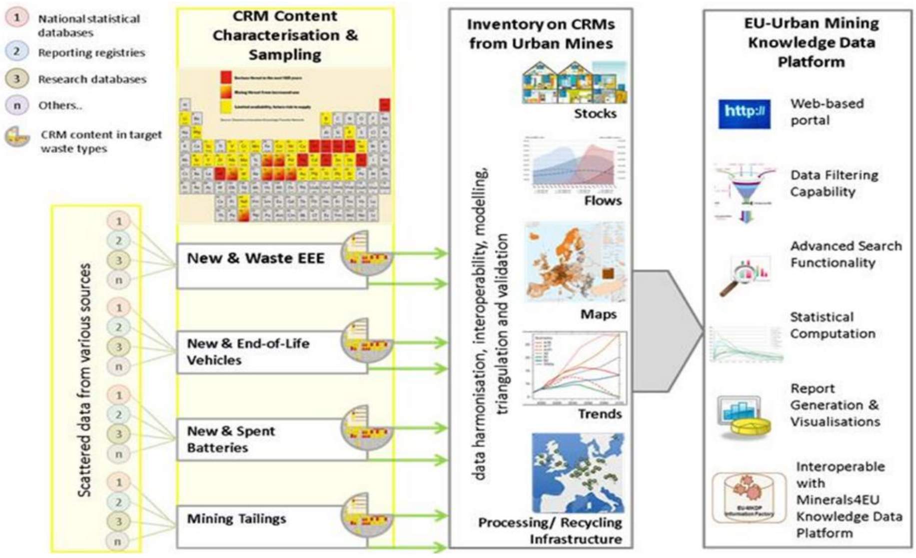
Protocol for Stocks & Flows Update and Quality Assessment

Deliverable D3.6



Project	Prospecting Secondary raw materials in the Urban mine and Mining waste
Acronym	ProSUM
Grant Agreement	641999
Funding Scheme	Horizon 2020
Webpage	http://www.prosumproject.eu/
Work Package	Work Package 3
Work Package Leader	Jaco Huisman
Deliverable Title	Protocol for Stocks & Flows Update and Quality Assessment
Deliverable Number	D3.6
Deliverable Leader	Jaco Huisman
Version:	3.0
Status	Final
Author	Jaco Huisman, Hina Habib, Michelle Wagner, Maria Ljunggren Söderman, Duncan Kushnir, Perrine Chancerel, Johanna Emmerich.
Reviewed by	Jaco Huisman
E-mail	Huisman@vie.unu.edu
Scientific Coordinators	Dr Jaco Huisman, Scientific Advisor, UNU – ViE - SCYCLE
E-mail	Huisman@vie.unu.edu

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 641999.

Document Control

Project Coordinator:	WEEE Forum
Work Package Leader:	UNU
Deliverable leader	UNU
Due date:	M34 13 October 2017
Date of submission:	M36 20 December 2017
Dissemination level:	PU (Public)

Version history			
Ver. no.	**Date**	**Reason for release**	**Responsible**
0	26.06.2017	Outline	MW
1.0	11.07.2017	Input from telco D3.4, new outline	UNU
1.1	20.10.2017	BATT related text	PC
1.2	21.11.2017	WEEE flow information added	MW, HH
1.3	24.11.2017	ELV related text has been added	DK
1.4	05.12.2017	Revised BATT text	PC
2	13.12.2017	PMT review	JH, SD
2.1	14.12.2017	Implementation of PMT review comments	all
2.2	15.12.2017	Finalisation	HH, JH
3, Final	20.12.2017	Uploaded version	PL, JH

Notice

The contents of this document are the copyright of the ProSUM consortium and shall not be copied in whole, in part, or otherwise reproduced (whether by photographic, reprographic or any other method), and the contents thereof shall not be divulged to any other person or organisation without prior written permission. Such consent is hereby automatically given to all members who have entered into the ProSUM Consortium Agreement, dated 1st December 2014, and to the Research Executive Agency / European Commission to use and disseminate this information.

The information and content of this report is the sole responsibility of the ProSUM Consortium members and does not necessarily represent the views expressed by the European Commission or its services. Whilst the information contained in the documents and webpages of the project is believed to be accurate, the author(s) or any other participant in the ProSUM consortium make no warranty of any kind with regard to this material.

Contents

List of Figures

List of Tables

PURPOSE

This report documents the activities undertaken in Task 3.4 within the ProSUM project to produce update protocols and quality assessment procedures for stocks and flows characterization.

The primary purpose of the report is to document the processes, templates and, where applicable, the scripts used to work with stocks and flows models in ProSUM, and thereby ensure the streamlined replication of the calculations for coming years as well as for when additional data sources become available. A secondary purpose is to provide recommendations to the scientific community and others reporting under the WEEE, ELV and Battery Directive on how to represent data in a way that facilitates sharing and inclusion in the EU Urban Mine Knowledge Data Platform (EU-UMKDP).

This report lays down quality assessment procedures for stocks and flows characterization, and complements the deliverables D2.7 on “CRM Product and Component Content and Quality Assessment” and D4.4 on “CRM Content in Waste Flows and Data Quality Assessment”.

EXECUTIVE SUMMARY

BATT

The main conclusions here are directly linked to the conclusions of D3.5 on data availability, harmonisation and evaluation. The main challenges, so far, are not related to technical issues to transfer the data to the portal, but to the data availability and harmonisation, which condition the feasibility and limit the automatization of updates. As long as the available data are not harmonised, the automatization and simplification of the update protocols are not possible and updates rely on manual work.

Specific recommendations for improving BATT stocks and flows information in the future are:

1. Improve data on the unknown whereabouts of (WEEE and) BATT through waste sampling.
2. Undertake household waste surveys to measure the stocks and residence times of (EEE and) batteries in households.
3. Collect data on the amount and types of batteries separated from WEEE and ELV at treatment facilities.

VEHICLES

Vehicle data sourcing still requires manual input at every stage of the process leading up to ingestion and running of the model. This could potentially be improved if the Eurostat databases provided an API (application programming interface) for remote checking and access of data values.

A number of countries are missing data for a large number of possible data points. At present, the model smoothly deduces and interpolates these values, but future improvements of the database are welcome and can be integrated automatically.

The most critical missing data is the large gap between, on the one hand, reported ELV treatment and exports of vehicles for use, and, on the other hand, the number of deregistrations. There are a number of possible explanations (temporary deregistrations, unauthorized treatment, unreported export, etc.) that should be directly addressed by harmonizing and better reporting of the national methods for recording deregistrations.

Specific recommendations for improving vehicle stocks and flows information in the future are:

1. Amend vehicle fleet statistics as reported by Eurostat to expand on vehicle characteristics that are crucial for CRM content, such as further specification of drivetrain types (especially electrified ones) and segment.
2. Develop reporting of vehicle fleets to minimise the number of vehicles of unknown whereabouts.
3. Amend the ELV Directive to include a requirement for reporting vehicle age and the destination of vehicles exported for recycling within the EU and Eurostat to publish .

EEE

The EEE datasets include four main components, with each a respective protocol for updates and deviating time and automatization levels. The POM, Stock and WEEE datasets are very much standardised. Here, the process of checking annually for new good codes and reviewing code changes will be the most consuming part. The complementary Waste flows databases are generally speaking not too elaborate, but need rather ad-hoc consolidation steps when being updated with very limited room for automatization. The consolidated compositions files from WP2 are extremely time-consuming to be updated as also described in Deliverable 2.7. The harvesting for combination with the EEE stocks and flows information is rather automatized. Creating the detailed results files from the UMD has been improved in speed considerably. Here, error-checking is always required consuming considerable updating time for new datasets.

Specific recommendations for improving EEE stocks and flows information in the future are:

1. Improve data on unknown whereabouts of WEEE by making adoption of EN 50625 standards legally binding.
2. Further sampling studies should be executed, particularly for large domestic appliances in light iron and mixed metal scrap.
3. Improve timeliness, frequency and granularity of WEEE (flows) reporting to Eurostat.

A more general recommendation for all three waste sectors combined is the following: It is advised to produce publicly available reports related to the quality of reported data under the WEEE, Batteries and ELV Directives. In general, it is not currently possible to assess the quality of reported data provided under the EEE, ELV or BATT Directives.

1 Introduction

1.1 Aim and scope of the Deliverable

According to the Description of Actions, this deliverable 3.6 addresses the update and quality assessment protocols, in order to determine the required sophistication level needed for BATT, vehicle and WEEE data, provided new datasets become available in the future. The data is collected, recorded, harmonised, and analysed in order to create consolidated data that is representative per country, year, key and/or subkey. This resulted in the two following steps to execute when new information becomes available:

1. Identification of data gaps and sanity checks of available data, and
2. Quality assured datasets for stocks and flows as representative as possible.

This Deliverable 3.6 Protocol for Stocks & Flows Update and Quality Assessment runs parallel to Deliverable 2.7 (CRM Product and Component Content and Quality Assessment) and Deliverable 4.4 (CRM Content in Waste Flows and Data Quality Assessment). The primary goal is explain how to include the new data in the EU-UKMDP, assess data quality and consolidate representative datasets. The stocks and flows models along with recording templates provide the formats and codes that allow direct use in the consolidation procedures and consequently the future update of the EU-UMKDP. Furthermore, the present report will provide product related specific and general recommendations for the future work with stocks and flows data to improve expand and expand the EU-UMKDP.

