Food and Agriculture Organization of the United Nations

FAO STATISTICAL DEVELOPMENT SERIES 20

ISSN 1014-3378

FOOD BALANCE SHEETS AND SUPPLY UTILIZATION ACCOUNTS RESOURCE HANDBOOK 2025

FAO STATISTICAL DEVELOPMENT SERIES 20

ISSN 1014-3378

FOOD BALANCE SHEETS AND SUPPLY UTILIZATION ACCOUNTS RESOURCE HANDBOOK 2025

Food and Agriculture Organization of the United Nations
Rome, 2025

Required citation:

FAO. 2025. *Food balance sheets and supply utilization accounts resource handbook 2025.* FAO Statistical Development Series, No. 20. Rome. https://doi.org/10.4060/cd4472en

ISSN 1014-3378 [Print]
ISSN 2664-746X [Online]

ISBN 978-92-5-139645-2

Contents

Tables

Figures

Equations

Preface

Acquiring accurate and timely statistical data is crucial for shaping effective strategies in agrifood systems, directly impacting living standards across nations. Since its establishment in 1945, the Food and Agriculture Organization of the United Nations (FAO) has prioritized the development of comprehensive food and agricultural statistics, recognizing data as the cornerstone for understanding agrifood systems and promoting necessary measures. Food balance sheets (FBS) are pivotal statistics that provide comprehensive insights into national food consumption patterns, levels and trends. Compiled by the FAO Statistics Division for over 180 countries and territories, FBS offer an in-depth view of countries' food supply and use patterns.

This handbook provides an essential description of the methodology used by FAO to develop FBS. It aims to familiarize Member Nations, food budgeting institutions, and other interested parties with the core processes involved in data collection and imputation strategies specifically for agricultural and livestock products. While offering a detailed overview of these internal FAO procedures, the handbook is not intended to be a manual for countries to compile their own FBS. Moreover, as this is not a technical manual, it does not make explicit reference to coefficient values or other numerical parameters mentioned; these fall outside the scope of this handbook and can be found on the FAOSTAT database. Additionally, this handbook does not address data sources or imputation procedures for fisheries or forestry. Innovative and refined techniques for compiling supply utilization accounts (SUAs) for more than 500 primary and derived products are presented. The SUAs undergo a rigorous process of transformation into over 90 FBS products and their related aggregates, a process that includes imputing missing data and ensuring that they balance. The ongoing review and continuous enhancement of these methodologies involve a dedicated effort to ensure the creation of replicable FBS for all nations, even with limited data availability. As advancements unfold, this handbook will be updated to incorporate the latest recommendations for FBS compilers, reflecting the evolution and advancements in this critical domain.

The handbook begins by offering a brief historical context and defining fundamental concepts pivotal to FBS compilation. It outlines the framework used and then delves into each component of SUA and FBS equations. A comprehensive examination of the methodological intricacies involved in compiling SUAs for various variables such as production, trade, loss, food, feed, seed, stocks, industrial utilization and tourist consumption is also provided. Furthermore, the handbook explains the methodology for computing commodity balances and provides examples of both SUA and FBS compilations.

Updated versions of the current methodology have been in use since 2014. The most current methodologies for the compilation of SUAs/FBS differ from those used in the past on some key features that will be further explained: the substitution of the FAOSTAT commodity list (FCL) with the United Nations Central Product Classification (CPC) and the absence of a balancing variable. Formerly, one of the components of the SUAs/FBS (often stocks, industrial utilization or feed) would compensate for unbalanced amounts, causing statistical errors, while the current balancing methodology assigns part of the imbalance to the stocks and to food in a manner that is in line with the historical data series and, for the remaining part, assigns the imbalance proportionally to the other uses through bounded maximum algorithms.

The FBS data have been updated in FAOSTAT, and the latest version of the methodology was applied backwards to the 2010–2013 period to produce a consistent time series from 2010. Calculations of per capita figures and other variables for all years were also updated with the latest population figures from the United Nations Population Division.

Acknowledgements

This handbook was developed by the Crop, Livestock and Food Statistics team of the Statistics Division of the Food and Agriculture Organization of the United Nations (FAO), with the overall coordination of Cristina Muschitiello, Editor of the handbook, and the active collaboration and contributions of Tomasz Filipczuk, Dominique Habimana, Irina Kovrova and Giulia Piva. The methodological developments, refinements, enhancements, fine-tuning and testing involved Kenneth Basham, Alfia Bonomo, Chiara Bordin, Rachele Brivio, Joshua Browning, Marcella Canero, Luigi Castaldi, Piero Conforti, Bernhard Dalheimer, Claudia De Vita, Carola Fabi, Gianluca Fiorentino, Michael Kao, Livia Lombardi, Christian Mongeau, Amsata Niang, Adam Prakash, Francesca Rosa, José Rosero Moncayo, Josef Schmidhuber, Aydan Selek, Sumeda Siriwardena, Charlotte Taglioni, Salar Tayyib, Cristina Valdivia and Bruno Vidigal. The team is also grateful to Iñaki Arto, Katherine Baldwin, Giulia Conchedda, Alexandra Fleischmann, Christophe Guyondet, Jim Hansen and Francesco Tubiello for their inputs during the peer review of the handbook. Nancy Chin and Olivier Lavagne d'Ortigue provided editorial support.

The various elements of the revised FBS methodology have been presented to the FAO Technical Data Coordination Group, whose members provided comments that are reflected in this handbook. Parts of the methodology were subject to external peer review, particularly the sections on trade, trade endorsement calculations, food losses and waste, and feed use. All methods were developed in close cooperation with FAO technical divisions. Specifically, collaboration with the following FAO bodies helped improve the supply utilization accounts/food balance sheets:

- the former Rural Infrastructure and Agro-Industry Division helped improve sections on post-harvest losses and waste and explored new sources of data on industrial use of food commodities;
- the Animal Production and Health Division helped improve feed estimates, using expert knowledge and information from various databases, notably the Global Livestock Environmental Accounting Model database;
- the Plant Production and Protection Division helped improve the section on seed rates and the methodology for imputing seed use;
- the African Commission on Agricultural Statistics, the Asia and Pacific Commission on Agricultural Statistics;
- the FAO/OEA-CIE/IICA Working Group on Agricultural and Livestock Statistics for Latin America and the Caribbean provided feedback, inputs and updates on methodological innovations and capacity development.

Collaboration with the United Nations Statistics Division (UNSD) and the World Customs Organization (WCO) facilitated implementation of a new classification system that is more in line with international standards.

Abbreviations

ARIMA	autoregressive integrated moving average
AUE	animal unit energy
AUP	animal unit protein
CPC	Central Product Classification
DES	dietary energy supply
Eurostat	Statistical Office of the European Communities
FAO	Food and Agriculture Organization of the United Nations
FBS	food balance sheets
FCL	FAOSTAT commodity list
FCT	food composition tables
GDP	gross domestic product
GLEAM	Global Livestock Environmental Assessment Model
HS	Harmonized Commodity Description and Coding System
kcal	kilocalorie
kg	kilogramme
MJ	megajoule
MJME	megajoule of metabolizable energy
NCT	nutrient conversion table
n.e.c.	not elsewhere classified
n.e.s.	not elsewhere specified
NSO	National Statistical Office
SDG	Sustainable Development Goal
SUA	supply utilization account
t	tonne
UNSD	United Nations Statistics Division
USDA	United States Department of Agriculture

INTRODUCTION

Overview

Food balance sheets (FBS) are a structured representation of a country's food availability, presented as an accounting of the supply and use of resources and food during a specified reference period. It also provides insights into the potential sufficiency of energy intake using the dietary energy supply (DES), a key indicator representing the average caloric availability for the entire population of a country or on a per capita basis.

The overall process for compiling FBS starts with the preparation of a supply utilization account (SUA) that provides data on the supply and utilization for each food product (primary and derived) in a country expressed in terms of the given primary product. To determine the available supply during a specific period, the total quantity of domestic production (adjusted for changes in stocks) is added to the total quantity of imports. On the utilization side, the main categories comprise exports, livestock feed, seed usage, manufacturing of food and non-food items, storage and transportation losses, and food available for human consumption.

FBS are calculated from SUAs to produce a harmonized set of variables that express all products in terms of primary commodities. The FBS therefore provide a clear and comparable picture of food production patterns and utilization trends that facilitates a comprehensive understanding of a country's food security status and may aid in global food policy formulation and evaluation.

The process of creating FBS often starts with a limited set of hard statistics. For many countries and many products, actual measurement of the variables in the account is entirely absent or, where available, is associated with large implicit or explicit measurement errors. Therefore, FBS require many, and often particularly complex, methods for the imputation of missing data, computation, and estimation of variables.[1]

The first attempts at preparing FBS date back to World War I. FBS were the major source of data when, in 1936, at the request of the League of Nations Mixed Committee on the Problem of Nutrition and its Subcommittee on Nutritional Statistics, a systematic international comparison of food consumption data was prepared. During World War II, the interest in FBS increased considerably and they were the main source of data used by FAO in the assessment and appraisal of the world food situation in the six *World Food Surveys* conducted between 1946 and 1996. Today, FBS are widely used for calculating indicators and as tools for policymaking.

Over time, FBS and SUAs have undergone methodological changes. They were first published in a standardized format in 1984. Since 2014 the FAO Statistics Division (ESS) has been regularly reviewing and revising the methodological approaches for all its products, including all the databases it maintains, their underlying and accompanying metadata and the approaches for imputing missing data.

The present handbook focuses on the methodology currently used by ESS for the compilation of SUAs and FBS. It builds upon the foundations laid by the *Guidelines for the compilation of food balance sheets* drafted in 2017 by the Global Strategy to Improve Agricultural and Rural Statistics. These guidelines were designed to provide countries with a user-friendly handbook facilitating the construction of country-level FBS for policy analysis. This handbook, while describing FAO's methodology for compiling FBS, is not a manual but takes a distinct approach and scope, drawing extensively from these earlier guidelines. It aims at national FBS compilers and also at a wider audience, including policymakers, researchers and stakeholders in the agricultural sector. By providing insights into the methodology and its significance, it seeks to enhance understanding of the processes involved in FBS compilation and their implications for food security and agricultural development.

[1] The difference between imputation, computation and estimation can be found in Section **2.3**.

Uses of food balance sheets

An increasing number of countries compile comprehensive FBS, although the exercise can be time-consuming and requires additional resources for agricultural statistics production and analysis. The rising interest in compiling FBS is due to the significant value of the data generated from this process, as well as the numerous potential applications of that information. Several potential applications showing the benefit of FBS to national statistical systems on food and agriculture are provided below.

Measuring and analysing overall food supply

One of the most important pieces of information for users of FBS is the measure of overall average calorie supply in a country provided using the DES. As previously defined, this is a measure of the amount of food available, even if it does not consider factors such as food waste, food loss or individual dietary preferences and habits. Although it cannot accurately reflect the actual amount of nutrients consumed by individuals, it is still a useful tool for identifying potential food scarcities or imbalances at the national level, which can inform policy and interventions to improve food security. Compiling FBS over several years will enable users to track changes in the food supply over time, including estimated total caloric availability, growth of consumption in new products, and general changes in dietary composition, providing some insights into overall trends in food consumption. While this is possible in principle, the observed changes refer only to the average diet and conclusions cannot be made about overall diet quality. While the average diet may appear in line with dietary guidelines in terms of volume and composition, the average may mask unhealthy over- and undernutrition.

Providing input for policy analysis aimed at ensuring food security

Uses of the FBS data include informing government policies and serving as foundational inputs for analysis. Using FBS, countries, international agencies and researchers can look for any shortfall or surplus in a nation's dietary energy and nutrient intake and can identify countries above or below certain standards (Pinstrup-Andersen, 1993). Published literature makes wide use of FBS (Jacobs and Sumner, 2002), most often citing DES and fat and protein intake (Grigg, 1993; Hopper, 1999; Pinstrup-Andersen, 1993; Svedberg, 1999; Trueblood, Shapouri and Henneberry, 2001; Smil, 1987).

Medical applications and research using food balance sheets

The medical community has also used FBS to explore the relationships between caloric intake, the types of protein, and amino acids in the diet (Hopper, 1999; Young and Pellett, 1990; Kazuo, 1991) or to investigate connections between diet and health, especially cardiac health, and cancers (Sasaki, 1992; Helsing, 1995).

Calculation of derived economic and food security indicators

The DES is a standard output of the FBS, and also a key input to the prevalence of undernourishment (PoU) that FAO computes to measure the percentage of individuals in the population "whose habitual food consumption is insufficient to provide, on average, the amount of dietary energy required to maintain a normal, active and healthy life" (FAO, IFAD, UNICEF, WFP and WHO, 2024). Measuring the PoU requires modelling the probability distribution of habitual dietary energy intake levels for the average individual; the DES is used to estimate the dietary energy consumption value, which is one element characterizing this distribution. Together with the minimum dietary energy requirement, food security analysts can calculate the number and percentage of people in a population without access to sufficient calories.

FBS are the basis for many other indicators of food security. The most straightforward indicators are simple ratios such as the self-sufficiency ratio, which compares a country's agricultural production to its domestic utilization, and the imports dependency ratio, which compares a country's imports to its domestic utilization. FBS also provide inputs into policy measures such as the producer and consumer support estimates, which are updated annually by the Organisation for Economic Co-operation and Development (OECD). The trade component of the FBS feeds into the aggregate measure of support used by the World Trade Organization (WTO).

Providing input for economic models

FBS are included in some prominent partial equilibrium models of the food and agriculture sector, such as the joint OECD–FAO AGricultural LINKage - COmmodity SImulation MOdel (Aglink-Cosimo) used to inform the *OECD-FAO Agricultural Outlook* series.

Food balance sheets and the monitoring of the Sustainable Development Goals

In September 2015, the Member States of the United Nations officially adopted the 2030 Agenda for Sustainable Development (UN, 2015), which encompasses 17 global Sustainable Development Goals (SDGs) and 169 targets. The Agenda applies to every country and is to be implemented by 2030. In fact, the comprehensiveness of the SDG indicators is such that they will most likely contribute to reference points for monitoring countries' progress even beyond 2030. While the many objectives span poverty alleviation, climate change mitigation, and natural resource conservation, food and agriculture occupy a central position in the 2030 Agenda. FBS underpin two of the 21 SDG indicators under FAO custodianship. First, as explained above, information on the DES is used to compute the PoU (Indicator 2.1.1), which is an essential component for the evaluation of Target 2.1, which is to "end hunger and ensure access by all people, particularly those in vulnerable situations, to safe, nutritious, and sufficient food all year round by 2030". Second, data on post-harvest losses are a critical input for the calculation of the global Food Loss Index (Indicator 12.3.1.a), used to evaluate progress towards Target 12.3, which calls for the "reduction of food losses along production and supply chains, including post-harvest losses, and halving of per capita global food waste at the retail and consumer levels by 2030"

Interpretation of food balance sheet estimates

FBS are the most comprehensive data related to food product supply and utilization available for many countries. The data are regularly revised, and methodologies refined to improve data quality and consistency. Overall, FBS provide a wealth of information and serve numerous uses. However, there are also limits to the applicability and usefulness of FBS estimates. FBS compilers and users must be aware of these limitations and of the potential errors in the estimates.

When analysing FBS, it is important to consider the definition of per capita dietary energy supply, i.e. the food supply (expressed in kcal/capita/day). This is an estimate of the per capita amount of energy (kcal) in food *available* for human consumption, which is sometimes confused with "*effective food consumption"*. Essentially, FBS estimates reflect the *availability* of food that is intended for human consumption and available for purchase at the point of sale. More appropriately, this is sometimes referred to as "*apparent food consumption*". This differs from "*effective food consumption"*, that is, the actual quantity of food purchased or subsequently ingested by an identified group of individuals. Because FBS estimates do not account for food waste at the retail or household levels, incorrect interpretation would imply much higher and unrealistic levels of purchase and consumption.

FBS results require careful interpretation; further details of the differences between food availability as represented in the FBS and food consumption as collected by household surveys are outlined. Food availability assessed through the FBS cannot be directly compared with food consumption data collected in household surveys, given that the latter refer to actual purchases of goods. Differences between the two also include the differences in coverage, with household surveys excluding collective consumption in hospitals, schools, the military or prisons, and not always covering away-from-home consumption in restaurants, street food, etc.

Surveys also report food consumption net of retail waste, while FBS include only food availability and include retail waste. Household surveys may also lack representative coverage over the full reference period of FBS, which is a calendar year. In this handbook, no systematic attempt is made to tally FBS results with food consumption or expenditures from household surveys. An FAO working paper systematically comparing FBS estimates with results from household surveys (Grünberger, 2014) should be consulted before using household survey results to guide FBS food estimates.

FBS can only provide estimates of average or per capita national food or nutrient availability. They do not offer any insights into the distribution of food and nutrient availability: users may draw inferences on the average diet, but not on the diet of food-insecure or poor people. FBS also do not provide information about the subnational distribution of food, the access of groups of households, or dietary habits. With these caveats in mind, FBS can and are being used as either a contributing factor or the sole basis for analysis of food demand and supply. They provide an approximate picture of the overall food situation in each country.

This FBS handbook offers methods for detecting inconsistencies, imputing missing data and filling data gaps. It must be kept in mind, when considering this information, that the imputation of missing data can never replace primary data collection, and that no model-based approach can substitute for, let alone produce more accurate information than, data collected directly. It is therefore imperative, for countries that decide to generate national FBS, to first take stock of the availability of food supply and utilization data; to consider the data available compared to the data needed, and to make decisions on whether robust, consistent FBS can be built given the remaining data gaps. Where data available for several products and variables are insufficient, it may first be necessary to improve the domestic data collection efforts, and resume efforts to compute FBS only once sufficient basic data are available. The Global Strategy to Improve Agricultural and Rural Statistics[2] provides the basis for such an endeavour by offering cost-efficient methods for data collection and providing capacity development support to facilitate access to these methods.

Even if most or all the data required to compile FBS are available, assembling such data may still entail considerable challenges, owing to the diversity of sources, coverage and quality. Often, the underlying accounting structure uses a mix of data from various sources, of diverse quality, with unknown errors. In practice, primary production and trade data are often the only official data regularly collected, while data on the production and utilization of derived products are virtually non-existent, or sparse at best.

This use of a variety of data sources thus requires caution when using, interpreting and verifying the results. It is necessary to carefully examine the outcomes for quality, make appropriate corrections, and accurate interpretations. This is especially important during the balancing process. In a supply/utilization system, the balancing component considers the measurement errors associated with all other elements in the equation. These measurement errors do not necessarily compensate for each other and may thus result in a biased assessment. This makes the balancing component the least reliable or the most variable element in the FBS. To address this challenge and improve data integration, the FAO Statistics Division has developed an approach taking into account that all variables are measured with some degree of inaccuracy. In doing so, it avoids attributing measurement errors solely to one element (see Section **14**), and the FBS estimates can be more robust.

Structure of this handbook

The handbook explains methodologies, processes and practical procedures involved in constructing FBS. All the stages of compiling a food balance sheet are built on the SUAs. Both FBS and SUAs encompass a range of variables describing the supply and utilization of each product. However, all variables presented in FBS are initially integrated into SUAs. The SUAs give a detailed breakdown of primary and derived products, presenting data for

[2] See https://www.fao.org/in-action/global-strategy-agricultural-statistics/en for more information.

the original form of each product. The more comprehensive and standardized format of the FBS is generated only during the final aggregation step, referred to as *standardization* (Section **15**).

Compiling information for SUAs requires many supplementary variables such as conversion factors and ratios. These are used as a basis to assign various derived products to their primary commodity counterparts. The relationships between derived and primary products are encapsulated in *commodity trees* (Section **2.6**), which serve as the foundation for coherent commodity aggregation and standardization.

Section **1** describes the methodology underpinning the compilation of FBS by FAO. Section **2** outlines the key concepts and definitions necessary for an informed understanding of the various aspects of SUA compilation. Section **3** explores the typologies and degrees of quality of data sources. Sections **4** to **13** outline the technical aspects of each variable or component of the SUA/FBS identity. This methodology considers the techniques and sources used, while also covering the imputation procedures for missing data used when compiling SUAs and FBS. Section **14** presents the balancing procedure; Section **15** presents the standardization methodology for compiling FBS from SUAs. Section **16** gives details on nutrient factors that can be extracted. Finally, Section **17** provides an example of FBS compilation.

1 METHODOLOGY

1.1 Basic identity and approach

The logical starting point of SUAs and FBS is that the *total supply* of a given food product must equal the *total utilization* of that product in a particular country and time period.

In SUAs, this principle is typically expressed through one identity that sets total supply equal to total utilization (**Equation 1**). When this occurs, the equation is balanced, and the process is referred to as balancing. The identity is:

Equation 1: Total supply = total utilization (identity at the SUA level)

$$production + imports - \Delta stocks = exports + food + food\ processing + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$$

All the variables of this equation will be defined in Section **2**.

SUAs include all the variables of **Equation 1** for most primary commodities and derived products that are produced or consumed in the country. Once SUAs for these commodities and products have been compiled, they need to be balanced to ensure that the equation holds.

A few constraints are applied to the balancing process that limit how imbalances are redistributed among elements of the equation to achieve a balance between supply and utilization. In principle, there are three types of constraints:

- The first constraint is that the equation holds for every commodity or product in the SUA, therefore supply is equal to utilization for every product in the SUA. Each commodity or product is shown individually (as a separate line or row) in the SUA.
- The second constraint reflects prior expectations regarding the DES, and forces the DES and changes to the DES to remain within empirically sound limits. This constraint addresses potential extremes that might arise when attempting to balance the equation but are unlikely to happen in practice, barring exceptional circumstances such as significant economic issues or disasters.
- The third constraint limits the level and amount of change allowed for a variable (such as feed, seed and stocks) during the balancing process. In balancing, the trend of the variable is considered, and constraints are applied so that the balancing does not assign all the amount to be balanced to one variable only nor for several consecutive years.

Constraints can also be applied during imputation procedures, for example in imputing feed, so that the cumulative nutritive values such as energy or protein derived from the feed products in the SUA meet the biological requirements of the livestock.

FBS are obtained after SUAs have been balanced: first by converting derived products to primary commodity equivalents, and then, for each variable in the SUA, aggregating to the relevant primary commodity. Primary equivalent products are synthetic values (including not only the primary commodity, e.g. wheat, but wheat and products derived from wheat such as macaroni and flour, converted to their equivalent as wheat) that are assumed to be homogeneous and comparable within and across different countries.

The identity of FBS takes the following form:

Equation 2: Total supply = total utilization (identity at the FBS level)

$$production + imports - \Delta stocks = exports + food + processed + \ feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$$

Although the identity appears to be similar to the SUA, the key difference is that the SUAs compile primary commodities and derived products in their original form while FBS standardize these to the equivalent of the primary commodities. Therefore, SUAs contain $food\ processing$ as a variable (see Section **5**), but in the final stages of FBS compilation, processed commodities are converted to their primary equivalents.

1.2 Data comparability

As emphasized in the introduction, FBS serve as the basis for comparing levels and trends in apparent food consumption. To ensure comparability, FBS must be consistent across three major areas: *item comparability*, *units of measurement* and *reference period*.

Item comparability

The FBS methodology assumes that the products compared are identical. For example, reporting production of rice as paddy rice would not be comparable to milled rice. For instance, this distinction is important when estimating food derived from rice as food is typically recorded based on milled rice. Depending on the milling rate, milled rice can be roughly two-thirds of the paddy equivalent. Using a common and structured classification of items is the best way to ensure that the same items are being referred to and are therefore fully comparable. FAO recommends compiling FBS with the Central Product Classification system (CPC version 2.1 expanded) (United Nations, 2015) for agricultural production and the Harmonized Commodity Description and Coding System (HS) for trade data (WCO, 2024). Both are fully described in Section **2.5**.

Units of measurement

Data comparability also requires the use of common measurement units across all products and countries. Products may be reported in different units (tonnes, thousand tonnes, quintals, etc.). FAO recommends using the measurement units described in Section **2.4** for all countries, ensuring full data comparability. These units of measurement are defined in the corresponding FAO statistical standard (FAO, 2023a).

Reference period

The third concept related to data comparability is the choice of the reference period of reporting, which can be defined as marketing year, fiscal year or calendar year. A common definition of the reference year should be decided for all variables in the FBS.

The *marketing year*, also known as crop year or agricultural year, refers to the period starting in the month the majority of the crop is harvested. The *fiscal year* is defined by governments for accounting purposes and varies by country. Finally, the *calendar year* runs from January to December in countries using the Gregorian calendar.

As the fiscal year is conceptually challenging and does not facilitate comparisons (since it varies by country), the marketing year or the calendar year may be used. The main advantage of the marketing year is that it follows the natural cycle of each season. However, it has several disadvantages: within a country, crops can be harvested at different times of the year; the same crop can be harvested at different times in different countries; the marketing year rarely lines up with the calendar year, complicating the comparability with trade data, which are frequently based on the calendar year.

FAO's preference is the calendar year. While the use of this period may sometimes be challenging, production is typically assigned to the calendar year in which the majority of the crop is harvested; when this is the case the calendar year provides a neutral reference.

1.3 New features of the current methodology

Over time, the approaches used for data harvesting and imputation, SUA compilation and FBS standardization and compilation have evolved. This section briefly outlines and highlights the ways in which the current methodology differs from those used previously.

International classifications adopted

The international classification considered appropriate for compiling FBS was changed from the FAOSTAT commodity list (FAO, n.d.a) for agricultural products to the CPC. The FCL is a classification system developed and used by FAO, originally based on the United Nations Standard International Trade Classification[3] (SITC). It has been in use since the 1960s with a revision in the 1990s. The FCL comprises approximately 700 products, organized in 20 chapters, including crops, livestock, food and other related products. The FBS data produced by FAO with the current methodology now use the CPC, which FAO recommends as it ensures comparability with other international datasets and data collection efforts (see Section **2.5.1**); countries using the CPC as a common classification will also allow comparability with the FBS data produced by FAO.

FAO's procedure for updating the SUA balance

A key difference between the current and the previous methodologies concerns the balancing procedure used for the SUAs. In the previous methodology, the balancing was implemented during FBS compilation, after the data were converted to primary equivalent products, and only one variable of the FBS would be used as a balancing variable. Either one of stocks, industrial use or feed were often used as balancing variables and would therefore absorb all the measurement and statistical errors of the other variables in the identity. With the current methodology, the balancing is implemented in the SUA and discrepancies are proportionally allocated to each of the variables of the identity. The imbalance is redistributed to various use variables based on the share of the relevant use variables in total utilization. Several iterations may be required to balance the equation.

Constraints are enforced, which are discussed fully in Section **14**. Examples include using a maximum number of iterations, upper and lower bounds for the different utilizations, maximum and minimum values for each time series. These constraints may also result in unfeasible solutions, when the equation cannot be balanced; in this case a residual variable is included that is equal to the unbalanced quantity. Therefore, the current methodology no longer forces the balancing of the accounts, thus reducing the risk of generating outliers in balancing variables.

Refining the imputation methods of the FBS components

All imputation methods for the different elements of the SUAs have been revised to harness a range of information from outside the FBS, thus resulting in improved accuracy of the SUA variables and reducing the need for manual imputation.

Updated conversion coefficients

Computing FBS uses several fixed coefficients as part of the standardization and conversion. Of particular importance are those that reflect the decrease (or increase) in weight that occurs in the production process when converting a derived product to a primary commodity. These "extraction rates" (see Sections **2.2** and **5.3.1**) are fixed on plausible values (or exogenous) in the current methodology, as opposed to implicit (or endogenous) in

[3] See https://unstats.un.org/unsd/classifications/Family/Detail/28.

the previous methodology. The coefficients are continuously improved, but the current approach reduces the need for the complex ex-post allocation of the item among processed products that was used in the previous methodology.[4]

Unique source for population data

Population figures affect the computation of the DES, as the total DES is divided by the number of people to obtain the quantity of nutrients available per person in a particular country and year. To ensure consistency in its international datasets and improve the consistency of FBS estimates, FAO only uses population data from the United Nations Population Division, which is based on a harmonized methodology. In principle, FBS can be computed using different population data, such as national population census data. This will, however, yield different results for the DES.

Expanded nutritional information

The FBS provide information about the dietary energy supply available for human consumption in terms of calories, proteins, and fats (macronutrients). The current FAO practice is to also generate levels of micronutrients from SUAs, considering a broad range of nutrients. As a result, FAOSTAT data derived from SUAs for 2010 onwards under the Food and Diet domain contain information on carbohydrates, minerals, macro-, and micronutrients, thereby increasing the available data on dietary energy supply (see Section **16**).

[4] To improve the extraction rates, ESS is collaborating with different institutions, including the International Fruit and Vegetable Juice Association (IFU) and the International Association for Cereal Science and Technology (ICC). As further described in Section 5.2, these collaborations have resulted in revised extraction rates and parameters describing losses that occur during the different processing stages, obtained from processing companies for selected fruits, cereals and derived products.

2 CONCEPTS AND DEFINITIONS

2.1 Main variables

Both SUAs and FBS present simple balances between supply and utilization (**Equation 1** and **Equation 2**). They include production, imports and stock withdrawals on the supply side, and exports, food (of the resident population), tourist consumption, feed, seed, post-harvest loss, industrial use and stock additions on the utilization side. These variables, based on the methodology adopted and recommended by FAO, are defined as follows.

Production

The Crop and Livestock Production and Utilization questionnaire (FAO, n.d.b) defines crop production data as "the actual harvested production from the field orchard or garden, excluding harvesting and threshing losses and that part of crop not harvested for any reason". Production therefore includes the quantities of the commodity sold in the market (marketed production) and the quantities consumed or used by the producers (auto-consumption). When the production data available refer to a production period spanning two successive calendar years[5] and allocation to each year is not possible, it is usual to attribute the data to the year where most of the production occurred. Crop production data are reported in tonnes (t). For livestock primary products, the questionnaire defines meat production in terms of the dressed carcass weight of "all animals of local and foreign origin slaughtered within the national boundaries" (FAO, n.d.b). For derived (processed) products, production refers to the total amount of output after processing.

Trade

Trade refers to all transboundary flows of food items destined for, or originating from, a given country. More precisely, imports are transboundary flows of goods destined for a given final destination country that are added to the total supply of goods available in that country, while exports are transboundary flows of goods from a given country of origin that take away from the total availability of goods in that country. As a general rule, trade quantity refers to net weight, excluding any sort of container. Trade includes exports, re-exports, imports and re-imports. Products that enter and exit the borders of a given country without undergoing any alteration are classified as re-imports and re-exports, respectively. Furthermore, the estimation of imports and exports figures should encompass both official and unofficial trade, including food aid deliveries. For specific countries and products, unrecorded trade flows might be considerable, and potentially have a significant impact on assessments of food availability.

During data collection, exports and imports of derived and primary products are often reported in their initial units of measurement. These can be heads, pieces or measures of volume such as *litres*, *hectolitres*, *bushels*, *barrels*, or *bales*. The FBS/SUA methodology requires all variables to be expressed in tonnes, and therefore all trade data are converted into tonnes.

Food supply

Food in the FBS definition refers to "all quantities available for human consumption, either direct by the producers, available for human consumption at the retail level or processed for food use." (FAO, n.d.b). This variable (also called *food availability* or *food supply*) presents the amount of food reaching the retail level and therefore includes all waste and losses that occur at or after retail. Hence, food amounts in the FBS are typically higher than the amounts of food reported by other methods, such as household income and expenditure surveys or nutrition surveys (which exclude certain amounts of waste). In contrast, food in the FBS includes all waste of edible products

[5] As mentioned in Section **1.2**, FAO uses the calendar year as the reference period for FBS compilation.

occurring in shops, supermarkets or households, such as during storage, in preparation and cooking, as plate waste, fed to domestic animals and pets, or otherwise discarded. It also includes food consumed in communal settings, including hospitals, schools, restaurants, military establishments and prisons.

Feed use

Feed use (thereafter referred to as *feed*) refers to "quantities fed to animals, whether direct or used to produce compound feed" (FAO, n.d.b). It excludes all feeds that could neither be used as foodstuffs nor destined for human consumption, including roughages and by-products, such as oilcakes, distillers dried grains with soluble (DDGS), and dregs. These feedstuffs are considered in the calculation of total feed use in the SUA but are excluded from the FBS calculations and presentation.

Seed use

FAO defines seed use (thereafter referred to as *seed*) to include all amounts of the product in question used during the reference period for reproductive purposes, such as seed, sugar cane planted, eggs for hatching, whether domestically produced or imported. This definition includes double or successive sowing or planting of the seed, whenever it occurs. Seed use also includes the quantities needed for the sowing or planting of crops for use as fresh fodder or food (e.g. green peas, green beans, maize for forage), at least when this information is available. On average, the amount of seed needed per hectare planted in a given country and for a given crop does not vary greatly from year to year but may vary over time as productivity-enhancing technologies are adopted.

Food losses

Crop and livestock product losses (thereafter referred to as *losses* or *loss*) cover "all quantity losses along the supply chain for all utilizations (food, feed, seed, industrial, other), up to the retail/consumption level. Losses of the commodity as a whole (including edible and non-edible parts) and losses, direct or indirect, that occur during storage, transportation and processing, also of relevant imported quantities, are therefore all included" (FAO, n.d.b). Food waste occurring at the retail and consumption level (FAO, 2019) is not part of this definition. **Figure 1** provides an overview of losses (and waste) along the entire value chain and distinguishes those covered in the FBS from those outside the FBS.

Industrial use

Industrial use refers to the quantities of the product used during the reference period for non-food purposes such as the production of biofuels, paints, detergents and cosmetics.

Tourist consumption

Tourist consumption (thereafter referred to as *tourist*) comprises the food allocated to non-resident visitors during their stay in the country. This variable covers food availability for all non-residents, including tourists, business travellers and non-resident migrants, when the latter are not counted as part of the country's population. In this case, the term consumption does not represent food consumed as measured by household or other surveys but represents the available supply.

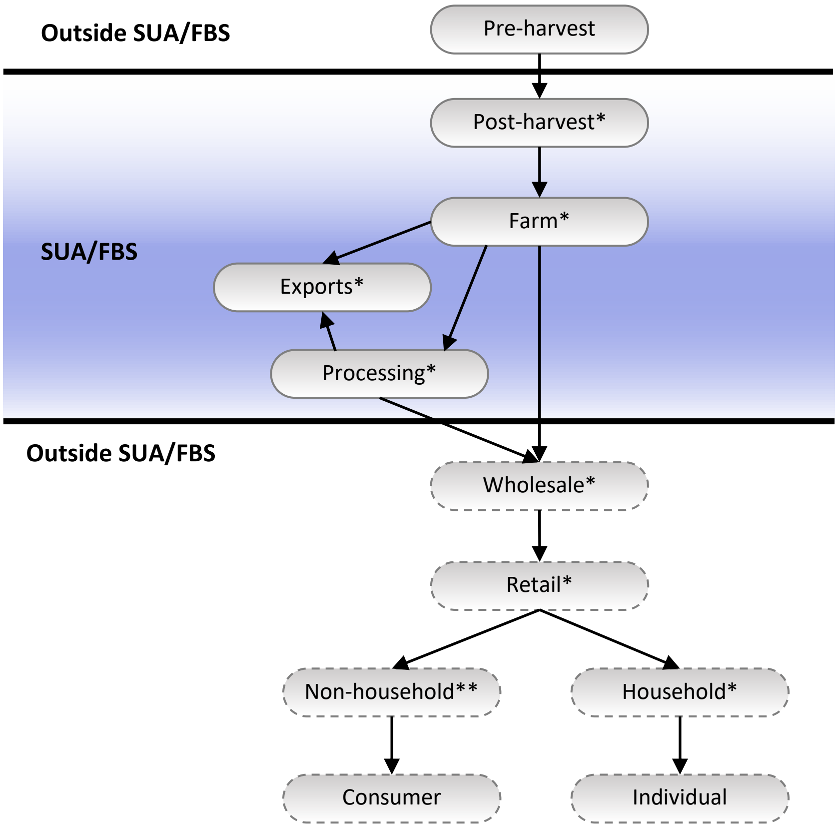
Stocks

Stocks refer to amounts of food allocated to or taken from storage for use at later stages in the food supply chain. This includes stocks held at all levels of the supply chain, from production up to, and including, retail.

Food processing

Food processing refers to manufacturing processes that transform the product used as an input into an edible product, distinct from the original. These processed products may belong to the same food group or may be in a different food group (e.g. beer made from barley). Quantities of food used to manufacture inedible products such as soap or biofuels should be classified under industrial use and not food processing.

