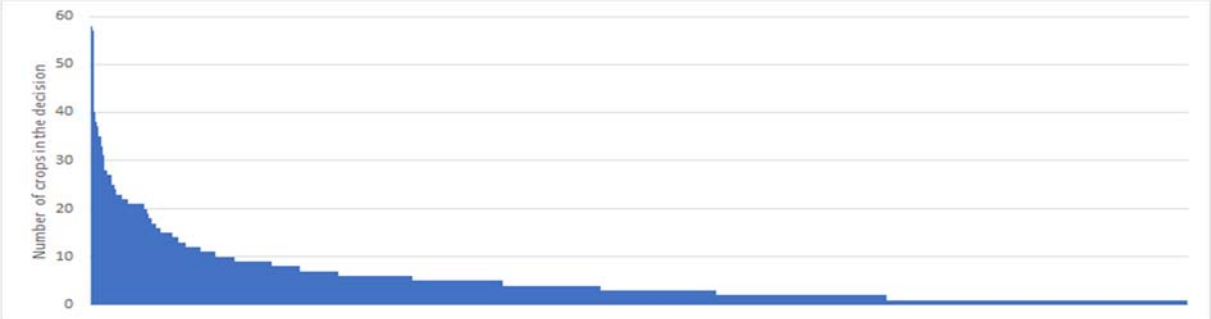


Fig. 7: *The number of crops that are involved in one authorized PPP (A) and for the same AS (B). Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out. Only PPP and AS with 20 or more crops are shown.*