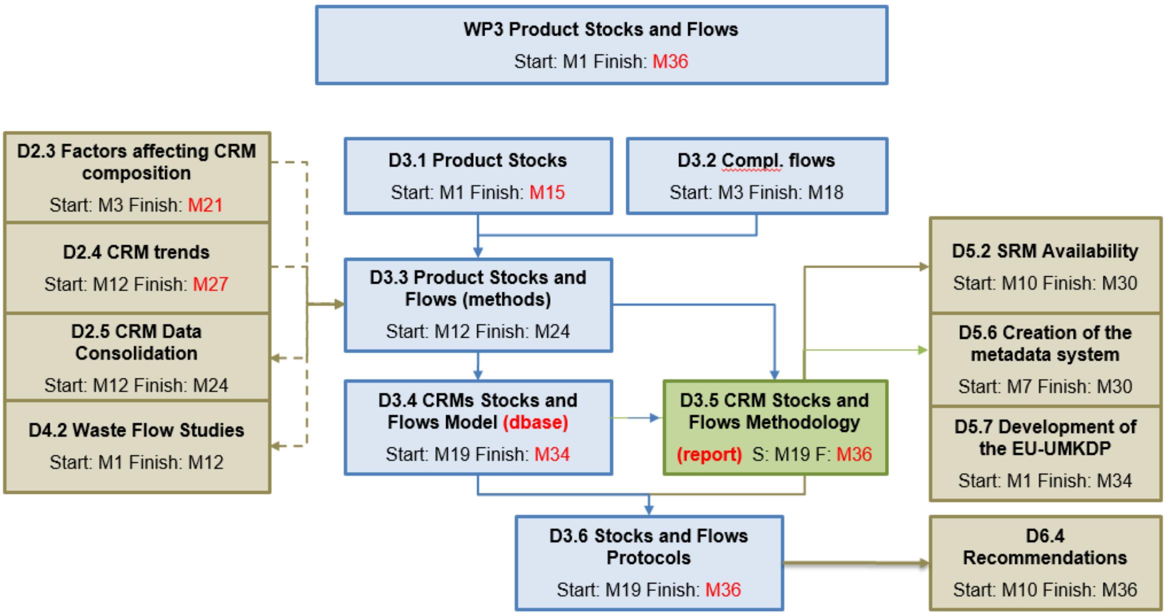


Figure 1. Pert chart positioning the relationship between D3.6 and other ProSUM deliverables

Figure 1 provides an overview of the various deliverables in WP3 and their connections to other Work Packages. The following linkages are of importance:

1. Deliverable 3.1 on product stocks and Deliverable 3.2 on complementary flows provides the basis for the existing data inventory, specific expansions from these are typically new annual information on stocks and flows. For market input revisions of the CN and Prodcom codes likely appear annually. Collection information traditionally is consolidated from Eurostat and the WEEE Forum/ Eucobat. Complementary waste flows data usually comes from specific studies and reports and remains of a rather unstructured nature.
2. Deliverable D3.3 evaluates different models to quantify historic and current stocks and flows of in-use and end-of-life batteries (BATT), vehicles, and electric and electronic equipment (EEE) in EU28+2. Here, typically future maintenance may focus on updating

and expanding the currently used model parameters for instance for lifespans and product average weights.

3. Deliverable 3.4 on CRMs Stocks and Flows Model, and D3.5 on CRM Stocks and Flows methodology are closely linked to D2.5 “CRM data consolidation” and D4.2 “Waste flow studies”. Here, the product data meets the compositional and waste data. As a consequence future updates may typically cover proper alignment of the basket of products describing the products in both domains and careful delineation of for instance the BATT content in WEEE respectively ELV.
4. Deliverable D2.7 and D4.4 are complementary to this deliverable D3.6 and address product composition and waste flows, respectively will feed D6.4 “Recommendations”. Here, typically for future updates, new information on scavenging levels and the count of products and components in various return channels needs to be linked as well as newer composition. On the latter, regularly more elaborate maintenance might be needed when new product types, components and materials are introduced into the market. This requires review of currently used code list and obviously, adaptation of the ProSUM Unified Data Model (UMD) from Deliverable 5.7

1.2 Overview of update protocols and quality assessment

A detailed overview of the procedure to include new data was developed for compositions and waste flow data in D2.7 and D4.4 respectively. An adaptation of the procedure for the stocks and flows is shown in Figure 2. Detailed overview of procedures to include new data in the UMKDP. New datasets will be first subject to data quality assessment and if meeting the criterion, will be included in existing datasets for stocks and flows. Following this, re-consolidation of data can be performed to update the representative stocks and flows datasets in the UMKDP. Depending on the type and source of new data sets, i.e. country studies, surveys, literature, ProdCom and Eurostat websites, etc., following steps will be taken:

In step 1, an initial quick assessment should be made to decide whether the data is to be further processed at all. Following this, in step 2, relevant data and metadata is recorded in the so-called “stocks and flows data templates” (Excel sheets used to store raw data for each of the three product categories, and detail the original source documents which are then stored together with their metadata in the UMKDP). In step 3, a data quality level is assigned to each datum based on the metadata recorded in the aforesaid templates.

Following this, in step 4, the data must be transferred to the consolidation files. Where applicable, step 4 comprises the description of how to: a) transfer new data to files for consolidation; b) (re-)consolidate the data; and c) assess the data quality/uncertainty of the updated stocks and flow data sets. Where data availability does not enable the consolidation of data, these sub-steps are merged and described together.

Following this, the updated data sets must be transferred to portrayals (step 5) from which they can be harvested to the UMKDP (step 6). If the parameters in question had been estimated before, and the update is merely an improved estimate, the old data in the UMKDP will be overwritten.

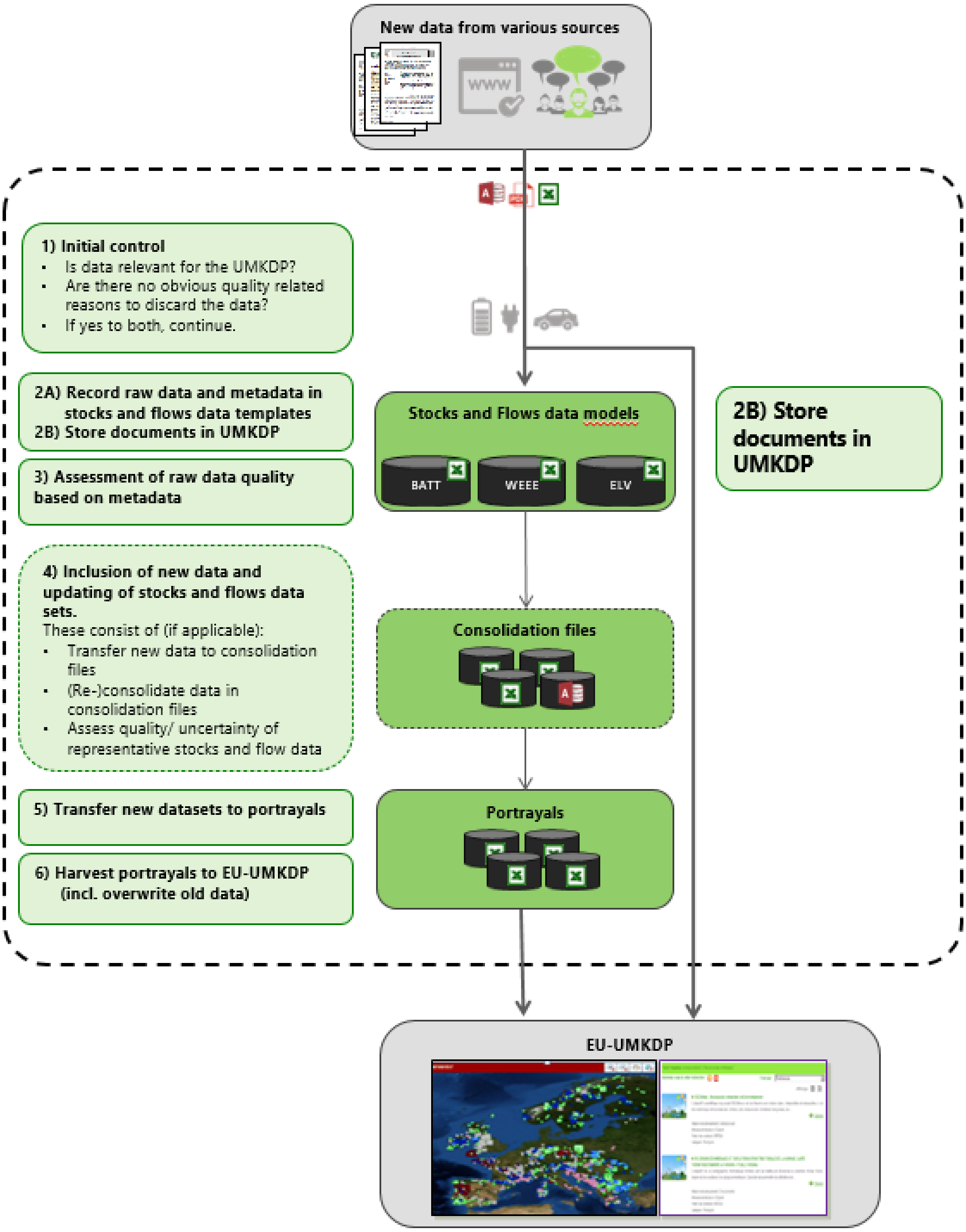


Figure 2. Detailed overview of procedures to include new data in the UMKDP

An initial control of new data should always be performed to avoid unnecessary work and cluttering of the aforesaid template with irrelevant or dubious data. The decision will be somewhat subjective, but should be guided by the following questions:

Are the data addressing products within the scope of the UMKDP?

The data should refer to a product within one of the three categories: batteries, EEE and vehicles and it should be possible to allocate it to one of the product keys in the respective code lists.