Figure 1: Food loss along the value chain



Note: *Storage occurs at all levels, as does wastage resulting from temperature, insect infestation, mold, deterioration, improper transport, handling and shrinkage. ** Non-household include restaurants, hospitals, prisons, schools, armed services and food services.
Source: Authors' own elaboration.

Other residual utilizations

FBS compilers, both at FAO and the national level, should aim to distinctly and comprehensively identify all individual uses, whenever feasible. However, some uses of a given product may not fit completely into the categories mentioned above. In these cases, a category labelled "Other residual utilizations" is used to account for quantities that have not been allocated elsewhere in the FBS methodology. This component is calculated retroactively and functions as a balancing element. This variable is used only when the final imbalance is lower than a small fixed percentage. As specified in Section 14, if the imbalance exceeds this threshold, the SUA remains unbalanced, necessitating further analysis to assess the consistency of the input data.

2.2 Auxiliary variables

To produce additional indicators such as *per capita* use, and to facilitate standardization, additional variables need to be collected by FBS compilers and are listed below.

Extraction rates (further defined in Section 5.3.1) are probably the most important. They are used to estimate the amount of a processed product obtained from a primary product. Extraction rates are typically expressed as a percentage, and are calculated as the amount of derived product that is produced using a given amount of input product. For example, when wheat is the input product, 75 percent is extracted into flour of wheat, the derived product.

Processing shares (further defined in Section **5.3.2**) are a second type of conversion coefficient to estimate the amount of a product used as an input to the processing for a derived product. They are a percentage of the quantity of a given product that is intended for a specific process. They are an essential component of the FBS, both because goods can be processed into a multitude of derived products, and the primary inputs used for creating processed goods is often unclear. For example, wheat may be processed into bread, pasta, cookies and breakfast cereals. At the same time, those products may be derived from multiple inputs. Breakfast cereals alone may be processed from wheat, corn, rice, oats or other grains.

The difference between extraction rate and processing shares is that the processing share gives the share of the product used as an input to be processed into a particular derived product, while the extraction rate shows the proportion of the input that is transformed (during processing) into the derived product. For example, from one tonne of wheat, 80 percent is the processing share (or the amount used as an input) when producing flour of wheat, with the remaining 20 percent used as input for other products. However, only 75 percent of this input is extracted or transformed into flour of wheat during processing as the remaining 25 percent of the wheat kernel goes to bran, waste, etc.

By using processing shares and extraction rates in tandem, FBS compilers can estimate the production of derived goods, even when information is limited. Processing shares are extremely dependent upon the structure of each country's particular product supply chains. As such, there are no international reference sources for data on processing shares.

Population

Population is defined by the United Nations Population Division as the "de facto population in a country, area or region as of 1 July of the year indicated".[6] The word *de facto* is important, because it indicates that citizens as well as all residents should be counted in the population, thus potentially including refugees or resident migrant workers. In addition, the people not counted as part of the population should be considered as *visitors*, so that their food availability can be properly captured under tourist food. Population estimates are used to convert total national nutrient supplies into *per capita* nutrient supplies. United Nations Population Division estimates are derived through a globally consistent methodology.

Agricultural activity and unit production rate

Supplementary information is often needed to impute the variables displayed in the FBS when source data are missing and to enhance accuracy. This information includes agricultural activity indicators such as *area harvested* (the area from which a crop is gathered), *area sown* (the area on which sowing or planting has been carried out, for the crop under consideration, on the soil prepared for that purpose) and the number of live or slaughtered animals for livestock imputation (FAO, n.d.b).

Moreover, unit production rates, which indicate the production quantity per unit, are also used. They include *crop yield* (both the measure of the yield of a crop per unit area of land cultivation, and the seed generation of the plant itself), *slaughtered weight* (the weight of animals slaughtered within national boundaries, irrespective of their origin) and *off-take rates* (the number of animals taken out from the national herd during the year to be slaughtered in the country or exported alive). These variables will be further discussed in Section **4**.

Energy and nutrient estimates

One of the primary motivations for compiling FBS is to derive estimates of food available for human consumption, but also the amounts of energy and key nutrients such as proteins, fats and various micronutrients available for consumption. These estimates are derived from the final food estimates in the balance sheet for each product by

[6] The Glossary of Demographic Terms can be found at https://esa.un.org/unpd/wpp/General/GlossaryDemographicTerms.aspx.

applying nutrient conversion factors to those quantities. As food products are typically eaten in their processed, rather than primary, form, nutrient conversion factors also apply to and are available for derived products. Given the relevance of nutrients estimates and the growing need to provide statistical information on the agricultural and food products at the country level, in terms of weight, but also in terms of quality from a nutritional point of view, an expanded global *nutrient conversion table* (NCT) has been generated to extend the quantity and quality of data on nutrients associated with products. Section **16** provides additional details on this process.

2.3 Imputation, estimation and computation

FBS compilation involves three distinct yet interconnected processes: imputation, estimation and computation.

This handbook refers to *imputation* as a process that addresses gaps or missing values in the data, ensuring a comprehensive representation of the required variables. Imputation maintains the integrity of the FBS data, especially when dealing with incomplete information. On the other hand, *estimation* is used to predict values based on available data. For example, estimation is used for trade data and primary crop production data, for which models are built to infer data from other available information. Finally, *computation* is the practical execution of mathematical and logical operations, facilitating the aggregation and calculation of various indicators, as will be seen in the calculation of quantities of derived products using conversion coefficients (Section **5**). In the context of FBS, these processes contribute to the generation of accurate, comprehensive, and representative food supply and utilization data, essential for informed policymaking and resource allocation.

2.4 Units of measure

After ensuring that the data compiled for the different SUA variables refer to the same products, it is important to harmonize the units in which values are reported, so that the supply–utilization identity of **Equation 1** can be balanced. FBS compiled by FAO report quantities in tonnes, nutritional values in kilocalories (kcal) and grams (g), persons in 1000 persons and DES in kilocalories per capita per day (kcal/capita/day). **Table 1** reports the complete list of measurement units applied in the current FBS methodology.

Table 1: Units of measure in the FBS methodology

COMPONENT NAME	UNIT
Quantity	t
Persons	1000 persons
DES	kcal/capita/day
Energy	kcal
Protein, total	g
Fat, total	g
Carbohydrate, available; calculated by difference	g
Fibre, total dietary	g
Calcium	mg
Iron	mg
Magnesium	mg
Phosphorus	mg
Potassium	mg
Zinc	mg
Vitamin A (expressed in retinol equivalents)	mcg
Vitamin A (expressed in retinol activity equivalents)	mcg
Thiamin (vitamin B1)	mg
Riboflavin (vitamin B2)	mg
Vitamin C	mg

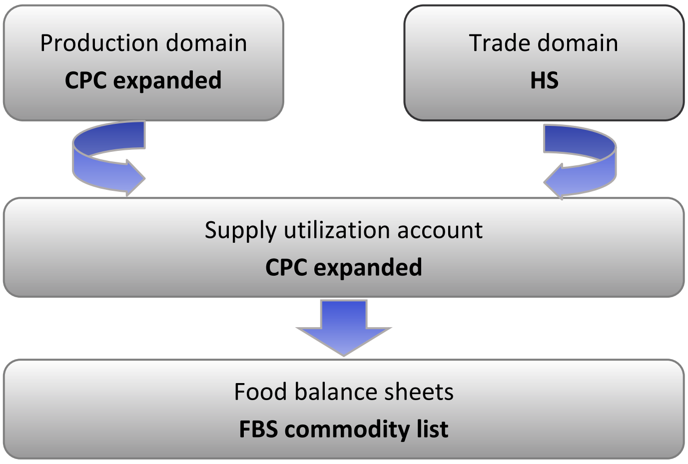
Source: Authors' own elaboration.

2.5 Product coverage and classifications

FBS include all potentially consumable products, irrespective of whether they are actually consumed or used for non-food purposes. Creating a comprehensive list of such products is challenging and involves many conceptual and statistical difficulties. Therefore, for practical reasons, it is necessary to adopt a pragmatic list of standardized statistical concepts that also allows comparability. While countries may develop their own statistical classification systems, several international statistical classifications can be adapted to meet the food and agricultural data requirements of a country.

FAO uses and recommends international classification systems for categorizing agrifood products. As previously mentioned, the current FBS methodology uses the Central Product Classification (CPC version 2.1) (United Nations, 2015). This replaces the previous FAOSTAT commodity list (FCL) (FAO, n.d.). The CPC was expanded from the standard 5-digit code to a maximum of 7 digits, to include all products available in the balance. In addition to the CPC, two other classification systems are used for the FBS. The first is the Harmonized Commodity Description and Coding System (HS), to classify and report trade in goods. Products categorized in the HS classification are then translated into the CPC (through correspondence tables) when creating the SUAs. The second is the FBS commodity list, which is used for reporting and disseminating final FBS data. The classification schemes used in the new framework are shown in **Figure 2**.

Figure 2: Classification scheme used and recommended by FAO for the FBS methodology



Source: Authors' own elaboration.

2.5.1 United Nations Central Product Classification

The CPC provides a coherent and consistent hierarchical structure to classify products (goods and services) based on a set of internationally agreed concepts, definitions, principles and classification rules.

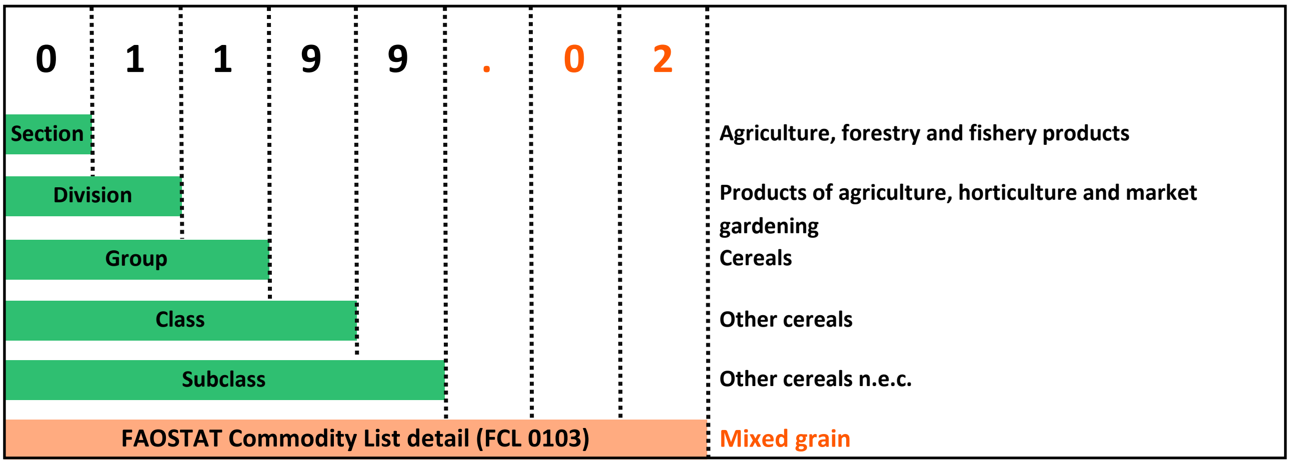
> As stated in the CPC:
> It serves as an international standard for assembling and tabulating all kinds of data requiring product detail, including statistics on industrial production, domestic and foreign commodity trade, international trade in services, balance of payments, consumption and price statistics and other data used within the national accounts (United Nations, 2015, p. iii).
> It provides a comprehensive framework within which data on products can be collected and presented in a format that allows for economic analysis supporting decision-taking and policy-making (United Nations, 2015, p. 3).
> The overall set of products is subdivided into a hierarchical, five-level structure of mutually exclusive

> categories, facilitating data collection, presentation and analysis at detailed levels of the economy in an internationally comparable, standardized way. The categories at the highest level are called sections, which are numerically coded categories. The sections subdivide the entire spectrum of products into broad groupings […]. The classification is then organized into successively more detailed categories, which are numerically coded: two-digit divisions; three-digit groups; four-digit classes; and, at the greatest level of detail, five-digit subclasses (United Nations, 2015, p.3).

A prerequisite that was imperative for the successful incorporation of the CPC into the new methodology involved mapping the CPC to the previously used FCL classification. This was necessary to maintain data comparability over time and to prevent any breaks in the data series. This required an increased level of detail in agricultural, forest and fishery products of CPC to produce the *expanded CPC for agricultural statistics* (CPC Expanded), which has been formally incorporated as an official annex to the CPC version 2.1 for primary products. This expanded form of the CPC provides an enhanced level of granularity for agricultural products, primarily focused on primary commodities. This involves adding an extra level, denoted by two digits, after the standard CPC hierarchy (Ramaschiello, 2015).

For illustrative purposes, **Figure 3** provides a representative example of a CPC description for mixed grain.

Figure 3: CPC expanded code 01199.02 for mixed grain



Source: Authors' own elaboration.

To facilitate the change of classification, back-cast data and to allow full alignment between the FCL and the CPC, the following solutions had to be adopted depending on the type of link (Ramaschiello, 2015):

- One-to-one cases were easily resolved, as old data were transferred to the new classification by assigning codes and definitions according to the new classification, while the data remained the same.
- For many-to-one cases (see example in **Table 2**), data conversion was also straightforward, as data in the FCL were aggregated into the CPC. Such aggregation entailed a loss of information, as the CPC is less detailed than the FCL. To avoid losing information in FAOSTAT, many-to-one cases have been turned into one-to-one correlations.

One-to-many and many-to-many links presented greater challenges. In these cases, data were converted based on the statisticians' best judgement according to the dominant correspondence and following the commodity description for countries that reported trade data using more than 6-digit HS codes. Coefficients of conversion were not calculated, as there was a risk of compromising data quality in the conversion. The conversion keys used are 1 and 0 exclusively:

- One-to-many relations between the FCL and the CPC (see example in **Table 3**) were managed by identifying the dominant correlation based on the statisticians' best judgement, and assigning the conversion key 1.

- In many-to-many cases (see example in **Table 4**), which represent a minority of FCL–CPC correlations, the target classification (CPC version 2.1 expanded) was modified and aligned to the source classification (FCL).

The examples below show the source FCL code, the relevant code in the CPC version 2.1 expanded, and the final mapped code as implemented in FAOSTAT (CPC version 2.1 expanded in FAOSTAT).

Table 2: FCL–CPC many-to-one link example

FCL		CPC version 2.1 expanded		CPC version 2.1 expanded in FAOSTAT	
Code	*Description*	*Code*	*Description*	*Code*	*Description*
0430	Okra	01239	Other fruit-bearing vegetables	01239.01	Okra
0463	Other vegetables fresh n.e.c.			01239.90	Other fruit-bearing vegetables n.e.c.

Source: Adapted from Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Paper presented at the Meeting of the Expert Group on International Statistical Classifications, 19–22 May 2015. New York, United Nations Statistics Division. https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-16.PDF

Table 3: FCL–CPC one-to-many link example

FCL		FCL to CPC conversion factor code	CPC version 2.1 expanded in FAOSTAT	
Code	*Description*		*Code*	*Description*
0577	Dates (fresh+dried)	1	01314 (agriculture)	Dates, fresh
		0	214190.03 (food products)	Dates, dried

Source: Adapted from Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Paper presented at the Meeting of the Expert Group on International Statistical Classifications, 19–22 May 2015. New York, United Nations Statistics Division. https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-16.PDF

Table 4: FCL–CPC many-to-many link example

FCL		CPC version 2.1 expanded		CPC version 2.1 expanded in FAOSTAT	
Code	*Description*	*Code*	*Description*	*Code*	*Description*
0603	Other fruit tropical, n.e.c.	01319	Other tropical and subtropical fruits, n.e.c.	01319	Other tropical fruits, n.e.c. (excluding subtropical fruit)
0619	Other fruits, n.e.c.	01359.90	Other fruits, n.e.c.	01359.90	Other fruit, n.e.c. (including subtropical fruit)

Source: Adapted from Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Paper presented at the Meeting of the Expert Group on International Statistical Classifications, 19–22 May 2015. New York, United Nations Statistics Division. https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-16.PDF

FAO uses the CPC version 2.1 expanded for its production surveys; production data are available using both the CPC and the FCL. This structure is well suited to FBS compilation, as products are aggregated at the primary equivalent level in the commodity trees. It is recommended that countries map from their national codes to the CPC version 2.1 expanded, so that national FBS can be harmonized with the international methodology used by FAO.

An additional benefit is that many statistical concepts and definitions listed in the CPC are based on the HS (see below), which greatly facilitates the comparison of production and trade data in the FBS context.

2.5.2 Harmonized Commodity Description and Coding System

The HS is a classification developed by the World Customs Organization. It is the most widely used classification for international trade, as it is used by over 200 countries and covers 98 percent of international merchandise trade (WCO, n.d.). The classification is updated every five years, and the latest version (HS 2022) entered into force on 1 January 2022. The HS consists of over 5 000 *commodity groups*, which are structured into 21 *sections*, 99 *chapters*, four-digit *headings* and six-digit *subheadings*. The six digits are included in the first 97 chapters and can be broken down into three parts. The HS harmonizes the codification of products in a six-digit scheme (HS-6). The first two digits (HS-2) identify the chapter the goods are classified in (for example 09 = coffee, tea, mate and spices). The next two digits (HS-4) identify groupings within that chapter (09.02 = tea, whether or not flavoured). The next two digits (HS-6) are even more specific (09.02.10 = green tea [not fermented]). Chapters 98 and 99 are for national use only. Most customs administrations, however, use ten or more digits in their commodity coding systems at the tariff line level, with the first six digits being the HS code. This ensures data comparability (given that more than 200 countries already use this classification) and ease of mapping with the CPC. The United Nations Statistics Division offers a correlation table between HS 2017 (up to HS-6 codes) and the CPC version 2.1[7] to streamline data mapping across these classifications.

2.5.3 FBS commodity list

The FBS commodity list (**Table 5**) is a hierarchical classification of the standardized commodities used in the FBS. It is not used when compiling the SUA. Therefore, the FBS has its own commodity code list that ensures the full comparability of data across time and countries. This classification includes primary crops, livestock and fish products that also include derived products after they have been standardized to primary equivalents. The terms in the classification refer to a primary product but also include the wording “& products” because the classification includes both primary products and the derived products converted to primary equivalents.

Table 5: FBS complete commodity classification tree

FBS GROUP	FIRST LEVEL OF AGGERGATION	SECOND LEVEL OF AGGREGATION	TOTAL
WHEAT & PRODUCTS; BARLEY & PRODUCTS; MAIZE & PRODUCTS; RYE & PRODUCTS; OATS & PRODUCTS; MILLET & PRODUCTS; SORGHUM & PRODUCTS; CEREALS, OTHERS & PRODUCTS; RICE & PRODUCTS (MILLED EQUIVALENT)	CEREALS & PRODUCTS EXCLUDING BEER	VEGETAL PRODUCTS	GRAND TOTAL
POTATOES & PRODUCTS; CASSAVA & PRODUCTS; SWEET POTATOES; ROOTS & TUBERS, OTHER & PRODUCTS; YAMS	STARCHY ROOTS & PRODUCTS		
SUGAR CANE; SUGAR BEET	SUGAR CROPS (EXCLUDING PRODUCTS)		
SUGAR, NON-CENTRIFUGAL; SUGAR (RAW EQUIVALENT); SWEETENERS, OTHER; HONEY	SUGAR & SWEETENERS		
BEANS, DRY; PEAS, DRY; PULSES, OTHER & PRODUCTS	PULSES & PRODUCTS		
NUTS & PRODUCTS	TREENUTS & PRODUCTS		
GROUNDNUTS (SHELLED EQUIVALENT); SOYABEANS & PRODUCTS; SUNFLOWERSEED; RAPE & MUSTARDSEED; COTTONSEED; COCONUTS & COPRA; SESAME SEED; PALM KERNELS; OLIVES; OILCROPS, OTHERS	OILCROPS (EXCLUDING PRODUCTS)		
SOYABEAN OIL; GROUNDNUT OIL; SUNFLOWERSEED OIL; RAPE & MUSTARD OIL; COTTONSEED OIL; PALMKERNEL OIL; PALM OIL; COCONUT OIL; SESAMESEED OIL; OLIVE & RESIDUE OIL; RICEBRAN OIL; MAIZE GERM OIL OILCROPS OIL, OTHER	VEGETABLE OILS & PRODUCTS		

[7] See https://unstats.un.org/unsd/classifications/econ. For a more detailed correspondence between HS 2017 (up to HS-10 codes) and the CPC version 2.1 expanded, researchers can refer to https://www.fao.org/statistics/caliper/en.

FBS GROUP	FIRST LEVEL OF AGGERGATION	SECOND LEVEL OF AGGERGATION	TOTAL
TOMATOES & PRODUCTS; ONIONS, DRY; VEGETABLES, OTHER & PRODUCTS	VEGETABLES & PRODUCTS		
ORANGES, TANGERINES, MANDARINES & PRODUCTS; LEMONS, LIMES & PRODUCTS; GRAPEFRUIT & PRODUCTS; CITRUS, OTHER & PRODUCTS; BANANAS; PLANTAINS; APPLES & PRODUCTS; PINEAPPLES & PRODUCTS; DATES; GRAPES & PRODUCTS (EXCLUDING WINE); FRUIT, OTHER & PRODUCTS	FRUITS & PRODUCTS (EXCLUDING WINE)		
COFFEE & PRODUCTS; COCOA BEANS & PRODUCTS; TEA & MATE	STIMULANTS		
PEPPER; PIMENTO; CLOVES; SPICES, OTHER	SPICES		
WINE; BEER; BEVERAGES, FERMENTED; BEVERAGES, ALCOHOLIC; ALCOHOL, NON-FOOD	ALCOHOLIC BEVERAGES		
INFANT FOOD; MISCELLANEOUS	MISCELLANEOUS		
BOVINE MEAT & PRODUCTS; MUTTON & GOAT MEAT & PRODUCTS; PIGMEAT & PRODUCTS; POULTRY MEAT & PRODUCTS; MEAT, OTHER & PRODUCTS	MEAT (SLAUGHTERED) & PRODUCTS		
OFFALS, EDIBLE	OFFALS		
FATS, ANIMALS, RAW; BUTTER, GHEE; CREAM; FISH, BODY OIL; FISH, LIVER OIL	ANIMAL FATS & PRODUCTS		
MILK & PRODUCTS (EXCLUDING BUTTER)	MILK & PRODUCTS (EXCLUDING BUTTER)	ANIMAL PRODUCTS	
EGGS & PRODUCTS	EGGS & PRODUCTS		
FRESHWATER FISH; DEMERSAL FISH; PELAGIC FISH; MARINE FISH, OTHER; CRUSTACEANS; CEPHALOPODS; MOLLUSCS, OTHER; AQUATIC ANIMALS, OTHERS	FISH, SEAFOOD & PRODUCTS		
MEAT, AQUATIC MAMMALS; AQUATIC PLANTS	AQUATIC PRODUCTS, OTHER & PRODUCTS		

Source: Authors' own elaboration.

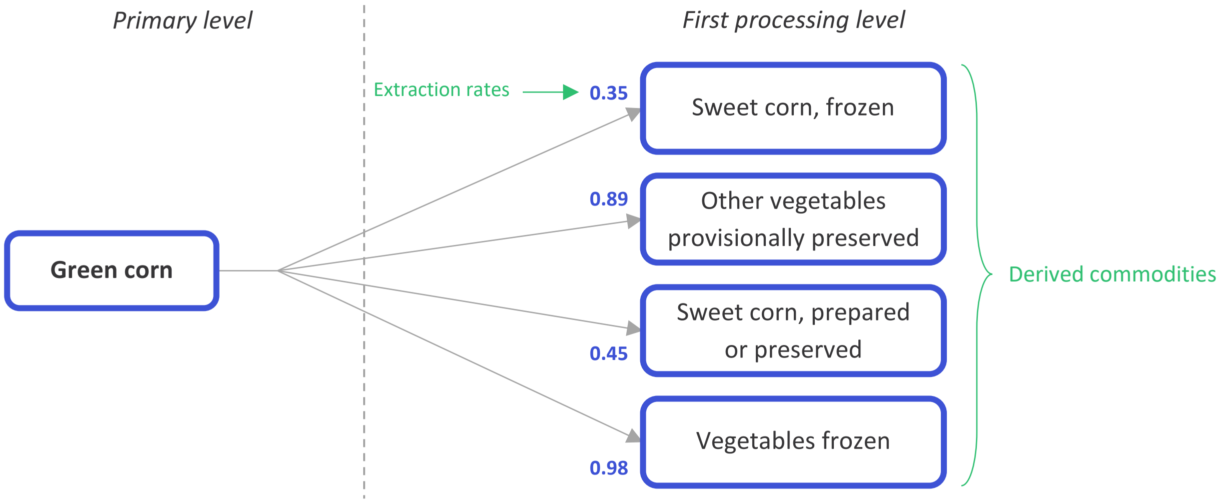
The column 'FBS group' shows the groups usually reported in FBS. These can be aggregated into the first level of aggregation, and further into the second level of aggregation and total. All four levels are reported in FAOSTAT. Yet, most countries choose to publish data only under 'FBS group' without publishing the more aggregated levels.

2.6 Commodity trees

The compilation of SUAs and FBS is based on *commodity trees*. A commodity tree is a "symbolic representation of the flow from a primary commodity to various processed products derived from it, together with the conversion factors from one commodity to another" (FAO, 2001). Commodity trees stem from a primary product and show all potential processing applications related to that product across multiple levels of processed items. Their complexity varies based on the scope of derived products, the depth of processing levels and whether co-products are generated throughout the processing stages.

For a better understanding of this concept, the tree reported in Figure 4 shows that green corn can yield four derived products: sweet corn frozen, other vegetables provisionally preserved, sweet corn prepared or preserved, and vegetables frozen. Each arrow in the figure represents an individual production process, meaning that the primary product allocated to one process does not contribute to any other process. The values in blue indicate the amount of derived product obtainable from an amount of the primary commodity with that production process; these are the extraction rates described in Section 2.2. For instance, an extraction rate of 0.35 implies that for every tonne of sweet corn, 0.35 tonnes of frozen sweet corn is produced.

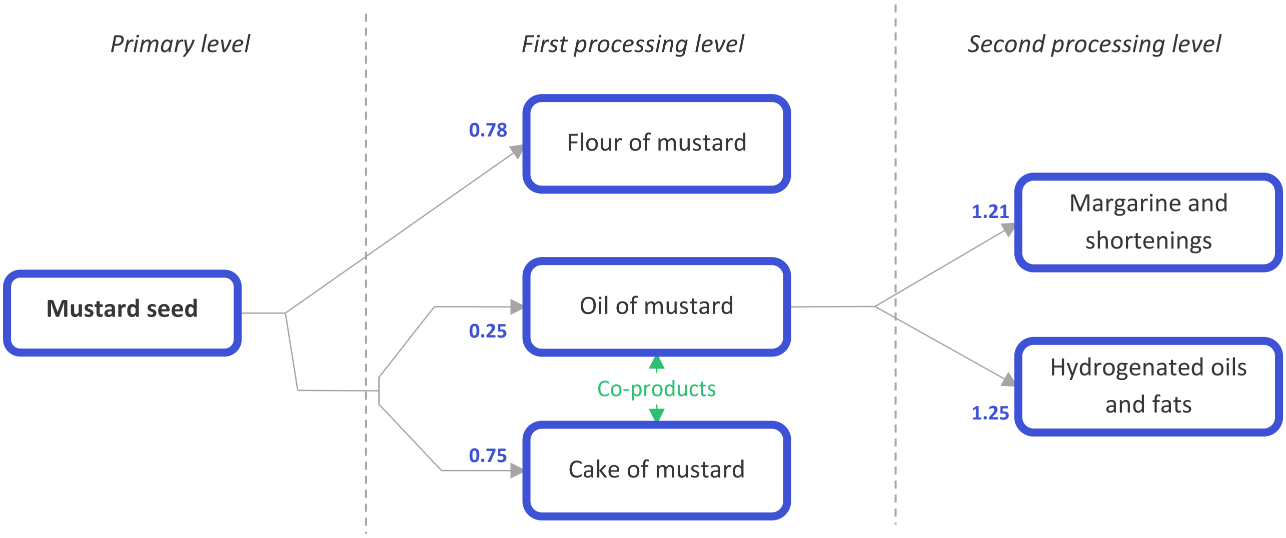
Figure 4: Green corn commodity tree: single product from each production process



Source: Authors' own elaboration.

While this example illustrates a single product from each production process, multiple products often arise from a single process, called co-products. An example of this is depicted in Figure 5.

Figure 5: Mustard seed commodity tree: multiple products and production processes



Source: Authors' own elaboration.

Figure 5 shows the example of mustard seed, which has multiple products and production processes. It is part of two production processes: one yielding flour of mustard, and the other producing both oil of mustard and cake of mustard, which are called co-products. The initial derived products (from the first processing level) are involved in subsequent production processes. For instance, at a second processing level, oil of mustard is a raw material for producing margarine and shortenings; and hydrogenated oils and fats. Section 15 on standardization provides further details about commodity trees and different commodity types.

3 DATA SOURCES AND QUALITY

3.1 Overview

Ideally, the data required for preparing FBS should come from one single source. A country therefore needs to have a comprehensive statistical system that records all information related to each component of the food balance sheet. As this kind of statistical system does not often exist in practice, the first step in FBS compilation is to gather data from a variety of sources.

To collect and improve information on crop and livestock production, regular production surveys using probability-based sampling are recommended. The FAO Statistics Division collects the majority of the statistical information needed for the FBS it produces via an annual crop and livestock production and utilization questionnaire (FAO, n.d.b) presented in Section **3.4**. It is sent annually to FAO Members and relevant counterparts and collects national data referring to the past three years.

The primary data source used by countries to complete the questionnaire is the official statistics on crop and livestock production generated from crop and livestock surveys, or a mix of administrative sources and surveys. Surveys are usually conducted by National Statistical Offices (NSOs) or Ministries of Agriculture. However, the NSO is considered to be the coordinator of the national statistical system in countries and should confirm as official statistics the data produced by line ministries. The Global Strategy on Agricultural and Rural Statistics (GSARS)[8] has produced a variety of publications providing guidance on methods for data collection on crop and livestock production.

For other data required for SUA/FBS purposes, FAO promotes the use of official statistics from the national statistical system as the central source (with data often available from the NSO or Ministry of Agriculture). However, if needed, unofficial data can also be consulted from specialized commodity international institutions, intergovernmental institutions or research papers.

When reliable data sources are unavailable, model-based imputation of missing data is an alternative, keeping in mind that the quality of imputed data depends on the quality of the source data. Separate imputation approaches are recommended for each variable in the basic identity.

All the possible sources of data mentioned above can be organized into a hierarchy of three categories: official, unofficial and imputed. These categories are considered sequentially in the compilation of FBS and are further described in this Section **3.2**. The FAO flag system for denoting data quality is presented in Section **3.3**.

3.2 Hierarchy of data sources

Official

Countries are encouraged to prioritize the use of their official statistics, as FAO does and recommends when compiling FBS. Official statistics are data produced by the national statistical system that have been confirmed as official by the coordinator of the national statistical system. Given that FAO is an intergovernmental organization with a core mandate to collect data from its Members, information received officially from them takes precedence over data from other sources. FAO considers data to be official when it is received through its questionnaires. The primacy of information obtained through official channels reflects FAO's commitment to respecting the authoritative status of data provided by its Members. Moreover, such data are most likely to have been compiled

[8] See https://www.fao.org/in-action/global-strategy-agricultural-statistics/en.

according to statistically sound methodologies. However, there are countries in which more than one institution produces the same information, resulting in conflicting figures. In this case, data differences are analysed, considering primarily the information provided by statistical authorities, notably NSOs, to explore the methodological reasons that may explain the discrepancies and determine which figures are the most appropriate for compiling the FBS.

Data is also used from Eurostat, or the United Nations Statistics Division Comtrade database. In these particular cases, even though the data is taken from an international organization, it is flagged as official by FAO as the primary source is official data from countries.

Unofficial

Where official data are not available – which is often the case for variables describing different types of utilization – alternative sources are considered. Unofficial data may be obtained from a variety of sources: some are international institutions specializing in specific markets, such as the International Sugar Organization;[9] others may be intergovernmental organizations or research outlets, producing journal articles or journals, such as Oil World.[10]

Other available sources may be administrative data, yearbooks or monthly bulletins (considered unofficial if they have not been confirmed as official statistics), industrial output surveys, FAO sources such as the Global Information and Early Warning System (GIEWS), or United Nations reports of missions to countries. In some cases, records of private firms and commodity organizations might also be helpful, which may be accessible through tax authorities or an agreement with an industry or commodity organization.

Imputation

In the absence of empirically collected or measured data, missing information can be imputed using a variety of statistical methods. This is a last resort, as no imputation method, no matter how sophisticated, can replace measurement; the widely diverse imputation methods can produce a broad range of different results, which can be considered appropriate under different circumstances. Sound imputation methods can also be applied to triangulate measured information, thus contributing to validate data or discover inconsistencies.

While computing FBS, as in all complex statistical processes, it is essential to apply quality checks on the input data. These include a range of operations such as vetting and adjusting units of measurement, removing outliers, identifying potential transcription errors and filling obvious gaps. A set of principles for improving data quality is laid out in the *FAO Statistics and Data Quality Assurance Framework* (FAO, 2023b).

3.3 Application of data quality flags

Incorporating data from various sources is quite prevalent in FBS compilation. However, not all the FBS data can be deemed of equivalent quality. In 2016, a unified and standardized system of codes (also named *observation status flags*) was established to increase the accessibility and clarity of data disseminated externally by FAO by providing a clear indication of the source of the figures (or missing figures) and their quality to potential users of FAO statistics, including international organizations that may use FAO data as input for deriving other statistical indicators. This system of flags was updated in 2023 (FAO, 2023c), and countries may wish to adopt a similar system of flags when compiling their FBS.

[9] See https://www.isosugar.org/.
[10] See https://www.oilworld.biz/.

Table 6: Observation status code list for secondary data (including missing values) disseminated by FAO

FLAG	NAME	DESCRIPTION
A	Official value	Value provided as official when the source agency assigns sufficient confidence that it is not expected to be dramatically revised.
B	Time series break	Observations are characterized as such when different content exists or a different methodology has been applied to this observation as compared with the preceding one (the one given for the previous period).
E	Estimated value	Observation obtained through an estimation methodology or based on the use of a limited amount of data (e.g. to produce a value at an early stage of the production stage while not all data are available). If needed, additional information can be provided through free text using the COMMENT_OBS attribute at the observation level or at a higher level (in SDMX-compliant environment). This code is also to be used when the estimation is done by a sender agency (and flagged as such). When the imputation is carried out by a receiver agency in order to replace or fill gaps in reported data series, the flag to use is I “Value imputed by a receiving agency”
F	Forecast value	Value deemed to assess the magnitude which a quantity will assume at some future point of time (as distinct from “estimated value” which attempts to assess the magnitude of an already existent quantity).
G	Experimental value	Data collected on the basis of definitions or (alternative) collection methods under development. Data not of guaranteed quality as normally expected from provider.
I	Value imputed by a receiving agency	Observation imputed by a receiving agency to replace or fill gaps in reported data series. This code is intended to cover all cases where a receiving agency publishes data about a sending agency that do not come from an official source in the sender agency’s reporting framework. When the estimation is done by the sender agency, the flag to use is E “Estimated value”
L	Missing value; data exist but were not collected	Used, for example, when some data are not reported/disseminated because they are below a certain threshold.
M	Missing value; data cannot exist	Used to denote empty cells resulting from the impossibility to collect a statistical value (e.g. a particular education level or type of institution may be not applicable to a given country’s education system).
N	Not significant	Used to indicate a value which is not a “real” zero (e.g. a result of 0.0004 rounded to zero).
O	Missing value	This code is to be used when the reasons why data are missing cannot be determined.
P	Provisional value	An observation is characterized as “provisional” when the source agency – while it bases its calculations on its standard production methodology – considers that the data, almost certainly, are expected to be revised.
Q	Missing value; suppressed	Used, for example, when data are suppressed due to statistical confidentiality considerations.
S	Strike and other special events	Special circumstances (e.g. strike) affecting the observation or causing a missing value.
U	Low reliability	This indicates existing observations, but for which the user should also be aware of the low quality assigned.
V	Unvalidated value	Observation as received from the respondent without further evaluation of data quality.
X	Value from international/mandated organization	Observation from an international or a supranational organization that does not use any flagging system in data sharing

Source: FAO. 2023. *Observation Status Code List – Version 3*. Statistical Standard Series. Rome. https://openknowledge.fao.org/handle/20.500.14283/cc6208en

Observation status flags denote an "attribute of a cell in a dataset representing qualitative information on the value of that cell" (FAO, 2023c). They are not applied to *primary microdata*, because these data are described in the metadata associated with them. They are applied to all *secondary data*, i.e. statistical outputs typically derived by national statistical agencies based on primary data to analyse some social or economic phenomena. The flag assigned to a data point affects the way that data are treated, in terms of tolerance intervals, estimations and imputation procedures for missing data. The most recent version of observation flags used in FBS compilation are listed in **Table 6**. These flags update the 2016 standard to account for the specific characteristics of the FAO integrated statistical system.