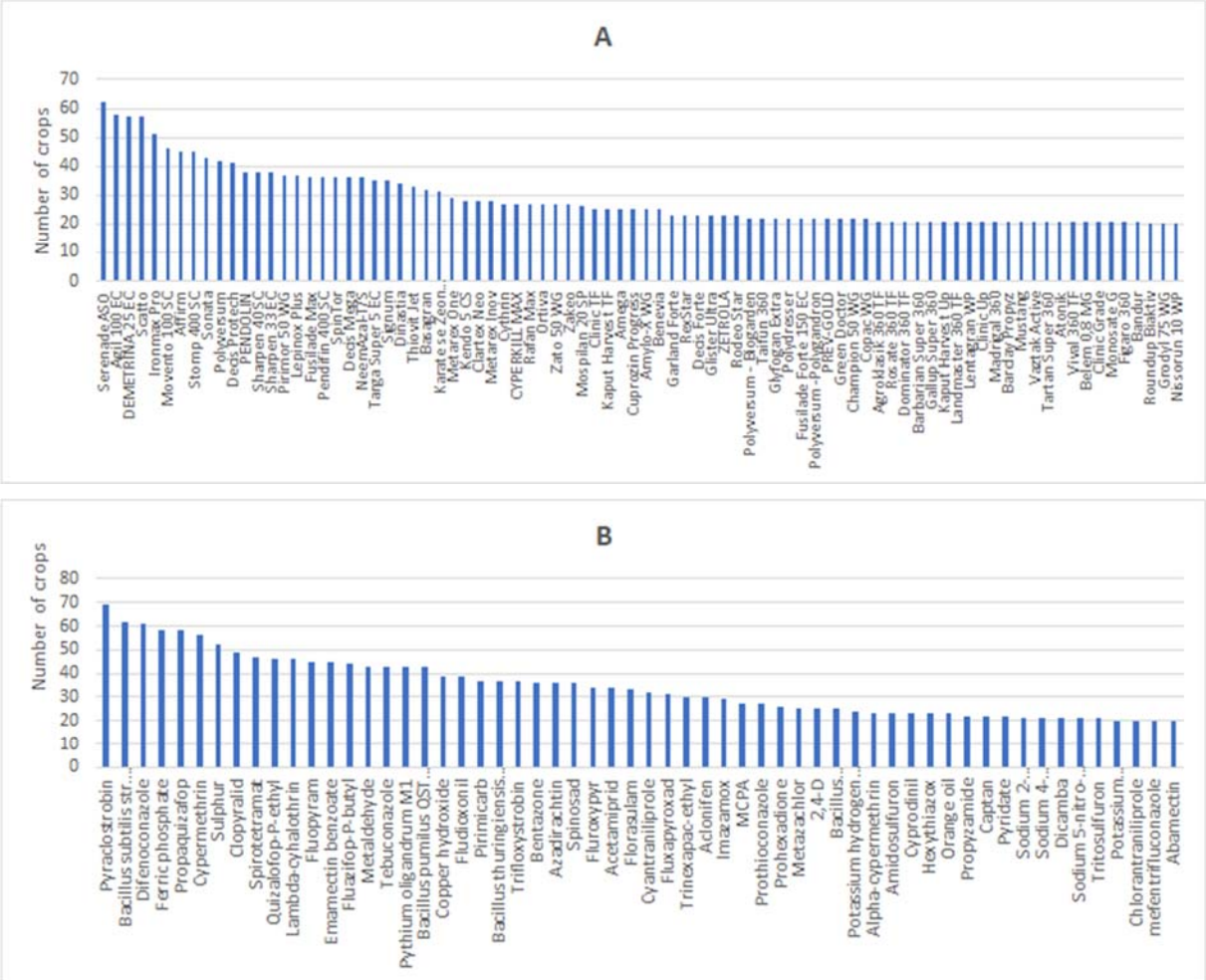
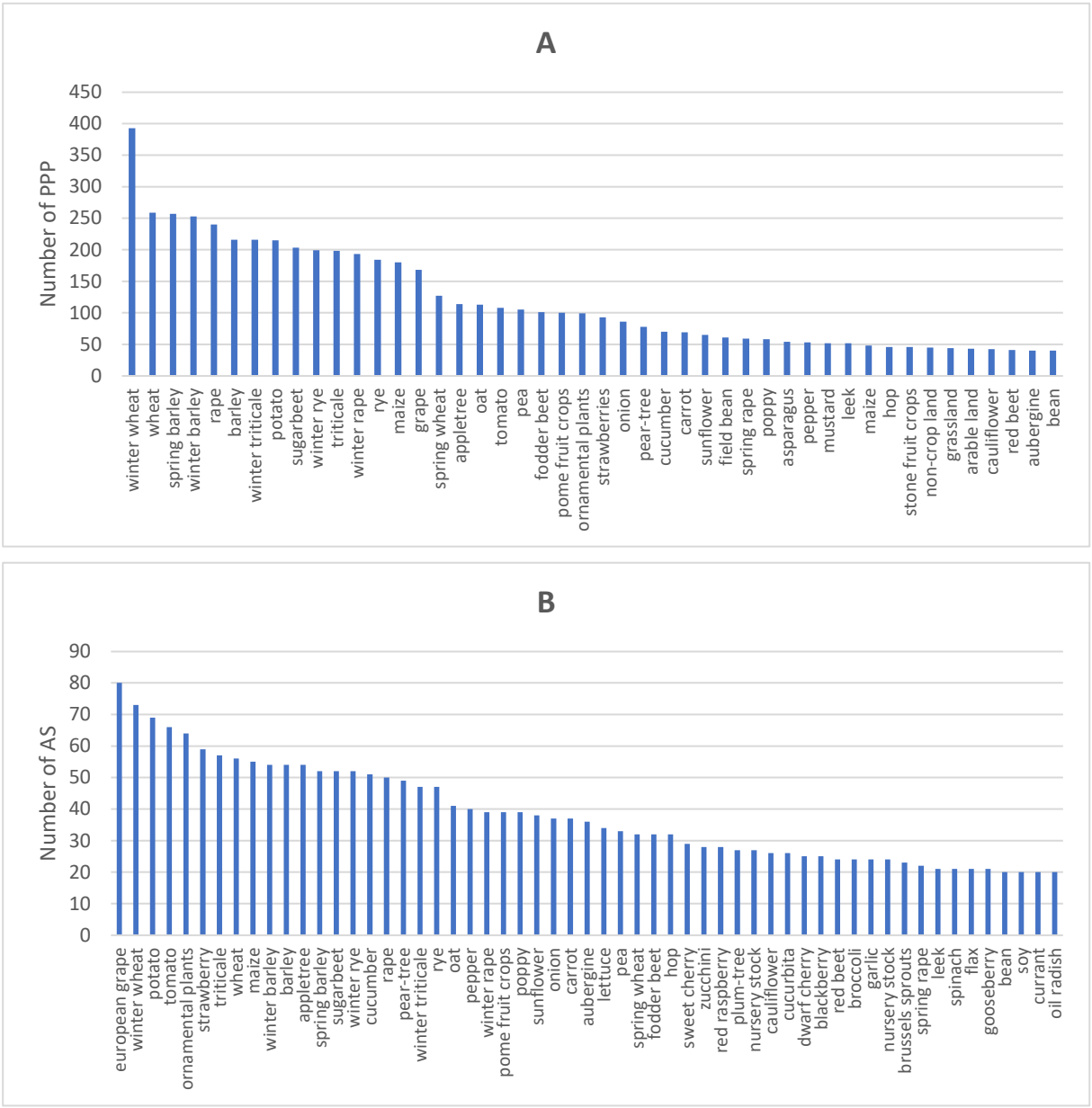