Are the data useful for calculating stocks and flows?

Data may be in the scope of the UMKDP but have to be excluded due to a too high level of aggregation or redundant data. These redundancies may occur due to multiple publications of data along the reporting chain, e.g. the same data on collected and reported WEEE is accessible from national registers, WEEE Forum, and Eurostat.

Are there obvious reasons to doubt the validity of the data?

If the data set is dubious or has extreme uncertainties, it should be filter out and should not be recorded.

1.2.1 BATT

New datasets can come in many different formats and sizes. Whereas the majority of sources for composition data (D2.7) have been journal articles or research reports available as pdfs, most data on stocks and flows of batteries were provided by third parties, for instance the industry associations RECHARGE and Eucobat, as Excel tables. Because the data on stocks and flows of batteries in the scientific literature are very scarce, getting data requires networking efforts and the willingness of stakeholders to cooperate. The cooperation does not only consist of getting the data, but also on information exchange on data selection, data quality and uncertainties, as well as different options for data consolidation to agree on one strategy.

New data will very likely not bring much new information on historic stocks and flows, but result from efforts to better detail and understand the current status and forecast the future. It is rather unlikely that the data gathered in the framework of ProSUM can be enriched or need to be overwritten. The inclusion of new data will primarily concern current and coming years. Two cases can be distinguished:

1. Inclusion of new data from the data sources identified and “mined” by ProSUM (e.g. inclusion of updated data for following years, data updates for other countries)
2. Inclusion of new data from new data sources

Based on the eight steps described in D2.7, the inclusion of new data consists of following actions:

1. Initial quick assessment to decide whether the data is to be further processed at all;
2. Relevant data and metadata is recorded in the parameter templates;
3. Assessment of data quality performed based on the metadata recorded in the template;
4. Comparison of the data with the ones from other data sources and manual checking of data;
5. Assessment of the data quality and uncertainty of the consolidated data;
6. Transfer to portrayals;
7. Harvesting to the UMKDP.

Basically, because the calculations of the stocks and flows of batteries are computed in the ProSUM portal, data updates mainly concern volumes of batteries put on the market, average weights of batteries and lifespan distribution. However, also new data on batteries in stocks and in complementary flows can be integrated in the future. So far, the data on batteries in waste bin are too scarce and too diverse to enable estimating by extrapolation the European and national amounts of waste batteries in the waste bin. They were not included into the portal.

When new information regarding batteries in waste bin is available from journals and/or country studies, these should be included into the portrayal files with an indication of the reference and a description of the unit of measure, of the data quality and a comment explaining for instance the type of residual waste sampled and the sample size. If in future it is decided to include European and national estimates of the amounts of waste batteries in the waste bin, the data need to be calculated in tonnes of waste batteries in residual waste per country or in grams per inhabitant and provided to harvesting using the portrayal tables.

New information may also exist and get integrated into the ProSUM portal in the future:

1. Batteries improperly discarded by users in other waste streams than waste bin, including batteries inside WEEE improperly discarded by users;
2. Batteries inside WEEE and ELV collected or included in de-registered used vehicles;
3. Batteries collected, not reported/improperly treated;
4. Batteries collected by manufacturers and distributors for repurposing/ refurbishment and replacement.

1.2.2 Vehicles

There are two modules to the vehicle calculations; the vehicle stock-flow model, and the vehicle element content model. Both parts have different update protocols. The element content model requires expert knowledge for even simple updates, and thus will follow a very similar process to BATT or EEE subgroups. The two halves of the model are independent however, and without new data for element content, the model will simply assume constant element content for all future years when combining the results with the stock flow model.

New data for the stock-flow model for the vehicle sector will come from Eurostat and COMEXT databases. In both cases, the online database tool must be used, with custom sheets exported for each of the model elements. This process must be repeated for each combination of vehicle keys.

For POM flows, the relevant Eurostat databases are from the "road_eqr" series, and for the stock, "road_eqs". COMEXT provides internal and external trade data, already broken down into motor type and motor size data. The data is in both cases structured in a country (row) x year (column) format. In theory, only the latest year need be considered, but historically, revisions have been made to past data and thus it is advantageous to retrieve all available data in each instance. Updates are infrequent (e.g. last update for alternative energy vehicles was 2015) and must be monitored manually.

Once collected, the data needs to be first annotated with the relevant metadata for model ingestion, and subsequently appended to the master input file. Again, manual input is required here as Eurostat does not allow export of all categories simultaneously, and thus the process must be repeated for each motor type, age distribution and weight distribution.

Finally, the model converts the input data into the harmonized data format, calculates the deregistrations, and finally uses the age distributions of the stocks and flows to calculate the element content vectors for each database entry.

1.2.3 WEEE

WEEE datasets consist of four independent databases:

1. The *Multivariate Sales-Stock-Lifespan method* is used to determine stock and flows of EEE. This is used in the Implementing Act of Article 7 of the EU-WEEE Directive. Here the market input component is determined by the *apparent consumption method* and subsequently the WEEE generated volumes are calculated by using the lifespan parameters in a sales-lifespan approach. Measured data on the stocks have been compared with published studies in Belgium, Romania, Italy, the Netherlands, France and Switzerland.
2. The (Complementary) Waste Flows database contains relatively speaking a limited number of records for the consolidated data per EU28+2 member state for WEEE officially collected, WEEE in mixed metal scrap, scavenged components, export (for reuse), complementary recycling and waste bin information.
3. The consolidated compositions files from Work Package 2. Each UNU key is having its own very elaborate XLS file. The EEE Compositions database contains information regarding all the elements and materials found in electronic products and in their

components. The quality assured representative datasets are then harvested as consolidated portrayals from these UNU key files per collection category and then linked to Unified Data model to directly feed EU-UMKDP. For more details see D2.5 and D2.7 for the update procedure of this part.

4. The resulting multiplication is performed in the data model and provided via GEUS to BRGM for populating the UMKDP and to UNU for cross-checking and detailed analysis on basis of the detailed consolidated data. The subsequent "weee_emc" files are organized per collection category and combine the stocks and flows model output to compositions portrayals per UNU Key.

The new data for stocks and flows model has two main source data files, e.g. the ProdCom data and International trade data to calculate the apparent consumption (= domestic production + imports – exports). The output data follows a rigorous scrutiny routine such as manual adjustments, removing of unreliable data, and outliers. Datasets are also calculated for past and future POM. 1995 is the first year with measured market input data and is used in the ProSUM project visualisations. Earlier market input between 1980 and 1995 is back casted to the market introduction year and contributing to the stocks and waste generation. Earlier market input prior to 1980 has not been included in the stock and flow analysis. A detail of procedure can be found at https://github.com/Statistics-Netherlands/ewaste. The output data then can be manually transferred to ProSUM portrayals which further access data quality, uncertainty and assign metadata to individual data set.

Data is manually linked to waste flows database. The main data sources include Eurostat, national registers, journals, country studies, recyclers, dismantling exercises. When new information regarding EEE waste flows is available, these should be included into the portrayal files with an indication of the metadata reference and a description of the unit of measure, of the data quality and a comment explaining the summarized procedure to consolidate the data point. A detail is provided here in section 2.2.3.

The compositions part of the "weee_emc" files requires expert knowledge for even simple updates and thus will follow steps described in D2.7 "CRM Product and Component Content and Quality Assessment".

Parallel to the *Multivariate Sales-Stock-Lifespan method* to calculate stocks and flows, the stocks and waste generated are also computed in the ProSUM portal, given the input files of volumes of EEE sales, average weights over time and dynamic lifespan distribution. This model is called the ProSUM Unified Data Model and it combines above mentioned four databases to provide robust calculations of CRMs Stocks and Flows. Once the data is collected, recoded, analysed, consolidate, re-consolidate, and assigned metadata, the data will be linked to EU-UMKDP. The process will be repeated for every single data point. The calculations are purposely run in parallel as an error check to filter out any issues that can occur when certain fields remain empty or when codelist additions are made.

2 Update protocols

2.1 Recording of new data from Stocks and Flows in ProSUM format, Sampling Protocol requirements and Consolidation of new Stocks and Flows with existing ProSUM data

2.1.1 BATT

For batteries, some issues related to the inclusion of new data according to the procedure described in chapter 1.2.1 require special attention. Table 1 shows, for different stocks and flows, what should be kept in mind during database update and expansion.