3.4 Crop and livestock production and utilization questionnaire

The crop and livestock production and utilization questionnaire (FAO, n.d.b) introduced in Section **3.1** is composed of a cover where details of national reporting contacts are required, and six sections for information on the following areas: *primary crop production*, *primary crop utilization*, *livestock* (animal numbers and livestock production), selected *derived agricultural commodities*, utilization of *oils* and *metadata*. Instructions, product and variables descriptions are also included. From 2018, commodities are reported in the CPC version 2.1 expanded.

FAO sends the questionnaire each year to Members, which are encouraged to return and complete it as fully as possible (even when its data are available on the website of the NSO or line ministries). This ensures that the data are provided in the format that can be fully used by FAO, and can be considered official, as agreed by the NSO as coordinator of the national statistical system.

4 PRIMARY CROP AND LIVESTOCK PRODUCTION

$\boldsymbol{production} + imports - \Delta stocks = exports + food + food\ processing + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$

4.1 Overview

The first element to be considered in the basic identity is the production of primary crop and livestock products, which is located on the supply side of the equation. Primary products together with the derived products (Section 5) comprise the $production$ variable.

4.2 Data imputation

After a quality assurance/quality control has been undertaken, the outliers removed and data gaps filled with information available from alternative sources, the remaining gaps in the production data may be filled using imputation procedures that rely on available time series data. The current methodology imputes production by combining robust econometric models with the expertise and knowledge of FBS compilers at the country level. This imputation technique considers all the information and knowledge derived from previous time series data and integrates the previous imputation methods used by other institutions. An in-depth examination of existing imputation methods revealed that no individual approach or model effectively produces satisfactory results across all countries and commodities.

Hence, a methodology based on the use of multiple models was developed, which relies on two key components: a shared theoretical framework based on the ensemble learning approach, and the concept of a “triplet” – three parameters linked by a functional relationship.

Ensemble learning makes it possible to leverage the strengths of different algorithms and construct a prediction model that surpasses individual procedures.

The triplet comprises the target variable ($output$) and two auxiliary variables, named $input$ and $productivity$, which guarantees that the values imputed by the ensemble approach align with the underlying relationships between variables.

The triplet provides the relationship among the variables, while the ensemble approach provides the method for imputing one or more of the variables in the triplet (and the relationship can then be used to estimate one or more of the missing variables).

The following subsections describe the approach that FAO uses and recommends for estimating the triplets and ensemble methods, which can also be followed by countries.

4.2.1 Triplets

The first key method in the overall process of imputing production data is the *triplet*. When estimating production for each product in the agricultural production domain, FAO produces three elements (three time series) linked by the following relationship:

Equation 3: Production triplet

$$output_t = input_t \times unit\ production\ rate_t$$

This is the basic equation to compute the missing production quantities, and to validate the output of the imputation since a priori information on feasible production rates and inputs associated with productive processes is generally available. When two of the three variables are available, the third can be computed as an identity, i.e. without statistical models. As this is not often the case, the ensemble approach needs to be used for estimation of the variables.

Equation 4 represents the triplet used when estimating the production of crops:

Equation 4: Production triplet for crop production imputation

$$production = area\ harvested \times yield$$

where:

- $production$ is the total quantity (t) of crops harvested or produced within a specific area during a given period;
- $area\ harvested$ is the total area (ha) from which a crop or crops have been collected or are intended for harvest in the previous period by the farmer; and
- $yield$ is the quantity of a crop harvested per unit of $area\ harvested$ (t/ha).

The *livestock production* imputation process consists of a set of six elements linked via the two triplets expressed in Equation 5 and Equation 6 for animal and meat estimation:

Equation 5: Production triplet for animals

$$slaughtered\ animals = livestock\ number \times off\text{-}take\ rate$$

Equation 6: Production triplet for meat

$$production = slaughtered\ animals \times carcass\ weight$$

where:

- $slaughtered\ animals$ is the number (heads) of animal slaughtered;
- $livestock\ number$ is the number of live animals expressed in heads;
- $off\text{-}take\ rate$ is a conversion factor representing the number of animals taken out from the national herd during the year to be slaughtered in the country or exported alive, divided by the total number of the herd for a given reference period; and
- $carcass\ weight$ identifies the quantity of meat, expressed in tonnes, that can be obtained from slaughtering one animal.

The triplet for egg production is reported in weight (usually tonnes) and is shown in Equation 7:

Equation 7: Production triplet for eggs

$$egg\ production = laying\ hens \times yield$$

where:

- $egg\ production$ refers to the quantity of eggs, in weight, produced during the year by all laying hens, whether in the subsistence or commercial sectors. It includes hatching eggs and eggs wasted at the farm;
- $laying\ hens$ refers only to the females kept primarily for egg production and data are shown in thousand units; and
- $yield$ refers to the weight of eggs laid by one laying hen during the year.

Egg production expressed in weight can be converted into egg production expressed in numbers by dividing the production weight by the average weight of one egg. The average weight of one egg is country-specific (likewise, the number of eggs may need to be converted to weight in missing years).

The triplet for milk production is reported in **Equation 8**:

Equation 8: Production triplet for milk

$$milk\ production\ = milking\ animals \times yield$$

where:

- $milk\ production$ indicates the quantity, in tonnes, of milk milked during the reference period from the milking animals of the considered species;
- $milking\ animals$ gives the number in heads of animals that, in the course of the reference period, have been milked (as opposed to the total number of lactating animals). The concept of milk production excludes the milk used for feeding young animals. The number of milking animals is lower than the number of reproducing females in the livestock herd account.
- $yield$ represents the average quantity of milk produced by a milking animal during the year and is reported in kg per head. It is generally not recorded but obtained by dividing milk production by the number of milking animals.

4.2.2 Ensemble learning

Values of the variables in each triplet are estimated via the “ensemble”. *Ensemble learning* refers to the process of building a collection of simple base methods that are later combined to obtain a composite prediction (Zhou, 2009). The technique consists of building multiple *learners* and then combining learner models to obtain final values. As shown by Dietterich (2000) the ensemble method reduces the risk of adopting an inaccurate value, as it averages multiple methods, which minimizes the potential for suboptimal imputations in specific subsets of data. Moreover, model selection is unnecessary, since all models are included in the final ensemble. Finally, from an implementation point of view, the algorithm is adaptive and does not require constant updating. For example, if the data-generating mechanism changes in the future, the next fit of the ensemble approach will shift weights to models that better represent the data. Accordingly, it will not be necessary to constantly monitor and manually update the methodologies/models.

In the production domain, the selection of methods is informed by the intention to capture the structural variations in data across different countries and commodities. Hence, the models display diverse levels of generality, with some geared towards capturing local variations and others aimed at representing global trends. This is obtained by properly constraining or expanding the training dataset. For instance, if a particular country or product has abundant data, a model will be developed by relying solely on data for this country or product. However, in cases where the available observations are scarce for a given country or product, global trends will be used to model the values in the country–product pair.

The model set comprises nine models, varying in complexity: *mean, linear, exponential, logistic, naive, autoregressive integrated moving average (ARIMA), locally estimated scatterplot smoothing (LOESS), splines, and multivariate adaptive regression splines (MARS)*. Missing observations are imputed using a weighted mean of the estimates of the ensemble learners for the missing time period.

During estimation, each model is associated with an extrapolation parameter, which controls the extent to which the model can be applied beyond the range of the data. Learners that exhibit poor extrapolation are omitted. Additionally, cross-validation is applied to evaluate the model’s capacity to explain the data.

4.3 Imputation of primary crop production

4.3.1 Overview

The imputation of primary crop production starts from the estimation of yields. Once this variable is estimated, area harvested and production are calculated to solve the triplet.

4.3.2 Imputation steps

Estimation of yield

Estimation of $yield$ is made using **Equation 4** but with $yield$ as the subject:

Equation 9: Production yield formula

$$yield = \frac{area\ harvested}{production}$$

where:

- $yield$ is expressed in tonnes per hectare (t/ha);
- $production$ is expressed in tonnes (t); and
- $area\ harvested$ is expressed in hectares.

If all three variables are available, any two variables can be used to cross-check the third variable. If the formula indicates an error for one of the given variables, a time series check should identify which variable in the triplet is not correct, for example, if the yield is higher than biologically feasible, the area harvested and production can be examined to see which one is too high or too low. If only two variables are available, the third is computed with **Equation 9**.

Estimation of production and area harvested

If data for only one of $production$ or $area\ harvested$ are available, $yield$ is imputed using its historical time series (using the ensemble approach described in Section **4.2.2**). The other missing element would then be calculated using the formula mentioned above.

If only the $yield$ variable is available, $production$ is imputed using the historical time series (using the ensemble approach); then using the formula as above, the area harvested would be calculated.

If all three variables are missing, $yield$ and $production$ data are imputed using the historical time series. The $area\ harvested$ would then be calculated using **Equation 4**.

4.3.3 Primary crop production example: wheat

As an example, in a given country in a given year, the official data reported in **Table 7** are available and the compiler needs to impute production.

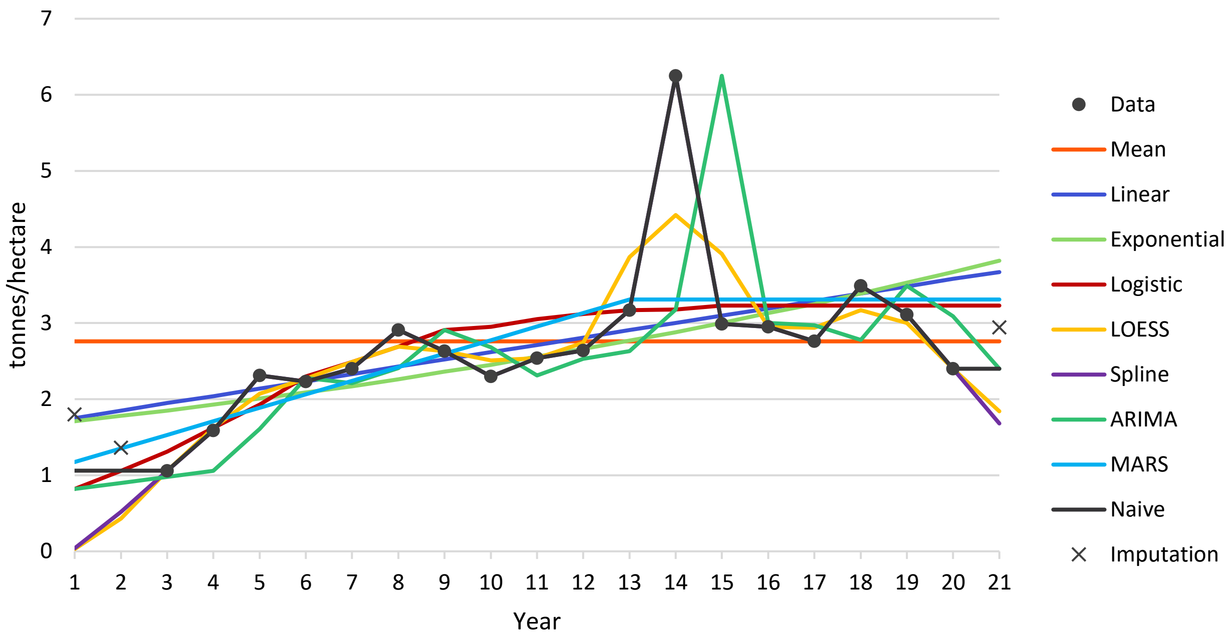
Table 7: Wheat production imputation example (part 1)

CPC CODE AND NAME	AREA HARVESTED (hectare)	YIELD (tonnes/hectare)	PRODUCTION (tonnes)
0111 - Wheat	18 500 000	-	-

Source: Authors’ own elaboration.

In this case $production$ and $yield$ are both missing for wheat. As indicated in Section 4.2, first the $yield$ value is imputed using the ensemble learning approach. In Figure 6 models are represented as lines and historical data as points). These models are combined in a weighted average (where the weights are determined by the ensemble learning algorithm based on how well the model fits the data) to form a final ensemble of models. This ensemble is used to predict the $yield$ value in the current year.

Figure 6: Ensemble modelling results for yield estimates



Source: Authors' own elaboration.

The final imputed value for $yield$ in the most recent year (shown in the graph above as the last "x" representing the ensemble imputation result) is 2.94 t/ha. This is a reasonable estimate compared to the historical time series. Some models fit the data well (such as the logistic regression, spline and ARIMA). Some of these models perform poorly and are omitted (in this case, the LOESS model), but by averaging together well-performing models, the approach produces a reasonable final estimate that is entered in Table 8:

Table 8: Wheat production imputation example (part 2)

CPC CODE AND NAME	AREA HARVESTED (hectare)	YIELD (tonnes/hectare)	PRODUCTION (tonnes)
0111 - Wheat	18 500 000	2.9422	-

Source: Authors' own elaboration.

With $yield$ estimated, there are enough data to compute $production$ for wheat using the triplet of Equation 4:

$$production = area\ harvested \times yield = \ 18\ 500\ 000 \times 2.9422 = 54\ 430\ 700$$

Table 9 shows the complete triplet:

Table 9: Wheat production imputation example (part 3)

CPC CODE AND NAME	AREA HARVESTED (hectare)	YIELD (tonnes/hectare)	PRODUCTION (tonnes)
0111 - Wheat	18 500 000	2.9422	54 430 700

Source: Authors' own elaboration.

4.4 Imputation of livestock production data

4.4.1 Overview

The livestock production imputation process involves six variables and aims to impute the meat and non-meat products produced from each animal. The process uses the animal triplet and the production of meat triplet as defined in **Equation 5** and **Equation 6**. The two functions have a common parameter, $slaughtered\ animals$, which is the output in the animal triplet (**Equation 5**) and the input in the production triplet (**Equation 6**).

4.4.2 Items involved

The livestock production imputation process is carried out for the 17 animal items reported in the CPC.

Table 10: Livestock items

CPC CODE	ITEM NAME
02111	Cattle
02112	Buffalo
02121.01	Camels
02121.02	Other camelids
02122	Sheep
02123	Goats
02131	Horses
02132	Asses
02133	Mules and hinnies
02140	Swine / pigs
02151	Chickens
02152	Turkeys
02153	Geese
02154	Ducks
02191	Rabbits and hares
02192.01	Other rodents
02194	Other birds

Source: Authors' own elaboration.

Several meat and non-meat items are derived from each animal. **Table 11** reports the items obtained from slaughtered cattle.

Table 11: Processed items derived from cattle

CPC CODE	ITEM NAME
21111.01	Meat of cattle with the bone, fresh or chilled
21151	Edible offal of cattle, fresh, chilled or frozen
21512	Cattle fat, unrendered
02951.01	Raw hides and skins of cattle

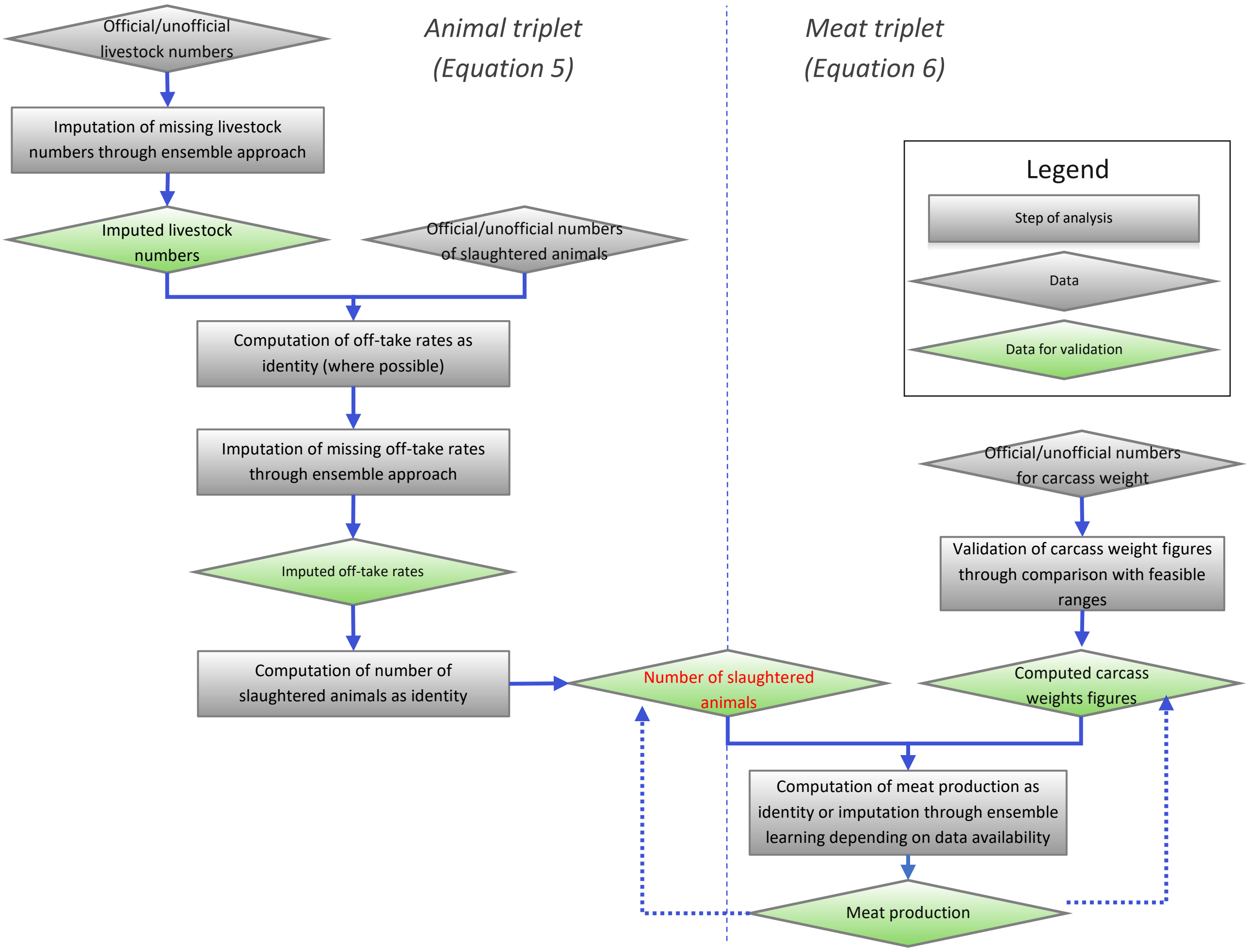
Source: Authors' own elaboration.

The non-meat items, such as fats, offals and skins, are considered as by-products and are less important than primary products like meat, as they undergo no further processing. For large animals such as cattle, these by-products are considered as primary products, and their values are calculated by applying conversion factors to the carcass weight. In contrast, for smaller animals such as poultry, the by-products are considered as "other processed commodities" in the FBS methodology, which does not use the imputation methodologies found in the triplets of **Equation 5** and **Equation 6**. Further details will be provided in Section **5**.

4.4.3 Workflow

The imputation process should be driven by the reliability of available information and it is highly recommended to start by imputing missing data for $livestock\ number$, as this is an input for the estimations of the production of items (such as meat, hides and skin) resulting from the slaughtering of animals. The complete workflow is reported in **Figure 7** while detail of the process follows for each of the triplets involved.

Figure 7: Livestock production imputation workflow



Source: Authors' own elaboration.

4.4.4 Animal triplet

Imputation of missing livestock number through ensemble approach

The livestock number is estimated first, using the ensemble approach because data regarding livestock are normally available and reliable, which provides for almost all countries a solid basis to produce imputations.

Computation of off-take rates as identity and imputation of missing off-take rates using ensemble approach

The off-take rate is not collected in surveys, and is therefore computed using the relationship shown in the triplet in **Equation 5** (it is defined as the number of animals taken out from the national herd during the year to be slaughtered in the country or exported alive, divided by the total number of the herd for a given reference period) is connected to the natural reproduction cycle of each animal (for example, chickens reach slaughtering age in 5–6 months, while cows need more than one year).

The resulting time series often contains many years of missing data that need to be imputed using the ensemble approach.

Computation of number of slaughtered animals as identity

Once both the livestock number and off-take rate have been imputed, the number of missing slaughtered animals can be computed using the triplet of **Equation 5**. This represents the output of the animal triplet workflow and will be the input of meat item triplet.

4.4.5 Meat triplet

Number of slaughtered animals

This value represents the output of **Equation 5** and is estimated using the methodology described in Section **4.4.4.** Further imputation is not needed at this stage.

Imputation and computation of the carcass weight

The variable $carcass\ weight$ represents the unit production rate in the meat item triplet of **Equation 5**. At an initial stage, missing values are imputed using the triplet with available official or unofficial data on slaughtered animals and production. The results are checked against a priori external data on feasible ranges for the imputed figures. FAO uses ranges based on data across all countries: if some data points lie outside the reported ranges, the values are replaced by the minimum or the maximum admitted value.

Missing data that cannot be imputed using external data because more than one value in the triplet is missing, would be imputed using the ensemble approach.

Imputation and computation of meat production

Meat production can be imputed through the ensemble approach or using the triplet depending on data availability: if more than one variable in the triplet is missing, $production$ is imputed using the ensemble approach simultaneously with $carcass\ weight$; if official and unofficial figures are available for $carcass\ weight$ and $slaughtered\ animals$, $production$ is computed using the triplet (**Equation 6**).

In very few cases, some carcass weight figures may be changed during the previous step; the triplet might then need to be rebalanced through an iterative procedure so that the figures for all items are readjusted, including the numbers of slaughtered animals (dashed arrows in **Figure 7**).

4.5 Eggs and milk

4.5.1 Eggs

Table 12 shows the correspondence between animals and their eggs.

Table 12: Animals and eggs item correspondence

ANIMAL CPC CODE	ANIMAL NAME	EGG CPC CODE	EGG ITEM
02151	Chickens	0231	Hen eggs in shell, fresh
02194	Other birds	0232	Eggs from other birds in shell, fresh, n.e.c.

Source: Authors' own elaboration.

While the match between chickens and hen eggs is straightforward, the generic animal item "other birds" cannot be matched to "eggs from other birds in shell, fresh, n.e.c." as it is not completely exhaustive. This is because the animal category "other birds" includes eggs from many types of birds (such as duck, geese and quails); therefore the number of *laying* animals starting from a specific series of livestock animals cannot be estimated. The only option is to use the time series associated to the item "other birds".

The key aspect of the imputation method adopted for eggs is the relationship between $livestock\ number$ and the number of laying hens (or other birds). The dependency between these numbers and $livestock\ number$ is captured through a hierarchical linear model (as opposed to the ensemble approach used for the other production items) with a single dependent variable ($laying\ hens$ or $other\ birds$) due to the need to have at least one explanatory variable, rather than a model where the time series is imputed autoregressively. In the model defined in **Equation 10**, **Equation 11** and **Equation 12**, $time$ is measured in years and is included to capture linear trends over time.

The model is defined on the following three levels:

Level 1: Observational model (response variable)

Equation 10: Hierarchical model for laying hens/other birds estimation – level 1

$$log\left(laying\ hens_{i,j,k}\right) = \beta_0 + \beta_1 time + \beta_{2,j,k}\left(log(livestock\ number)\right)_i + \varepsilon_{i,j,k}$$
$$log\left(other\ birds_{i,j,k}\right) = \beta_0 + \beta_1 time + \beta_{2,j,k}\left(log(livestock\ number)\right)_i + \varepsilon_{i,j,k}$$

where:

- $log\left(laying\ hens_{i,j,k}\right)$ and $log\left(other\ birds_{i,j,k}\right)$ are the outcome variables and represent the log of the number of laying hens or other birds for country i, CPC group j and the unique country–code combination k;
- β_0 is the overall intercept;
- β_1 is the fixed effect of $time$ on the outcome;
- $\beta_{2,j,k}$ is a coefficient that varies at levels 2 and 3 by CPC group j and the specific country–code combinations k, representing the effect of the log of $livestock\ number$;
- $livestock\ number$ is the average estimated number of poultry livestock or other birds' livestock of country i; and
- $\varepsilon_{i,j,k}$is the residual error term at the observation level.

Level 2: Random coefficient model for $\beta_{2,j,k}$

The coefficient $\beta_{2,j,k}$, which represents the effect of $livestock\ number\ i$ within each CPC group j and specific country–code combination k, is modelled as:

Equation 11: Hierarchical model for laying hens/other birds estimation – level 2

$$\beta_{2,j,k} = \gamma_0 + \gamma_{1,j,k}(CPC{:}\,Country)_{j,k} + \delta_{j,k}$$

where:

- γ_0 is the average effect of $log(livestock\ number)$ across all CPC groups j and specific country–code combinations k;
- $\gamma_{1,j,k}$ is a coefficient that varies by group (j,k) and models the influence of the $(CPC{:}\,Country)_{j,k}$ predictor on $\beta_{2,j,k}$; and
- $\delta_{j,k}$ is a random error term at level 2.

This level allows the $livestock\ number$ effect $\beta_{2,j,k}$ to vary by both CPC group and specific country–code combinations, influenced by country-level CPC characteristics.

Level 3: Hierarchical structure for $\gamma_{1,j,k}$

The coefficient $\gamma_{1,j,k}$, which represents the influence of the $CPC{:}\,Country$ variable on the $livestock\ number$ effect is modelled as:

Equation 12: Hierarchical model for laying hens/other birds estimation – level 3

$$\gamma_{1,j,k} = k_0 + k_1(CPC)_j + \phi_j$$

where:

- k_0 is the overall mean effect of CPC group at the country level;
- k_1 captures how the predictor $(CPC)_j$ influences the slope $\gamma_{21,j}$;
- ϕ_j is a random effect for each CPC group j, capturing group-specific deviations from the overall mean.

The model estimates the number of laying animals proportionally to the livestock numbers. The model also accounts for changes over time and differences among countries; the latter are captured by $livestock\ number$ in the country. If data for a particular country–code combination are sparse, then γ coefficients will likely be estimated close to 0. Thus, β coefficients will all have similar values close to their mean, and the model will account only for availability within CPC groups. However, if data are available for a country or commodity, the estimates of k will be greater than 0, and the model can reflect the individual characteristics of a single country.

4.5.2 Milk

Table 13 reports the correspondence between animals and their milk.

Table 13: Animals and milk item correspondence

ANIMAL CPC CODE	ANIMAL NAME	MILK CPC CODE	MILK NAME
02111	Cattle	02211	Raw milk of cattle
02112	Buffalo	02212	Raw milk of buffalo
02122	Sheep	02291	Raw milk of sheep
02123	Goats	02292	Raw milk of goats
02121.01	Camels	02293	Raw milk of camel

Source: Authors' own elaboration.

The key aspect of the imputation method adopted for milk is the relationship between $livestock\ number$ and the number of $milking\ animals$. The dependency between $milk\ production$ and $milking\ animals$ is captured by a hierarchical linear model, following the same approach mentioned for eggs: this choice depends on the need to have at least one explanatory variable instead of building a model (using the ensemble approach) where the time series is imputed by only looking at past data. The equations and variables for estimating egg and milk production are the same, with the exception that $milking\ animals$ replace $laying\ hens$. The model estimates the number of milking animals proportionally to the livestock numbers.

5 PRODUCTION OF PROCESSED CROP AND LIVESTOCK GOODS

$$\mathbf{production} + imports - \Delta stocks = exports + food + \mathbf{food\ processing} + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$$

5.1 Overview

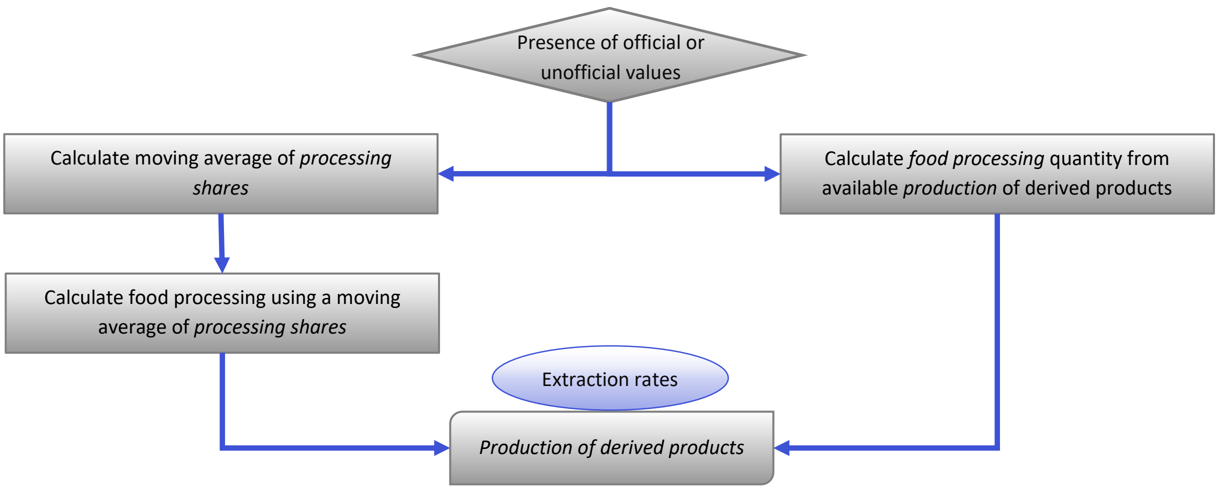
This chapter covers the food processing variable of **Equation 1** and the *production* of derived products computed using this variable. The production of primary products (Section **4**) and derived products together make up the $production$ variable.

In the FBS methodology, processed products, or *derived products* are assumed to be those originating from a physical transformation of primary commodities. As described in Section **2.6** on commodity trees, more than one derived product can be obtained from each primary commodity and more than one level of processing may be involved. Furthermore, edible processed products can be involved in production processes that produce a second level of derived products and so on. Countries often do not report data on the production of processed products, therefore computations in the FBS methodology are based on the availability of primary commodities data and conversion coefficients.

As these computations require data on each primary commodity available, this step is performed after all other variables of the equation have been imputed or estimated. However, this Section is placed after Section **4** on the imputation of production of primary commodities to put together all imputation procedures connected to production. The overall FBS example presented in Section **17** shows the calculation of processed goods in its correct position, after all other variables have been imputed.

The computation process for the production of derived commodities has two main steps (**Figure 8**): first the quantity of primary commodities to be processed (*food processing*) is calculated and then, the *production* of derived products is computed by applying *extraction rates* applied to the *food processing* quantities. This process can differ depending on the presence or absence of official or unofficial data of production of derived products. When no official or unofficial data are available, an additional step is required, in which *processing shares* (the percentage of a given commodity used as input in the production process of a derived product) are estimated and used to calculate the *food processing* variable.

Figure 8: Main steps in the production of processed products



Source: Authors' own elaboration.

5.2 Imputation workflow for processed products.

5.2.1 Imputation of food processing

As introduced in Section **2.1**, $food\ processing$ represents the quantity of a parent product to be used as input in the production process of derived products. Food processing data may come from the FAO production questionnaire or be collected from unofficial sources such as industrial output surveys. When no official or unofficial data are available (either official or unofficial), the amount of food processing for a commodity needs to be imputed. Conversely, the imputation of food processing can be fairly simple when data on the production of derived goods exist.

No official or unofficial data available

When there are no official or unofficial available data on production of derived products, the $food\ processing$ variable is calculated using the historical ratio of the processed quantity to the quantity of the available product ($availability$ in **Equation 13**). This ratio is projected using an average estimation over a number of years. In the second step, the $food\ processing$ variable to be computed for the following years is calculated by multiplying the $availability$ variable by the $processing\ share$.

Equation 13: Food processing – version 1

$$food\ processing_P = availability_P \times processing\ share_P$$

where:

- $availability_P$ consists of production plus net trade and captures the transformation of raw agricultural products into consumable goods through processing, considering both what is domestically produced and what is obtained through trade before processing;
- $processing\ share_P$ represents the percentage of a given commodity that is used as input in the process for making a derived product (see Section **5.3.2**).

Official or unofficial production available for one or more child products

If official or unofficial data for one or more production quantities of the derived products exist, the $food\ processing$ variable is calculated using the official data for the quantity produced for the derived product and the corresponding bottom-up shares (defined in Section **5.3.4**), which represent the percentage of a good (primary or derived) used as an input in a transformation process. For example, for flour of millet (child) and millet (parent), a $bottom\text{-}up\ share_{C-P}$ of 1 indicates that all the production of flour of millet comes from millet; an extraction rate eR_{C-P} of 0.85 indicates that 85 percent of the millet sent for producing flour of millet, is actually transformed to flour, while the remaining 15 percent is either lost during the transformation process or transformed into other co-products of the flour.

Equation 14: Food processing – version 2

$$food\ processing_P = \sum_C \frac{production_C}{eR_{C-P}} \times bottom\text{-}up\ share_{C-P}$$

where:

- $production_C$ is the production of derived, or child, product C;
- eR_{C-P} is the extraction rate implied in transformation of the parent product P to the child C; it indicates the share of the parent product that is transformed into the child product; and
- $bottom\text{-}up\ share_{C-P}$ is the part of the production of a derived product C coming from a parent product P.

5.2.2 Imputation of derived production

Once the $food\ processing$ variable for each parent is calculated, the production of child products can be calculated (**Equation 15**). The production for those derived products is the multiplication of (1) the quantity of the parent product used in processing of other products ($food\ processing$), (2) the top-down shares (defined in Section **5.3.6**) between the derived product and the parent, and (3) the extraction rate of the child from the parent. If a child has many parents, the calculation below is performed for each parent and then summed. For example, for millet (parent) and bran of millet (child), an eR_{C-P} of 0.14 indicates that 0.14 tonne of bran of millet is obtained from each tonne of millet, and a $top\text{-}down\ share_{C-P}$ of 1 indicates that all of the millet going to $food\ processing$, goes to bran of millet.

Equation 15: Production of derived products imputation formula

$$production_C = \sum_P food\ processing_P \times eR_{C-P} \times top\text{-}down\ share_{C-P}$$

where:

- $production_C$ is the production of derived product C;
- eR_{C-P} is the extraction rate implied in transformation of the parent product P to the child C; it indicates the share of the parent that is transformed to the child product; and
- $top\text{-}down\ share_{C-P}$ is the proportion of the $food\ processing$ of a parent commodity P that goes into the production process of a specific child product C.

This calculation is performed for all production levels in all commodity trees of each country-year combination.

5.2.3 Derived product imputation example

Consider the following SUA for the primary commodity millet in a given country and a given year.

Table 14: SUA for millet

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Millet	204 990	460	0	167 977	20 830	8 290	4 100	-	/	4 100	/

Source: Authors' own elaboration.

And suppose the following data to be available for the millet commodity tree:

Table 15: Millet values for the calculation of the production of derived products

PARENT	CHILD	EXTRACTION RATE	BOTTOM-UP SHARE	TOP-DOWN SHARE	PRODUCTION OF CHILD
Millet	Flour of millet	0.85	1.000	1.000	680
Millet	Bran of millet	0.14	1.000	1.000	-
Millet	Beer of millet	0.00	-		-

Source: Authors' own elaboration.

The country does not produce beer of millet. Flour and bran are co-products so both have a top-down share of 1 and official data are available for production of flour of millet.

The variable $food\ processing$ has to be calculated first by applying **Equation 14**:

$$food\ processing_{millet} = \frac{production_{flour\ of\ millet}}{eR_{flour\ of\ millet-millet}} \times bottom\text{-}up\ share_{flour\ of\ millet-millet} = \frac{680}{0.85} \times 1 = 800$$

$production$ of bran of millet uses the $food\ processing$ variable, which has provided the amount of millet used as input for all its derived products. This uses **Equation 15**:

$$production_{bran\ of\ millet} = food\ processing_{millet} \times eR_{millet-bran\ of\ millet} \times top\text{-}down\ share_{bran\ of\ millet-millet} = 800 \times 0.14 \times 1 = 112$$

Thus, out of the available millet (204 990 − 167 977 + 460), 800 units are sent for food processing. The production quantity of the derived bran of millet is calculated with a top-down share of 1, so all of the 800 units sent for food processing are sent to be transformed into bran of millet, but only 14 percent of that millet can be extracted to bran of millet.