Fig. 8: *The number of distinct PPP (A) and AS (B) that are related to specific crops. Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out. "Crop" is not necessarily the crop name, but the term specifying the targeted use of PPP. Only crops involved in more than 40 PPP and AS with 20 or more crops are shown.*



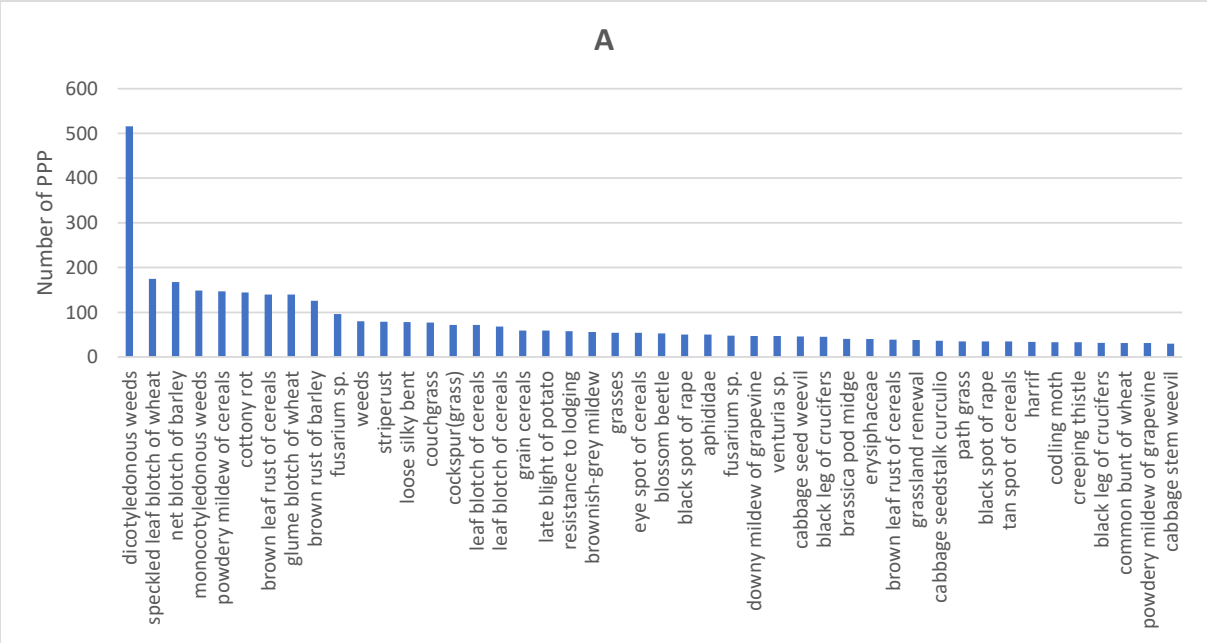
05. Pests

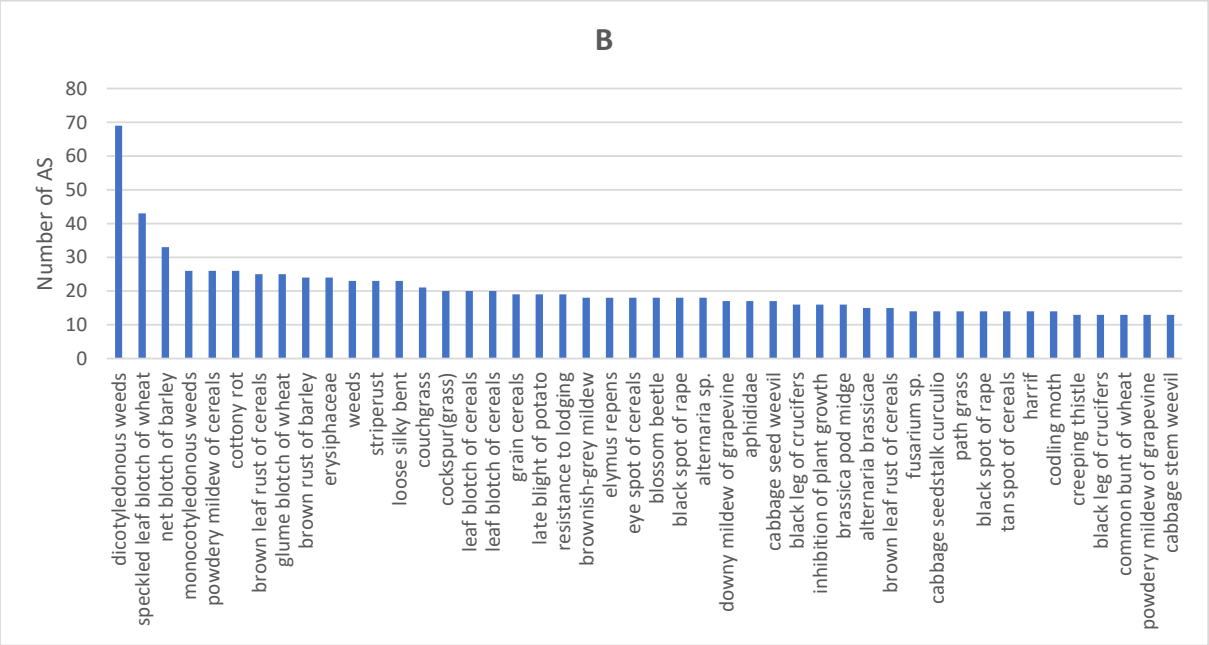
The **pests** for PPP use are defined either as species names or as various terms like "undesired trees", "keeping fruits quality". The application of PPP is always relevant for the given use which is authorized only for the given pest(s). There are 922 distinct pests (687 for the valid and UOS decisions filtered as described above) in the register.