Table 1. Specificities of the inclusion of new data on different stocks and flows of batteries

Step of the inclusion of new data	POM volumes	Lifespan distribution	Average weight	Measured stocks	Complementary flows
1. Initial quick assessment to decide whether the data is to be further processed at all	Description of the scope in terms of BATT keys and subkeys, applications, countries and years. New codes may need to be defined in the code list to reflect market changes, e.g. new products containing batteries (e.g. electric buses) or new battery technologies. Also new combinations of BATT keys and applications may appear. A rough expert assessment of the data credibility, reliability and representativeness is needed.				
	So far, the collection of POM data is very complex due to diverse data sourcing. In general, it is better if the POM data show POM flows both in pieces and in weight	The representativeness of the sampling needs to be checked.	The representativeness of the sampling needs to be checked.	The geographical and temporal coverage are limited. The representativeness of the sampling needs to be carefully checked.	The nature and representativeness of the sampling needs to be carefully checked to avoid extrapolations leading to misestimations.
2. Relevant data and metadata is recorded in the parameter templates	The raw data may not be perfectly matching the ProSUM classifications, so that conversions and other operations may be needed before recording of the data in the templates. The recording may, for large datasets, require automation. This was done by macros in Microsoft Excel. If necessary, the code lists need to be expanded to cover new battery technologies (electrochemical systems –subkeys- or families of electrochemical systems –keys-), applications or combinations of applications and keys/subkeys, as well as new flows and stocks.				
		The Weibull distribution scale and shape need to be calculated based on the raw data on the age of waste batteries.			New complementary flows may need to be defined in the codelists (batteries inside collected WEEE or ELV)
3. Assessment of data quality (DQ) performed based on the metadata recorded in the template	Expert knowledge is needed to assess the DQ, which is supported by the comparison of redundant data if available. The DQ assessment is documented in the metadata.				
	Key issues for the DQ assessment are: representativ	Key issues for the DQ assessment are: consideration	Key issue for the DQ assessment is: sample representat	Key issues for the DQ assessment are: representativeness of the	Key issues for the DQ assessment are: representativeness of the

	eness of the data, definition of the scope in terms of application, electrochemical systems, year and geographical area	of batteries that are not brought to collection but stay in the stocks, sample representativeness in terms of application, electrochemical systems, year and geographical area	iveness in terms of application, electrochemical systems, year and geographical area	sample, used survey method to make sure that all stocked batteries are considered	sample in terms of type of flow, year and geographical area
4. Comparison of the data with the ones from other data sources and manual checking of data	The scope of the data in terms of battery chemistries, applications, year and geographical area has to be well defined to enable the comparison between different data sources. In most cases, no other comparable data are available.				
5. Assessment of the data quality and uncertainty of the consolidated data	Like in step 3, expert knowledge is needed to assess the DQ and the uncertainty. The DQ assessment is documented in the metadata.				
6. Transfer to portrayals	Macros programmed in Visual Basic for Applications enabled completion of the portrayal templates. The macros automatically copy the data from the tables in which the raw data are recorded into the right column of the portrayal template (year, country, BATT key, value). For new flows, new macros need to be written				
7. Harvesting to the UMKDP	See WP5.				

2.1.2 Vehicles

Inclusion of new data in the vehicle stock-flow model is not subject to the same stringent procedure as per the other groups because it is a direct count of the vehicles by motivated national statistics agencies. Therefore, questions of sampling and representativeness are minimal.

As reported in other deliverables, e.g. D3.3 and D3.4, the model itself handles questions of incomplete data internally and features a smooth handoff between the incomplete Eurostat and COMEXT data and the interpolated / projected results.

Table 2. Specificities of the vehicle new data inclusion process

Inclusion Step	Eurostat	COMEXT	Notes
1. Detection of new data	Manual	Manual	Both databases are infrequently updated
2. Download and conversion to ingestion format	A number of tables need to be manually downloaded from the online database tool. Units and formats are inconsistent and require manual attention	A number of tables need to be manually downloaded from the online database tool.	
3. Ingestion and model run	This aspect is handled autonomously by a dedicated python program		
4. Harvesting to the UKMDP	The program outputs directly into harvestable templates.		

2.1.3 WEEE

For WEEE, some issues related to the inclusion of new data according to the procedure described in section 2.1 require special attention. The input data is ProdCom and trade data. New data is available every year. Therefore updates on POM, WEEE and Stock data can be done every year. In Annex, Table 13 provides the breakdown of hours required for future datasets. Table 3 shows, for different stocks and flows, what should be kept in mind during database update and expansion.

Table 3. Specificities of the WEEE new data inclusion process

<table>
<tr><th>Step of the inclusion of new data</th><th>POM volumes</th><th>Lifespan distribution</th><th>Average weight</th><th>Measured stocks</th><th>Complementary flows</th></tr>
<tr><td rowspan="2">1. Initial quick assessment to decide whether the data is to be further processed at all</td><td colspan="5">The inclusion of new data must match UNU Keys, sub keys or device type, countries and years. A rough expert assessment of the data credibility, reliability and representativeness is needed.</td></tr>
<tr><td>The collection of POM data is rather complex, i.e. ProdCom, International trade and Solar Panel data.

When units and formats are inconsistent they require manual attention.</td><td>The representative-ness of the sampling needs to be checked.</td><td>The representative-ness of the sampling needs to be checked.

i.e. Convert ProdCom data and International trade data to weight (kg) with the use of average weights data</td><td>The scope of ProSUM should be matched, i.e. temporal and geographical coverage.

The granularity and representative-ness of sampling should be checked.</td><td>Product return stream data should be carefully checked to determine Relative presence (RELP) of the UNU key in the collection category in comparison with its presence in WEEE generated.</td></tr>
<tr><td>2. Relevant data and metadata is recorded in the parameter templates</td><td colspan="5">The raw data may not be matching the ProSUM format, i.e. conversion of units is needed. Since the data for POM, Stock and WEEE are large datasets and hence need automation. The automation is already been done by CBS as specified at the GitHub repository. However, manual transfer of data is required to transfer them in ProSUM format.</td></tr>
<tr><td>3. Assessment of data quality (DQ) performed based on the metadata recorded in the template</td><td colspan="5">When new information provided by the user has been extrapolated for other collection categories or different years and it's not original, the sheet "Lists" should be updated specifying that for that country its data consolidation are coherent estimates and that the data was reconsolidated, on the other hand if the information is original specify that it's an original/consolidated data point.

Expert knowledge is needed to assess the DQ, which is supported by the comparison of redundant data if available. The DQ assessment is documented in the metadata. Key issues for the DQ assessment are: definition of the scope in terms of UNU Keys or collection categories, representativeness of the data, representativeness of sample, used survey methods, sample in terms of waste flow type, etc.</td></tr>
<tr><td>4. Comparison of the data with the ones from other data sources and manual checking of data</td><td colspan="5">This takes substantial time. It involves analysing data for outliers and finding a way to correct them. Each correction also has impact on other data points, so there are often many corrections/analysing cycles needed. Often errors in the underlying data are only revealed after the flows versus composition multiplication. This will be speed up considerably in the future due to the large amount of error checking already performed. Extrapolation methods also take quite some effort. When extrapolations have been connected to the last available data point, the same will be done in the next year to the last available data point at that time.</td></tr>
<tr><td>5. Assessment of the data quality and uncertainty of the consolidated data</td><td colspan="5">This needs a check if all is still looking good. When a specific connection year has been chosen in the past for instance due to bad data in later years, the same connection year keeps being used. So when new data becomes available this will be overwritten with extrapolation estimates. To prevent this, values have to be checked in detail to see if the connection year can be changed.</td></tr>
<tr><td>6. Transfer to portrayals</td><td colspan="5">Once all above steps are completed, only then data can be manually transferred in ProSUM format.</td></tr>
<tr><td>7. Harvesting to the UMKDP</td><td colspan="5">See WP5.</td></tr>
</table>

When data is being updated, the following should be taken into account:

1. The Multivariate Sales-Stock-Lifespan models contains a source file listing that should be updated when new information is available for the respective years, more detail is available at https://github.com/Statistics-Netherlands/ewaste/blob/master/README.md.
2. Similarly, the databases for the WEEE complementary flows (collected and reported WEEE, waste bin, export, metal scrap and the complementary recycling products and Scavenged products) contain a content sheet which guides the researcher to the sheets that should be updated when new information is encountered for the respective years.
3. In case Eurostat has updated the population of European countries, the sheet "POP" should also be updated for the corresponding country and year in the respective database. Currently estimation until 2021 is provided for Waste bin. For Export, Metal scrap and the complementary recycled and scavenged products the population is provided until the year 2016 as no future extrapolation are regarded reliable in these cases.
4. When new information provided by the user has been extrapolated for other collection categories or different years and it's not original, the sheet "Lists" should be updated specifying that for that country its data consolidation are coherent estimates and that the data was reconsolidated, on the other hand if the information is original specify that it's an original/consolidated data point.
5. Once new information has been recorded in the different databases, all calculations from the sheets "READ INTO DATABASE" and "READ INTO MASTER" will be automatically updated and ready to be harvested.
6. Any new codes should travel along with the harvested data to update the ProSUM UMD accordingly.

A short description on how to update each database will be explained subsequently.

Collected and reported WEEE

WEEE collected data is reported regularly to national registers by EPR (Extended Producer Responsibility) organizations and EEE producers. From August 2018 or January 2019 onwards, the data will later be provided by Member States to Eurostat in the 6 collection categories of recast Directive (2012/19/EU). Besides Eurostat, the Key Figures platform managed by the WEEE Forum also reports the WEEE collected data in collected B2C (consumer electronics), collected B2B (business) and collected undifferentiated in the 6 collection categories of the recast Directive (2012/19/EU). Once collected, the data needs to be first annotated with the relevant metadata "tmp_metadatID" in order to have a record were the information came.

Updating Waste bin

When new information regarding waste bin is available from journals and/or country studies this should be included in the raw data sheet in a structure manner and its reference should be added in the "tmp_metadatID" in order to have a record were the information came. Approximately every two years Eurostat provides new waste collected updated version and therefore the table contained in the sheet "WEEE Dir" should be updated as well as the national Purchasing Power Parity values in sheet "PPP".

In the "overview_totalsEu" sheet are highlighted in colour orange which columns are necessary to be updated in order for the calculations to be done. Most of the columns are updated automatically by adding the new information in the corresponding sheet (for example by adding new population data in the "POP" sheet will automatically update the information in column D). The following columns should be updated manually: WEEE for the year with most confident information (Column G) and the formula in column H should be updated depending on the collected year were the most confident information was provided.

Updating Export

When new information regarding export is available from journals, statistical information, registers, country studies among others, this should be included in the raw data sheet in a structure manner and its reference should be added in the “tmp_metadatID” in order to have a record were the information came.