The SUAs resulting for these products are reported in **Table 16**:

Table 16: Production of millet-derived products after calculations

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Millet	204 990	460	0	167 977	20 830	8 290	4 100	800	/	4 100	/
Flour of millet	680	/	/	/	/	/	/	/	-	/	/
Bran of millet	112	/	/	/	/	/	/	/	-	/	/

Source: Authors' own elaboration.

5.3 Conversion coefficients

The computation of the production of processed products is done with conversion coefficients, which connect parent products (P) to derived products that are also referred to as child products (C), namely: *extraction rates*, *processing shares*, *bottom-up shares* and *top-down shares*. The inclusion of such shares is required in FBS compilation because commodities are often processed into a multitude of derived goods, and the input used when producing these items is rarely known.

5.3.1 Extraction rates

As mentioned in Sections **1.3** and **2.2**, the extraction rate (eR_{C-P}) reflects the quantity of output obtained in the processing of one product into another. For instance, the average global extraction rate of wheat flour is about 0.79. This means that 1 tonne of wheat that goes through a country's flour milling industry renders on average 790 kg of flour. Flour milling also produces certain co-products: for each tonne of wheat milled, on average 180 kg of bran and 20 kg of wheat germ are produced, and the remaining 10 kg are losses that occur in the milling process.

5.3.2 Processing shares

As discussed in Section **2**, *processing shares* represent the percentage of a given commodity that is used as input in the production process of a derived product. Processing shares are obtained as a moving average of the ratio (reported in **Equation 16**) between the amount of the variable $food\ processing$ of the product P and its $availability$ in a given year:

Equation 16: Processing share

$$processing\ share_P = \frac{food\ processing_P}{availability_P}$$

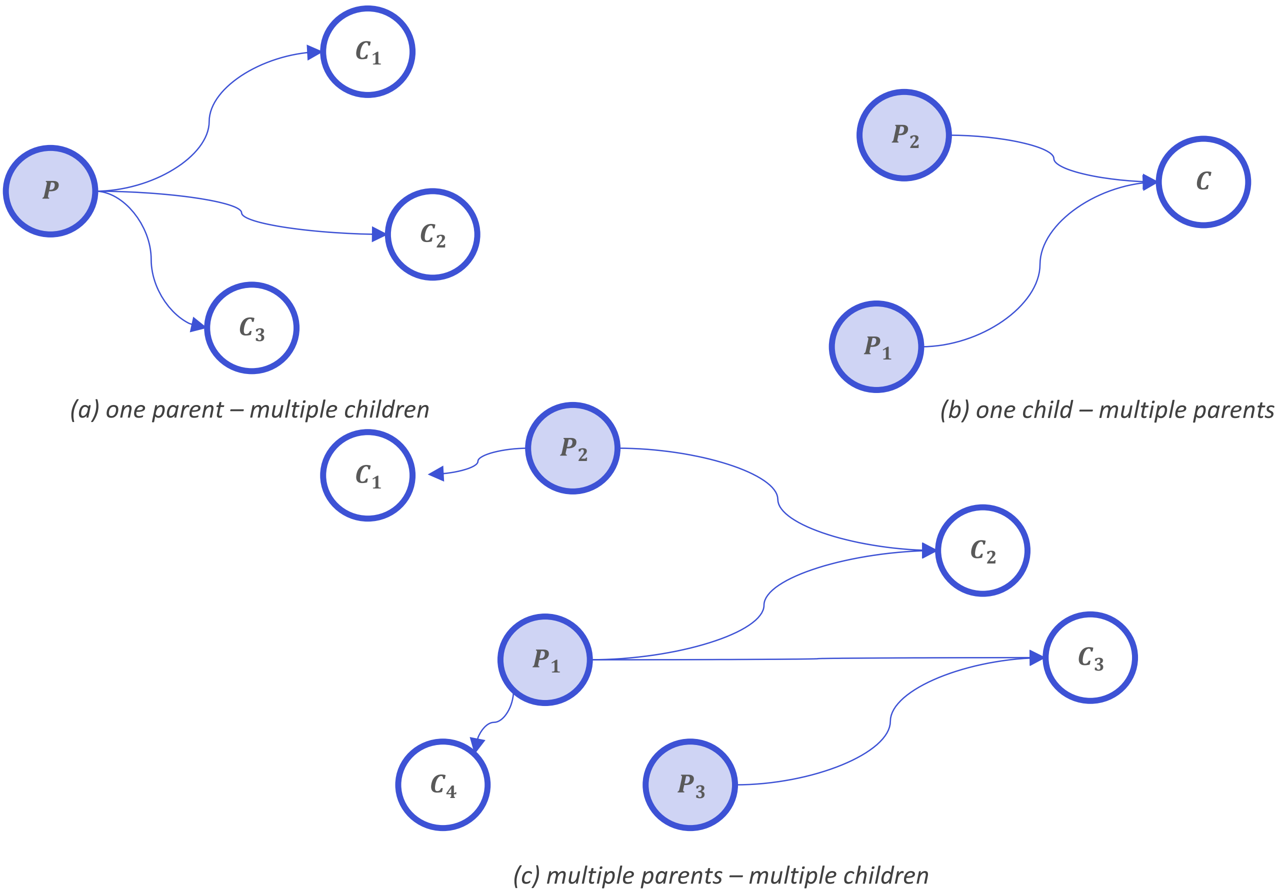
5.3.3 Bottom-up shares: the mirror imputation loop

Bottom-up shares represent the quantity of processed products produced from a parent commodity. In other words, a bottom-up share is the part of the production of a derived product – which is also referred to as a child product (C) – that comes from a parent commodity (P).

Each bottom-up share accounts for all the child commodities that can be produced from a parent and all parents that can be potentially processed into each child product. For a better understanding of this concept, consider the following relations between parents and children that might exist in a commodity tree:

- One parent – multiple children (**Figure 9a**);
- Multiple parents – one child (**Figure 9b**); and
- Multiple parents – multiple children (**Figure 9c**).

Figure 9: Possible relations between parent (P) and child (C) products



(a) one parent – multiple children

(b) one child – multiple parents

(c) multiple parents – multiple children

Source: Authors' own elaboration.

Bottom-up shares reflect all the existing parent and child relations in all commodity trees. They are calculated every year for every country based on previous values of food processing and production values of parent commodities. The bottom-up shares are computed using the workflow shown in **Figure 10** and described below.

To calculate bottom-up shares, first, all the quantities of variables $food\ processing$ (for parent commodities) and $production$ (for child products) are computed from existing available values in a SUA, starting from the one-to-many parent–child relations represented in **Figure 9a** and **b**. Then, a proportional approach is used for the production of parent commodities that remains unassigned.

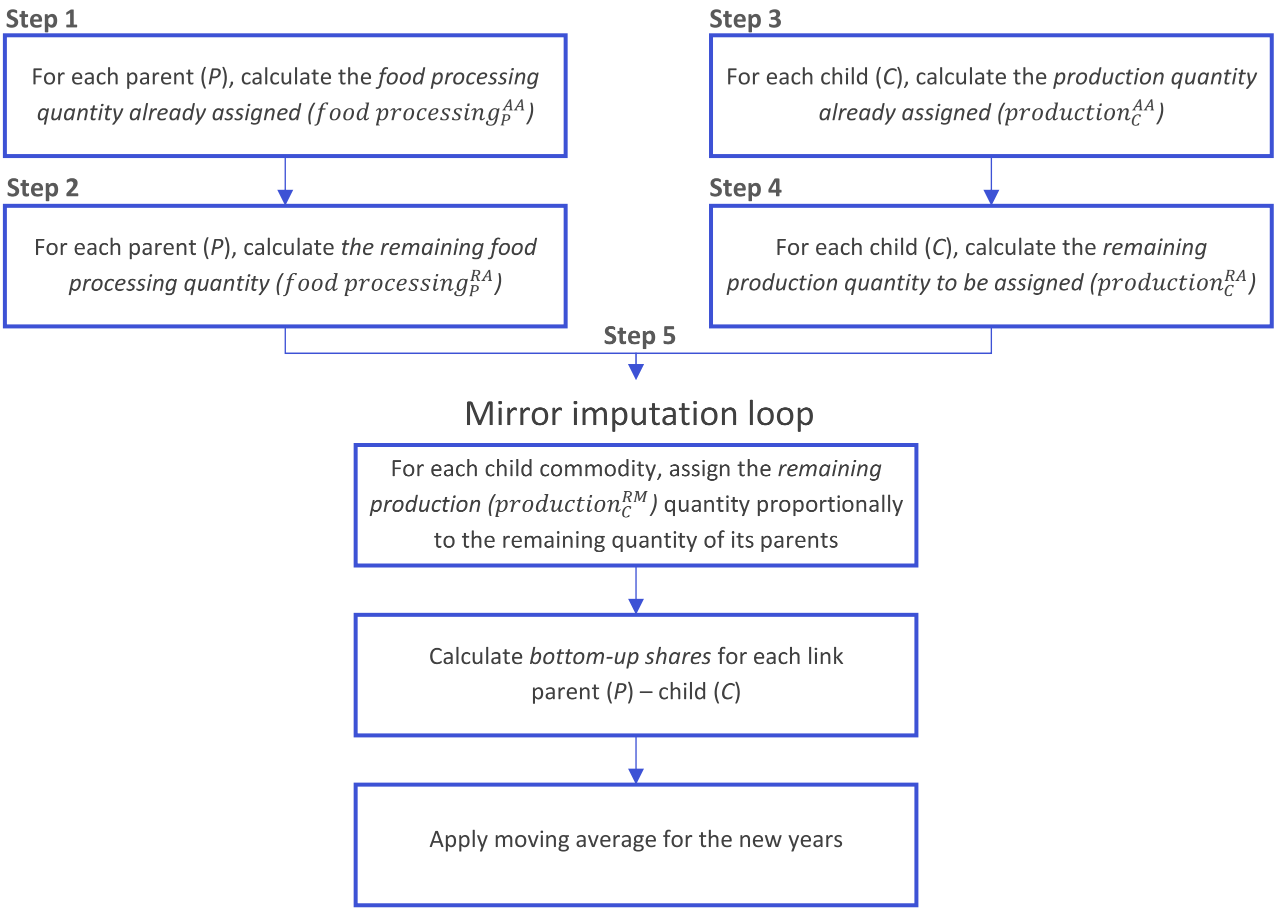
In the process, for each derived product C, several quantities are involved. The previously mentioned food processing p refers to the total amount of the parent sent as inputs to all derived products and production c refers to the total quantity of the child product. The quantities below are subsets of these totals split by their destination and origin:

- *food processing quantity already assigned* ($food\ processing_P^{AA}$). Food processing of the parent P of derived product C – i.e. the amount of P sent as input to the production process of derived product C. As

C might be not the only child produced from that parent, a second food processing has to be considered, which is the next element.

- *remaining food processing quantity* ($food\ processing_P^{RA}$). This is the amount of the variable $food\ processing$ not involved in the production process of derived product C, but in the processes of other derived products that might be obtained from the same parent P.
- *production quantity already assigned* ($production_C^{AA}$). The production quantity of derived product C that is produced from a parent commodity P, when P is the only parent of C and no other derived commodities are obtained from P. This amount is considered "already assigned", meaning that it is deterministic.
- *remaining production quantity to be assigned* ($production_C^{RA}$ and $production_C^{RM}$). These values reflect the production of the derived product C in scenarios such as the one represented in Figure 9 c where the child has multiple parents and its parents have multiple children. $production_C^{RA}$ is assigned through a procedure known as the mirror imputation loop, while $production_C^{RM}$ is the amount left after the mirror imputation loop. This is assigned through a proportional approach.

Figure 10: Bottom-up shares calculation flow



Source: Authors' own elaboration.

5.3.3.1 Step 1

For each parent, the *food processing quantity already assigned* ($food\ processing_P^{AA}$) is computed. This is the sum of the ratio of the production of children with only one parent ($production_{C^1}$) and their extraction-rates (eR_{C-P}):

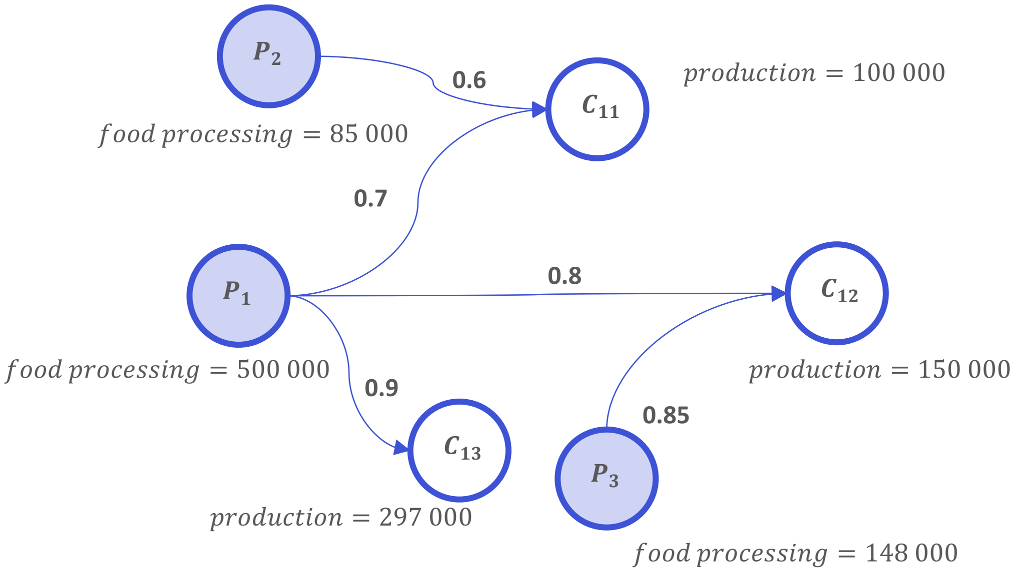
Equation 17: Food processing quantity already assigned

$$food\ processing_P^{AA} = \sum_C \frac{production_{C^1}}{eR_{C-P}}$$

If all child products of a given parent have only one parent, the quantity already assigned corresponds to the entire food processing quantity of that parent, i.e., food processing of parent P. If this is not the case and derived products have more than one parent, a second typology of food processing is computed for each parent (see Step 2).

Consider the scenario reported in **Figure 11**. A parent commodity P_1has three derived products C_{11}, C_{12} and C_{13} with C_{11} and C_{12} having multiple parents. Besides P_1, child C_{11} has another parent P_2, which has only one child (C_{11}), and C_{12} has another parent P_3. The extraction rates of C_{11}, C_{12} and C_{13} for the parent P_1 are respectively 0.7, 0.8 and 0.9. The extraction rate of C_{11} for parent P_2 is 0.6 and the extraction rate of C_{12} for parent P_3 is 0.85.

Figure 11: Example of derived production



Source: Authors' own elaboration.

Food processing quantity already assigned for C_{13} can be calculated, as this child only has one parent, P_1:

$$food\ processing_{P_1}^{AA} = \frac{production_{C_{13}}}{eR_{C_{13}-P_1}} = \frac{297\ 000}{0.9} = 330\ 000$$

5.3.3.2 Step 2

The *remaining food processing quantity* ($food\ processing_P^{RA}$) is calculated by subtracting the food processing already calculated in the previous step ($food\ processing_P^{AA}$) from the total food processing quantity of the parent ($food\ processing_P$). This will be equal to zero for the parents of children with only one parent.

Equation 18: Remaining food processing quantity

$$food\ processing_P^{RA} = food\ processing_P - food\ processing_P^{AA}$$

The following calculates the *remaining food processing quantity* ($food\ processing_P^{RA}$) to be assigned for parent P_1 because it has a child with one parent (C_{13}):

$$food\ processing_{P_1}^{RA} = food\ processing_{P_1} - \frac{production_{C_{13}}}{eR_{C_{13}-P_1}} = 500\ 000 - \frac{297\ 000}{0.9} = 170\ 000$$

5.3.3.3 Step 3

The *production quantity already assigned* ($production_C^{AA}$) is calculated for each child product. This is the sum of the product of food processing of the parent with only one child ($food\ processing_{P^1}$) and the extraction rate of a parent with multiple children (eR_{C-P}).

Equation 19: Production quantity already assigned

$$production_C^{AA} = \sum_P (food\ processing\ _{P^1} \times eR_{C-P})$$

In the example $production_C^{AA}$ for C_{11} can be obtained as:

$$production_{C_{11}}^{AA} = processed_{P_2} \times eR_{C_{11}-P_1} = 85\ 000 \times 0.6 = 51\ 000$$

5.3.3.4 Step 4

Finally, for each child product, the *remaining production quantity to be assigned* ($production_C^{RA}$) is calculated by subtracting the *production quantity already assigned* ($production_C^{AA}$) from the total production $production_C$.

Equation 20: Remaining production quantity to be assigned

$$production_C^{RA} = production_C - production_C^{AA}$$

In the example, the remaining production quantity for child C_{11} is obtained as follows:

$$production_{C_{11}}^{RA} = production_{C_{11}} - processed_{P_2} \times eR_{C_{11}-P_1} = 100\ 000 \text{ - } 85\ 000 \times 0.6 = 49\ 000$$

5.3.3.5 Step 5

The mirror imputation loop is used to iteratively determine the following quantities: (i) the *food processing* quantity of a parent with multiple children and (ii) the part of the production of a parent in one of its multiple children.

For each child product, assign the *remaining production* ($production_C^{RM}$) proportionally to the remaining $food\ processing$ quantity of its remaining parents ($food\ processing\ _P^{RA}$).

Equation 21: Remaining production

$$production_C^{RM} = production_C \times \frac{food\ processing\ _P^{RA} \times eR_{C-P}}{\sum_P food\ processing\ _P^{RA} \times eR_{C-P}}$$

Calculate the bottom-up shares for each parent–child connection as the ratio between the production of each child product from a parent and the total production of the child.

Equation 22: Bottom-up share

$$bottom\text{-}up\ shares_{C-P} = \frac{production_{C-P}}{\sum_C production_{C-P}}$$

When missing, these shares are imputed using a moving average.

In the example mirror imputations are used to calculate true values for $production_{C_{12}-P_1}$ and $production_{C_{12}-P_3}$.

The mirror imputation in this case starts by identifying whether a child product with multiple parents has only one parent for which production need to be assigned. If yes, the remaining production will be assigned to that parent.

In the example above, it is the case for the child C_{11} with multiple parents: the parent P_1 is the only one left. Therefore:

$$production_{C_{11}-P_1} = production_{C_{11}}^{RA} = 49\ 000$$

Having assigned this production, the remaining $food\ processing$ quantity of the parent P_1 will be updated by subtracting the equivalent quantity of the remaining production of child C_{11}.

$$food\ processing_{P_1}^{updated} = food\ processing_{P_1}^{RA} - \frac{production_{C_{11}}^{RA}}{eR_{C_{11}-P_1}} = 170\ 000 - \frac{49\ 000}{0.7} = 100\ 000$$

Subsequently, the mirror imputation loop checks if a given child has only one of his multiple parents left for which $food\ processing$ needs to be assigned. If this is the case, the updated or remaining $food\ processing$ of the parent will be assigned to that child with multiple parents; and hence the part of that parent in the production of that child with multiple parents can be calculated exactly. In the example below, it is the case of the parent P_1. After the update of the remaining $food\ processing$ of the parent P_1, the child C_{12} with multiple parents is the only one for which $food\ processing$ need to be assigned. In this case, the part of the parent P_1 in the production of the child product C_{12} with multiple parents is:

$$production_{C_{12}-P_1} = food\ processing\ _{P_1}^{updated} \times eR_{C_{12}-P_1} = 100\ 000 \times 0.8 = 80\ 000$$

After the production is assigned, the remaining production of the child C_{12} with multiple parents will be updated.

$$production_{C_{12}}^{updated} = production_{C_{12}} - production_{C_{12}-P_1} = 150\ 000 - 80\ 000 = 70\ 000$$

The first search algorithm described above can be applied again to search if a child product with multiple parents has only one parent left. This will be true for the child product C_{12} with multiple parents, which has P_3 as the only parent left. Therefore, the remaining production of C_{12} (70 000) will be assigned to P_3 and $food\ processing$ of P_3 can be updated before reapplying the second search algorithm. This is called the mirror imputation loop because it can be run many times in order to find all the possible mirror calculations.

After the mirror imputation loop is finished, there may be production remaining of some child products with multiple parents that still needs to be assigned. Those quantities will be assigned to the corresponding parents proportionally to their remaining $food\ processing$ quantity.

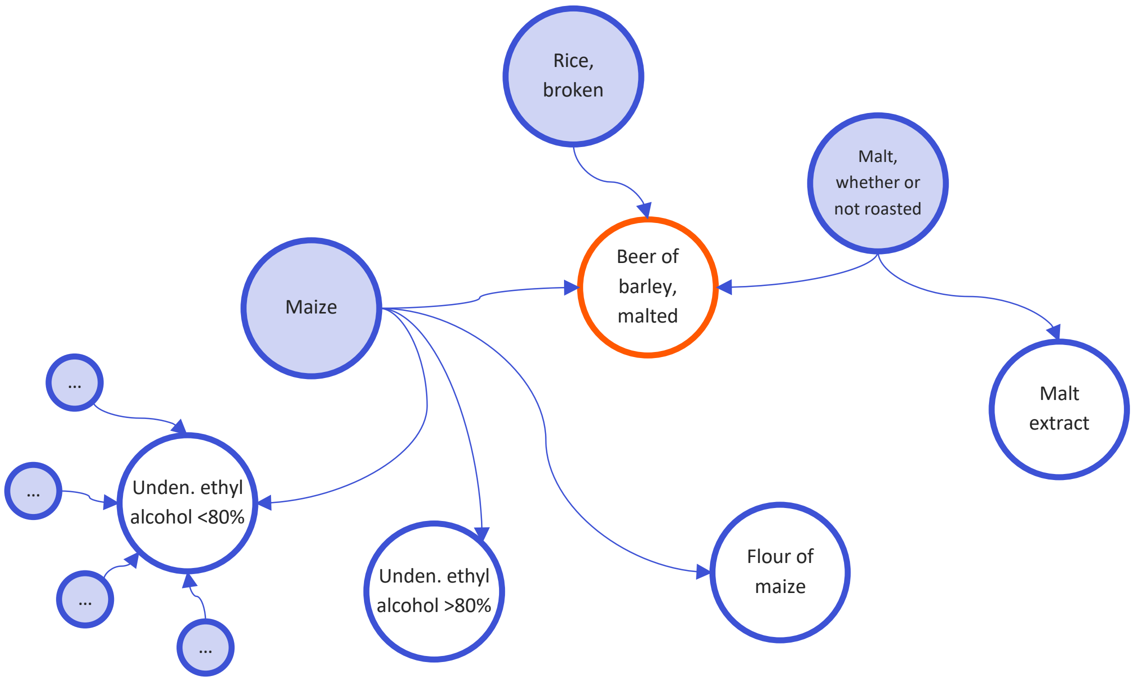
$$production_{C}^{RM} = production_{C} \times \frac{food\ processing\ _{P}^{RA} \times eR_{C-P}}{\sum_{P} food\ processing\ _{P}^{RA} \times eR_{C-P}}$$

After this allocation, both remaining $food\ processing$ quantities of parent products and production of child products are equal to zero.

5.3.4 Bottom-up shares example

This example shows the computation of bottom-up shares for the beer of barley, malted and its different parents: maize; malt, whether or not roasted; and rice, broken (**Figure 12**). Maize has three other derived products: flour of maize, undenatured ethyl alcohol (greater than 80 percent) and undenatured ethyl alcohol (less than 80 percent) (which is a multiple parent commodity). Rice, broken has a unique child product (beer of barley malted) and malt, whether or not roasted has another child product (malt extract).

Figure 12: Beer of barley, malted example graph



Source: Authors' own elaboration.

In the table below, the different child products of parents maize, rice broken and malt whether or not roasted are presented. From Figure 12 the corresponding number of parents can be seen: the child products have one parent, apart from beer of barley, which has three parents (maize; rice broken and malt, whether or not roasted).

Table 17: Maize and its child products

PARENT	CHILD	eR	NUMBER OF PARENTS	PRODUCTION OF CHILD
Maize	Flour of maize	0.8000	1	24 493 436
Maize	Undenatured ethyl alcohol (> 80%)	0.2500	1	197 420
Maize	Undenatured ethyl alcohol (< 80%)	0.3681	5	1 725 000
Maize	Beer of barley, malted	8.0425	3	22 527 000

Source: Authors' own elaboration.

Table 18: Malt, whether or not roasted and its child products

PARENT	CHILD	eR	NUMBER OF PARENTS	PRODUCTION OF CHILD
Malt, whether or not roasted	Malt extract	0.8000	1	50 000
Malt, whether or not roasted	Beer of barley, malted	8.0425	3	22 527 000

Source: Authors' own elaboration.

Table 19: Rice, broken and its child commodities

PARENT	CHILD	eR	NUMBER OF PARENTS	PRODUCTION OF CHILD
Rice, broken	Beer of barley, malted	8.0425	3	22 527 000

Source: Authors' own elaboration.

Table 20: Beer of barley, malted and its parent products

PARENT	CHILD	eR	PROCESSED QUANTITY OF PARENTS
Maize	Beer of barley, malted	8.0425	34 212 558
Rice, broken	Beer of barley, malted	8.0425	488 364
Malt, whether or not roasted	Beer of barley, malted	8.0425	2 237 667

Source: Authors' own elaboration.

For the calculation of bottom-up shares, the *remaining food processing quantity* ($food\ processing\ _{P}^{RA}$) is calculated first, to be eventually assigned for maize and malt, whether or not roasted using **Equation 17** and **Equation 18** (Steps 1 and 2 of Section **5.3.3**).

Table 21: Remaining food processing quantities to be assigned in the example of beer of barley, malted

PARENT	PROCESSED	ALREADY ASSIGNED	PROCESS TO ASSIGN
Maize	34 212 558	31 406 475	2 806 083
Malt, whether or not roasted	2 237 667	62 500	2 175 167

Source: Authors' own elaboration.

where:

$$\begin{aligned} food\ processing^{RA}_{Maize} &= food\ processing_{Maize} - food\ processing^{AA}_{Maize} \\ &= food\ processing_{Maize} - \left(\frac{production_{Flour}}{eR_{Flour-Maize}} + \frac{production_{alc>80}}{eR_{alc>80-Maize}}\right) \\ &= 34\ 212\ 558 - \left(\frac{24\ 493\ 436}{0.8} + \frac{197\ 420}{0.25}\right) = 34\ 212\ 558 - (30\ 616\ 795 + 789\ 680) \\ &= 34\ 212\ 558 - 31\ 406\ 475 = 2\ 806\ 083 \end{aligned}$$

$$\begin{aligned} food\ processing^{RA}_{Malt} &= food\ processing_{Malt} - food\ processing^{AA}_{Malt} \\ &= food\ processing_{Malt} - \frac{production_{MaltExt}}{eR_{MaltExt-Malt}} = 2\ 237\ 667 - \frac{50\ 000}{0.8} = 2\ 237\ 667 - 62\ 500 \\ &= 2\ 175\ 167 \end{aligned}$$

The preassigned values are assigned to flour of maize and undenatured ethyl alcohol (greater than 80 percent), which have only maize as parent; and to malt extract, which only has malt, roasted or not as parent.

Then, the *remaining production quantity to be assigned* to beer of barley, malted ($production^{RA}_{Beer}$) is calculated (see Steps 3 and 4 of Section 5.3.3). Beer of barley, malted has a parent that only has one child product (rice, broken). In this case, the part of rice, broken in the production of beer of barley, malted can be calculated with the formula of Equation 20:

$$\begin{aligned} production^{RA}_{Beer} &= production_{Beer} - production^{AA}_{Beer} \\ &= production_{Beer} - food\ processing_{Rice} \times eR_{Beer-Rice} \\ &= 22\ 527\ 000 - (488\ 364 \times 8.0425) = 22\ 527\ 000 - 3\ 927\ 667 = 18\ 599\ 333 \end{aligned}$$

This remaining 18 599 333 is assigned proportionally to maize and malt, whether or not roasted. However, before applying the proportional assignment, the mirror imputation loop should be performed to calculate all the possible values that do not require expert judgement (Step 5 of Section 5.3.3).

After the calculation of the remaining $food\ processing$, malt, whether or not roasted has only one child left: beer of barley, malted. Therefore, the remaining $food\ processing$ (2 175 167) of malt, whether or not roasted corresponds to what is sent to the production of beer of barley, malted. And the part of malt, whether or not roasted in the $production$ of beer of barley, malted is:

$$production_{Beer-Malt} = food\ processing^{updated}_{Malt} \times eR_{Beer-Malt} = 2\ 175\ 167 \times 8.0425 = 17\ 493\ 780$$

The remaining production of beer of barley, malted to be assigned is:

$$production^{updated}_{Beer} = production^{RA}_{Beer} - production_{Beer-Malt} = 18\ 599\ 333 - 17\ 493\ 780 = 1\ 105\ 552$$

After all $food\ processing$ quantities have been computed, bottom-up shares are easily computed through Equation 22 and final results are reported in Table 22.

5.3.5 Top-down shares

Top-down shares represent the percentage of the quantity of a good (primary or derived) used in a particular transformation process. Put differently, they indicate the proportion of the $food\ processing$ of a parent commodity (P) that goes into the production process of a specific derived product (C). These differ from bottom-up shares in that they are computed on the parent commodity (top), instead of being a percentage of the child product (bottom).

Equation 23: Top-down share

$$top\text{-}down\ share_{P-C} = \frac{\frac{production_C \times bottom\text{-}up\ share_{C-P}}{eR_{C-P}}}{food\ processing_P}$$

where:

- $production_C$ is the production value of the child product C;
- $bottom\text{-}up\ share_{C-P}$ is the bottom-up share that links the child and the parent;
- eR_{C-P} is the extraction rate between the parent and the child products;
- $\frac{production\ _C \times bottom\text{-}up\ share_{C-P}}{eR_{C-P}}$ is, therefore, the part of the production of the parent product P involved in the production of the child product C expressed in terms of the child product C; and
- $food\ processing_P$ is the quantity of parent product P that is processed.

Given the way they are constructed, the top-down shares of products that have a unique child are equal to 1, while for multi-child products, these shares are lower than 1. The sum of all top-down shares of a parent product across its different child products is equal to 1.

5.3.6 Top-down shares example

Using the beer of barley, malted example presented in **Figure 12**, the top-down share between beer of barley and its three parents (maize, malt whether or not roasted and rice broken) are calculated as follows:

$$top\text{-}down\ share_{Maize-Beer} = \frac{\frac{production\ _{Beer} \times bottom\text{-}up\ share_{Beer-Maize}}{eR_{Maize-Beer}}}{food\ processing_{Maize}} = \frac{\frac{22\ 527\ 000 \times 5\%}{8.0425}}{34\ 212\ 558} = 0.004\%$$

$$top\text{-}down\ share_{Malt-Beer} = \frac{\frac{production\ _{Beer} \times bottom\text{-}up\ share_{Beer-Malt}}{eR_{Malt-Beer}}}{food\ processing_{Malt}} = \frac{\frac{22\ 527\ 000 \times 78\%}{8.0425}}{2\ 237\ 667} = 98\%$$

For the specific and illustrative country–year combination of **Figure 12**, 0.004 percent of the processed quantity of maize is sent to the production of beer of barley, malted and 98 percent of the processed quantity of malt, whether or not roasted is sent to the production of beer of barley, malted. It can be deduced that the 2 percent of malt, whether or not roasted that is left goes to the production of malt extract.

Rice, broken has only beer of barley, malted as child, so all the $food\ processing$ of rice, broken goes to the production of beer of barley.

$$top\text{-}down\ share\ _{Rice-Beer} = 100\%$$

Final results are reported in **Table 22**.

Table 22: Final result of the estimation of top-down shares

PARENT	PART IN THE PRODUCTION OF BEER	BOTTOM-UP SHARE	TOP-DOWN SHARE
Maize	1 105 552	5%	0.004%
Malt, whether or not roasted	17 493 781	78%	98%
Rice, broken	3 927 667	17%	100%
Total production of beer of barley, malted	22 527 000	100%	-

Source: Authors’ own elaboration.

6 TRADE

$$production + \mathbf{imports} - \Delta stocks = \mathbf{exports} + food + food\ processing + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$$

6.1 Data sources

Trade (composed of imports and exports) is the next variable to be considered from **Equation 1**. The trade data used by FAO for FBS compilation consist of raw trade data from Eurostat for European countries and official data from the United Nations Commodity Trade Statistics Database (UN Comtrade) for all other countries reporting to the United Nations Statistics Division (UNSD).

Where available, mirror data[11] are used for non-reporting countries. Importantly, in addition to being a source of data, mirror data can also be used to reconcile values reported by the reporting country with the same trade flows reported by the trading partner. This is especially useful for exports, as imports are typically documented more thoroughly and verified more rigorously than exports (UNSD, 2013).

When data from countries are not available, other sources include specialized publications (e.g. Oil World and publications of the International Sugar Organization), statistical yearbooks of agencies and other online statistical databases (e.g. Trade Map[12] of the International Trade Centre).

Out of the variables used in FBS compilation, international trade data on imports and exports are among the most reliable, given that trade data usually originate from national customs offices that usually collect data on transactions of all cross-border goods for taxation purposes. Moreover, collecting these data is necessary to ensure compliance with World Trade Organization (WTO) and World Custom Organization (WCO) guidelines.

However, official reported trade flows may not include all cross-border transactions in the trade of agricultural goods. Food aid transactions, for example, may be excluded from official trade flows, even though they may represent a significant proportion of a country's food supply. It is therefore important to separately measure and include all the food aid obtained from relief agencies.

In some countries, agricultural goods are traded informally outside of formal customs procedures. These unrecorded trade flows, also referred to as unreported trade, can be critical sources of both household income and food security. Therefore, it is essential that food balance sheets include these transactions when relevant. In particular, these flows are vital for estimating livestock headcounts, especially in countries with large populations of nomadic herders who frequently cross the national borders with their herds. While official trade data are generally available, incorporating data from other sources may be necessary to provide more accurate estimates of aggregate imports and exports for use in food balance sheets.

Additional sources of data for such international transactions include records from central banks and other entities involved in trade. For instance, financial flows associated with cross-border transactions between residents and non-residents can be monitored by a country's central bank. Additionally, shipping manifests are valuable sources of information, detailing the cargo transported by ships, aircraft and vehicles across a country's borders. Other documents, such as records of parcel and letter post, are typically maintained in most countries based on regulations outlined in the acts of the Universal Postal Union. Finally, corporate surveys can also be used to reconcile national trade data.

[11] Mirror data refers to information from the trade partners of a country.
[12] See https://www.trademap.org/Index.aspx.

6.1.1 UNSD tariff-line dataset

NSOs report detailed trade flow data to UNSD, which publishes the data in the UN Comtrade database (United Nations, 2025) at the tariff line[13] level. The level of commodity detail is country-specific, with some countries reporting at the basic 6-digit level of the HS, while others report more detailed data at the 12-digit HS level.

Table 23 shows an example of filtered data from Comtrade (variable names have been modified) for a given year. It contains information on the reporter, partner, flow (1 = imports; 2 = exports; 3 = re-exports; 4 = re-imports), HS codes (variable length), monetary value (in USD), net weight excluding packaging (in kilograms), supplementary quantity (which is the measurement related to the item e.g. litre for liquids), the unit in which the supplementary quantity is given (e.g. 7 = litres, 8 = kilograms), and the HS chapter. Reporters and trade partners are represented with three-digit numerical codes used by the UNSD, based on the international standard M49 codes.