There are often **multiple pests** (up to 14) **for the same application scenario** as defined by use/application record. However, 65 % of use/application records have only one pest.

One specific pest can be, of course, targeted by high number (hundreds) of decisions and affiliated use/application records. Of course, the given pest, is targeted by many different PPP and by multiple AS. The number of distinct PPP and AS that are related to specific pest are shown in **Fig. 9** (this is done only for valid and UOS decisions with filtering as described above).

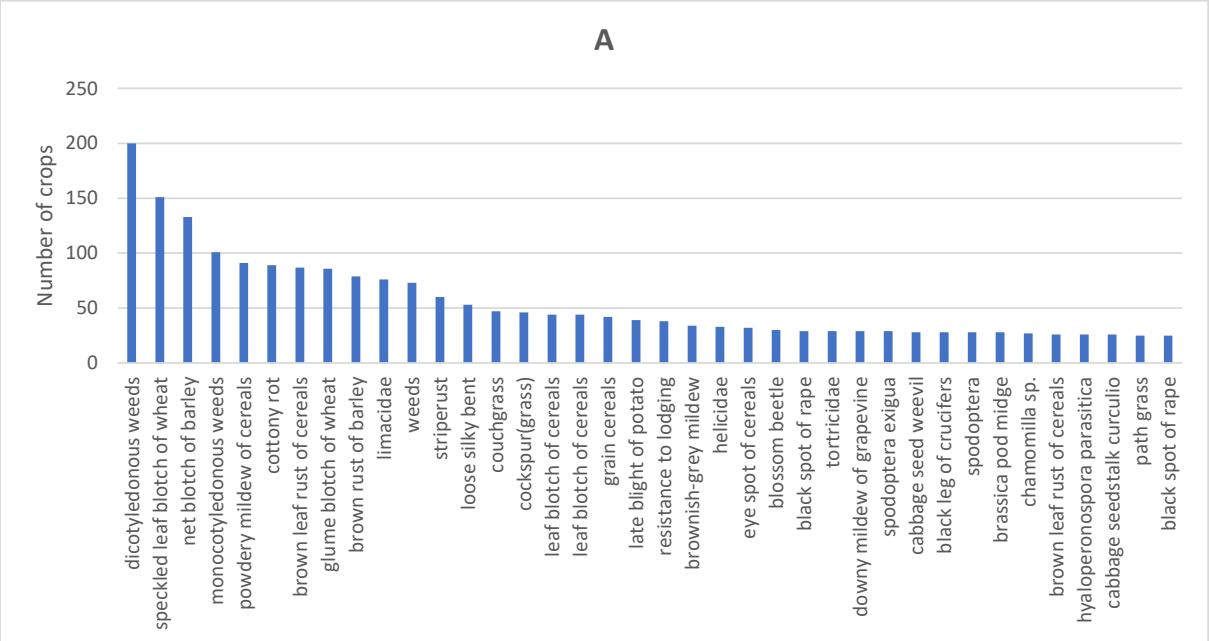
Fig. 9: *The number of distinct PPP (A) and AS (B) that are targeting to specific pest. Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out. "Pest" is not necessarily the pest name, but the term specifying the targeted use of PPP. Only pests involved in more than 40 PPP and AS with more 10 pests are shown.*