Metal Scrap

When new information regarding Metal scrap is available from journals, Eurostat, country studies among others, this should be included in the “WstatR” sheet in a structure manner for the corresponding year and its reference should be added in the “tmp_metadatID” in order to have a record were the information came. In the “WstatR” sheet, a new table in columns C-H should be made for the subsequent years containing information from Metal wastes ferrous and nonferrous, metal wastes that contain mixed ferrous and non-ferrous and if possible information that provides WEEE in metal scrap. Normally this information can be obtained approximately every two year from the EUROSTAT website, but in the case the information’s origin is not EUROSTAT it can be added for the corresponding country and year in the tables displayed in this work sheet. Once this information is recorded in the corresponding columns (C-H), the tables in columns J-R with their formulas should be copied as well in order to make the calculation process automatically. It is important that the user be capable of analysing the information in order to detect possible overestimations done by Eurostat or by country studies.

Complementary Recycled and Scavenged Products

The Project dbase “Compl Recycl Products and Scavenged” workbook contains two databases in one, the “Complementary Recycled” dbase and “Scavenged Product” dbase. In both databases the data is provided as a total value of WEEE and in the case of complementary Recycled products is provided in 6 categories, in the case of scavenged product it is only provided as a total value and for Category I (Cooling and Freezing appliances).

Scavenged Products

In the case of the scavenging of products database, it is important to update WEEE data in the WEEE_ProSUM for the desire year to expand. This WEEE data can be extracted from the Project dbase POM Stock WEEE file.

In the sheet “Read Scavenging into MASTER”, the years the user wishes to expand should be included and simply drag the formula from column U until the last country and its corresponding year (for example for the year 2016 expand until Sweden 2016).

Complementary Recycled

When new information regarding complementary recycled data is available it should be included in the “RawDataComplProducts” sheet in a structure manner and its reference should be added in the “tmp_metadatID” in order to have a record were the information came.
In the case Eurostat has updated the population of European countries, the sheet “POP” should be updated and include information for the corresponding country and year. Currently estimation from 2016 onwards is provided.

In the case were the information had to be extrapolated for other collection categories and it’s not original it should be as well updated specifying for that country its data consolidation is an original/consolidated data and in case the information is original otherwise specify that it’s a coherent estimates and that the data was reconsolidated.

When a new update from D2.5 “Consolidate data into CRM database” containing information of e-p, e-m, m, m-p, c, c-p is made available, the updated UNU Key portrayal will be provided to BRGM and GEUS to update the “Colcat_emc” datasets in ProSUM database. A detailed procedure is given in D3.5 “CRMs Stocks and Flows” section 1.2.6.

2.2 Data quality assessment of consolidated versus new information from Stocks and Flows

2.2.1 BATT

Expert knowledge is needed to assess the DQ and the uncertainty of each new data point or dataset. The DQ assessment will be documented and added to the metadata in the same way as currently done. As explained in D3.5, the available data do not present a comprehensive view of the batteries POM with the level of detail required for the modelling, so that different sources of data were manually combined. Each data source brings uncertainties, which were estimated by expert assessment. Lower uncertainties were allocated to measured data than to estimates, depending on the measurement methods, the sample representativeness and other factors specific to the considered flow or stock, country, type of battery and year. The combination of several data sources leads to uncertainties higher than the uncertainty of the used data, due to the error propagation.

2.2.2 Vehicles

There is no formalized procedure for assessing new data, as it is a direct count of vehicles, and subject to already extensive checks. Uncertainties are calculated directly inside the model from formal error propagation methods, and data confidence is assigned based on the relative uncertainty of each data point. The critical input value to these methods is the overall assumed uncertainty for the total stock or flow as reported by Eurostat, and this value may need to be changed based on expert assessment.

2.2.3 WEEE

All parameters used in modelling are subject to a range of uncertainties, i.e. ProdCom and International trade codes, average weight per unit, lifespan distribution, European placed on the market and waste generated data, and product disposal age and stock age. Data quality has been improved by introducing dedicated “Flag values”. A flag value states if the resulting value is taken without alterations from the apparent consumption method or not (described in Deliverable 3.1). If not, the value specifically also states which assumptions specific actions, assumptions or imputations are applied. A detailed description is available in Section 2.3 of Deliverable 3.5. CRM Stocks and Flows Methodology.

2.3 Transfer new consolidated data to portrayals

2.3.1 BATT

For batteries, the transfer of new consolidated data is realised with Excel files having a format aligned with the portrayal templates used by the WEEE and Vehicle teams. Table 4 lists the fields in the portrayal harvesting template for BATT POM data. Similar templates are used for the data on the average weights of batteries and the lifespan parameters.

Table 4. Fields in the portrayal harvesting template for BATT POM data

field name	*input type*	*Instructions*
ID_Country	txt	ISO-2 country code
Year	#	Year of data point
BATT Key	txt	Battery key as defined in 3.2
BATT Subkey	txt	Battery subkey as defined in 3.2
Application	txt	Application (e.g. product) in which the battery is used. Linked to the UNU keys and subkeys as defined in 3.2
Value	#	Numerical value (in this case, # of vehicles)
UOM	txt	Parameter subscript (in this case, “pieces” or “tonnes”)
Reference	txt	Direct reference to a data source or reference to one or more data sources and description of calculations

UNC Min	#	Minimum value of uncertainty interval (2 standard deviations, in %)
UNC Max	#	Maximum value of uncertainty interval (2 standard deviations, in %)
Data Quality	txt	Data quality, automatically determined by model from the formal uncertainty of the item
Data Quality comment	txt	comments to the data quality beyond the indicated level
Metadata ProSUM Creator	Txt	Person who prepared the data
Measurement or Estimate	#	Source reference or shorty description

Table 5 lists the fields needed for the creation of new data on BATT flows in waste bin in the portrayals.

Table 5. Fields in the portrayal harvesting template for data on BATT in waste bin

field name	*input type*	*Instructions*
Year	#	Year of data point
ID_Country	txt	ISO-2 country code
BATT Key	txt	Battery key as defined in 3.2., or "BATT" if the scope of the data does not differentiate the battery keys
Value	#	Numerical value (in this case, # of vehicles)
UOM	txt	Parameter subscript (in this case, "Percent of batteries in waste bin" or "tonnes")
Reference	txt	Direct reference to a data source
UNC Min	#	Minimum value of uncertainty interval (2 standard deviations, in %)
UNC Max	#	Maximum value of uncertainty interval (2 standard deviations, in %)
Data Quality	txt	Data quality, automatically determined by model from the formal uncertainty of the item
Data Quality comment	txt	comments to the data quality beyond the indicated level
Measurement or Estimate	#	Source reference or shorty description

The inclusion of new information may require the expansion of the code lists to cover new flows and stocks (section 1.2.1) that need to be carefully defined and linked to the ones already covered in the portal. The portrayal can be realised by using the fields listed in Table 5.

2.3.2 Vehicles

For vehicles, the transfer to portrayals is handled directly by the model software. All stocks and flows share the same format.

Table 6. Fields in the portrayal harvesting template for vehicles stock-flows

field name	*input type*	*Instructions*
ID_Country	txt	ISO-2 country code
Year	#	Year of data point
Vehicle Key	txt	Vehicle key as defined in 3.2
Stock/Flow Name	txt	Which aspect of the model the data point belongs to (stock, pom, etc.)
Value	#	Numerical value (in this case, # of vehicles)
UOM	txt	Parameter subscript (in this case, "# of vehicles")

UNC Min	#	minimum value of uncertainty interval (2 standard deviations, in %)
UNC Max	#	maximum value of uncertainty interval (2 standard deviations, in %)
Measurement or Estimate	#	Source reference or shorty description
Data Quality	txt	Data quality, automatically determined by model from the formal uncertainty of the item
Data Quality comment	txt	comments to the data quality beyond the indicated level
Metadata ProSUM Creator	Txt	Person who prepared the data

An example extract of the portrayal is detailed below:

ID_Country	Year	Vehicle Key	Stock/Flow Name	Value	UoM	Uncertain quantity (min %)	Uncertain quantity (max %)	Measured or Estimate	Data quality	Data quality comment	Person who entered data
FI	2000	V01000000	pom	3	pieces	-0.09761	0.097608	Estimated	Confident	Automatically entered based on model confidence. Derived from EUROSTAT	Duncan Kushnir
BE	2000	V01000001	pom	9033	pieces	-0.00122	0.001219	Estimated	Highly Confident	Automatically entered based on model confidence. Derived from EUROSTAT	Duncan Kushnir
CH	2000	V01000001	pom	2	pieces	-0.66364	0.663645	Estimated	Less Confident	Automatically entered based on model confidence. Derived from EUROSTAT	Duncan Kushnir

Figure 3. Extract of the portrayal for vehicles

In addition, because the e-p vectors for all stocks and flows in the model (excepting POM) are a complex blend of previous POM values, a resource file with the blended e-p vectors for each database entry is also generated.

field name	*input type*	*Instructions*
ID_Country	txt	ISO-2 country code
Year	#	Year of data point
Vehicle Key	txt	Vehicle key as defined in 3.2
Sheetname	txt	Which aspect of the model the data point belongs to (stock, pom, etc.)
Element [Ag,Au, etc. x28]	# [x28]	Blended e-p content for this database entry (in kg/vehicle average)

An example extract of this companion file is shown below:

Vehicle Key	Country	Year	Sheetname	Ag	Al	Au	Ce	Co	Cu	Dy	Fe
V01000000	BG	2000	stock	0.007718	53.47715	0.001094	0.008691	0.00749	27.92475	0.000342	800.7688
V01000000	BG	2001	stock	0.007861	55.73769	0.001115	0.009464	0.007628	27.92475	0.000412	797.4461

Figure 4. Extract of the companion e-p vector supplement

2.3.3 WEEE

The aforementioned databases in GitHub and ColCat_emc are automated by CBS and GEUS respectively, the data received from GitHub extraction files can easily be manually transferred to portrayals.