Table 23: Example of Comtrade tariff-line data

YEAR	REPORTER	PARTNER	FLOW	HS CODE	VALUE	WEIGHT	QUANTITY	UNIT	CHAPTER
2014	12	699	1	38089119	109 821.79	7160.22	7 160.22	5	38
2014	600	380	2	8140000000	24 456.00	15870.00	15 870.00	8	8
2014	398	276	1	400259	532.12	0.20	0.20	8	40
2014	703	616	1	5119910	39.83	6.00	6.00	8	5
2014	251	76	1	20081912	1 933.60	148.00	148.00	8	20
2014	702	156	1	382311	117 527.41	60000.00	60 000.00	8	38
2014	203	682	1	100620	143.00	20.00	NA	1	10
2014	344	840	3	16055900	17 245.07	907.00	907.00	8	16
2014	48	682	1	19059093	753 403.97	368416.00	368 416.00	8	19
2014	616	428	1	401693	362.00	20.00	20.00	8	40
2014	384	466	2	19021900	901 389.92	1238270.00	1 238 270.00	8	19
2014	764	152	2	400942	108.10	NA	NA	1	40
2014	414	380	1	38089990	574.72	2.00	2.00	8	38
2014	158	324	2	19059090006	462.59	103.00	103.00	8	19
2014	690	784	1	18069090	1 210.97	28.80	15.00	5	18

Source: Authors' own elaboration.

6.1.2 Eurostat dataset

A data sharing and exchange agreement exists between FAO and Eurostat; therefore FAO uses the Eurostat dataset[14] for European countries. This arrangement aims to promote the consistent use of data between both organizations and alleviate the reporting burden on European Union Member States.

Table 24 presents an example of filtered Eurostat data (variable names have been modified). The key distinctions from the tariff-line datasets are: reporter and partner codes are *geonomenclature* codes,[15] the maximum HS code length is eight digits, monetary values are denominated in thousands of euros, weight is expressed in tonnes and the supplementary quantity unit of measure is always specific to the product (unlike Comtrade tariff-line data, where the same HS code for a supplementary quantity can be reported in various units that need to be converted to tonnes). Additionally, certain HS codes may contain letters, which EU countries are permitted to include instead of numbers for confidentiality reasons.

[13] As stated in the WTO glossary (https://www.wto.org/english/thewto_e/glossary_e/tl_e.htm) tariff line "is a product as defined in lists of tariff rates. Products can be sub-divided, the level of detail reflected in the number of digits in the Harmonized System (HS) code use to identify the product."
[14] See https://ec.europa.eu/eurostat/data.
[15] Geonomenclature codes are a set of codes used in geographical and statistical classifications to represent and organize geographic regions, areas, or entities. These codes are often standardized and used by various organizations, including statistical agencies, governments, and international bodies, to categorize and reference geographic entities for data collection, analysis, and reporting purposes.

Table 24: Example of Eurostat trade data

YEAR	REPORTER	PARTNER	FLOW	HS CODE	VALUE	WEIGHT	QUANTITY
2014	3	346	2	18069060	6.88	2.2	NA
2014	8	1	2	40119200	26.65	14.8	1 225
2014	68	5	2	10MMM000	42.17	0.0	NA
2014	5	6	2	52053400	0.11	0.0	NA
2014	6	647	2	52085100	1 361.25	17.7	180 134
2014	17	4	2	9096100	7.45	1.2	NA
2014	1	11	2	2042300	4.76	0.6	NA
2014	60	32	2	19049080	17.55	3.1	NA
2014	8	32	2	2031955	1 947.23	438.4	NA
2014	17	212	2	33051000	282.50	11.4	NA
2014	1	690	2	3061410	9.05	0.3	NA
2014	8	1	1	35030080	1 002.17	7.4	NA
2014	18	6	1	52094300	0.33	0.0	93
2014	17	604	2	22011019	0.09	0.1	70
2014	6	91	2	38089110	1.71	1.6	NA

Source: Authors' own elaboration.

6.2 Overall approach and workflow

In the new FBS methodology, trade data sourced from UNSD and Eurostat for individual HS codes are mapped to the CPC (see Section **2.5.1** for a description of the HS and CPC). Data used for the SUAs/FBS fall under the 34 HS chapters reported in **Table 25**. Not all items within those chapters are used, as only certain HS codes in the chapters related to food or agricultural goods are considered. For both UNSD and Eurostat data, codes that contain non-numeric characters are dropped. An example of this can be seen in **Table 24**, which contains the 10MMM000 code. The only information of the transaction reported with that code is that it corresponds to some type of cereal (as Chapter 10 is for cereals), but it is otherwise not possible to assign it to a specific cereal type. Therefore, given the difficulty in assigning codes with non-numeric characters to detailed items, and given that these cases are relatively marginal, these codes are removed from the raw data and no attempt is made to assign them.

Raw data obtained by the UNSD and Eurostat are commonly tariff-line data that can have from 6 to 12 digits depending on the detail reported. The minimum required for mapping to the CPC is 6 digits. These codes are not directly comparable across countries, as each country can add these digits independently and may have a description for such codes that differ from those of other countries. As for Eurostat data, it uses *Combined Nomenclature* (CN) codes at 8 digits (CN8), which are EU extensions of HS codes. This is why a mapping between Eurostat and Comtrade data is required. Countries can use their own national trade data, but may choose to use either or both sources as mirror data, or if national trade data are not available to the FBS compiler.

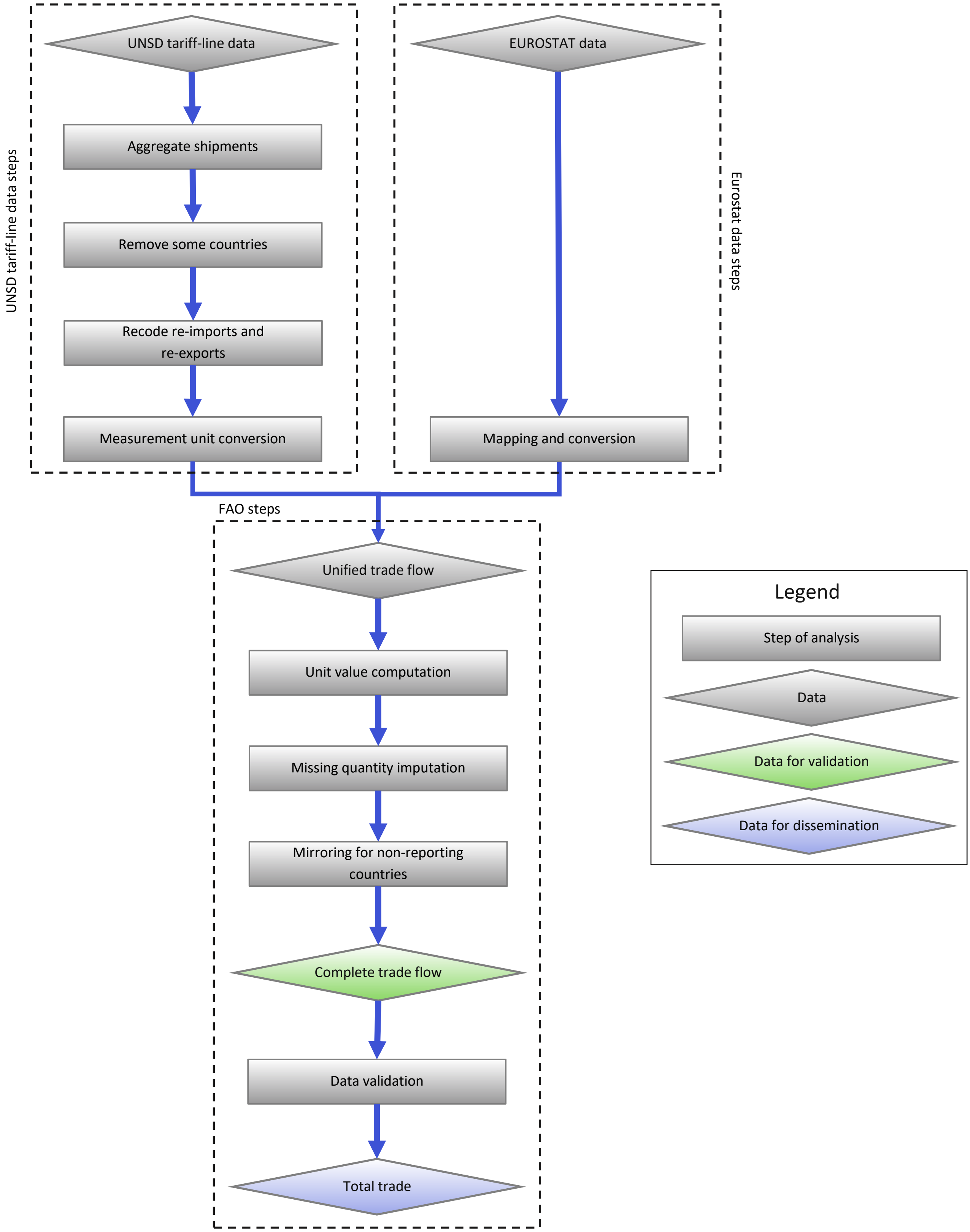
The complete workflow applied in the FBS methodology is reported in **Figure 13**, which shows that total trade values are obtained after various steps. Starting from the two separate flows (UNSD tariff-line data and Eurostat data), unified and validated values are generated, which are further processed by FAO to obtain the final total trade values used in FBS. The steps to compute total trade are described below.

Table 25: HS chapters used in the FAO trade domain

SECTION	CHAPTER	CHAPTER DESCRIPTION
I	1	Live animals
	2	Meat and edible meat offal
	3	Fish and crustaceans, molluscs and other aquatic invertebrates
	4	Dairy produce; birds' eggs; natural honey; edible products of animal origin, not elsewhere specified or included
	5	Products of animal origin, not elsewhere specified or included
II	6	Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage
	7	Edible vegetables and certain roots and tubers
	8	Edible fruit and nuts; peel of citrus fruit or melons
	9	Coffee, tea, mate and spices
	10	Cereals
	11	Products of the milling industry; malt; starches; inulin; wheat gluten
	12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruits; industrial or medicinal plants; straw and fodder
	13	Lac; gums, resins and other vegetable saps and extracts
	14	Vegetable plaiting materials; vegetable products not elsewhere specified or included
III	15	Animal or vegetable fats and oils and their cleavage products prepared edible fats; animal or vegetable waxes
IV	16	Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates, or of insects
	17	Sugars and sugar confectionery
	18	Cocoa and cocoa preparations
	19	Preparations of cereals, Flour, starch or milk; bakers' wares
	20	Preparations of vegetables, fruit, nuts or other parts of plants
	21	Miscellaneous edible preparations
	22	Beverages, spirits and vinegar
	23	Residues and waste from the food industries; prepared animal feed
	24	Tobacco and manufactured tobacco substitutes
VI	33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations
	35	Albuminoidal substances; modified starches; glues; enzymes
	38	Miscellaneous chemical products
VII	40	Rubber and articles thereof
VIII	41	Raw hides and skins (other than fur skins) and leather
	43	Fur skins and artificial fur; manufactures thereof
XI	50	Silk
	51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric
	52	Cotton
	53	Other vegetable textile fibres; paper yarn and woven fabric of paper yarn

Source: Based on WCO. 2022. HS Nomenclature 2022 edition. In: WCO. [Cited 7 February 2025]. https://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs-nomenclature-2022-edition/hs-nomenclature-2022-edition.aspx

Figure 13: Complete trade workflow



Source: Authors' own elaboration.

6.2.1 Data content assessment

A series of preprocessing steps are carried out during the validation procedure to ensure that the data quality meets the minimum requirements for further processing. Low-quality data will negatively affect the overall reliability of the trade flow information and the quality of the global totals. The checks that FAO conducts, which are also applicable to countries, include:

- Verifying the completeness of data. For instance, if transactions show a noticeable reduction in one year, this may indicate that the file is incomplete. Therefore, it may be decided to treat this flow as non-reported (so that mirroring will be used for imputing all its transactions in that year) until its data are more complete. The definition of such thresholds needs to be based on the observation of the time series and the experience of the analyst;
- Detecting and marking reporting countries for each year;
- Detecting and marking non-reporting countries;
- Detecting and marking the minimum and maximum number of HS digits reported by the country for the year and trade flow;
- Checking whether both imports and exports data (including re-imports and re-exports) have been reported by a country in a given year; and
- Detecting missing values.

6.2.2 UNSD tariff-line data steps: aggregate shipments

The tariff-line data from the UNSD contains multiple rows with identical reporter/partner/HS code/flow/year/unit combinations. For example, the data reported in **Table 26** show 15 recorded transactions with HS code 22071090 for the same reporter/partner/flow/year/unit. Multiple transactions (with weight or quantity) can be aggregated; and transactions with missing weight or quantity can be aggregated separately with other transactions that are also missing these variables. **Table 27** shows the results of the aggregation of the example reported in **Table 26**. In this case, even though there is a unique combination of reporter/partner/HS code/flow/year/unit, the aggregated data include two rows because the weight variable was reported for some transactions, and not for others.

Table 26: Example of multiple transactions by reporter/partner/flow/year/HS code

YEAR	REPORTER	PARTNER	FLOW	HS CODE	VALUE	WEIGHT	QUANTITY	UNIT	CHAPTER
2014	508	710	1	22071090	99 570	NA	126 000	7	22
2014	508	710	1	22071090	126 530	NA	168 000	7	22
2014	508	710	1	22071090	87 950	NA	77 141	7	22
2014	508	710	1	22071090	194 740	NA	190 719	7	22
2014	508	710	1	22071090	69 580	116332	116 261	7	22
2014	508	710	1	22071090	1 050	NA	2 871	7	22
2014	508	710	1	22071090	109 770	NA	126 000	7	22
2014	508	710	1	22071090	30 050	NA	40 000	7	22
2014	508	710	1	22071090	147 840	NA	210 000	7	22
2014	508	710	1	22071090	252 100	NA	230 538	7	22
2014	508	710	1	22071090	300	500	23	7	22
2014	508	710	1	22071090	28 690	NA	40 000	7	22
2014	508	710	1	22071090	36 360	37847	42 020	7	22
2014	508	710	1	22071090	2 240	3197	2 500	7	22
2014	508	710	1	22071090	75 540	101700	126 000	7	22

Source: Authors' own elaboration.

Table 27: Results of the aggregation of multiple transactions

YEAR	REPORTER	PARTNER	FLOW	HS CODE	VALUE	WEIGHT	QUANTITY	UNIT	CHAPTER	NROWS
2014	508	710	1	22071090	184 020	259 576	286 804	7	22	5
2014	508	710	1	22071090	1 078 290	NA	1 211 269	7	22	10

Source: Authors' own elaboration.

6.2.3 UNSD tariff-line data steps: remove some countries

UNSD tariff-line data report the area code using the M49 standard codes[16] and codes for undefined or not elsewhere specified areas. European countries (as reporters) are removed from the UNSD dataset, as trade data for European countries are obtained directly from Eurostat. Undefined or not elsewhere specified areas are mapped to a single code.

6.2.4 UNSD tariff-line data steps: recode re-imports and re-exports

Re-imports (code 4) are recoded into imports (code 1) and re-exports (code 3) into exports (code 2). This procedure is applied following UNSD standards. Exports are classified into two categories in the Comtrade tariff-line database: exports of domestic goods and exports of foreign goods (re-exports). Re-exports are foreign goods that are exported in the same condition as they were imported and, therefore, are considered part of the overall country exports. For analytical and cross-checking purposes, it is recommended that re-exports be recorded separately, which may require additional information to identify their origin, to ensure that they are truly re-exports and not domestically produced goods. Similarly, re-imports are goods that are imported in the same condition as they were previously exported. They are included in the country's overall imports and should be recorded separately for analytical purposes, with supplementary sources of information used to identify their origin. The return of exported goods to their country of origin can occur for various reasons, including defects in the product, cancellation of orders by importers, imposed imports barriers and increased demand or prices for the good in the country of origin.

6.2.5 UNSD tariff-line data: measurement unit conversion

Units of measurement are converted to meet FAO standards: all weights are reported in tonnes, animals in heads or 1000 heads and, for some products, just the value is provided. For example, if the originally reported quantity is "units" and the FAO unit is "1000 heads", the quantity is divided by 1000. Some of these conversions are simple mathematical conversions that involve multiplying by a fixed factor known a priori. Other conversions are less trivial and use information derived from the data. In particular, non-livestock and livestock product-specific conversions are added.

For non-livestock products, if a quantity is originally expressed in "units", and weight is not available, a conversion factor (qw_u) to convert from units to kilograms is required. In order to obtain these conversion factors, for a particular non-standard unit, all transactions with that non-standard unit that also report weight are used to obtain a weight in kilograms.

This factor (qw_u) is obtained as the median of the ratio between weight and quantity across all tariff-line transactions:

Equation 24: Non-livestock commodity unit conversion formula

$$qw_u = median(weight/quantity_u)$$

The non-standard units are then converted to weight by multiplying by the conversion factor.

Equation 25: Non-livestock commodity unit weight formula

$$weight = quantity_u \times qw_u$$

[16] See https://unstats.un.org/unsd/methodology/m49/.

where:

- $quantity_u$ is the quantity expressed in a given non-standard unit of measurement u; and
- qw_u is the conversion factor of quantity and represents the weight in kilograms per non-standard unit.

Suppose that country X does not report the weight of eggs, but reports numbers of eggs. In this case, the factor qw_u is computed for all countries where both weight and quantity are reported. This median conversion factor is then applied to the units reported by country X to estimate the weight of eggs for country X in tonnes.

A different conversion is applied for livestock. In this case the conversion from weight to the final unit (head or 1000 heads, depending upon the species) where the reported unit is either missing or different from the final unit, is made by dividing the reported (or imputed unit as shown in Section **4.4**) weight by an average livestock weight by country and species. For these purposes, an FAO team of experts has compiled a table of average livestock weight by country and species. Using these average weights denominated as aw, the final unit in heads (or 1000 heads) will be obtained as:

Equation 26: Livestock commodity unit conversion formula

$$quantity = \frac{weight}{aw}$$

where aw is the average livestock weight.

6.2.6 Eurostat data steps: mapping and conversion

Eurostat classifies areas using the geonomenclature coding system (Eurostat, 2024). These codes are converted to M49 country codes. Any area codes that cannot be mapped to the M49 are flagged, and the corresponding records are removed from the dataset. Countries without identified M49 codes are reassigned to a single code for unspecified areas.

Information on FAO units is then added, which is straightforward since Eurostat and FAO units are mostly the same.

Finally, values are converted from EUR to USD using the average exchange rate of the year of the transaction.

6.2.7 FAO steps: unified trade flow

After the data processing described above, the UNSD and Eurostat data are expressed in the same unit and country measurement units. This "unified trade flow" contains trade data for all countries and is then further processed.

6.2.8 FAO steps: unit value computation

For each record having both quantity and value, the unit value (uv), which gives a value per unit, is computed as follows, which will be used in the subsequent steps:

Equation 27: Trade unit value

$$uv = \frac{value}{quantity}$$

where:

- $quantity$ is the quantity;
- $value$ represents the value; and
- uv in the unit value.

6.2.9 FAO steps: missing quantity imputation

As FBS compilation is based on physical quantities, additional imputations need to be performed for any traded good for which data are only available in value and not in volume. For records showing only the value, the corresponding quantity is imputed by dividing the value at time t by the unit value at time $t-1$, adjusted for the annual change in the global average unit value for the product in question.

Equation 28: Imputed quantity

$$quantity_{\ t,x} = \frac{value_{t,x}}{uv_{t-1,x} \times \frac{uv_{t,global\ average}}{uv_{t-1,global\ average}}}$$

where:

- x represents the country;
- t represents the year for the data in question;
- $value_{t,x}$ is the available value for country x and year t;
- $uv_{t-1,x}$ is the unit value for country x in year $t-1$.
- $uv_{t,global\ average}$ is the global average unit value in year t; and
- $uv_{t-1,global\ average}$ is the global average unit value in year $t-1$.

To ensure that this change in the global average unit value is as accurate as possible, the countries with missing quantities are validated last among the reporter countries by using a validation routine described in Section **6.3**.

6.2.10 FAO steps: mirroring for non-reporting countries.

Trade flows that are missing for non-reporting countries need to be estimated. This is done using the *mirroring* routine: the corresponding trade of the non-reporting countries is extracted from the partners, inverting the flows. The quantities are kept the same while the values are corrected by a fixed factor to account for the CIF/FOB (cost, insurance and freight/free on board) conversion. This factor is retrieved from the literature and often set to 12 percent (i.e. imports of the partner are divided by 1.12, while exports of the partner are multiplied by 1.12).[17] "The CIF unit values rely on importers' declarations, and include all trade costs (except tariffs and domestic taxes after the border). The FOB unit value is a proxy for the trade price at the factory gate, relying on exporters' declarations, and does not include trade costs" (Berthou and Emlinger, 2011).

Table 28 shows an example of the mirroring procedure. Country A is a non-reporter in year 3, so country B's exports will be used to impute country A's imports. First, quantities are copied as reported by country B (890); values reported by country B are then adjusted by the CIF/FOB margin of 12 percent ($1\,400 \times 1.12 = 1\,568$).

Table 28: Mirroring example

COUNTRY A				COUNTRY B	
	IMPORTS			EXPORTS	
YEAR	VALUE	QUANTITY		VALUE	QUANTITY
1	1 450	800		1 250	800
2	1 550	950		1 390	925
3	1 568	890	←	1 400	890

Source: Authors' own elaboration.

[17] The choice of a fixed CIF/FOB markup is simply for convenience, as these margins are neither time- nor distant-constant, i.e. they evolve over time and are different depending on geographic proximity of reporters and partners. Using different CIF/FOB margins will be explored in the future.

As shown in the example, when both countries report trade flows, the reported quantities are not necessarily the same (they are different in year 2). These discrepancies may occur for a variety of reasons, including the use of tariff-line descriptions that can affect how items are recorded as they are aggregated into the CPC; time differences in recording transactions (e.g. a transaction can be exported in December but arrives at the destination in January); transactions reported with non-standard units by at least one reporter; or the use of net weight by one reporter, while gross weight was used by the other. Moreover, the CIF/FOB margin can vary. As indicated above, the 12 percent margin can be adjusted over time, as these margins evolve and can vary depending on geographic proximity of reporters and partners.

The mirroring procedure is an approximation of the actual trade flows, as they are inferred from the declarations made by reporting partner countries. In some cases, flows are structurally similar, and the mirror gives a very good approximation. In some instances however, the mirroring mechanism results in data that are not in line with the time series of the mirrored country. For instance, if a main partner for a given product is also a non-reporter for a given year, then the resulting total flow may be incomplete. Thus, mirrored flows are checked for completeness, and if they are flagged as inconsistent with the historical time series, they are considered separately and potentially supplemented with external data or manual imputation during the validation stage (Section 6.3).

6.3 Data validation

6.3.1 Overview

During imputation, automatic outlier imputation is not carried out, as preliminary tests demonstrated that any automatic process was likely to introduce errors to the dataset. For this reason, once the data have been aggregated and saved for all years, a semi-automatic (guided) workflow for corrections is performed through an external validation routine. The validation process differs depending on whether the validation needs to be carried out for reporting or non-reporting countries.

6.3.2 Outlier detection criteria

The process identifies which reporter/product/flow/year combination is likely to be an outlier and computes various outlier detection routines on the data. It assigns scores based on how many times a transaction has been found to be an outlier and the analyst can manually select time series with outliers and use an interactive validation tool that allows analysis of the composition of total trade flows using bilateral data. The tool displays the outliers, offers a variety of methods for correcting them, and stores the corrections that the analyst deemed most appropriate in the unified trade flows dataset. When no statistically based imputation seems appropriate, the analyst can "force" some values (e.g. obtained by consulting external sources) through a manual overwrite.

Outlier detection checks are carried out at the item level on the trade flow of each product.

Outlier detection criteria include checks on the following values, which are compared with appropriate ranges and thresholds to find and correct outliers:

- the ratio between quantity (in a given year) and a moving average of each item(i)/flow(f)/quantity(q) combination;
- the unit value; and
- the *total merchandise trade*, the *total value of agricultural products* and the ratio between these two values.

6.3.3 Data validation for reporter countries with an example

For reporter countries, the FAO trade data validation process offers the following options for data flagged as outliers. These corrections will be applied at the item level to the trade flow of the product:

Measurement factor

If the quantity or value are determined to have been reported with an incorrect measurement unit, for example a value can be off by a factor of 10, values can be corrected according to that factor.

Mirror flow

As described above, the flow of the partner (if it exists) is used, adjusted for CIF/FOB differences.

Outlier detection and imputations

The value outside of the limits (outlier) is deleted and replaced using one of the following three alternatives. The outlier check is per product:

- Median partners: Median value with respect to all partners.
- Median world: Median value of the unit value of all reporters to all partners.
- Moving average: Moving average of the preceding three years.

Publication

When no other correction is applicable, data may be found in external publications.

Estimation

If no other correction is deemed appropriate, figures are assigned based on the literature or expert knowledge.

For example, **Table 29** represents a possible outcome of the FAO validation process for traded values. The table shows, for country A an outlier in year 6 (in bold) compared to the time series over years 1 to 5.

Table 29: Subset of the trade validation plugin

				YEAR					
AREA	FLOW	CPC	ITEM	1	2	3	4	5	6
A	Import	0111	Wheat	4 990 865	5 781 712	10 004 830	9 659 186	8 877 310	8 907 417
A	Import	0141	Soya beans	2 340 974	2 660 352	2 741 886	3 040 452	2 493 107	3 038 983
A	Import	0112	Maize (corn)	2 055 543	2 122 733	4 347 475	2 299 614	2 303 899	3 401 947
A	Import	39120.01	Bran of wheat	1 700 848	1 335 005	1 396 264	1 579 597	1 561 956	1 369 992
A	Import	0115	Barley	384 109	655 988	562 777	889 319	2 383 928	1 585 200
A	Import	01921.02	Cotton lint	914 377	751 703	946 099	1 064 782	1 191 084	1 148 397
A	Import	21910.07	Cake of sunflower seed	965 137	834 489	1 091 154	1 043 748	763 104	681 973
A	Import	01445	Sunflower seed	640 442	712 122	1 239 492	1 206 590	734 179	**5 563 693**
A	Import	21910.03	Cake of soya beans	384 109	655 988	562 777	889 319	2 383 928	1 585 200

Source: Authors' own elaboration.

The product sunflower seed has a value that is notably higher in comparison to its historical data. This requires a review of the bilateral transactions for the import quantity, import value and unit values. This check (**Table 30**) indicates that the unit value for the specified product is substantially lower and the quantity is substantially higher compared with the historical trend.

Table 30 shows the imports for reporter A from all of its partners.

Table 30: Total sunflower seed trade of country A

AREA	ITEM	CPC	ELEMENT	YEAR 1	2	3	4	5	6
A	Sunflower seed	01445	Import quantity [t]	640 442	712 122	1 239 492	1 206 590	734 179	**5 563 693**
A	Sunflower seed	01445	Import value [1000 $]	356 471	361 115	568 306	628 373	542 090	647 689
A	Sunflower seed	01445	Import UV [$/t]	557	507	458	521	738	**116**

Source: Authors' own elaboration.

Looking further at the countries trading with country A, the issue has been identified in the flow between country A and country B reported in **Table 31**. The unit value appears to differ from the historical series by a factor of 1000, while the reported quantity is unusually high.

Table 31: Example of bilateral trade flow of country A and country B

REPORTER	PARTNER	FLOW	CPC	ELEMENT	YEAR 1	2	3	4	5	6
A	B	Import	01445	Import quantity [t]	6 196	4 549	4 254	4 314	5 064	**4 742 000**
A	B	Import	01445	Import value [1000 $]	23 200	18 199	18 162	18 002	21 543	19 025
A	B	Import	01445	Import UV [$/t]	3 744	4 001	4 270	4 173	4 254	**4 012**

Source: Authors' own elaboration.

A verification against country B's exports data has been conducted (**Table 32**) using country B's reported data. This indicates that data for both quantity and unit value are consistent with the historical values.

Table 32: Example of bilateral trade flow of country B and country A

REPORTER	PARTNER	FLOW	CPC	ELEMENT	YEAR 1	2	3	4	5	6
B	A	Export	01445	Export quantity [t]	7 889	5 869	6 241	5 617	7 605	**5 762**
B	A	Export	01445	Export value [1000 $]	28 300	19 162	20 002	20 520	22 670	23 678
B	A	Export	01445	Export UV [$/t]	3 587	3 265	3 205	3 653	2 981	**4 109**

Source: Authors' own elaboration.

Given that the data for these flows suggest a possible digit error in the quantity, the analyst may choose to use the measurement factor approach to correct the flow. By applying the measurement factor, the reported quantity of the product flow from country A to country B will be recalibrated and divided by 1000, resulting in a flow that is in line with the time series (**Table 33**). Once this flow is corrected, the revisions are reflected in the total trade values (**Table 34**).

Table 33: Example of bilateral flow of country A and country B with the measurement factor applied

REPORTER	PARTNER	FLOW	CPC	ELEMENT	YEAR 1	2	3	4	5	6
A	B	Import	01445	Import quantity [t]	6 196	4 549	4 254	4 314	5 064	**4 742**
A	B	Import	01445	Import value [1000 $]	23 200	18 199	18 162	18 002	21 543	19 025
A	B	Import	01445	Import UV [$/t]	3 744	4 001	4 270	4 173	4 254	**4 012**

Source: Authors' own elaboration.

Table 34: Total sunflower seed trade of country A after correction

AREA	ITEM	CPC	ELEMENT	YEAR 1	2	3	4	5	6
A	Sunflower seed	01445	Import quantity [t]	640 442	712 122	1 239 492	1 206 590	734 179	**826 435**
A	Sunflower seed	01445	Import value [1000 $]	356 471	361 115	568 306	628 373	542 090	647 689
A	Sunflower seed	01445	Import UV [$/t]	557	507	458	521	738	**780**

Source: Authors' own elaboration.

6.3.4 Data validation for non-reporting countries with an example

As mentioned, trade values for non-reporting countries are obtained from their reporting trade partners through mirroring. While the quantities are retained, the values are adjusted by a factor to account for the CIF/FOB price differences.

For example, in the following scenario, a missing value was identified in year 6 for country A (**Table 35**).

Table 35: Example of data validation for non-reporting countries

				YEAR					
AREA	FLOW	CPC	ITEM	1	2	3	4	5	6
A	Import	0111	Wheat	99 326	114 384	101 017	110 982	124 047	122 507
A	Import	23161.02	Rice, milled	75 505	73 463	59 570	105 986	57 803	43 345
A	Import	21121	Meat of chickens	61 275	64 961	66 633	71 830	73 025	79 176
A	Import	24320	Malt	19 074	22 105	20 031	21 788	27 471	19 533
A	Import	02122	Sheep	20 000	20 000	20 000	20 000	20 000	20 000
A	Import	21124	Meat of turkeys	14 044	14 487	13 170	12 551	12 503	7 496
A	Import	21151	Edible offal	9 161	9 329	8 538	8 776	10 371	8 819
A	Import	23520	Refined sugar	8 113	8 328	9 974	9 593	4 939	**NO DATA**
A	Import	23710	Uncooked pasta	8 357	9 451	5 950	9 266	5 468	6 983

Source: Authors' own elaboration.

Refined sugar has no data in year 6. This means that neither country A nor any of its partners reported any data in year 6 for this product. Under such circumstances, the analyst first needs to verify if the true value is actually zero or if data is missing. If, after examination, it is confirmed that there is a genuine absence of data, the analyst should verify the information with alternative data sources. If alternative data are found, these should be included by labelling them with the appropriate flag, and accompanying them with metadata that report the source.

7 FOOD SUPPLY

$production + imports - \Delta stocks = exports + \mathbf{food} + food\ processing + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$

7.1 Data sources

In the FBS methodology, the variable $food$ (food supply) is probably the most important, as it is the basis for the dietary energy supply calculation. As already stressed in Section **2.1** and in the introduction, this represents the food *available* for human consumption to the resident population and not the real food *consumption* pattern of a country.

Directly measured data on food supply, as defined in Section **2.1**, may be difficult to obtain. FAO obtains food data from countries primarily via its annual production questionnaire and missing data are estimated using a procedure that uses consumption data collected by household surveys. The procedure assumes that food available is either consumed or wasted. However, a number of caveats and adjustments need to be made before such data can be used to inform estimates of food supply because of some limitations in the data provided. Indeed, such surveys are usually carried out for small periods of time and rather infrequently – typically once every three or five years, and sometimes less frequently.

7.2 Approach followed in food supply estimation

The approach proposed starts with the definition of a reference period, which is a time range. This parameter is composed of two fields: the min reference year $t_{r,min}$ and the max reference year $t_{r,max}$. The time range selected is crucial for accurate results, as it determines when estimations are started.

The food supply value for the reference year t_r is then computed using the average of values for all years in the specified time range. This establishes a starting point for the imputation, working both forwards and backwards from this reference year, as elaborated below.

Estimation works differently for products classified as *food estimates* or *food residuals*: food estimates are products that can have several utilizations (e.g. food and feed), while *food residuals* are goods for which food is the only utilization and can, therefore, be calculated as the balance of production minus net trade.

If some figures in the reference year are official, these are considered protected and never overwritten.

For food estimates the value for $food$ is estimated as described below. However, food residuals are computed during the balancing procedure. As will be seen in Section **14**, for the food residual group of products, the amount of production net of trade goes to $food$ only when greater than zero. When net trade is negative, $food$ is recorded as zero.

Imputation for food estimates

These steps apply to food estimates. Once the reference year t_r is set, imputation takes place for any following (and preceding) year t as a function of a known level of food supply in the reference year, adjusted for changes due to income and population (captured by a trend factor). In a stylized form, imputed food supply is:

Equation 29: Imputed food supply – general function

$$food_{t_{r+1}} = f(food_{t_r})$$

where t_{r+1} is the year to be imputed, t_r is the reference year for forward imputation (for backward imputation, the year to be imputed is t_{r-1}) and f is the functional form chosen that has the following parameters: population, gross domestic product (GDP), elasticity of demand for the particular food product, and previous food supply figures.

Population estimates used by FAO come from the United Nations Population Division and include, to the extent known, migrants, guest workers and refugees. However, they do not include tourists. Indeed, in the current methodology, tourists are accounted for as a separate variable (see Section 12).

GDP data are retrieved from the World Bank and the United Nations Conference on Trade and Development (UNCTAD).

The approach is based on Engel's law, an empirical regularity in economics first described by the statistician Ernest Engel in 1857: as family income increases, the share of income spent on food decreases (Perthel, 1975). Engel's law stems from the fact that the income elasticity of the demand for food is normally positive but smaller than 1. This regularity is verifiable at different scales. At the national level, it shows that countries with higher average per capita income have on average a lower proportion of income spent on food.

FAO has estimated elasticities for all food items for internal use using models made on each country–product combination based on the work of Salathe (1979). Elasticity estimates are constant over country–product combinations. Four different functional forms have been identified: linear specification (**Equation 30**), a log-log specification (**Equation 31**), a semi-log (**Equation 32**) and a log-inverse (**Equation 33**) function.

All equations for all products and all countries have been parameterized with an income elasticity ε_j for every country i, product j and a time factor t_j.

Equation 30: Linear specification

$$food_{t,i,j} = food_{t-1,i,j} \times \frac{population_{t,i}}{population_{t-1,i}}$$

Equation 31: Log-log specification

$$food_{t,i,j} = food_{t-1,i,j} \times \frac{population_{t,i}}{population_{t-1,i}} \times e^{\varepsilon_{i,j} \times log\left(\frac{GDP_{t,i}}{GDP_{t-1,i}}\right)}$$

Equation 32: Semi-log specification

$$food_{t,i,j} = food_{t-1,i,j} \times \frac{population_{t,i}}{population_{t-1,i}} \times \left(1 + \varepsilon_{i,j} \times log\left(\frac{GDP_{t,i}}{GDP_{t-1,i}}\right)\right)$$

Equation 33: Log-inverse specification

$$food_{t,i,j} = food_{t-1,i,j} \times \frac{population_{t,i}}{population_{t-1,i}} \times e^{\varepsilon_{i,j} \times \left(1 - \frac{GDP_{t-1,i}}{GDP_{t,i}}\right)}$$

where:

- $food_{t,i,j}$ is food in country i for year t and product j to be estimated;
- $food_{t-1,i,j}$ is food in country i for year $t-1$ and product j;
- $population_{t,i}$ is the population in country i for year t;
- $population_{t-1,i}$ is the population in country i for year $t-1$;
- $\varepsilon_{i,j}$ is the elasticity for the $i; j$ country–product combination;
- $GDP_{t,i}$ is the GDP per capita in country i for year t; and
- $GDP_{t-1,i}$ is the GDP per capita in country i for year $t-1$.