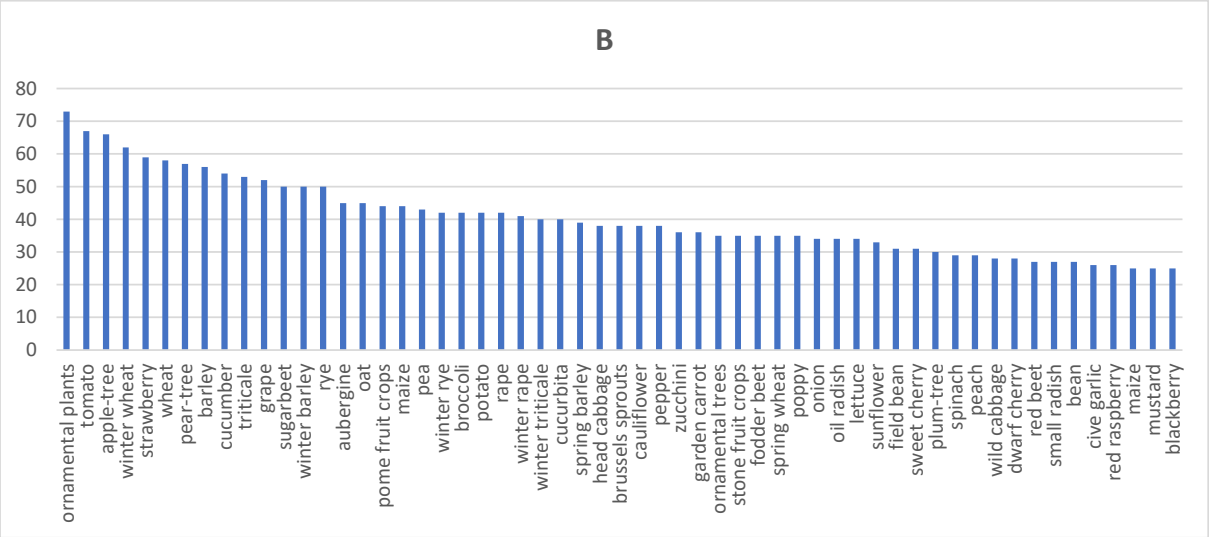




Interesting output is the relation between crops and pests. Number of crops that are related to the individual pests and number of pests that are related to individual crops in the Czech PPP register are shown in **Fig. 8** (this is done only for valid and UOS decisions with filtering as described above).

Fig. 8: *The number of crops that are linked to each pest (A) and the number of pests that are linked to each crop (B) in the Czech PPP register. Only valid or use-of-stocks decisions (2023-02-21) considered with decisions for parallel trade, bioagens, basic substances, auxiliary products and non-professional users filtered out. "Crop" and "Pest" are not necessarily the crop or pest names, but the term specifying the targeted use of PPP. Only the pests related to more than 25 crops and the crops related to more than 50 pests are shown.*




B
80
70
60
50
40
30
20
10
0
ornamental plants
tomato
apple-tree
winter wheat
strawberry
wheat
pear-tree
barley
cucumber
triticale
grape
sugarbeet
winter barley
rye
aubergine
oat
pome fruit crops
maize
pea
winter rye
broccoli
potato
rape
winter rape
winter triticale
cucurbita
spring barley
head cabbage
brussels sprouts
cauliflower
pepper
zucchini
garden carrot
ornamental trees
stone fruit crops
fodder beet
spring wheat
poppy
onion
oil radish
lettuce
sunflower
field bean
sweet cherry
plum-tree
spinach
peach
wild cabbage
dwarf cherry
red beet
small radish
bean
cive garlic
red raspberry
maize
mustard
blackberry