For WEEE, the transfer to portrayals is handled manually in Excel files having a format aligned with the portrayal templates used by the BATT and Vehicles teams. All stocks and flows share the same format. Table 7 lists the fields in the portrayal harvesting template for WEEE POM data. Similar templates are used for the data on the average weights and life span parameters.

Table 7. Fields in the portrayal harvesting template for WEEE stock-flows

field name	*input type*	*Instructions*
ID_Country	txt	ISO-2 country code
Startum	text	Stratum as defined in 3.1
Year	#	Year of data point
UNU Key	txt	UNU key as defined in 3.1
Collection category	txt	As defined in recast Directive (2012/19/EU)
Stock/Flow Name	txt	Which aspect of the model the data point belongs to (stock, POM, etc.)
Value	#	Numerical value (# of EEE)
UOM	txt	Parameter subscript (kg/inh of EEE)
ID_metadata and ProSUM creator	#, txt	Person who prepared the data
UNC Min	#	Minimum value of uncertainty interval (2 standard deviations, in %)
UNC Max	#	Maximum value of uncertainty interval (2 standard deviations, in %)
Data Quality	txt	Data quality, automatically determined by Uncertainty and flag values
Data Quality comment	txt	comments to the data quality beyond the indicated level
Data consolidation	txt	Data quality, automatically determined by model from the formal uncertainty of the item
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.
Additional fields	txt	Population and households data for conversion to different units, i.e. tonnes, pieces, pieces per inhabitant, pieces per households, etc.

For WEEE, the transfer of new consolidated data is split in the different complementary flows. Once the products in the flow were we had information have been analysed and the uncertainties have been defined the transfer of this information to the portrayals should be made as follows:

For the countries which information was provided its original data is allocated to the respective country and for the respective collection category. In the case of the uncertainties, as previously described in Deliverable 3.2 for the countries were there was information the data was computed and the countries were we didn't have information a qualitative assessment was done. As seen in Figure 5 the portrayals made for complementary flows are simple and user friendly and only require to fill in the information described in Table 8.

Table 8. Fields of the portrayal template and instructions for filling them. Not all fields are harvested.

field name	*input type*	*Instructions*
Year	#	Year being analysed
ID_Country	txt	Name of the country which the value correspond to.
UNU Key Collection Category	#	Product key Category for which the data applies. Taken from code lists for EEE.
Compl. Recycl Products	#	Value of the data set being analysed.
UOM	txt	Parameter subscript (in this case p-f).
ID_Metadata	#	Source reference or shorty description
Metadata ProSUM Creator	#	Source reference from which the reference comes from
UNC Min	#	minimum value of uncertainty interval (confidence interval)

UNC Max	#	maximum value of uncertainty interval (confidence interval)
Data Quality	txt	data quality, as determined by the procedure described in D3:2
Data Quality comment	txt	comments to the data quality beyond the indicated level
Consolidation comment	txt	Comments to how the data consolidation was performed. It is recommended to avoid long text, as the details of the consolidation can be explained elsewhere (i.e. in the metadata catalogue).
Data consolidation	txt	Comments on how data was consolidated.
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.

An example extract of the WEEE p-f waste flow portrayal is shown in Figure 5.

Year	COUNTRY	UNU-KEY Collection Category	Compl Recycl Products & Scavanged (kg/inh)	UoM	ID_metadata	Metadata ProSUM Creator	Uncertain quantity (min)	Uncertain quantity (max)	Data quality	Data quality comment	Data consolid ation	Original or estimate?	Estimation type
2012	FIN	0	0.09	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2013	POL	0	0.31	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	ESP	0	0.10	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	ITA	0	0.01	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	FIN	I	0.02	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2013	POL	I	0.00	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	ESP	I	0.00	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	ITA	I	0.00	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	FIN	II	0.02	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2013	POL	II	0.00	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat
2012	ESP	II	0.04	Individual reporting fron	9	CWIT D4.2	-15%	15%	Confident	Data from re	None	Original/ consolidated d	Original/ consolidated dat

Figure 5. Extract of the WEEE waste flow portrayal

Collected and reported WEEE

A detail procedure to update collected and reported WEEE is provided in section 2.6 of D4.4 ‘Protocols for CRM Content in Waste Flows and Data Quality Assessment”.

Scavenging flows

As mentioned in section 2.2.3, in the case of scavenged flows the database only contains a total value and values for Category I corresponding to cooling and freezing appliances. No consistent data is available for other collection categories yet. For the countries which information was provided its original data is allocated to the respective country and the uncertainties, as previously described in Deliverable 3.2 for the countries were there was information the data was computed. As seen in Figure 6 the portrayals made for the scavenging flows are simple and user friendly and only require filling in the information described in Table 9.

Table 9. Fields of the scavenging flow portrayal template and instructions for filling them. Not all fields are harvested.

field name	*input type*	*Instructions*
ID_Country	txt	Name of the country which the value correspond to.
Year	#	Year being analysed
UNU Key Collection Category	#	Product key Category for which the data applies. Taken from code lists for EEE.
Value	#	Value of the data set being analysed.
UOM	txt	Parameter subscript (in this case p-f).
ID_Metadata	#	Source reference or shorty description
Metadata ProSUM Creator	#	Source reference from which the reference comes from
UNC Min	#	minimum value of uncertainty interval (confidence interval)
UNC Max	#	maximum value of uncertainty interval (confidence interval)
Data Quality	txt	data quality, as determined by the procedure described in D3:2

Data Quality comment	txt	comments to the data quality beyond the indicated level
Consolidation comment	txt	Comments to how the data consolidation was performed. It is recommended to avoid long text, as the details of the consolidation can be explained elsewhere (i.e. in the metadata catalogue).
Data consolidation	txt	Comments on how data was consolidated.
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.

An example extract of the database from the scavenge flow portrayal is shown in Figure 6.

ID_Country	Year	UNU_Key, Colcat6 or Colcat 1	Value	UoM	ID_metadata	Metadata ProSUM Creator	Uncertain quantity (min)	Uncertain quantity (max)	Data quality	Data quality comme	Data consolidation	Original or estimate?	Estimation type
AUT	2010	I	0.13	kg/inh	9	CWIT D4.2	-90%	130%	Less confid	Data repre	Average EU	Coherent e	Estimated b
BEL	2010	I	0.16	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
BGR	2010	I	0.08	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
CHE	2010	I	0.14	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
CYP	2010	I	0.13	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
CZE	2010	I	0.11	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
DEU	2010	I	0.13	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated
DNK	2010	I	0.18	kg/inh	9	CWIT D4.2	-90%	130%	Less confic	Data repre	Average E	Coherent	Estimated

Figure 6. Extract of the WEEE scavenging flow portrayal

Export flows

As described in D3.2, the available data and expert opinion only allows for qualitative assessments since there are a limited number of country studies estimating the proportion of WEEE mixed with pre-shredder residue, therefore for countries were no information was available estimations were made. The uncertainties and qualitative assessment were done as described in 2.1.3. As seen in Figure 7 the portrayals made for the export flow are simple and user friendly and only require filling in the information described in Table 10.

Table 10. Fields of the export flow portrayal template and instructions for filling them. Not all fields are harvested.

field name	*input type*	*Instructions*
Code	txt	Code used to automatize the excel sheet
Country & Colcat	txt	Code used to produce graphs in the Master file
ID_Country	txt	Name of the country which the value correspond to.
Year	#	Year being analysed
UNU Key Collection Category	#	Product key Category for which the data applies. Taken from code lists for EEE.
Value Metal Scrap (kg/inh)	#	Value of the data set being analysed.
UOM	txt	Parameter subscript (in this case p-f).
ID_Metadata	#	Source reference or shorty description
Metadata ProSUM Creator	#	Source reference from which the reference comes from
UNC Min	#	minimum value of uncertainty interval (confidence interval)
UNC Max	#	maximum value of uncertainty interval (confidence interval)
Data Quality	txt	data quality, as determined by the procedure described in D3:2
Data Quality comment	txt	comments to the data quality beyond the indicated level
Consolidation comment	txt	Comments to how the data consolidation was performed. It is recommended to avoid long text, as the details of the consolidation can be explained elsewhere (i.e. in the metadata catalogue).
Data consolidation	txt	Comments on how data was consolidated.
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.

An extract of the database from the export flow portrayal is shown in Figure 7.