If there is any country or product with a missing elasticity value, the elasticity is imputed by taking the average $\bar{x}$ elasticity of that product across countries that are in the same income group.[18] Moreover, elasticities are tested for outliers with the following criteria:

Equation 34: Elasticity outlier detection criteria

$$x \text{ is an outlier if } \begin{cases} x < \bar{x} - 2\sigma & \text{lower outlier} \\ x > \bar{x} + 2\sigma & \text{upper outlier} \end{cases}$$

where:

- $\bar{x}$ is the average elasticity; and
- σ is the standard deviation of the elasticity of a specific product across countries that belong to the same income group.

When an outlier is detected, it is replaced by the average elasticity for that income group.

[18] The World Bank updates the income group classification each year in July, based on the gross national income per capita of the previous calendar year. For more details, see https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html.

8 STOCKS VARIATION

$production + imports - \Delta \boldsymbol{stocks} = exports + food + food\ processing + feed + seed + tourist\ consumption + industrial\ use + loss + residual\ use$

8.1 Data sources

Stocks of agricultural products are held by a variety of economic agents, including households, farms, processing or trading companies, government agencies and international relief organizations such as the WFP. A number of factors determine the amount of stocks held, including economic conditions, demand and supply fluctuations, trade policies, and seasonal variations. Stocks are generally held for smoothing consumption, but also sometimes for speculative purposes. They also constitute strategic reserves against food shortages.

Medium-term stockholding is confined to a few non-perishable items that are considered critical for a country's food security. These are usually grains, sugar, pulses and some oilseeds. Short-term stockpiling, from one marketing year to the next, can be done for various other items, including horticultural products (e.g. apples and potatoes), processed horticultural products (e.g. frozen concentrated orange juice and canned tomatoes), or processed dairy products (e.g. butter and cheese). The accurate measurement of total stockholdings among all actors, especially for primary food products, is a critical policy priority because stock levels can both influence and be influenced by food prices. Nevertheless, accurate estimates of total stocks are not frequent due to difficulties in determining levels that may exist at any point in the supply chains.

The limited collection of data on stocks at the national level is directly reflected in the poor availability of stocks data. To compute FBS, estimates of stocks for primary commodities and a list of named "stockable items" are retrieved as far as possible from authoritative sources, such as the Agricultural Market Information System (AMIS) database,[19] which estimates closing stocks levels for maize, wheat, rice and soyabeans for more than 20 of the world's largest producer and consumer countries of those commodities. Similar stocks data are also available from the Production, Supply and Distribution (PS&D) online data provided by the United States Department of Agriculture's (USDA) foreign agricultural service.[20] Similarly, estimates of global sugar stocks can be accessed from F.O. Licht[21] and stocks estimates for numerous oils and fats can be sourced from Oil World.[22] Data from the USDA PS&D database are collected according to specific selection criteria (see below) to be consistent with existing production and trade time series in the FAO database. All data are analysed and selected using the imputation methodologies detailed below.

8.2 Estimation model

Stocks variation $(\Delta stocks)$ enters the SUA identity of **Equation 1** on the supply side but it can be positive or negative. Stocks values that are missing after data harvesting and selection criteria applied are estimated.

Assuming that stocks are maintained mostly to smooth consumption levels between harvests (FAO, 2017a), it is possible to estimate stocks variation for year t as a "supply smoother". The hypothesis is that stakeholders have an interest in keeping fluctuations in the available supply within certain thresholds for various reasons (e.g. decision-makers will promote withdrawals from or increases in stocks as necessary to prevent shortages of food staples, particularly when considering traded quantities).

[19] See https://app.amis-outlook.org/#/market-database/supply-and-demand-overview.
[20] See https://apps.fas.usda.gov/psdonline/app/index.html#/app/home.
[21] See http://www.fo-licht.de/.
[22] See https://www.oilworld.biz/.

Therefore, estimating stocks variation requires a smoothed representation of supply. The current FBS methodology uses a regression method that links available data on production, imports and exports.

In particular, a linear model is estimated between *supply including stocks* (**Equation 35**) and *supply excluding stocks* (**Equation 36**). Supply including stocks can be expressed with a regression equation allowing for a linear $trend$ as in **Equation 37**.

Equation 35: Supply including stocks

$$supply_{inc} = production + imports - exports - \Delta stocks$$

Equation 36: Supply excluding stocks

$$supply_{exc} = production + imports - exports$$

Equation 37: Supply linear trend

$$supply_{inc} = \alpha + \beta \times supply_{exc} + \delta \times trend + \varepsilon$$

where:

- $supply_{inc}$ is the supply including stocks and is defined only when variation is available;
- $supply_{exc}$ is the supply excluding stocks;
- $trend$ is the linear trend;
- α, β and δ are, respectively, the intercept and slope parameters to be estimated by the model; and
- ε is the error term of the equation.

The estimate of supply including stocks, is forecast using $trend$ and $supply_{exc}$, which can be estimated from available production, imports and exports data (either unofficial or imputed). The predicted value ($supply_{inc}^{pred}$) will be obtained as :

Equation 38: Estimation of supply including stocks

$$supply_{inc}^{pred} = \hat{\alpha} + \hat{\beta} \times supply_{exc} + \hat{\delta} \times trend$$

where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\delta}$ are the parameters estimated for **Equation 37**. It follows that stocks variation can be obtained as:

Equation 39: Estimation of stocks variation

$$\Delta stocks^{pred} = supply_{exc} - supply_{inc}^{pred}$$

8.3 Stocks variation: an example

Consider an example based on wheat production, imports, exports and stocks variation data over the 2000–2017 period for a hypothetical country. Values are reported in **Table 36**. It is assumed that no data for stocks variation are available for the years 2014–2017, and production, imports, exports and stocks of the previous years have already been validated by country analysts in a previous FBS compilation cycle.

Table 36: Stocks variation example for wheat

YEAR	PRODUCTION	IMPORTS	EXPORTS	$supply_{inc}$	$supply_{exc}$	$supply_{inc}^{pred}$	$\Delta stocks^{pred}$
2000	220	103	68	220	255	329	35
2001	532	54	64	522	522	374	0
2002	537	1	14	393	523	383	130
2003	715	5	204	387	517	390	130
2004	800	0	319	386	481	394	95
2005	620	0	613	352	7	335	-345
2006	800	1	251	379	549	420	170
2007	800	13	274	379	539	427	160
2008	800	0	580	390	220	390	-170
2009	1 067	0	869	438	198	396	-240
2010	1 402	0	1 104	388	298	418	-90
2011	1 464	0	651	483	813	499	330
2012	1 561	1	1 455	437	107	408	-330
2013	1 430	69	440	559	1 059	550	500
2014	840	54	357	468	537	485	52
2015	1 139	1	872	490	268	455	-187
2016	1 144	0	806	494	339	474	-135
2017	1 020	3	351	523	673	529	144

Source: Authors' own elaboration.

First, the parameters of **Equation 38** are estimated. The following model is obtained:

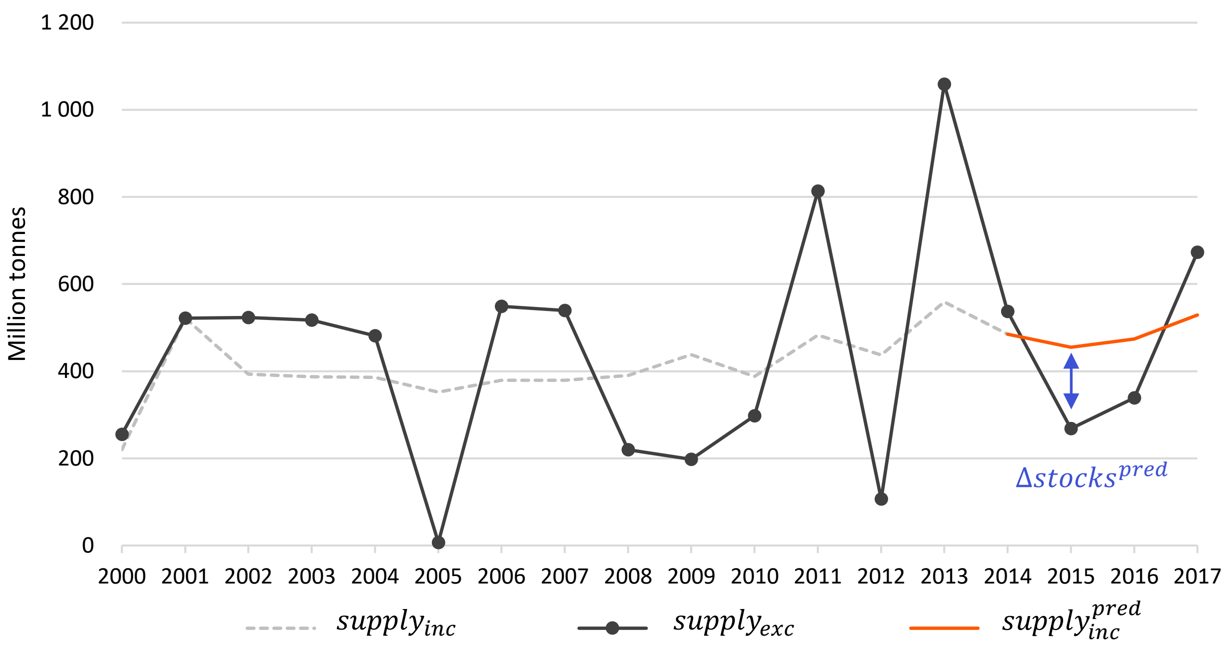
$$supply_{inc}^{pred} = 284.43 + 0.14 \times supply_{exc} + 8.34 \times trend.$$

Then, $\Delta stocks^{pred}$ is estimated using **Equation 39**. For example, in 2015, this value is obtained as:

$$\Delta stocks^{pred} = supply_{exc} - supply_{inc}^{pred} = 268 - 455 = -187.$$

This value is shown in **Figure 14**.

Figure 14: Example of supply including and excluding stocks



Source: Authors' own elaboration.

9 FEED USE

$production + imports - \Delta stocks = exports + food + food\ processing + \boldsymbol{feed} + seed + tourist\ consumption + industrial\ use + loss + residual\ use$

9.1 Data sources

Feed use in FBS is the quantity of a given commodity prepared from an edible agricultural product that is fed to live animals or fish. For some crops such as soyabeans and maize, feed use can be the main use. Feed use may be subject to fluctuations following changes in the relative prices of feed products and livestock products. However, the total nutrient availability across all feed sources is expected to remain relatively stable, at least with reference to livestock production units.

In the current FBS methodology, feed use is computed based on livestock production. The model integrates data on the country's livestock population, the nutritional requirements for each output, and the expected share of feed in the rations. This feed proportion, in turn, is influenced by the characteristics of the livestock production systems. The model also considers information on the supply of primary commodities used as feed ingredients as reflected in the FBS at both the country and product levels.

9.2 Estimating feed use

The approach is *demand-driven*, i.e. feed use results from feed requirements of herds measured in energy and protein units, as these are the two main drivers of feed composition. Energy requirements tend to be primarily satisfied through forage, roughage and cereals, while protein requirements tend to be primarily satisfied through legumes and oil crops. Given data availability, the model cannot take into account micronutrients, such as amino acids and vitamins requirements.

The largest part of the feed use commonly arises from the energy and protein requirements of livestock and poultry. In addition, feed used in aquaculture tends to be significant (especially in some Asian countries), and is increasing in volume, which is why it is taken into account. Feed for livestock and poultry and feed for aquaculture are computed separately, and then added up to obtain total feed:

Equation 40: Feed balance

$$D_{j,t} = D_{j,t}^{livestock} + D_{j,t}^{aqua}$$

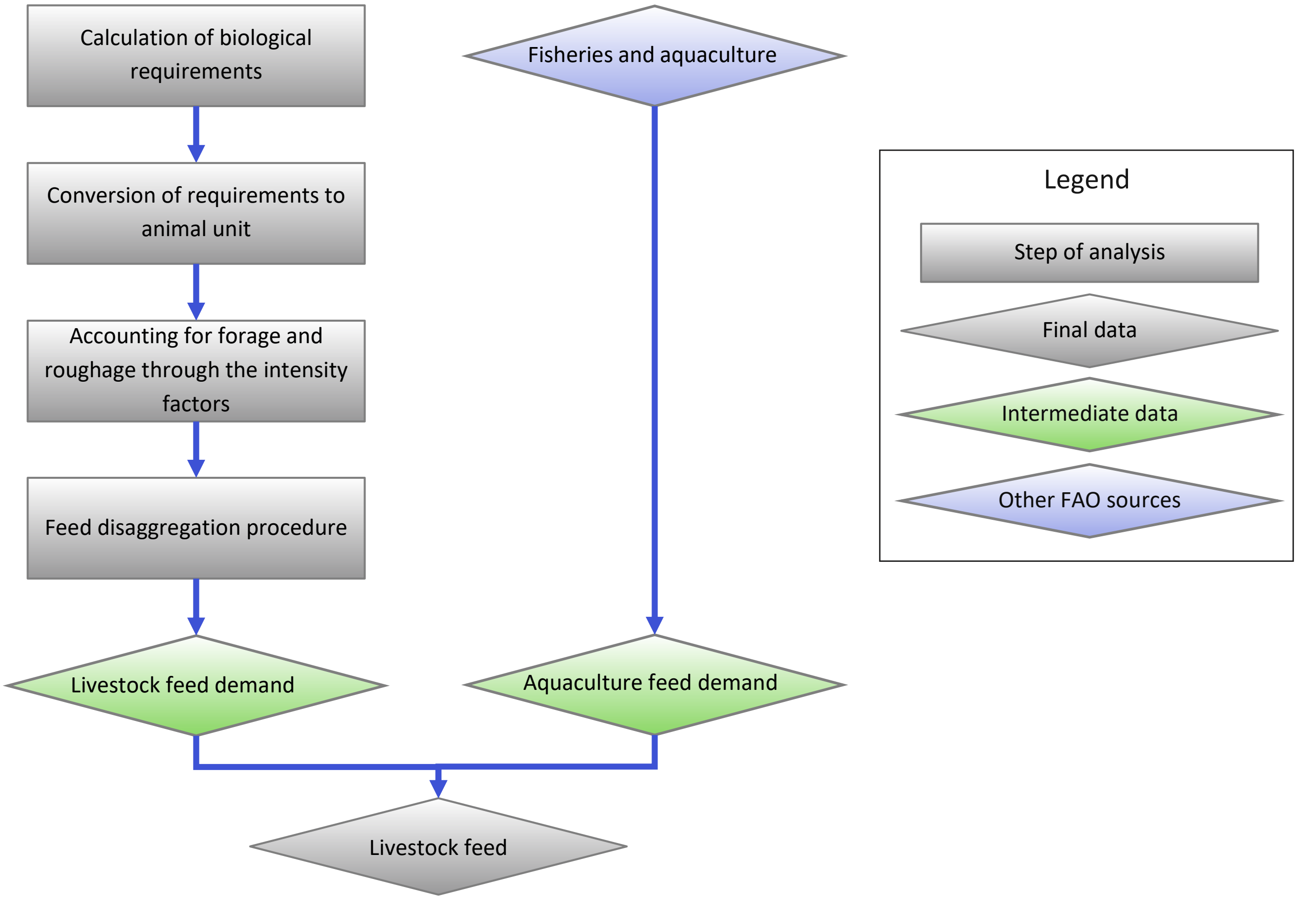
where:

- $D_{j,t}^{livestock}$ is the feed demand for livestock in country j and reference period t; and
- $D_{j,t}^{aqua}$ is the feed demand for aquaculture in country j and reference period t.

The model compiles estimates of the annual total feed use of a country in *energy units* (measured in thousands of megajoules [MJ]) and *protein units* (measured in tonnes).

The overall flow is described in **Figure 15**. First, the biological requirements for livestock are estimated in terms of protein and energy needs. These requirements are then standardized to a common unit (animal unit) for all species and adjusted by an intensity factor, which accounts for compound feed. Next, a procedure is applied to calculate the final figures for livestock feed demand. This process takes into account the proportion of each potential feed type in the total available feed production and allocates the total requirement accordingly. Feed for aquaculture is retrieved from the FAO fisheries and aquaculture website.[23] The feed values for livestock and poultry are then combined with those for aquaculture to determine the total feed requirement.

Figure 15: Feed overall workflow



Source: Authors' own elaboration.

9.3 Livestock feed demand

9.3.1 Biological requirements

Biological requirements consist of energy and protein requirements.

Energy requirements defined

In general, energy available from feedstuffs has been classified as gross energy (GE, digestible energy (DE), metabolizable energy (ME) and net energy (NE) (National Research Council, 1994, 1998). GE does not depend on the species that takes it in, but entirely on the composition of nutrients in feedstuffs, as it includes energy from indigestible ingredients that are excreted by the animal after consumption. The digestible component of GE, i.e. DE, varies across feed species as it depends on the individual animals' digestive systems. Subtracting the urinary and gaseous energy losses from DE, gives ME. NE represents the energy that remains after subtracting energy released as heat as a result of digestive and metabolic processes.

ME is chosen as the preferred category for the measurement of energy requirements, as it represents the energy intake that is fully used by the animal, i.e. not excreted unused, and is at the same time species-specific.

[23] See https://www.fao.org/fishery/en.

Protein requirements defined

Protein requirements are usually indicated as a percentage of the dry matter intake of the feedstuff. Animals require specific components of protein, which are amino acids (National Research Council, 1998). To simplify the feed requirement estimation, the amino acid requirements are expressed in corresponding amounts of crude protein measured in g or kg, although the exact amount of required protein also depends on the quantity of relevant amino acids contained in each unit of crude protein fed.

Estimation of biological requirements

These estimations are driven by prediction formulas tailored to specific species, considering their relative importance and the availability of pertinent data. To ensure practicality and applicability, average expected requirements are incorporated into these formulas, providing a standardized basis for calculation. Energy requirements are computed as ME and are primarily expressed in the preferred unit of MJ per year. Additionally, protein requirements are approximated in grams per day, using the concept of crude protein (CP) needs.

In a general form, protein and energy requirements (R) of the animals of species i in country j in year t can both be estimated as the sum of the number of animals (N) of each species (i) multiplied by the average annual requirements (r):

Equation 41: Biological requirements by species – version 1

$$R_{j,t} = \sum_{i=1}^{I} N_{i,j,t} \times r_{i,j,t}$$

where:

- $R_{j,t}$ refers to requirements in country j in year t for all species i;
- $N_{i,j,t}$ is the number of animals of species i in country j in year t; and
- $r_{i,j,t}$ is the average annual requirements of animals of species i in country j in year t.

The model accounts for the following seven animal species:

1. Cattle (buffalo and camels)
2. Sheep and goats
3. Pigs
4. Poultry (chickens, geese, ducks and turkeys)
5. Horses
6. Asses and mules
7. Rabbits

The average biological requirements (r) are compiled using a method based on models from the National Research Council of the Academy of Science of the United States of America (NRC)[24] and the Livestock Development Planning System tool (LDPS2) developed by FAO.[25] The final model is based on linear regressions and biological requirements are based on the need to produce the main livestock products (namely milk, eggs, and meat production), which in particular is included through the approximate weight gain. Accordingly, the average biological requirements are calculated as:

[24] The NRC has issued guidelines on nutrient requirements available at https://animalnutrition.org/nrc_reports. These have been published by the National Academies Press for a range of species, including beef (2016), dairy cattle (2001), dogs and cats (2006), fish and shrimp (2011), horses (2007), poultry (1994), small ruminants (2007) and swine (2012).
[25] See https://www.fao.org/4/x5878e/X5878e01.htm.

Equation 42: Average biological requirements

$$r^n_{i,j,t} = f(w^n_{i,j,t}, p^n_{i,j,t})$$

Equation 42 states that the requirements ($r^n_{i,j,t}$) in terms of macronutrient n, for the animal species i in country j and year t are a function of weight $w^n_{i,j,t}$ and production $p^n_{i,j,t}$.

9.3.2 Animal units

To facilitate the calculation of biological requirements and to make different animal species comparable, an *animal unit* is used. Animal units link the different classes of animals mentioned above to one chosen base unit in terms of a property that is to be captured (e.g. weight, feed consumption, energy requirement, monetary value) so that heads can be expressed in one single unit. Any given animal type may then be expressed as a share or multiple of the base unit or property that livestock share across species (e.g. one cow is equal to ten sheep or one cow is equal to two tonnes of feed).

For the purpose of estimating energy and protein demands per country and year, two units, the *energy animal unit* (AUE) and *protein animal unit* (AUP) are established for each set of species, country and year (Gerber *et al.*, 2013).

Estimation of animal units

First, the base unit is established using the Livestock Development Planning System base unit, which represents a dairy cow weighing 500 kg, producing 3 500 kg milk per year and calving every 13 months. This hypothetical cow is expected to require 35 600 MJ of metabolizable energy and 319 079 kg of crude protein per year. Those values represent the base unit consumption r^n_0. In the feed demand model, the animal unit index for species i and for macronutrient n (energy or protein) takes the following form:

Equation 43: Animal unit index

$$u^n_{i,j,t} = \frac{r^n_{i,j,t}}{r^n_0}$$

where:

- $u^n_{i,j,t}$ is the animal unit index of species i in country j in year t for macronutrient n;
- $r^n_{i,j,t}$ is the average annual requirements of species i in country j in year t for macronutrient n; and
- r^n_0 is the base unit consumption for macronutrient n.

These values are used in **Equation 41** and **Equation 42** yielding:

Equation 44: Biological requirements by species – version 2

$$R^n_{j,t} = r^n_0 \sum_{i=1}^{I} N_{i,j,t} \times u^n_{i,j,t}$$

where:

- $R^n_{j,t}$ refers to requirements in country j in year t for all species i for macronutrient n;
- r^n_0 is the base unit consumption for macronutrient n, i.e. 35 600 MJ and 319 079 kg of crude protein respectively;
- $N_{i,j,t}$ is the number of animals of species i in country j in year t; and
- $u^n_{i,j,t}$ is the animal unit index of species i in country j in year t for macronutrient n.

9.3.3 Accounting for forage and roughage: the intensity factor

Types of feed

In practice, animals are fed a variety of plant-based materials. In traditional pasture- and backyard-based production systems grazing, hay, table scraps and other waste products account for a large proportion of the animals' consumption, while in modern industrialized feed lots, oilmeals and grains have become increasingly important.

For the construction of SUAs/FBS, forage products and roughage are not of interest and have therefore not been included in the definition of feed use. However, it is necessary to determine the amount satisfied through compound feeds.

Intensity ratio and its relation to feed

To estimate the actual feed demand, the biological requirement is deflated by a factor representing the proportion of the requirements satisfied by consumption of forage or roughage, called the *intensity factor*. It can vary across countries and over time due to differences in animal characteristics, methods of animal husbandry and technology.

Levels of feeding intensity depend first on the type of livestock. Ruminants require larger shares of forage and roughage in their diet than pigs and chickens. Feeding intensity also changes over time, especially in economies with fast-growing per capita income and rising demand for animal products, where the intensity of livestock production tends to increase across all reared species. In contrast, in countries where the demand for animal products is more stable, and mostly saturated, little increase in intensity should be expected.

Modelling the intensity ratio in practice

With the Global Livestock Environmental Assessment Model (GLEAM),[26] data on the size of different production systems and the representative feed baskets of these systems, broken down by animal species and country, were collected for the first time in 2005. Therefore, to model a complete annual time series of the feeding intensity, the data provided by GLEAM for 2005 have been used as a starting point, as these allow the calculation of the proportion of forage and roughage in the total consumption of animals.

Additionally, for earlier years, namely 1981–1983 and 1991–1993, and for much broader geographic units (*geoeconomic regions*), a global study on the interaction between livestock and the environment (Seré, Steinfeld and Groenewold, 1996), provided estimates of the feeding intensity factors for global livestock production systems.

The two datasets have been harmonized, and the GLEAM dataset aggregated for each geoeconomic region to generate three cross-sections for 1981–1983, 1991–1993 and 2005, each consisting of the same seven regions. Using these data as a base, coefficients associated with livestock density for cattle and labour productivity in agriculture for other species are estimated using a (pseudo-) panel regression. The estimated coefficients have been used to extrapolate the country-level GLEAM data of 2005 to other years, based on the developments of the auxiliary variables in the individual countries.

Adjusting demand for feed by intensity ratio

Recalling **Equation 40** and **Equation 41**, the livestock demand for feed can be expressed for energy and protein as follows:

Equation 45: Demand for feed

$$D_{j,t}^{n,livestock} = r_0^n \sum_{i=1}^{I} N_{i,j,t} \times u_{i,j,t}^n \times IR_{i,j,t}$$

[26] See https://www.fao.org/gleam/en/.

where:

- $D_{j,t}^{n,livestock}$ is the demand for livestock of macronutrient n for feed in country j and year t;
- r_0^n is the base unit consumption for macronutrient n;
- $N_{i,j,t}$ is the number of animals of species i in country j in year t;
- $u_{i,j,t}^n$ is the animal unit index of species i in country j in year t for the macronutrient n; and
- $IR_{i,j,t}$ is the intensity factor of animals of species i in country j in year t and reflects the intensity of animal farming. It can take values between 0 and 1, where 1 indicates fully intensified production, in which animals are raised entirely on feeds (typically compound feed), and 0 indicates full use of forage and roughage.

9.3.4 Feed disaggregation procedure

Estimating feed use for a specific product also requires the overall domestic availability of feed products in a country. The model assumes that relatively scarce products are less likely to be used as feed than abundant ones. This notion is driving the *disaggregation of total feed dema*nd in the feed model.

This procedure does not apply to official figures, which are left unchanged and never overwritten. In practice, if a value for feed use is reported in a questionnaire, it will be subtracted from the total demand for feed (computed on the basis of the procedure described above), and only the remaining demand for feed is allocated to the other feed products on the basis of the energy and protein requirements.

Products and their potential for feed

FAO classifies the products of the SUAs into three categories depending on their potential to be used as feed:

- *not for feed* products, are all products that are evidently not used for livestock-feeding purposes, such as most processed food products, for example, butter or flour, and that have already been ruled out by the application of the feeding intensity factor;
- *feed-only* products, are all products that are exclusively used as feed, in particular all oilseed meals, oilmeal cakes that are generated as by-products in the vegetable oil production, as well as the by-products of grain processing, most importantly dried distillers grains and brans, except breakfast brans; and
- *potential feed* products, are all products that may be used, at least to some extent, also for other purposes than feed, in particular grains, tubers, pulses and fruits.

The *not for feed* products are not taken into account in the disaggregation of the estimated total feed demand. On the other hand, the amounts allocated to the *feed-only* products are equal to their respective domestic availability, as their utilization for purposes other than feed is considered to be negligible.

Determining the total feed demand to be allocated

Accordingly, the total energy and protein demand to be allocated among *potential feed* products is reduced by the available quantities of all *feed-only* products, transformed into energy and protein units.

In addition, the quantity of feed use coming from official figures is subtracted in this step, which gives the residual demand for energy and protein to be satisfied through the *potential feed* products:

Equation 46: Residual demand

$$RD_{j,t}^n = D_{j,t}^n - \sum_{fo} q_{fo,j,t} \times n_{fo} - OD_{j,t}^n$$

where:

- $RD_{j,t}^n$ is the residual demand in country j and year t for macronutrient n (energy or protein);

- $D_{j,t}^{n}$ is the feed demand in country j and year t for macronutrient n;
- n_{fo} represents the macronutrient conversion factor;
- $q_{fo,j,t}$ is the quantity of *feed-only* items in country j and year t; and
- $OD_{j,t}^{n}$ is the demand resulting from official figures in country j and year t for macronutrient n.

In **Equation 46**, the figures for feed use coming from official sources are first subtracted from $RD_{j,t}^{n}$, thus giving a second form of residual demand, for which the actual use of feed should be proportional to the relative availability of feed, measured in energy and protein units. For example, in a country where maize is relatively more abundant than wheat, more maize than wheat is expected to be fed. More precisely, available nutrients in potential feeds can be accounted for due to the dualistic demand estimation. This implies that a country will use relatively abundant energy-rich products more than energy-poor products. The same holds for protein and leads to a more accurate calculation of feed availability considering the macronutrient structure of animal diets.

Allocating potential feed demand among potential feed items (1/3): estimating availability shares

Potential feed items are allocated proportionally to the share of their availabilities; which is the availability of each potential feed product to the total availability of potential feed products in each country and year. In particular:

Equation 47: Availability share of a potential feed item

$$a_{pf,j,t}^{n} = \frac{q_{pf,j,t} \times n_{pf}}{\sum_{pf} q_{pf,j,t} \times n_{pf}}$$

where:

- $a_{pf,j,t}^{n}$ is the availability share of a *potential feed* item (pf) in country j and year t for macronutrient n (energy or protein);
- $q_{pf,j,t}$ is the available quantity of the given *potential feed* item (pf) in country j and year t for macronutrient n; and
- n_{pf} represents the macronutrient conversion factor for each *potential feed* item (pf).

The availability share of a potential feed product for macronutrient n ($a_{pf,j,t}^{n}$) is equal to the availability of this nutrient from the given potential feed product ($q_{pf,j,t} \times n_{pf}$) divided by the whole macronutrient supply generated from all potential feeds. The energy and protein nutrient conversion factors are expressed in megajoules of metabolizable energy (MJME) per kg and in percent of dry matter for energy and protein, respectively.

Equation 47 provides two different sets of availability shares: one for energy and one for protein availability. These can be applied to the residual feed demand to calculate the amount of feed demand met by each feed product.

Allocating potential feed demand among potential feed items (2/3): selecting a set of availability shares

As feed use of a product is determined using only one availability share, a decision rule to select between energy and protein availability shares is required. In that context, three different cases can be distinguished:

(i) one set of availability shares leads to sufficient energy and protein supply to cover the demand for both types of nutrients;
(ii) both sets of availability shares lead to sufficient energy and protein supply to cover the demand for both types of nutrients; and
(iii) none of the two sets of availability shares leads to sufficient energy and protein availability to cover the demand for both types of nutrients.

Case (i) is straightforward, and the allocation that meets both demands is chosen.

Cases (ii) and (iii) are more difficult as both sets of shares are equal in their ability to meet (or not) the demand for both types of nutrients.

For case (ii), since each set of availability shares meets the demand for its respective type of nutrient, i.e. the energy set meets the energy demand and the protein set meets the protein demand, a set of availability shares represented by the maximum values of each pair of shares must at least satisfy demand for both types of nutrients. Likewise, the set of the minimum values of each pair of shares will not satisfy both demands.

Therefore, the range of possible solutions is bounded by these sets of minimum and maximum shares. At this point, an allocation algorithm is applied that starts from the minimum amount and calculates the additional quantities of feed needed to satisfy the requirements, in accordance with the relative availability of nutrients in feedstuffs. This allocation is performed via a linear optimization mechanism that gives preference to nutrient-rich feeds to arrive at the optimal allocation that meets both demands by minimizing actual feed quantities.

For case (iii) the same rationale as for (ii) is applied, subtracting the maximum amount of feed to obtain the optimal allocation that meets both demands.

Allocating potential feed demand among potential feed items (3/3):final allocation

Having established a set of availability shares, the final allocation procedure is applied by multiplying the shares by the residual feed demand. These amounts are then converted into actual quantities of feed by dividing by the respective macronutrient content factor. The result is directly assigned to the *potential feed* as final utilization.

Equation 48: Feed quantity assignment

$$Fe_{pf,j,t} = \frac{a_{pf,j,t} \times RD_{j,t}^{n}}{n_{pf}}$$

where:

- $Fe_{pf,j,t}$ is the final quantity of feed for the *potential feed* item pf in country j and year t;
- $a_{pf,j,t}^{n}$ is the availability share of the *potential feed* item pf in country j and year t for macronutrient n (energy or protein);
- $RD_{j,t}^{n}$ is the residual demand in country j and year t for macronutrient n (energy or protein); and
- n_{pf} is the specific nutrient content factor of the *potential feed* item pf.

Example of feed disaggregation

In the example in **Table 37**, cattle has the highest demand. Camels and rabbits are not farmed, and the number of goats and mules is negligible. Therefore, their requirements are equal to zero when converted.

Table 37: Example of feed demand estimation

ANIMAL	NUMBER	AUE	AUP	IR	ENERGY DEMAND D^E (MJ)	PROTEIN DEMAND D^P (Mt)
Cattle	11 535 000.00	1.1323885	0.2868086	0.4822851	224 267 789 506	508 984.12
Buffalo	813 900.00	0.0650027	0.0741583	0.2356970	443 921 667	4 538.12
Goats	30 026.39	0.0000000	0.0000000	0.2356970	0	0.00
Pigs	13 935 000.00	0.2202864	0.1195400	0.9851761	107 661 025 314	523 509.74
Chickens	168 608 522.99	0.0129327	0.0321377	0.9645101	74 872 892 714	1 667 217.63
Turkeys	5 696 165.36	0.0757045	0.0796998	0.9645101	14 806 786 686	139 680.96
Geese	328 539.87	0.0375192	0.1059744	0.9645101	423 251 299	10 712.39
Ducks	1 459 196.93	0.0201131	0.0461497	0.9645101	1 007 743 656	20 719.54
Horses	399 345.27	0.6155621	0.4102391	0.4720046	4 130 633 317	24 667.35
Mules	4 000.00	0.0000000	0.0000000	0.4720046	0	0.00
				TOTAL DEMAND	427 614 044 159	2 900 029.85

Source: Authors' own elaboration.

Once the total energy and protein demand have been calculated, energy and protein from *feed-only* products and from official demand have to be subtracted from the total demand to obtain the residual demand in country j and year t expressed in energy and protein. For the given country and the given year, the amounts of residual and official energy and protein availability are reported in **Table 38**. The residual energy and protein demand are obtained from **Equation 46** as:

$$RD^{MJ} = 427\,614\,044\,159.00 - 65\,292\,460\,386.85 - 37\,003\,320\,000.00 = 325\,318\,263\,772.03$$

$$RD^{g} = 2\,900\,029.85 - 2\,130\,673.37 - 337\,856.40 = 431\,500.076$$

Table 38: Example of feed demand estimation

FEED-ONLY ENERGY AVAILABILITY	FEED-ONLY PROTEIN AVAILABILITY	OFFICIAL FEED ENERGY AVAILABILITY	OFFICIAL FEED PROTEIN AVAILABILITY	RESIDUAL ENERGY DEMAND	RESIDUAL PROTEIN DEMAND
65 292 460 386.85	730 673.37	37 003 320 000.00	137 856.40	325 318 263 772.03	2 031 500.08

Source: Authors' own elaboration.

These amounts are allocated to *potential feed* products proportionally to a share of their availabilities, through the formula expressed in **Equation 47** and via an optimization mechanism that gives preference to nutrient rich feeds in order to arrive at the optimal allocation that satisfies both demands by minimizing actual feed quantities.

Table 39: Example of feed demand estimation for potential feed products – final values

ITEM	FEED ALLOCATED (t)	ENERGY ALLOCATED (MJ)	PROTEIN ALLOCATED (Mt)
Maize (corn)	8 953 950.24	128 041 488 498.86	940 164.78
Sorghum	572.39	8 185 226.02	61.82
Barley	4 918 610.75	68 860 550 524.63	580 396.07
Rye	40 521.37	551 090 654.76	4 173.70
Oats	1 575 879.69	22 062 315 616.25	173 346.77
Millet	11 568.05	165 423 086.21	1 446.01
Triticale	32 179.43	440 858 208.75	3 764.99
Buckwheat	1 739.54	24 875 362.73	187.87
Mixed grain	1 626.93	23 265 141.34	162.69
Sugar cane	889.76	12 278 755.65	36.48
Buttermilk, dry	15.94	194 510.76	1.99
Dairy products n.e.s.	41.19	502 532.68	5.15
Germ of wheat	71 317.14	1 112 547 350.59	21 038.56
Germ of maize	308 199.34	4 438 070 432.94	46 538.10
Cereal preparations	1 044.16	14 931 455.13	104.42
Bran of oats	78 103.47	1 171 552 097.70	15 073.97
Bran of pulses	79 680.92	1 139 437 125.36	17 131.40
Cocoa husks and shells	8 860.53	131 135 849.64	832.89
…	…	…	…
TOTAL	23 629 374.52	325 318 263 772.03	2 031 500.08

Source: Authors' own elaboration.

10 FOOD LOSSES

$$production + imports - \Delta stocks = exports + food + food\ processing + feed + seed + tourist\ consumption + industrial\ use + \boldsymbol{loss} + residual\ use$$

10.1 Data sources

In the FBS context, the term *food losses* (or *loss*) refers, as defined in Section **2.1**, to post-harvest/post-slaughter loss; that is, to the "all quantity losses, direct and indirect, that occur during storage, transportation and processing, up to the retail/consumption level" (FAO, n.d.b). The reason is that food availability is defined at this point of the production chain. Accurately measuring or estimating loss is crucial, as not doing so could result in overestimating food availability or any other utilization. Furthermore, understanding losses can help countries identify problems in the production or supply chains and take action to maximize resource efficiency. Incorporating losses in the equations might lead to a revision of other variables, which may have been estimated net of losses.

Collecting data on food loss can be a complex and costly process. In fact, few countries report official data, as requested in the FAO production questionnaire. When official data are not available, information is estimated with a model that complements official information with alternative data sources and auxiliary data.

The alternative sources can be divided into two categories: (i) data collected from regional food loss databases managed by governments or research projects and (ii) data harvested from the available literature (peer-reviewed articles, grey literature, etc.). FAO standardizes, stores and regularly updates these data in the Food Loss and Waste database.[27] These data refer to food losses along the whole supply chain or at a particular stage of the food supply chain and constitute, with the official data, the core input values of the estimation model.

Auxiliary data provide additional information essential to improve food loss estimation and constitute the explanatory variables in the model. They refer to several causes of food loss and affect distinct parts of the food supply chain. The most important categories of auxiliary data include climate- and human-related factors. A country's climate affects yields and farm activities, storage duration, the need for special storage, and transport facilities. Human-related factors also play a crucial part in preventing or favouring post-harvest losses. Social, trade and income indicators provide good proxy measurements to help estimate structural food loss in a country. The choice of auxiliary variables has been driven by the connection to food loss and waste causes systematically identified by the High Level Panel of Experts in 2014 (HLPE, 2014), which identifies three levels of causes: micro-, meso- and macro-level.

10.2 Estimation

In the current FBS methodology, food losses are imputed based on the methodology for SDG 12.3.1a (Food Loss Index) (Taglioni, Rosero Moncayo and Fabi, 2023). Losses are modelled for all countries and specific products, taking into account a large set of potential explanatory variables. The model applies a regression tree to select a subset of eight explanatory variables. FAO runs the model annually to update the FBS and for SDG monitoring; each year a new set of explanatory variables is selected. The pre-processed loss data and explanatory data are the inputs to the food loss model. Preprocessing is a quality check and fills data gaps, which includes outlier detection and gap filling of the auxiliary (explanatory) variables. Another step in preprocessing includes the estimation of losses at stages in the food supply chain where no data are available and then the aggregation of the losses at each stage to the whole

[27] See https://www.fao.org/platform-food-loss-waste/flw-data/en/.

supply chain. The model applied is a random-effect model with the random effect by food group assuming that loss levels are similar among products of a similar nature. The food groups considered are: cereals and pulses; roots, tubers, and oil-bearing crops; fruits and vegetables; meat and animal products; and other products.

If a country has more than three data points available for the food group, the model estimates losses at the country level; otherwise, estimates are based on a global model that uses all available data from any country (**Figure 17**), plus a number of explanatory variables. Estimates that exceed three standard deviations from the median loss are re-estimated in a restricted model, which only includes the country, the product, and time as covariates. Resulting losses that exceed identified thresholds for a given food group are then winsorized.[28]

The model is based on pseudo-panel data, which controls for state- and time-invariant variables and reduces multicollinearity. It includes random effects capturing series-specific characteristics in the error component that are not explicitly included in the covariates. The random effect is specified as:

Equation 49: Random effect model

$$y_{i,j,t} = \alpha + x_{i,j,t}^T\beta + z_{i,j}^T\gamma + u_{i,j,t}$$

where:

- $y_{i,j,t}$ is the logit transform of the percentage of food losses for the country i, for a given product j, at time t;
- α is the intercept;
- $x_{i,j,t}^T$ is a row vector of time, country and product varying explanatory variables;
- $z_{i,j}^T$is a row vector of time-invariant dummy variables based on the indices i, j (country, product); and
- $u_{i,j,t}$is the composite error term.

Within the current specification, the error component u_{ijt} is defined as:

Equation 50: Error component of the random effect model

$$u_{i,j,t} = \mu_{i,j} + \nu_{i,j,t}$$

where:

- $\mu_{i,j} \sim N(0, \sigma_\mu^2)$ is the unobservable, time-invariant, country-product effect; and
- $\nu_{i,j,t} \sim N(0, \sigma_v^2)$ is the remainder disturbance or idiosyncratic error term.

The x_{ijt}^T are assumed independent of the error components μ_{ijt} and ν_{ijt}.

The logit transformation of the loss ratio to map the loss ratio naturally bounded in [0, 1] to (-∞, +∞).

Equation 51: Logit transformation of the random effect model

$$y = f(l) = ln\left(\frac{l}{1-l}\right) \quad and\ its\ inverse,\ \ l = \frac{e^y}{1 + e^y}$$

where l is the loss ratio, $l \in [0\ 1]$.

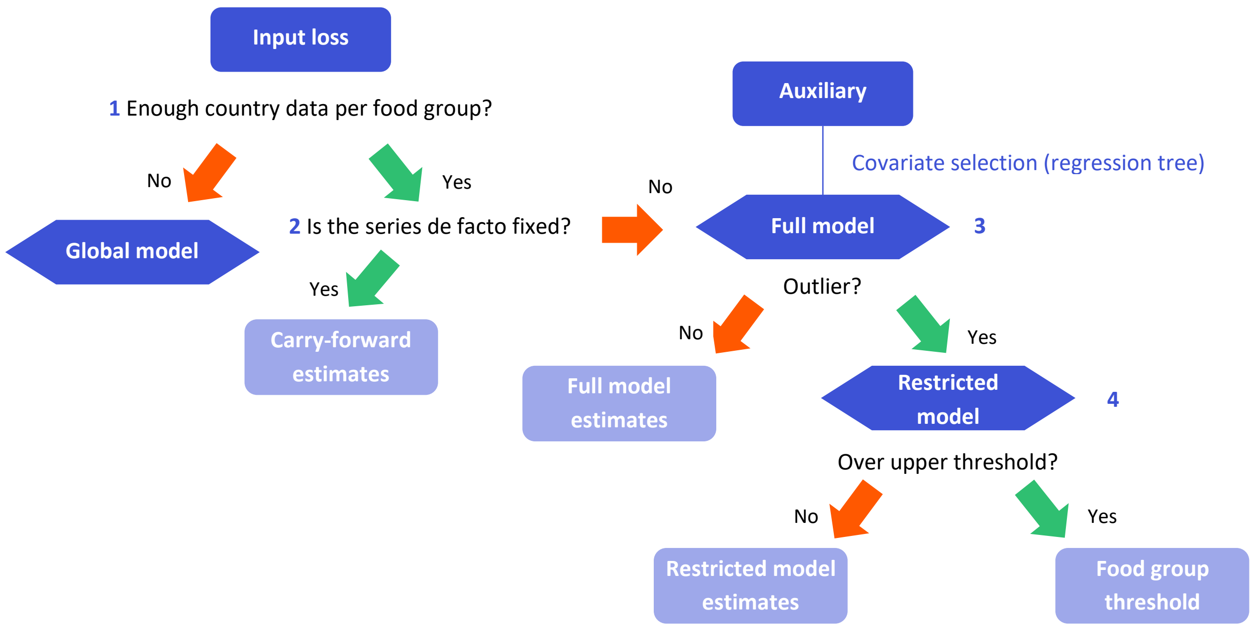
The model specified in **Equation 49** has the same form if the model runs at the country or the global level except for the coefficient referring to the geographical variable that refers, respectively, to the relevant country or SDG region. At the country level, product series are often almost fixed or constant loss factors over time (e.g. around

[28] Winsorization is a statistical technique used to mitigate the impact of outliers in a dataset by capping extreme values to a specified percentile. Instead of removing outliers outright, winsorizing adjusts these extreme values to a predetermined threshold, reducing their influence on statistical analysis without discarding them entirely.

5 percent loss percentages for the whole time series). In these cases, a simple carry-forward of the last value of the time series is applied, whereas for all the other series, the random effect model applies. Carry-forward is not applied in the global model.

At the end of the estimation process, final checks apply particularly for series containing official data. Imputed data for officially reported series must be consistent with the available official data; therefore, an a posteriori correction is applied if the difference between imputed losses and the latest official figure is higher than 5 percent in absolute value. In such a case, the model estimates are discarded and the series is re-imputed with ARIMA models or carry-forward/backward method, if ARIMA cannot be applied. In so doing, estimates and official figures are reconciled. Additional cross-checking and outlier detection are performed during the compilation of SUAs and FBS.

Figure 16: Food estimation model workflow at the country level



Source: Taglioni, C., Rosero Moncayo, J.R. & Fabi, C. 2023. *Food loss estimation: SDG 12.3.1a data and modelling approach.* FAO Statistics Working Paper Series, No. 23-39. Rome, FAO. https://doi.org/10.4060/cc9173en

The strength of this model relies on the wide range of data sources, the use of additional explanatory variables used as proxy for the country level of losses, the estimation by food group, and the possibility of adapting the geographical area considered (country or SDG region) depending on the data availability.

11 SEED

$production + imports - \Delta stocks = exports + food + food\ processing + feed + \mathbf{seed} + tourist\ consumption + industrial\ use + loss + residual\ use$

11.1 Data sources

FAO collects data for $seed$ and the area sown through the FAO production questionnaire. However, although overall response rates to the questionnaire have been rising, not all countries provide seed use estimates for all products. Where no official seed use information is available, seed use is imputed.

The current approach has been developed to properly exploit all the sources of information to produce a coherent and consistent output. Thus, the imputation methodology starts from an evaluation of the area sown.

11.2 Overall approach for imputation

The overall process can be summarized by two steps:, the area sown is first imputed starting from the area harvested, then seed use is estimated with a hierarchical linear model that uses as covariates the area sown, time and temperature. The rationale behind the inclusion of these covariates will be introduced and described below.

11.2.1 Imputation of area sown

Estimation occurs after the crop–production triplet of **Equation 4** has already been calculated as $production = area\ harvested \times yield$. For this reason official or imputed data on $area\ harvested$ are always available at this stage of the process. As discussed in Section **1.2**, the reference period chosen by FAO in compiling SUAs and FBS is the calendar year, therefore seed estimates are based on this time unit.

If previous values of the area sown and the area harvested are available, then an average ratio of the area sown in year t to the area harvested in year $t+1$ is computed.

Equation 52: Average ratio of the area sown

$$\overline{average\ ratio_t} = \frac{area\ sown_t}{area\ harvested_{t+1}}$$

If the area sown is unavailable in one year, it is imputed by multiplying the area harvested in the following year by the average ratio:

Equation 53: Imputation of the area sown

$$area\ sown_t = area\ harvested_{t+1} \times \overline{average\ ratio_t}$$

The link between $area\ sown$ and $area\ harvested$ ensures that the final seed use estimations are implicitly linked to the crop-production and yield. The idea is that the area sown is always equal to or greater than the area harvested. The ratio is multiplied by the available area harvested to obtain the area sown. The average is computed for country–product combinations. The ratio of the area sown and area harvested strictly depends on the product, taking into account environmental factors such as the climate and soil type.

When no prior information on the area sown is available, or the ratio is erroneously lower than 1, the $\overline{average\ ratio_t}$ is set equal to 1 and consequently, the area sown is assumed to be equal to the area harvested. This is an approximation that leads to feasible final imputations for the seed use component.

11.2.2 Seed use: the model

The imputation of $seed$ uses a hierarchical linear model. The rationale for this model is that it can capture and model complicated trends when data are available. Moreover, the hierarchy of the model allows accurate imputation for countries with very sparse data by pooling together global data. The model can be written as follows:

Level 1: Observational model (response variable)

The outcome variable $log(Se_{i,j,k})$ represents the log of the seed variable in country i, CPC group j and the unique country–code (i.e. country–product) combination k. This outcome is modelled as:

Equation 54: Seed use hierarchical model – level 1

$$log(Se_{i,j,k}) = \beta_0 + \beta_1 temp_i + \beta_2 time + \beta_{3,j,k} log(area\ sown_{i,j,k}) + \varepsilon_{i,j,k}$$

where:

- β_0 is the overall intercept;
- β_1 is the fixed effect of temperature $temp_i$ (provided by the World Bank) in country i;
- β_2 is the fixed effect of $time$;
- $\beta_{3,j,k}$ is a coefficient that varies by CPC group j and unique country–code combination k, representing the effect of the log of $area\ sown$; and
- $\varepsilon_{i,j,k}$ is the residual error term specific to each observation.

This level models the outcome variable as influenced by the country temperature, time and area sown.

Level 1: random coefficient model for $\beta_{3,j,k}$

The coefficient $\beta_{3,j,k}$ that captures the effect of $area\ sown$ in each CPC group j and unique country–code combination k is modelled as:

Equation 55: Seed use hierarchical model – level 2

$$\beta_{3,j,k} = \gamma_{30} + \gamma_{31,j,k}(CPC\ code\text{:}\ country)_{j,k} + \delta_{j,k}$$

where:

- γ_{30} is the average effect of $area\ sown$ across all CPC group j and unique country–code combination k;
- $\gamma_{31,j,k}$ is a coefficient that varies by CPC group j and unique country–code combination k, capturing the influence of the specific $(CPC\ code\text{:}\ country)_{j,k}$ on $\beta_{3,j,k}$; and
- $\delta_{j,k}$ is a random error term specific to each (j,k) combination, accounting for additional variability at this level.

This level allows the $area\ sown$ effect $\beta_{3,j,k}$ to vary depending on characteristics associated with each unique product and country combination.

Level 3: Hierarchical structure for $\gamma_{31,j,k}$

The coefficient $\gamma_{31,j,k}$, representing the influence of the $CPC\ code\text{:}\ country$ variable on the $area\ sown$ effect in each (j,k) combination, is modelled as:

Equation 56: Seed use hierarchical model – level 3

$$\gamma_{31,j,k} = \kappa_{310} + \kappa_{311}(CPC\ group)_j + \phi_j$$

where:

- κ_{310} is the overall mean effect of the CPC group predictor across all groups;
- κ_{311} represents how the $(CPC\ group)_j$ influences $\gamma_{31,j,k}$; and
- ϕ_j is a random effect term for each $(CPC\ group)_j$, that captures group-specific deviations.

The model estimates seed use proportional to the area sown. The model also accounts for changes over time and differences across countries; the latter are captured by the annual temperature variable, assuming that seed rates need to be higher where production conditions are difficult, with potential late and frequent frosts, but can be lower where production conditions are more favourable.

If data for a particular country and product are sparse, then κ_{310} and κ_{311} will likely be estimated to be close to 0. Thus, $\gamma_{31,j,k}$ will be close to its mean value, and the model will only account for availability in product groups. However, if data are available for a country or product, the estimates of κ_{310} or κ_{311} will be far from 0, and thus the model can adapt to the individual characteristics of a particular country–product combination.

12 TOURIST CONSUMPTION

$$production + imports - \Delta stocks = exports + food + food\ processing + feed + seed + \boldsymbol{tourist\ consumption} + industrial\ use + \text{loss} + residual\ use$$

Tourist consumption can constitute a significant part of the use of food products, particularly in small island states that are popular holiday destinations, such as Maldives or Seychelles. The measurement of tourist consumption should therefore not be neglected.

Data sources include surveys collecting detailed information on the different types of food consumed by tourists, which offer an efficient way of measuring food consumed by tourists. Other valuable data sources include the guest lists of hotels and records of the food purchased by the hotels. The quantities of food purchased when multiplied by the proportion of days spent by guests from other countries may be used as a second-best estimator of tourist consumption. Some countries conduct international passenger surveys that provide information on international tourism to the country and its trends. Information on the number of days spent by international tourists may also be obtained from travel agencies, airlines, railways and shipping companies.

The World Tourism Organization (UN Tourism)[29] collects data on the number of tourists entering a country, including their origins and the durations of their stays. The UN Tourism data set therefore provides another useful starting point for the compilation of data on tourist food consumption.

In the current FBS methodology, tourist consumption data are only computed for countries where this is a major component of the food consumption pattern. The current practice performs a manual imputation of tourist consumption based on past available data for countries. A refined methodology has been developed based on the number of day visitors and overnight visitors (N_O) to and from each country, and information on the average number of nights of stay per visitor in each country.

[29] See https://www.unwto.org/tourism-data/unwto-tourism-dashboard.

13 INDUSTRIAL USE

$$production + imports - \Delta stocks = exports + food + food\ processing + feed + seed + tourist\ consumption + \boldsymbol{industrial\ use} + loss + residual\ use$$

In recent decades, there have been significant increases in the use of agricultural products for industrial use. One key example is the use of feedstocks to produce energy, notably in the case of sugar crops, maize and oilseeds; moreover, starch-rich and oil-rich products including coconuts or palm kernels are channelled into the production of cosmetics, paints, soaps and other detergents. Some starch-based products have become increasingly important as construction materials.

Unfortunately, information about these uses is scarce, and there is no straightforward way to estimate or impute the amount of products involved. Information on industrial uses is asked to FAO Members through the FAO questionnaire on production; however, very few countries seem to have the information available, except on the use of agricultural feedstocks for biofuels production.

Some data can be available in product-specific bulletins and publications, and from the USDA PS&D database mentioned in Section **8.1**.

14 BALANCING METHODOLOGY

14.1 Introduction

Chapters 4 to 13 explained the variables of the base identity introduced in **Equation 1**. The next step in generating FBS involves balancing the SUAs that have just been produced. This balancing process is carried out on the SUAs before initiating the FBS compilation processes of standardization and aggregation to the primary commodity level. Analysts need to ensure that the SUAs, including both primary and derived product accounts, are balanced for each product (or line). To achieve this, a calculation is performed at this stage to verify that the 'supply equals utilization' identity holds true for each product. If discrepancies are identified, a balancing mechanism is applied. Balancing the equation may not be straightforward for various reasons. Most countries only collect data directly for the supply-side variables, while the demand-side variables are estimated using statistical models or by subject matter experts. This can result in an accumulation of measurement errors affecting the accuracy of and increasing uncertainty in demand-side estimates. Even in rare cases where all variables are measured independently, discrepancies in data sources, collection methods, reference periods and measurement errors could result in an imbalanced equation. As a result, an overall strategy for balancing the equation needs to be found.

In the current FBS methodology, balancing is performed through an iterative approach that systematically allocates imbalances to several variables, either entirely or in part, through a proportional assignment.

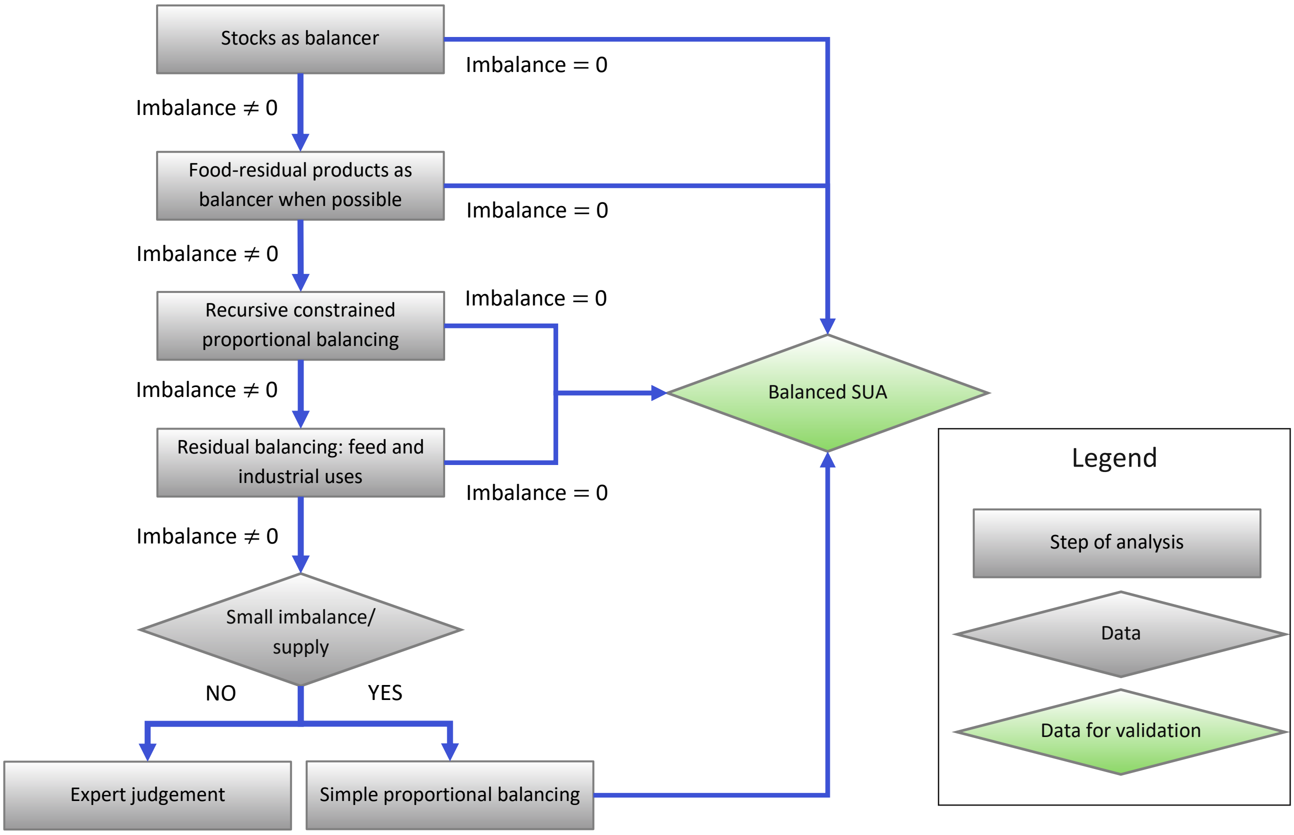
This marks a significant change and improvement to the methodology, by spreading the imbalance among several variables, rather than allocating it only to a single variable, thus avoiding the accumulation of measurement errors in the variable that is chosen as the balancing item.

14.2 Balancing workflow

The overall workflow for balancing is described in **Figure 17**.

After calculating the initial imbalance, the process begins by using stocks to balance the equation. The use of stocks as a balancing element is constrained by certain limitations (see Section **14.2.1**). When a balance is still not achieved, estimated quantities for food-residuals products (i.e. products that are consumed but not processed, as described in Section **7.2**) are adjusted to ensure that figures remain positive. For any remaining imbalance, estimates of other variables in the SUA equation can be proportionally adjusted based on their initial values, following specific rules that will be explained in Section **14.2.3**. If a residual imbalance persists, estimated figures for industrial and feed use may also be adjusted according to additional rules. As a final step, the ratio between the imbalance and total supply is checked. When this ratio falls below a predefined threshold, considered to be minimal, a process known as s*imple proportional balancing* is applied (see Section **14.2.5**). Conversely, if the ratio is deemed significant, expert judgment is required to decide whether the imbalance can remain unresolved or if further adjustments are necessary.

Figure 17: Balancing workflow



Source: Authors' own elaboration.

14.2.1 Stocks as a balancing item

Stocks variation is the first element the balancing procedure uses to assign an imbalance between supply and utilizations. Stocks variation is the difference in stocks between one period and another.

Equation 57: Stocks as a balancing Item

$$\Delta stocks^{IMB} = \Delta \widehat{stocks} + imbalance$$

where:

- $\Delta \widehat{stocks}$ is the stocks variation as estimated via the methodology described in Section 8.2; and
- $imbalance = total\ supply - total\ utilization.$

To avoid extreme variations in $\Delta stocks$, a threshold is introduced, defined as:

Equation 58: Threshold for estimation of stocks variation during balancing

$$\Delta stocks^{IMB}_{max} = max\left(\frac{\left|\Delta stocks^{validated}\right|}{mean\left(supply^{validated}_{inc}\right)}\right)$$

where:

- $\left|\Delta stocks^{validated}\right|$ is the absolute value for validated stocks variations;
- $supply_{inc} = production + imports - exports - \Delta stocks_t$ (Equation 35); AND
- $mean\left(supply^{validated}_{inc}\right)$ is the average $supply_{inc}$ for the last three validated years.

The final estimation of $\Delta Stocks$ will be:

Equation 59: Final estimation of stocks variation during balancing

$$\Delta\widehat{stocks} = \begin{cases} \left|\Delta stocks^{validated}\right| \times mean\left(supply_{inc}^{validated}\right) \times sign(\Delta stocks^{pred}) & if \;\; \Delta stocks^{pred} > \Delta stocks_{max}^{validated} \\ \Delta stocks^{pred} & otherwise \end{cases}$$

Moreover, two constraints on the estimated stocks variations are implemented:

- If $\Delta stocks$ is negative (i.e. stocks withdrawal), it cannot exceed the accumulated opening stocks; and
- If $\Delta stocks$ is positive (i.e. stocks accumulation), a limit to the amount accumulated is set.

14.2.2 Food-residual products as balancing item

After stocks have been assigned to all or part of the imbalance, some imbalance could remain. Food-residual products are goods for which food is the only utilization (and generally products that cannot be stocked), as defined by the FAO methodology (e.g. wheat is not a food-residual product but bread is). If $food$ is the only utilization, and the item is a food item, the imbalance is assigned to $food$ with an amount that keeps food positive:

Equation 60: Food residuals computation during balancing

$$food = \begin{cases} net\ supply & if\ net\ supply > 0 \\ 0 & if\ net\ supply < 0 \end{cases}$$

where $net\ supply = production + imports - exports$.

If the imbalance is not completely absorbed after assigning the imbalance to food, proportional balancing is used.

14.2.3 Recursive constrained proportional balancing

Proportional balancing is a procedure that apportions $supply - utilizations$ imbalances to different utilizations (e.g. food supply, industrial utilization, other uses). Only some utilization variables can be modified by the balancing, notably those obtained as estimates through statistical models. Other data are "protected" in this step; in particular:

- Official and unofficial figures;
- The variables $food\ processing$, $seed$ and $loss$:
  - $food\ processing$ is not modified as it is linked to the supply of parent commodities, which at the balancing level is considered as given;
  - $seed$ remains as estimated since it is linked to the production of the country: it would be inconsistent to unlink it just for balancing the SUA; and
  - $loss$ figures, because they are connected to production (similarly to $seed$); and
- imports (for import-dependent countries), for the same reason adopted for $seed$.

Proportional balancing considers that if the SUA is unbalanced, and it is not evident which of the estimated components is more reliable (food, feed, industrial use, tourist consumption), the imbalance is distributed proportionally across all estimated variables that are not protected. The proportion of the imbalance assigned to each variable is equal to the variable's share of the total of utilizations that can be used for balancing. For instance, if a SUA has three utilizations, among which one of them is protected (for example because the data are official), the imbalance is going to be assigned to the non-protected variables with shares equal to:

Equation 61: Share of a single element in simple proportional balancing

$$s_i = \frac{variable_i}{\sum_{n=1}^{i} variable_n}$$

where:

- $variable_i$ is a non-protected variable; and
- n are the utilizations that can be used for balancing.

Thus, the new (post-balancing) value for a non-protected element is:

Equation 62: Post balancing value for a generic non-protected variable in simple proportional balancing

$$variable_i^{IMB} = variable_i + (s_i \times imbalance)$$

Setting constraints

The *simple proportional balancing* provides a resulting value that is compared to constraints using the *constrained proportional balancing* methodology. In this methodology, constraints are imposed on the maximum and minimum values of $variable_i^{IMB}$ so that the resulting allocations remain in line with the historical trends (i.e. avoiding outliers).

Moreover, the shares used in balancing are not defined based on the current year, but as a moving average (frequently set as a 3-year moving average) $\tilde{s}_i$ of the lag of the ratios r^a of each element i over the total availability of non-protected variables in validated years t:

Equation 63: Share of a single variable in constrained proportional balancing

$$\tilde{s}_i = \frac{\sum_{j=1}^{3} r_{i,t-j}^a}{3}$$

with

Equation 64: Non-protected availability ratio of a single variable in constrained proportional balancing

$$r_{i,t}^a = \frac{variable_{it}}{\sum_{n=1}^{i} variable_{nt}}$$

where i indicates a non-protected element and $t \in (recent\ validated\ years)$.

Constraints for $variable_i^{IMB}$ are obtained by first deriving the maximum and minimum values of the ratio to supply:

Equation 65: Minimum and maximum ratios of a single variable in constrained proportional balancing

$$r_i^{s,min} = \min_t \left(\frac{variable_{i,t}}{supply_t}\right)$$
$$r_i^{s,max} = \max_t \left(\frac{variable_{i,t}}{supply_t}\right)$$

where:

- i indicates a non-protected element; and
- $supply_t$ is the total supply over recent validated years (t), thus the maximum or minimum only uses validated data from recent years as given by $production + imports$.

Then, to set the maximum or minimum value of $variable_i$, the shares are multiplied by $supply_T$ (i.e. the year for which the FBS are compiled: as explained above, the shares are not calculated on the current year):

Equation 66: Minimum and maximum values in constrained proportional balancing

$$variable_{i,T}^{min} = r_i^{s,min} \times supply_T$$
$$variable_{i,T}^{max} = r_i^{s,max} \times supply_T$$

Given the thresholds above, the new value for $variable_i$ will be:

Equation 67: Post-balancing value for a generic non-protected variable in constrained proportional balancing

$$variable_{i,T}^{IMB} = \begin{cases} variable_{i,T}^{min} & if\ variable_{i,T} + \tilde{s}_i \times imbalance_T < variable_{i,T}^{min} \\ variable_{i,T}^{max} & if\ variable_{i,T} + \tilde{s}_i \times imbalance_T > variable_{i,T}^{max} \\ variable_{i,T} + \tilde{s}_i \times imbalance_T & otherwise \end{cases}$$

The constrained proportional balancing just discussed is *recursive* and stops if the imbalance reaches zero or if all variables have reached their minimum or maximum values or after ten loops. In that case some imbalance could still remain, but there is no more space to modify the variables via proportional balancing, so the SUA is kept unbalanced. This is another improvement in the new methodology: the different variables are not forced to be set to unlikely values in order to remove the imbalance.

If after these checks, an imbalance still remains, it is assigned to a variable called *residual other utilizations*. However, before leaving the residual, other attempts are made to balance, depending on the characteristics of the item or the magnitude of the remaining imbalance. These will be discussed after an example of the recursive constrained proportional balancing.

14.2.4 An example of proportional balancing

A working example is presented with fictional data reported in **Table 40**. For simplicity, exports and stocks variations are considered to be zero, while there are only three utilizations: food, feed and food losses.

Table 40: Pre-balancing values in constrained proportional balancing example

YEAR	SUPPLY		TOTAL SUPPLY	PRE-BALANCING UTILIZATION			TOTAL UTILIZATION	IMBALANCE
	PRODUCTION	IMPORTS		FOOD	FEED	FOOD LOSSES		
1	**1 200**	**300**	1 500	1 200	225	**75**	1 500	0
2	**1 100**	**250**	1 350	945	324	**81**	1 350	0
3	**980**	**350**	1 330	931	333	**67**	1 330	0
4	**1 100**	**200**	1 300	1 040	156	**104**	1 300	0
5	**1 500**	**250**	1 750	1 575	53	**123**	1 750	0
6	**1 450**	**300**	1 750	1 225	403	**123**	1 750	0
7	**1 800**	**350**	2 150	1 505	495	**151**	2 150	0
8	**1 900**	**200**	2 100	1 680	294	**126**	2 100	0
9	**1 850**	**250**	2 100	1 680	273	**147**	2 100	0
10	**2 000**	**300**	2 300	1 300	100	**161**	1 561	739
11	**2 000**	**300**	2 300	1 000	500	**161**	1 661	639
12	**2 000**	**300**	2 300	1 000	200	**161**	1 661	639

Source: Authors' own elaboration.

In this example years 10 to 12 are the years for which the FBS are compiled, while years 1 to 9 are validated years. The "recent validated years" are the previous five years with respect to the first year of the new FBS compilation, i.e. years 5 to 9. In the table, protected data are displayed in bold and correspond to production, imports and losses. For simplicity, production and imports are set equal for years 10 to 12, but in real data this is not the case.

The balancing for years 10 to 12 requires, first of all, the calculation of the moving averages $\tilde{s}_i$ of the ratios of each variable over the total non-protected variables for uses (**Equation 63** and **Equation 64**). The values of $\tilde{s}_i$ for the current example are in **Table 41**. For food these shares are:

$$\tilde{s}_{food} = \frac{r_{food,7}^{a} + r_{food,8}^{a} + r_{food,9}^{a}}{3} = \frac{\frac{1\,505}{1\,505+495} + \frac{1\,680}{1\,680+294} + \frac{1\,680}{1\,680+273}}{3} = \frac{0.75+0.85+0.86}{3} = 0.82$$

Initial balancing values for non-protected variables can now be imputed as the product of the moving averages $\tilde{s}_i$ and the imbalance:

$$variable_{i,T}^{IMB*} = variable_{i,T} + \tilde{s}_i \times imbalance_T$$

Where the "*" indicates that these are initial values to be tested against thresholds. The initial balancing values for food and feed in year 10 are:

$$food_{10}^{IMB*} = food_{10} + \tilde{s}_{food} \times imbalance_{10} = 1\,300 + 0.82 \times 739 = 1\,907$$

$$feed_{10}^{IMB*} = feed_{10} + \tilde{s}_{feed} \times imbalance_{10} = 100 + 0.18 \times 739 = 232$$

All values for the current example are reported in **Table 41**:

Table 41: Calculation of balancing shares in constrained proportional balancing example

		PRE-BALANCING					AVAILABILITY RATIOS		MOVING AVERAGE		INITIAL BALANCING VALUES	
YEAR	SUPPLY	FOOD	FEED	FOOD LOSSES	UTILIZATION	IMBALANCE	$r_{food,t}^{a}$	$r_{feed,t}^{a}$	$\tilde{s}_{food}$	$\tilde{s}_{feed}$	$food_T^{IMB*}$	$feed_T^{IMB*}$
1	1 500	1 200	225	**75**	1 500	0	0.84	0.16	-	-	-	-
2	1 350	945	324	**81**	1 350	0	0.74	0.26	-	-	-	-
3	1 330	931	333	**67**	1 330	0	0.74	0.26	-	-	-	-
4	1 300	1 040	156	**104**	1 300	0	0.87	0.13	0.77	0.23	-	-
5	1 750	1 575	53	**123**	1 750	0	0.97	0.03	0.78	0.22	-	-
6	1 750	1 225	403	**123**	1 750	0	0.75	0.25	0.86	0.14	-	-
7	2 150	1 505	495	**151**	2 150	0	0.75	0.25	0.86	0.14	-	-
8	2 100	1 680	294	**126**	2 100	0	0.85	0.15	0.82	0.18	-	-
9	2 100	1 680	273	**147**	2 100	0	0.86	0.14	0.79	0.21	-	-
10	2 300	1 300	100	**161**	1 561	739	-	-	0.82	0.18	1 907*	232*
11	2 300	1 000	500	**161**	1 661	639	-	-	0.82	0.18	1 525*	614*
12	2 300	1 000	200	**161**	1 661	639	-	-	0.82	0.18	1 525*	314*

Source: Authors' own elaboration.

To compare the balancing values against thresholds, first the ratios over supply are computed and then, their minimum and maximum values are identified (**Equation 65)**. **Table 42** reports all the ratios for non-protected utilizations, with the ratio for $food$ in year 2 computed as:

$$r_{food}^{s} = \frac{food_2}{supply_2} = \frac{945}{1350} = 0.7$$

Table 42: Calculation of supply ratios in constrained proportional balancing example

	SUPPLY		PRE-BALANCING						SUPPLY RATIOS	
YEAR	PRODUCTION	IMPORTS	SUPPLY	FOOD	FEED	FOOD LOSSES	UTILIZATION	IMBALANCE	r_{food}^{s}	r_{feed}^{s}
1	**1 200**	**300**	1 500	1 200	225	**75**	1 500	0	0.80	0.15
2	**1 100**	**250**	1 350	945	324	**81**	1 350	0	0.70	0.24
3	**980**	**350**	1 330	931	333	**67**	1 330	0	0.70	0.25
4	**1 100**	**200**	1 300	1 040	156	**104**	1 300	0	0.80	0.12
5	**1 500**	**250**	1 750	1 575	53	**123**	1 750	0	0.90	0.03
6	**1 450**	**300**	1 750	1 225	403	**123**	1 750	0	0.70	0.23
7	**1 800**	**350**	2 150	1 505	495	**151**	2 150	0	0.70	0.23
8	**1 900**	**200**	2 100	1 680	294	**126**	2 100	0	0.80	0.14
9	**1 850**	**250**	2 100	1 680	273	**147**	2 100	0	0.80	0.13

Source: Authors' own elaboration.