Code	Country&Colcat	ID_Country	Year	UNU_Key, Colcat6 or Colcat 10	ValueMetal Scrap (kg/inh) (f)	UoM	ID_metadata	Metadata ProSUM Creator	Uncertain quantity (min)	Uncertain quantity (max)	Data quality	Data quality comment	Data consolidation	Original or estimate?	Estimation type
2010aut0	aut0	aut	2010	0	2.25	kg/inh	46	CWIT D4.2	-15%	75%	Less confident	Extrapolation is less confide	Original data/estimates	Original/ consolidated data	Original/ consolidated data
2010bel0	bel0	bel	2010	0	0.83	kg/inh	12	CWIT D4.2	-15%	75%	Less confident	Extrapolation is less confide	Original data/estimates	Original/ consolidated data	Original/ consolidated data
2010che0	che0	che	2010	0	0.00	kg/inh	0	CWIT D4.2	0%	0%	0%	0	0	0	
2010deu0	deu0	deu	2010	0	1.14	kg/inh	47	CWIT D4.2	-15%	75%	Confident	Extrapolation is less confide	Original data/estimates	Original/ consolidated data	Original/ consolidated data
2010dnk0	dnk0	dnk	2010	0	0.00	kg/inh	0	CWIT D4.2	0%	0%	0%	0	0	0	
2010fin0	fin0	fin	2010	0	0.00	kg/inh	0	CWIT D4.2	0%	0%	0%	0	0	0	
2010fra0	fra0	fra	2010	0	0.00	kg/inh	36	CWIT D4.2	-15%	75%	0%	0	0	0	
2010gbr0	gbr0	gbr	2010	0	0.65	kg/inh	0	CWIT D4.2	0%	0%	Less confident	Extrapolation is less confide	Original data/estimates	Original/ consolidated data	Original/ consolidated data
2010irl0	irl0	irl	2010	0	0.00	kg/inh	0	CWIT D4.2	0%	0%	0%	0	0	0	

Figure 7. Extract of the WEEE export flow portrayal

Metal Scrap flows

Deliverable 3.2 describes the calculation methodology used for the metal scrap flows in Europe as well as the criteria to evaluate the quality and uncertainty of this flow. Figure 8 illustrates the portrayals made for metal scrap and

Table 11 describes the information required to produce the portrayals.

Table 11. Fields of the metal scrap portrayal template and instructions for filling them. Not all fields are harvested.

field name	*input type*	*Instructions*
Code	txt	Code used to automatize the excel sheet
Country & Colcat	txt	Code used to produce graphs in the Master file
ID_Country	txt	Name of the country which the value correspond to.
Year	#	Year being analysed
UNU Key Collection Category	#	Product key Category for which the data applies. Taken from code lists for EEE.
Value Metal Scrap (kg/inh)	#	Value of the data set being analysed.
UOM	txt	Parameter subscript (in this case p-f).
ID_Metadata	#	Source reference or shorty description
Metadata ProSUM Creator	#	Source reference from which the reference comes from
UNC Min	#	minimum value of uncertainty interval (confidence interval)
UNC Max	#	maximum value of uncertainty interval (confidence interval)
Data Quality	txt	data quality, as determined by the procedure described in D3:2
Data Quality comment	txt	comments to the data quality beyond the indicated level
Consolidation comment	txt	Comments to how the data consolidation was performed. It is recommended to avoid long text, as the details of the consolidation can be explained elsewhere (i.e. in the metadata catalogue).
Data consolidation	txt	Comments on how data was consolidated.
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.

An example extract of the database from the export flow portrayal is shown in Figure 8.

Code	Country&Colc	ID_Country	Year	UNU_Key, Colcat6 or Colcat 1	ValueMetal Scrap (kg/inh) (f	UoM	ID_metadata	Metadata ProSUM Creator	Uncertain quantity (min)	Uncertain quantity (max)	Data quality	Data quality comment	Data consolidation	Original or estimate?	Estimation type
2010aut0	aut0	aut	2010	0	2.66	kg/inh	19	CWIT D4.28	-15%	15%	Confident	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous
2010bel0	bel0	bel	2010	0	2.68	kg/inh	19	CWIT D4.27	0%	100%	Less confident	Allocation to collection cate	Data from national survey	Original/ consolidated data	0
2010che0	che0	che	2010	0	0.00	kg/inh	19	CWIT D4.26	-100%	100%	Dubious	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous
2010deu0	deu0	deu	2010	0	1.57	kg/inh	19	CWIT D4.25	-100%	100%	Dubious	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous
2010dnk0	dnk0	dnk	2010	0	4.16	kg/inh	19	CWIT D4.24	-100%	100%	Dubious	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous
2010fin0	fin0	fin	2010	0	0.00	kg/inh	19	CWIT D4.23	-100%	100%	Dubious	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous
2010fra0	fra0	fra	2010	0	3.19	kg/inh	19	CWIT D4.22	0%	100%	Confident	Allocation to collection cate	Data from national survey	Original/ consolidated data	0
2010gbr0	gbr0	gbr	2010	0	4.35	kg/inh	140	CWIT D4.21	-35%	35%	Less confident	Allocation to collection cate	Data from sampling light ir	Original/ consolidated data	0
2010irl0	irl0	irl	2010	0	4.16	kg/inh	19	CWIT D4.20	-100%	100%	Dubious	Allocation to collection cate	Average data for stratum :	Coherent estimate/ reconciliated data	Average value of 1.4% in total ferrous+non-ferrous

Figure 8. Extract of the WEEE metal scrap flow portrayal

Waste bin flow

As previously described in Deliverable 3.2 for the countries were there was information the data was computed and the countries where we didn't have information an extrapolation was made taking into consideration the countries purchasing power. Figure 9 illustrates the portrayals used for waste bin flow and

Table 12 describe what is needed for the creation of the portrayals.

Table 12. Fields of the waste bin portrayal template and instructions for filling them. Not all fields are harvested.

field name	*input type*	*Instructions*
Code	txt	Code used to automatize the excel sheet
Country & Colcat	txt	Code used to produce graphs in the Master file
ID_Country	txt	Name of the country which the value correspond to.
Year	#	Year being analysed
UNU Key Collection Category	#	Product key Category for which the data applies. Taken from code lists for EEE.
Value Metal Scrap (kg/inh)	#	Value of the data set being analysed.
UOM	txt	Parameter subscript (in this case p-f).
ID_Metadata	#	Source reference or shorty description
Metadata ProSUM Creator	#	Source reference from which the reference comes from
UNC Min	#	minimum value of uncertainty interval (confidence interval)
UNC Max	#	maximum value of uncertainty interval (confidence interval)
Data Quality	txt	data quality, as determined by the procedure described in D3:2
Data Quality comment	txt	comments to the data quality beyond the indicated level
Consolidation comment	txt	Comments to how the data consolidation was performed. It is recommended to avoid long text, as the details of the consolidation can be explained elsewhere (i.e. in the metadata catalogue).
Data consolidation	txt	Comments on how data was consolidated.
Estimation type	txt	Comments on whether the dataset was a coherent estimation original.

An example extract of the database from the export flow portrayal is shown in Figure 9.

Code	Country&Colc	ID_Country	Year	UNU_Key, Colcat6 or Colcat 1	Value Waste Bin (kg/inh)	UoM	ID_metadata	Metadata ProSUM Creator	Uncertain quantity (min)	Uncertain quantity (max)	Data quality	Data quality comment	Data consolidation	Original or estimate?	Estimation type
2010AUT0	AUT0	AUT	2010	0	1.36	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT01	AUT01	AUT	2010	01	0.03	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT02	AUT02	AUT	2010	02	0.41	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT03	AUT03	AUT	2010	03	0.29	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT04	AUT04	AUT	2010	04	0.29	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT05	AUT05	AUT	2010	05	0.17	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT06	AUT06	AUT	2010	06	0.07	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT07	AUT07	AUT	2010	07	0.07	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries
2010AUT08	AUT08	AUT	2010	08	0.00	kg/inh	9	CWIT D4.2	-15%	15%	Confident	Allocation to collection categ	Average data for stratum 1	Coherent estimate/ reconciliated data	Proxy for similar countries

Figure 9. Extract of the WEEE waste bin flow portrayal

2.4 Harvest data to the UMKDP

The harvesting of data to the UMKDP is covered in chapter 3.4 of deliverable report D5 (Cassard et al. 2016).

2.5 Update frequency

For WEEE, the input data is ProdCom and trade data. New data is available every year. Therefore updates on POM, WEEE and Stock data can be done every year. Below is breakdown of hours required for future datasets. Also for batteries, new data is available every year for most of the stocks and flows. Therefore updates on POM, WEEE and Stock data can be done either every year or for two years every second year. Beside the time required to realise the steps listed in Table 1, communication efforts and time are required to ask the different stakeholders for the data, understand the tables they provide and discuss with them the data scope, representativeness, the data quality assessment etc.

3 Conclusions, Recommendations and Outlook

3.1 BATT Conclusions

The main conclusion regarding the BATT flows is that the data sourcing is too complex to enable an automatization of the ProSUM data. Both the update of the data to cover more years and the inclusion of new data to cover more stocks and (complementary waste) flows require manual data processing that was done, so far, using Microsoft Excel. One exception is the fully automatized calculation by the portal of the stocks and waste generation based on the lifespan parameters and the POM data. These main conclusions are directly linked to the conclusions of D3.5 on data availability, harmonisation and evaluation. The main challenges, so far, are not related to technical issues to transfer the data to the portal, but on the data availability and harmonisation, which condition the feasibility and limit the automatization of updates. As long as the available data are not harmonised, the automatisation and simplification of the update protocols are not possible.

3.2 Recommendations for BATT Stocks and Flows

The representation of the data should ensure that the uncertainties are clearly stated and that no conclusions are drawn without keeping in mind the uncertainties. Depending on the users, the interests related to data representation may lay on specific electrochemical systems (BATT keys and subkeys), applications, flows, countries and years. For example,

- Battery producers may be interested in the volumes of a certain type of batteries put on the market;
- National authorities in data from their own country;
- Battery collectors in waste generation, i.e. the theoretical availability of waste batteries for collection;
- Battery recyclers in the amounts and types of secondary resources waste batteries contain;
- Producers of electric vehicles on forecasts of the CRM consumptions to manufacture batteries.

The gap between waste batteries generation and collection is very relevant for many users of the battery data, so that it is recommended to create representations of the data that enable visualising it, differentiating the electrochemical systems (BATT keys and subkeys), applications, countries and years. The representations of the CRM flows should focus on the elements that are relevant for batteries: cobalt, lithium, graphite and rare earth elements. These visualisations in turn may trigger data providers to align future data along the classifications developed and applied in ProSUM now.

Specific recommendations are the following:

1. Improve data on the unknown whereabouts of (WEEE and) BATT through waste sampling. More work is needed to substantiate the flows described as ‘unknown whereabouts’. For batteries, the complementary flow data is largely for waste batteries in municipal waste. No data are available for batteries disposed of in packaging, batteries not properly treated in Member States, or batteries disposed of in WEEE and vehicles. For these flows a feasibility study should be developed to derive the necessary methodology. Hence it is proposed to create one specific case study on this under the umbrella of the H2020 project ORAMA.
2. Undertake household waste surveys to measure the stocks and residence times of EEE and batteries in households. Surveys have previously been undertaken at a Member State level, usually by market research organisations, and are, therefore, not harmonised or consistent in scope. Standardised consumer surveys would allow for more precise determination of stocks of WEEE and batteries, to increase the reliability of total stocks and derived lifespan information, and consequently reduce error propagation for stocks and waste generated. Such surveys should be based on the protocols developed in

ProSUM which provide guidance in issues such as whether the product is still working, whether it was purchased second-hand, and consumer hoarding and disposal behaviour. Further work is needed to determine efficient and effective survey distribution channels.

3. Collect data on the amount and types of batteries separated from WEEE and ELV at treatment facilities and sent to battery recyclers: It is a legal requirement that batteries embedded in WEEE and ELVs are separated at treatment facilities and sent for recycling. Very limited data are available on these batteries. There is no data on the batteries that are in deregistered or exported used vehicles, collected by manufacturers for replacement, refurbishment or repurposing, inside WEEE or improperly discarded by users. Data is required for these sources and those batteries removed at certified and authorised treatment. This data would improve waste flow information in the UMKDP.

3.3 Vehicle Conclusions

Vehicle data sourcing still requires manual input at every stage of the process leading up to ingestion and running of the model. This could potentially be improved if the Eurostat databases provided an API (application programming interface) for remote checking and access of data values.

A number of countries are missing data for a large number of possible data points. At present, the model smoothly deduces and interpolates these values, but future improvements of the database are welcome and can be integrated automatically.

The most critical missing data is the large gap between, on the one hand, reported ELV treatment and exported vehicles for use, and, on the other hand, the number of deregistrations. There are a number of possible explanations (temporary deregistrations, unauthorized treatment, unreported export, etc.) that should be directly addressed by harmonizing and better reporting of the national methods for recording deregistrations.

3.4 Recommendations for Vehicle Stocks and Flows

Again, the representation of the data should ensure that the uncertainties are clearly stated and that no conclusions are drawn without keeping in mind the uncertainties.

For vehicle data, there are three primary aspects that should be kept in mind: 1) the differing orders of magnitude in uncertainty between the vehicle figures and the element content model. 2) The fact that the long residence time of vehicles means that significant stocks exist from before there was detailed data. 3) The fact that compositions data fully specified to all vehicle keys is still incomplete.

The implications of these observations are as follows: The effective uncertainty of element stocks and flows for most vehicle keys arises almost entirely from the vehicle composition model, so long as the vehicle figures are sufficiently aggregated. While some vehicle keys contain small numbers of vehicles, and thus high uncertainty, all aggregations of vehicle key data at the drivetrain energy level have low uncertainty. Representations should thus avoid drilling down to the finest grained resolution on vehicle keys. Aggregations by drivetrain energy and by country are accurate, but the motor size and weight distribution contain omissions and nonlinearities that primarily arise from missing data and how it is treated in the model.

We thus recommend viewing vehicle statistics aggregated by drivetrain energy and by country, and using the uncertainties associated with the composition (e-p) model to convey accuracy.

Specific recommendations for vehicle stocks and flows information are:

1. Amend vehicle fleet statistics as reported by Eurostat to expand on vehicle characteristics that are crucial for CRM content, such as further specification of drivetrain types (especially electrified ones) and segment.
2. Develop reporting of vehicle fleets to minimise the number of vehicles of unknown whereabouts.
3. Amend the ELV Directive to include a requirement for reporting vehicle age and the destination of vehicles exported for recycling within the EU and Eurostat to publish.

3.5 WEEE Conclusions

The main conclusions regarding the various parts of the EEE datasets are:

1. The POM, Stock and WEEE datasets are created by CBS and are very much standardised when it comes to the processing of the data itself and also the back and forecast work has become rather automatized. However, the process of checking annually for new good codes and reviewing code changes will always take some time. In total, the estimate for an annual update of this part is about 4 to 6 weeks of work in total.
2. Complementary Waste flows databases are generally speaking not too elaborate, but need rather ad-hoc consolidation steps when being updated with very limited room for automatization. The conversion of very diverse studies, when available at all, is considerable. The same counts for consolidating data on reported collection amounts between different sources.
3. The consolidated compositions files from WP2 are extremely time-consuming to be updated as also described in Deliverable 2.7. Such updates should not be done on a very regular basis, but rather batch-wise covering multiple years of new data. The process of harvesting them into consolidated portrayals and linking them to the EEE stocks and flows information is already rather automatized and can basically be performed in few hours.
4. Creating the detailed results files from the UMD has been improved in speed considerably. Still, the computation of these files and export from GEUS to UNU specifically takes considerable computing time due to the huge amount of records and resulting file sizes involved. Also, running error-checks, in particular when there are code lists additions or omitting fields, like a missing or additional UNU key, component, country or year for example, will always require a dedicated and manual checking round when new data is taken on board.

3.6 Recommendations for EEE Stocks and Flows

To improve future EEE stocks and flows information, the following can be considered:

1. Improve data on unknown whereabouts of WEEE by making adoption of EN 50625 standards legally binding. There is evidence to suggest that mandatory conformity of WEEE treatment operations with EN 50625 standards on WEEE treatment and handling, specifically in France, Ireland, Lithuania and the Netherlands, leads to improvements in the amount and quality of reporting. If actors on the market were legally required to handover WEEE to certified plants, more data of a higher quality on WEEE generated would be available.
2. Further sampling, particularly for large domestic appliances in light iron and mixed metal scrap is required. Sampling using standard methods would improve data on the amount of small WEEE and batteries actually disposed of in municipal waste. Some Member States already undertake periodic waste stream sampling. Such sampling methodologies could be harmonised and undertaken more frequently.
3. Improve timeliness, frequency and granularity of WEEE (flows) reporting to Eurostat. Amending the Member State reporting requirements of the WEEE Directive and statistics related to EU legislation to report on both the hazardous and CRM content of WEEE should be considered. It is already a requirement that hazardous and specified components are removed from WEEE under Annex VII of the WEEE Directive. It is recommended that the removal of such components should be reported at Member State level. This could be implemented under Article 16 of the WEEE Directive. In addition, the European Commission could consider taking initiatives to improve the timeliness and quality of Eurostat waste statistics.

A more general recommendation for all three waste sectors combined is the following:

- Produce publicly available reports related to the quality of reported data under the WEEE, Batteries and ELV Directives: In general, it is not possible to assess the quality of reported data provided under the EEE, ELV or BATT Directive. The data reports provided by reporting member states should be made publicly available.

Key references

Van Straalen, V.M, Roskam, A.J., & Baldé, C.P. (2016). Waste over Time [computer software]. The Hague, The Netherlands: Statistics Netherlands (CBS). Retrieved from: http://github.com/Statistics-Netherlands/ewaste

Cassard, D., Tertre, F., Schjøth, F., Heijboer, T., Čápová, D., Hallberg, A., Šinigoj, J., (2016) ProSUM EU-Urban Mining Knowledge Data Platform Specifications.

Heijboer, T., Schjøth, F., Podboj, M., Cassard, D., Tertre, F., Hallberg, A., (2016) ProSUM Data models and code lists (D5.5).

Annex - Specification of the POM, Stock and WEEE data update

Table 13. CBS annual update of EEE data

Task	Task detail	Hours	Comment
Getting data ready for first analysis	Collecting data if data sources are not changed	1	This is download of the new year and of all revised previous years.
	Merging all data together	2	There is a script that does this task completely automatically. It takes a long time to run, but this can be done at night.
	Collecting data if data sources are changed	24	Depends greatly on the way data will be made available.
	Update ProdCom and CN links to UNU_Keys	4	Based on number of code changes in the last few years. This can be more when lots of code changes take place.
	Updating average weights without research	1	This involves just using the same weights for ProdCom/CN and UNU_Keys as in the previous year. No new research.
	First run of the scripts	1	
	Total time	*33*	Less than sum of parts because automatic procedure does not take labour time.
Cleaning up the data	Analyzing all data, correcting outliers manually and choosing extrapolation methods	80	This takes a lot of time. It involves analysing data for outliers and finding a way to correct them. Each correction also has impact on other data points, so there are often many corrections/analysing cycles needed. Extrapolation methods also take a lot of time. When extrapolations have been connected to the last available data point, the same will be done in the next year to the last available data point at that time. This needs a check if all is still looking good. When a specific connection year has been chosen in the past for instance due to bad data in later years, the same connection year keeps being used. So when new data becomes available this will be overwritten with extrapolation estimates. To prevent this those values have to be checked in detail to see if the connection year can be changed.
Dissemination	GitHub	4	Cleaning up all files and making sure the right versions are all together before uploading
	shinyapps.io	1	App linked to GitHub for data security. Each Shiny application runs in its own protected environment and access is always SSL encrypted
	In case of new additions to the program, writing documentation	8	
	Total time	*13*	on average when not too much new documentation needs to be written.
Total hours		*126*	*CBS steps to update data*