Once all the supply ratios for validated years 1 to 9 are computed, it is possible to compute the maximum and minimum values for utilizations in years 10 to 12. Given that supply is the same for these years, the minimum and maximum values are also the same. For instance, the minimum value for $food$ in year 10 from **Equation 66** is given by:

$$food_{10}^{min} = r_{food}^{s,min} \times supply_{10} = 0.7 \times 2\,300 = 1\,610$$

where:

- $r_{food}^{s,min}$ is the minimum value of the column $r_{food}^{s,}$ of **Table 42**; and
- $supply_{10}$ is the supply for year 10 as reported in **Table 41**.

The value for $feed$ is obtained with the same approach and these values are reported in **Table 43**.

Table 43: Minimum and maximum values of utilizations in constrained proportional balancing example

	MINIMUM		MAXIMUM	
	SHARE	VALUE	SHARE	VALUE
FOOD	0.70	1 610	0.90	2 070
FEED	0.03	69	0.23	529

Source: Authors' own elaboration.

Finally, the initial balancing values reported in **Table 41** are tested against the thresholds. For year 10, given that 1 610 < 1 907 < 2 070 and 69 < 232 < 529, i.e. $food$ and $feed$ all fall within the minimum and maximum thresholds, the values obtained can be used as the values for the variables.

This does not happen in years 11 and 12, where at least one element, once a proportion of the imbalance is assigned, falls outside of the thresholds. This is shown in rows marked (a) for both years in **Table 44**.

When the constrained proportional balancing is activated in rows marked (b), variables are set to the minimum or maximum (highlighted values). In year 11 – line 11(b) – the new values were enough to completely remove the imbalance.

Table 44: Post-balancing example

		PRE-BALANCING					MA		POST-BALANCING			
YEAR	SUPPLY	FOOD	FEED	FOOD LOSSES	UTILIZATION	IMBALANCE	$\tilde{s}_{food}$	$\tilde{s}_{feed}$	$food_T^{IMB*}$	$feed_T^{IMB*}$	FOOD LOSSES	IMBALANCE
…	…	…	…	…	…	…	…	…	…	…	…	…
10	2 300	1 300	100	**161**	1 561	739	0.82	0.18	1 907	232	**161**	0
11(a)	2 300	1 000	500	**161**	1 661	639	0.82	0.18	1 525	614	**161**	0
11(b)	2 300	1 000	500	**161**	1 661	639	0.82	0.18	1 610	529	**161**	0
12(a)	2 300	1 000	200	**161**	1 661	639	0.82	0.18	1 525	314	**161**	300
12(b)	2 300	1 000	200	**161**	1 661	639	0.82	0.18	1 610	314	**161**	215
12(c)	2 300	1 000	200	**161**	1 661	639	0.82	0.18	1 610	529	**161**	0

Note: Availability ratios are omitted. MA: moving average

Source: Authors' own elaboration.

In year 12 – line 12(b) – an imbalance of 215 still remains. The constrained proportional balancing is re-activated; but now the variable $food$ is one of the protected variables (as its threshold has been reached). Therefore the only utilization that can be used is $feed$, so the imbalance will be assigned to $feed$. The resulting value is 529 (314 + the remaining imbalance of 215 that is at the allowed maximum), and the SUA is balanced.

However, if the maximum allowed for $feed$ was 400, instead of 529, the SUA would have remained with an imbalance of 129, and it would have been absorbed by the variable *residual other utilizations*.

14.2.5 Residual balancing

The workflow for dealing with imbalances, as described above, can either eliminate the imbalance or leave some non-null residual, which is assigned to *residual other utilizations*. There are specific cases for which a variable can still be used as a balancer to absorb an imbalance left after the proportional balancing. First, if the median of the previous years for $industrial\ use$ is greater than zero, the imbalance can be assigned to this variable so that the new value is positive. Second, if the median of the previous years for $feed$ is greater than zero, and the other utilization variables are zero, the imbalance can be assigned to this variable so that the new $feed$ value is positive.

14.2.6 Balancing of remaining small imbalances

Simple proportional balancing "without limits" is done for the small remaining imbalances, that is for imbalances that account for a small amount of supply. This amount is set to a threshold based on a percentage that can be considered small and would thus not create much change to the variables. Imbalances greater than this percentage remain and analysts need to review the SUA to check for data consistency or external sources of information.

15 STANDARDIZATION

15.1 Overview

This Section focuses on standardization, which is the expression of products in equivalent primary commodities. It is the final step in the FBS methodology, during which SUAs are transformed into FBS so that commodities can be presented at an aggregated level.

Before standardization, the SUA needs to be balanced as described by the identity in Section **1.1**, which ensures that total supply equals total utilization. In essence, the various processing steps in the food chain generate numerous derived or processed products, which must be "rolled up", that is, converted into their primary equivalents for each SUA component (the variables in the identity). There are some exceptions: the variables under $production$ are not standardized as they are already represented as primary commodities and derived products do not have uses of seed and losses. The other components and variables need to be expressed in the same units in order to be standardized.

For example, wheat and products in the SUA come in many forms: imports or exports of wheat may be in the form of wheat, but also in the form of different wheat products such as flour (first level of processing), bread or pastry (second level of processing) or even more processed forms. Some products and variables of the SUA identity may be exclusively available in their processed forms: $food$ of wheat, for example, only exists in the form of flour, flour products such as bread, noodles, pastry or biscuits, and other derived products such as bulgur or cracked wheat. Wheat is very rarely eaten as such, and the same holds for all other cereals and many other primary products. However, in the food balance for "wheat and products" these items are reported in terms of wheat as a primary commodity and therefore need to be standardized.

The standardization reports different products in a common unit, which allows processed products to be added up and expressed in their primary product equivalents. The methodology varies depending on the SUA variable to be standardized, as shown below.

15.2 Commodity types in the standardization process

The process of "rolling up" products into their primary equivalent is based on the commodity tree for each primary product, as introduced in Section **2.1**. This process entails removing the processing level, computing and combining extraction rates (Section **5.3.1**) and shares (Section **5.3.2**) along each branch, and appropriately defining the weight of products.

The weight parameter is crucial. It can be equal to 0 or 1: for each process only one derived product can have a weight equal to 1. Therefore, when more than one product can be obtained from a process, these products are called *joint-products* (Section **15.2.2**) and only one of the joint-products will have its weight equal to 1, while all other products will be assigned a weight equal to 0. The rationale behind this value is connected to the standardization process, as it indicates whether the derived product should undergo standardization or not. Products with weight equal to zero are called *zero-weight* products and represent one of the special cases of the standardization process.

Another special case is represented by *proxy-primary* commodities (Section **15.2.1**). Proxy-primary commodities are not standardized and are treated as primaries commodities. For example, soyabean oil at the SUA level is a derived product that it is treated as a proxy-primary and represented in the FBS separately as a commodity.

15.2.1 Proxy-primary commodities

Proxy-primary commodities are derived products for which FBS are compiled because of their importance as staple foods. Indeed, while FBS are mainly produced only for primary commodities, products such as vegetable oils, raw sugar, alcohols and beverages are treated as if they were primary commodities and are, therefore, not converted to primary commodity equivalents. These commodities are "cut" from the tree of the primary commodities and, if they can be processed in other products, have their own commodity tree. However, one variable, $production$, is an exception; the production quantity of proxy-primary commodities will be converted into primary equivalent and then assigned to the food processing component of their corresponding parents. For example, soyabean oil production is converted into the primary equivalent (quantity of soyabean); however this quantity is assigned to $food\ processing$ of soyabeans not to soyabean oil. Other variables are not standardized.

15.2.2 Joint-products

Joint-products are processed products that are derived from the same process that generates one or more other products. The standardization of all products derived from the same production process would be a multiple counting of the amounts of primary product used in that process. Therefore, these products are assigned a weight of zero in the standardization, and are called zero-weight products so that their (already counted) corresponding "parent equivalent" quantity is not counted more than once. While their quantities are not considered, their calories are, because all calories have to be considered in the final FBS. For this reason, zero-weight products are sometimes referred to as "calories-only" products.

Some examples of zero weight products are reported in **Table 45**.

Table 45: Examples of zero-weight products

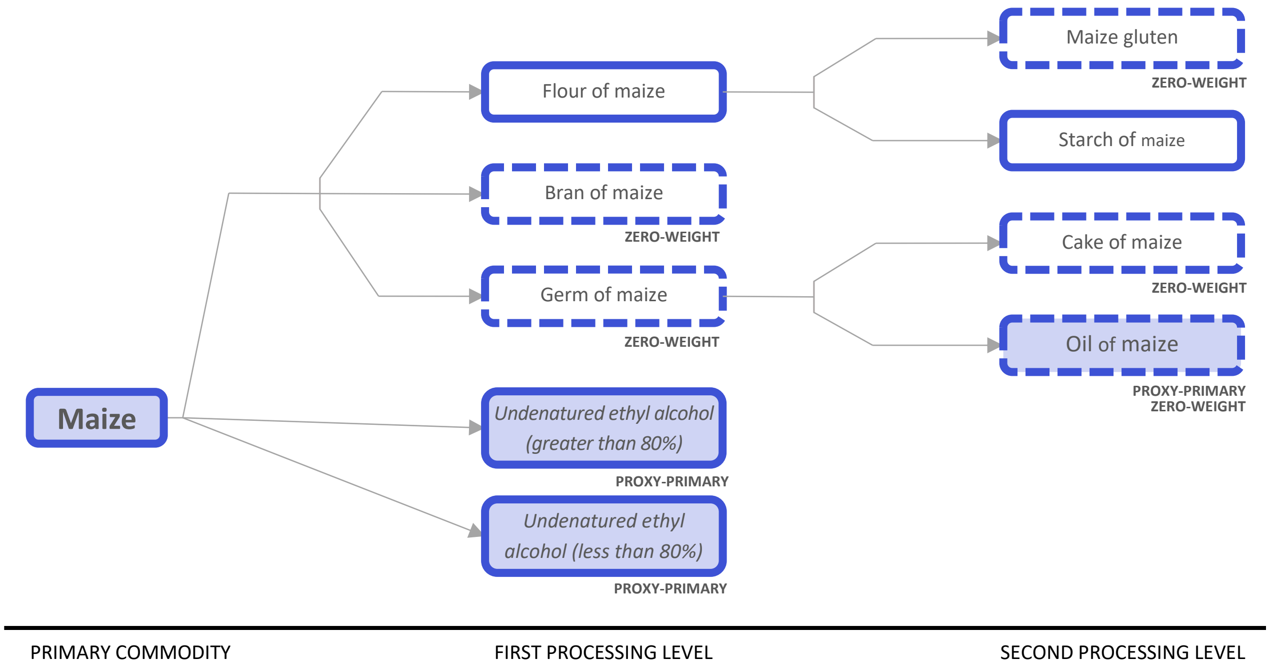
ZERO-WEIGHT PRODUCT	CO-PRODUCT
Barley flour and grits	Barley, pearled
Bran of barley	Pot barley
Bran of maize	Flour of maize
Bran of millet	Flour of millet
Bran of mixed grain	Flour of mixed grain
Bran of oats	Oats, rolled
Bran of pulses	Flour of pulses
Bran of rice	Rice, milled
Bran of rye	Flour of rye
Bran of sorghum	Flour of sorghum
Cake of cottonseed	Cottonseed oil
Cake of groundnuts	Groundnut oil
Cake of linseed	Oil of linseed
Cake of palm kernel	Oil of palm kernel
Cake of rapeseed	Rapeseed or canola oil, crude
Cake of sesame seed	Oil of sesame seed
Cake of soya beans	Soyabean oil
Cake of sunflower seed	Sunflower-seed oil, crude
Cake of copra	Coconut oil
Cake of maize	Oil of maize
Cocoa butter, fat and oil	Cocoa powder and cake
Cotton lint, ginned	Cottonseed
Cotton linters	Cottonseed oil
Marc of grape	Must of grape
Olive residues	Olive oil
Rice, gluten	Starch of rice
...	...

Source: Authors' own elaboration.

15.2.3 An example: the standardization of maize

The commodity tree of maize is used as an example of standardization and shows how the derived products (at the processing level) relate to the primary commodity (**Figure 18** and **Table 46**). In this tree, the primary commodity is maize and there are two processing levels.

Figure 18: Maize commodity tree



Source: Authors' own elaboration.

Table 46: Maize commodity tree with standardization coefficients

PARENT	CHILD	PROCESSING LEVEL	EXTRACTION RATE	BOTTOM-UP SHARES	WEIGHT	NON-STANDARDIZED
Maize	Flour of maize	1	0.82	1	1	FALSE
Maize	Bran of maize	1	0.11	1	0	FALSE
Maize	Undenatured ethyl alcohol (< 80%)	1	0.35	0.7	1	TRUE
Maize	Undenatured ethyl alcohol (> 80%)	1	0.65	0.4	1	TRUE
Maize	Germ of maize	1	0.06	1	0	FALSE
Germ of maize	Oil of maize	2	0.45	1	0	TRUE
Germ of maize	Cake of maize	2	0.52	1	0	FALSE
Flour of maize	Starch of maize	2	0.75	1	1	FALSE
Flour of maize	Maize gluten	2	0.10	1	0	FALSE

Source: Authors' own elaboration.

At the first processing level for the derived products of undenatured ethyl alcohol, the bottom-up shares of undenatured ethyl alcohol (less than 80 percent) and of undenatured ethyl alcohol (greater than 80 percent) are less than 1, therefore these two products are "multiple-parent": they have more than one parent and the other parent is a product other than maize. As stated in Section 5.3.3 a bottom-up share is the part of the production of a derived, or child, product from a particular parent commodity; these shares are equal to 1 when the derived product is produced from only one parent commodity and less than 1 when there is more than one parent. Therefore, only a part of the quantity of these alcohols will be considered when standardizing and aggregating to maize. Moreover, both products are considered proxy-primary commodities because of their importance as staple food. Therefore, they will not be standardized to maize (TRUE in the column Non-standardized of **Table 46**), and they will be cut from this commodity tree and considered separately.

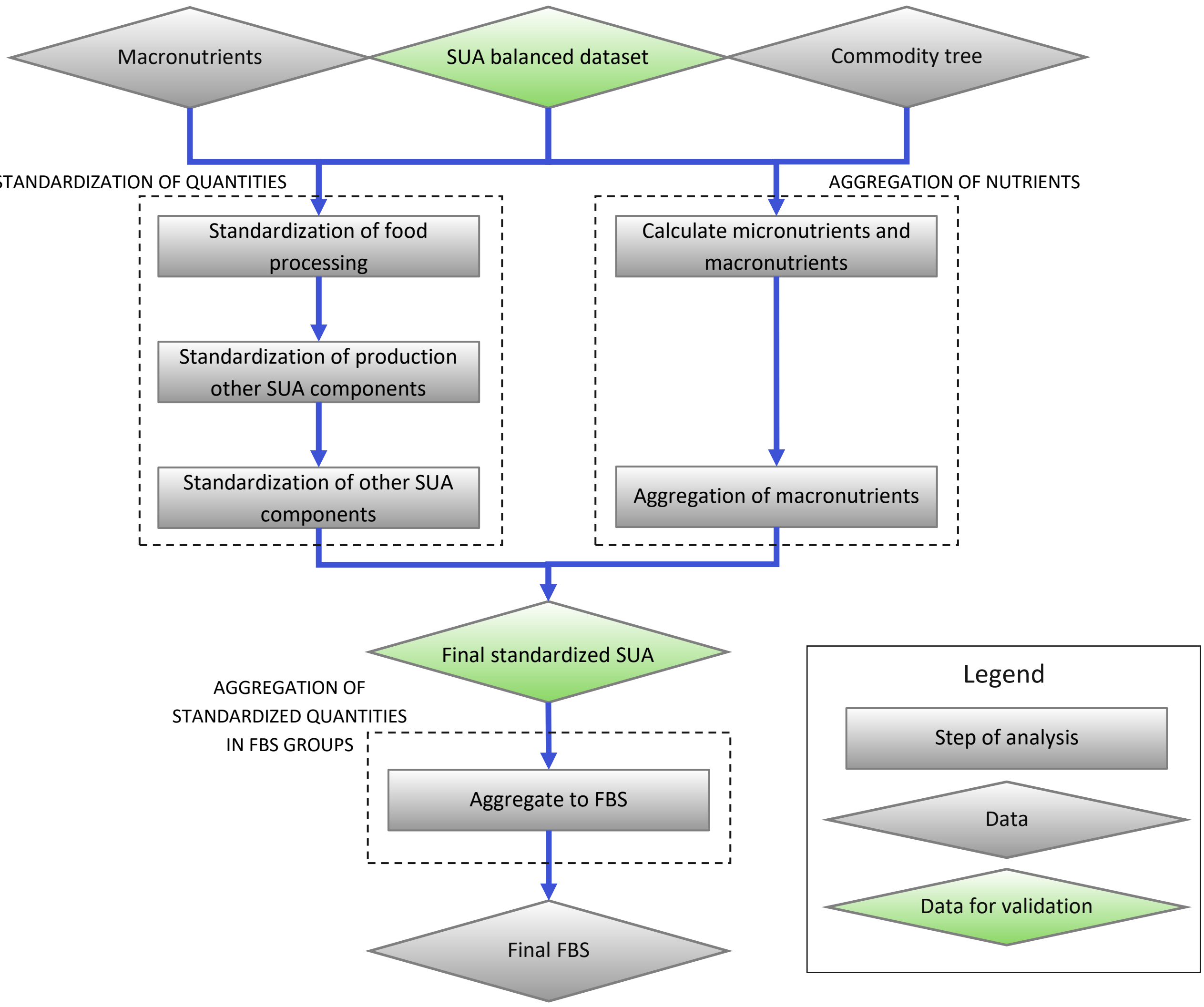
At the first processing level, flour of maize, bran of maize and germ of maize are derived from the same production process, therefore only one of them (flour of maize) has a weight equal to 1, while the other two products are set as zero-weight products.

In a similar way, at the second processing level, maize gluten and cake of maize are considered zero-weight products. Maize gluten because it shares a production process with starch of maize. Cake of maize because it shares a production process with oil of maize. Oil of maize is not standardized because it is a proxy-primary commodity and is cut from the commodity tree of the maize primary commodity. It is also zero-weight because it is derived from a zero-weight commodity (germ of maize), therefore its final weight is equal to 0.

15.3 Standardization flow

The complete standardization workflow is represented in **Figure 19**.

Figure 19: Complete standardization flow



Source: Authors' own elaboration.

Standardization consists of three main steps:

1. Standardization of quantities. Quantities of the variable $food\ processing$ are standardized first, while $production$ and other variables follow. This process is performed by "rolling up" the edges of the commodity tree, which means that products resulting from further processing are converted into primary level equivalents using extraction rates, shares and weights.

2. Aggregation of nutrients. To derive micronutrients and macronutrients, extraction rates, shares and weight are not used. Instead, the nutrient values of all products are applied to the standardized quantities of the previous step, independently from their weight.
3. Aggregation of standardized quantities in the FBS group. In this final step, the quantities and nutrients of each FBS group are aggregated.

15.3.1 Standardization of quantities (maize example, continued)

Standardization of food processing quantities

In the FBS, the food processing quantity of an FBS group is the sum of the primary equivalent of the production of all the derived products of each commodity tree belonging to that FBS group. Recalling **Equation 14**, $food\ processing$ is given by:

$$food\ processing_P = \sum_C \frac{production_C}{eR_{C-P}} \times bottom\text{-}up\ share_{C-P}$$

where:

- $production_C$ is the production of derived product C;
- eR_{C-P} is the extraction rate implied in the transformation of the parent product P into the Child C; and
- $bottom\text{-}up\ share_{C-P}$ is the bottom-up share of child C from parent P and represents the quantity of processed products C produced from the parent commodity P.

In this process, when a derived product is a proxy-primary, the corresponding *food processing* quantity will not be assigned to the parent commodity, but to the proxy-primary commodity and will be standardized to the FBS group of the proxy-primary. For example, the derived product beer in the SUA is part of the barley commodity tree but in the FBS, it is under FBS group beer because it is a proxy-primary. In these cases, $food\ processing$ is standardized through the formula below:

Equation 68: Standardization of food processing quantities

$$food\ processing_P = \sum_{C_{nsd}} \left(\frac{production_{C_{nsd}}}{eR_{C_{nsd}-P}} \times bottom\text{-}up\ share_{C_{nsd}-P} \times weight_{C_{nsd}} \right)$$

where:

- C_{nsd} denotes the child products that are not standardized to the parent product as in the commodity tree;
- $production_{C_{nsd}}$ is the production of a derived product not standardized to their parent product C_{nsd};
- $eR_{C_{nsd}-P}$ is the extraction rate implied in transformation of the parent product P in the child C_{nsd}; and
- $bottom\text{-}up\ share_{C_{nsd}-P}$ is the bottom-up share of children C_{nsd} from parent P and represents the quantity of processed products C_{nsd} produced from the parent commodity P.

In the maize example reported in **Table 46**, for the parent product maize, C_{nsd} is defined as:

$$C_{nsd} = \{undenatured\ ethyl\ alcohol\ (< 80\%), undenatured\ ethyl\ alcohol\ (> 80\%), oil\ of\ maize\}.$$

Undenatured ethyl alcohol (greater than 80 percent) and undenatured ethyl alcohol (less than 80 percent) do not have any child, therefore their food processing quantity in FBS will be 0.

The first version of the “rolled up” commodity tree in this case (i.e. standardizing from child to parent) merges the production data of processed products with the value FALSE in the Non-standardized column of **Table 47** and then applies **Equation 68** to the proxy-primary commodities. In the table, all child products are reported to maize even when they are not processed directly from maize, but from a product that is derived from maize (e.g. starch of maize, which is derived from flour of maize). This happens because their food processing quantities have been expressed in terms of their parent commodity.

Table 47: "Rolled up" edges of the commodity tree of maize to the FBS group of maize

PARENT	CHILD	PROCESSING LEVEL	EXTRACTION RATE	BOTTOM-UP SHARES	WEIGHT	NON-STANDARDIZED
Maize	Maize	1	1	1	1	-
Maize	Flour of maize	1	0.82	1	1	FALSE
Maize	Bran of maize	1	0.11	1	0	FALSE
Maize	Germ of maize	1	0.06	1	0	FALSE
Maize	Undenatured ethyl alcohol (< 80%)	1	0.35	0.7	1	TRUE
Maize	Undenatured ethyl alcohol (> 80%)	1	0.65	0.4	1	TRUE
Maize	Oil of maize	1	0.027	1	0	TRUE
Maize	Cake of maize	1	0.0312	1	0	FALSE
Maize	Starch of maize	1	0.615	1	1	FALSE
Maize	Maize gluten	1	0.082	1	0	FALSE

Source: Authors' own elaboration.

Standardization of production

The production of derived products is not standardized back to the primary commodity, as these are already accounted for in the food processing quantity. Only the production quantities of primary and proxy-primary commodities are included. In the example, these are maize, oil of maize and the ethyl alcohols.

Standardization of the other SUA variables

The other variables of the SUA ($imports$, $exports$, $food$, $feed$, $seed$, $industrial\ uses$ and $other\ uses$) are standardized in the same way as $food\ processing$, using the formula below:

Equation 69: Standardization of SUA components

$$utilization_P = \sum_C \left(\frac{utilization_C}{eR_{C-P}} \times bottom\text{-}up\ share_{C-P} \times weight_{C-P} \right)$$

where:

- $utilization_P$ is the utilization ($imports$, $exports$, $food$, $feed$, $seed$, $industrial\ uses$ or $other\ uses$) expressed for the parent P in the "rolled-up tree" (it corresponds to primary or proxy-primary);
- $utilization_C$ is the utilization ($imports$, $exports$, $food$, $feed$, $seed$, $industrial\ uses$ or $other\ uses$) expressed for a child C;
- eR_{C-P} is the extraction rate implied in transformation of the parent product P into the child C;
- $bottom\text{-}up\ share_{C_{nsd}-P}$ is the bottom-up share of children C from parent P and represents the quantity of processed products C produced from the parent commodity P; and
- $weight_{C-P}$ is the weight of each child.

In order to include them in the calculation, all primary commodities are considered, by convention, as both parent and child, with the extraction rate, bottom-up share and weight set to 1. In this way, the primary commodity maize is still included in the "rolled up" commodity tree, together with all the "rolled up" children.

In the maize example, this step results in the values reported in Table 48, where weights and extraction rates are set to 1 for these products (highlighted cells).

Table 48: "Rolled-up" edges of the commodity tree of maize to the other SUA components

PARENT	CHILD	PROCESSING LEVEL	EXTRACTION RATE	BOTTOM-UP SHARES	WEIGHT	NON-STANDARDIZED
Maize	Maize	1	1	1	1	-
Maize	Flour of maize	1	0.82	1	1	FALSE
Maize	Bran of maize	1	0.11	1	0	FALSE
Maize	Germ of maize	1	0.06	1	0	FALSE
Undenatured ethyl alcohol (< 80%)	Undenatured ethyl alcohol (< 80%)	1	1	1	1	TRUE
Undenatured ethyl alcohol (> 80%)	Undenatured ethyl alcohol (> 80%)	1	1	1	1	TRUE
Oil of maize	Oil of maize	1	1	1	1	TRUE
Maize	Cake of maize	1	0.0312	1	0	FALSE
Maize	Starch of maize	1	0.615	1	1	FALSE
Maize	Maize gluten	1	0.082	1	0	FALSE

Source: Authors' own elaboration.

15.3.2 Aggregation of macronutrients (maize example, continued)

The concepts of extraction rate and weight are not applicable when standardizing micro- and macronutrients because calories, proteins and fats are inherently expressed in standardized units across all SUA items. As a result, the aggregation of nutrients is done by aggregating to the FBS groups and applying the following formula, where extraction rates and weights are both set to 1:

Equation 70: Standardization of nutrients

$$macronutrient_P = \sum_C \left(\frac{macronutrient_C}{eR_{C-P}} \times bottom\text{-}up\ share_{C\text{-}P} \times weight_{C-P}\right)$$

where:

- $macronutrient_{iP}$ represents calories, proteins or fats;
- $eR_{C-P} = 1$ for all children C;
- $weight_{C-P} = 1$ for all children C; and
- P denotes the parent in the "rolled up" tree (it corresponds to primary or proxy primary).

In the maize example, for nutrients, child products are reported as in Table 49, where all extraction rates are set to 1 and all weights of primary and proxy-primary commodities are also set to 1.

Table 49: "Rolled-up" edges for the aggregation of macronutrients

PARENT	CHILD	PROCESSING LEVEL	EXTRACTION RATE	BOTTOM-UP SHARES	WEIGHT	NON-STANDARDIZED
Maize	Maize	1	1	1	1	-
Maize	Flour of maize	1	1	1	1	FALSE
Maize	Bran of maize	1	1	1	0	FALSE
Maize	Germ of maize	1	1	1	0	FALSE
Undenatured ethyl alcohol (< 80%)	Undenatured ethyl alcohol (< 80%)	1	1	0.7	1	TRUE
Undenatured ethyl alcohol (> 80%)	Undenatured ethyl alcohol (> 80%)	1	1	0.4	1	TRUE
Oil of maize	Oil of maize	1	1	1	1	TRUE
Maize	Cake of maize	1	1	1	0	FALSE
Maize	Starch of maize	1	1	1	1	FALSE
Maize	Maize gluten	1	1	1	0	FALSE

Source: Authors' own elaboration.

In the new FBS methodology, some macronutrients have been added to those calculated in this step. A more detailed description follows in Section **16**.

15.3.3 Aggregation in the FBS tree (maize example, continued)

After the standardization of *food processing* quantities, production quantities, quantities of other SUA variables and the macronutrients, these values are aggregated into the different FBS groups and aggregates by using the FBS tree presented in Section **2.5.1**.

The FBS commodity tree for maize is reported in **Table 50** where zero-weight products are identified through the "(calories only)" statement and the SUA products include also proxy-primary and processed products. The calories only means that the quantities of that commodity are not standardized to the primary equivalent e.g. bran of maize is not standardized to maize, but the calories of bran of maize are included in the FBS group of maize.

Table 50: Maize FBS aggregation table

FAMILY	GROUP	ITEM	(PROXY-)PRIMARY AND PROCESSED
VEGETAL PRODUCTS	CEREALS (EXCLUDING BEER)	MAIZE & PRODUCTS	MAIZE GERM MAIZE (calories only) FLOUR MAIZE BRAN MAIZE (calories only) MAIZE GLUTEN (calories only) STARCH MAIZE GLUTEN FEED & MEAL

Source: Authors' own elaboration.

16 NUTRIENT ESTIMATES

As mentioned in Section **2.1**, FAO has developed an updated global *nutrient conversion table* (NCT), based on newly released *food composition tables* (FCTs). This table will be integrated in the new FBS methodology to enrich the nutrient estimates from the FBS.

Besides updating information published by FAO on existing nutrients, namely energy expressed in kcal, proteins and fats, the global NCT includes a set of additional macro- and micronutrients. These new nutrient datasets allow for wider analysis and provide a comprehensive picture of the quality of agricultural and food products from a nutritional point of view.

The main steps followed for creating the NCT are summarized below.

FAO sourced 29 FCTs available at the national or regional level, with different structures, product coverage, nutrient coverage and methods used to measure nutrients.

At this stage existing FCTs were evaluated, and those with comprehensive and up-to-date data on food composition selected. This exercise produced a list of 13 out of the 29 initial FCTs as potential sources of data to be included in the NCT for agrifood data, shown in **Table 51**.

Table 51: Food composition tables sources

COUNTRY/REGION	FOOD COMPOSITION TABLE AND SOURCE
Australia	Australian food composition database (Food Standards Australia New Zealand, 2022)
Bangladesh	Food composition table for Bangladesh (Shaheen *et al.*, 2013)
Brazil	Brazilian food composition table (University of São Paulo, 2022)
Denmark	Danish food composition database – Frida (National Food Institute, Technical University of Denmark, 2024)
India	Indian food composition database (Longvah, Ananthan, Bhaskarachary and Venkaiah, 2017)
Japan	Standard tables of food composition in Japan (MEXT, 2015)
Kenya	Kenya food composition tables (FAO/Government of Kenya, 2018)
New Zealand	New Zealand Food composition database (New Zealand Institute for Plant & Food Research Limited and Ministry of Health, 2024)
Uganda	Food composition table for Central and Eastern Uganda (Hotz *et al.*, 2012)
United Kingdom of Great Britain and Northern Ireland	Composition of foods integrated dataset (CoFID) (Public Health England, 2019)
United States of America	USDA National nutrient database for standard reference (Legacy Release) (Haytowitz *et al.*, 2019)
Western Africa	FAO/INFOODS Food composition table for Western Africa (Vincent *et al.*, 2020)
Global	FAO/INFOODS Global food composition database for pulses (uPulses1.0) (FAO, 2017b)

Sources: See References.

The 13 selected sources cover different lists of nutrients. In order to identify the macro- and micronutrients to be included in the global NCT, two criteria were adopted: public health relevance and whether data were available for the selected FCTs.

The final list of selected nutrients includes energy, macro- and micronutrients (**Table 52)**. The components were compiled from the selected FCTs; except for energy, available carbohydrates and vitamin A that were not compiled but recalculated from the new FCTs to enable standardization. For the other nutrients, to ensure standardization of the components, the comparison across the different FCTs was assessed considering FAO/INFOODS component identifiers (also known as tag names), and all items were coded in the CPC version 2.1 expanded classification to allow use in the FBS.

Table 52: Nutrients included in the nutrient conversion table

COMPONENT GROUP	COMPONENT NAME	INFOODS TAG NAMES	UNIT	DENOMINATOR
Edible portion	Edible portion	EDIBLE	-	-
Energy	Energy	ENERC	kcal	/100g EP
Macronutrients	Protein, total	PROTCNT	g	/100g EP
Macronutrients	Fat, total	FAT and FATCE	g	/100g EP
Macronutrients	Carbohydrate, available; calculated by difference	CHOAVLDF	g	/100g EP
Macronutrients	Fibre, total dietary	FIBTG	g	/100g EP
Minerals	Calcium	CA	mg	/100g EP
Minerals	Iron	FE	mg	/100g EP
Minerals	Magnesium	MG	mg	/100g EP
Minerals	Phosphorus	P	mg	/100g EP
Minerals	Potassium	K	mg	/100g EP
Minerals	Zinc	ZN	mg	/100g EP
Vitamins	Vitamin A (expressed in retinol equivalents)	VITA	µcg	/100g EP
Vitamins	Vitamin A (expressed in retinol activity equivalents)	VITA_RAE	µcg	/100g EP
Vitamins	Thiamin (vitamin B_1)	THIA and THIAHCL	mg	/100g EP
Vitamins	Riboflavin (vitamin B_2)	RIBF	mg	/100g EP
Vitamins	Vitamin C	VITC	mg	/100g EP

Source: Authors’ own elaboration.

17 COMPILATION OF FOOD BALANCE SHEETS: A STEP-BY-STEP EXAMPLE

17.1 Overview

This section presents examples of SUAs for a hypothetical country and year. In particular, the primary and derived products pertaining to the FBS groups of wheat and products, barley and products, and sweeteners, other and products are presented, and will be standardized and aggregated to these FBS groups.

First, the procedure for generating the FBS for wheat and the derived product wheat and meslin flour is presented, followed by the complete procedure for generating the FBS group Wheat and products. Then the commodity tree, and the compilation process for the SUAs for all products derived from wheat are presented. Finally, the balancing process is applied.

Second, the balanced SUAs for the products of two other FBS groups are presented: barley and products; and sweeteners, other and products, along with their conversion into primary equivalents as explained in Section **15**.

Third, the example shows how to standardize all the SUAs into FBS groups and aggregates.

In order to demonstrate how a SUA table can be produced through imputation, it is assumed that information available to FBS compilers is very limited.

17.2 Example: SUA for wheat

17.2.1 SUA compilation

Table 53 represents an initial SUA for wheat for a given country and year. In this example, official figures are available for production and trade values, while all other values have to be imputed via estimation. A slash ("/") sign is reported for the elements $food$ and $tourist\ consumption$ because these two values will not be computed for wheat. In particular, $food$ is not required as wheat is not eaten in primary form and $tourist\ consumption$ is excluded because, in this example, the hypothetical country is not one for which it is a major component of the food supply pattern (Section **12**). On the other hand, the dash ("-") sign indicates that the value is currently unknown.

Table 53: Initial SUA for wheat

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	-	22 874 184	-	-	-	-	/	-	/

Source: Authors' own elaboration.

The total imports and exports for each commodity in this example are obtained by aggregating the respective trade flows by partner. A typical trade dataset with wheat data is represented in **Table 54** (the dataset has been simplified for this example).

Table 54: Wheat trade flow information at the HS6 disaggregation level

REPORTER	PARTNER	HS6 CODE	FLOW	WEIGHT (kg)	VALUE (USD)
124	932	100110	1	3 350 000	502 500 000
124	899	100110	1	1 200 000	264 000 000
124	961	100190	2	870 000	113 100 000

Source: Authors' own elaboration.

In the table above, the country codes refer to a specific reporter and three different partners. The HS codes refer to wheat and meslin durum wheat (100110) and wheat and meslin other (100190). Flows (1) and (2) indicate imports and exports, respectively. The quantity weights are in kilograms and the values are, in this case, in US dollars. The totals for wheat imports would be obtained by summing up all the imports flows, and similarly for the total exports (a typical trade dataset would have many more flows than the simple example above). For the compilation of the FBS, only the quantities are of interest, not the monetary values which are only used to impute missing quantity data (see Section 6).

The first estimation to be computed is the stocks variation (DSTOCK in the tables below). The number reported in Table 55 represents the estimated stocks variation for the example. The stocks imputation methodology (Section 6.3.4) estimates the stocks variation in the current year as a linear regression on the cumulative stocks variations in the previous years. In this case, the estimate represents an increase (hence the positive sign) in the stocks held. The idea behind this assumption is that an increase in stocks will be more likely if high stocks withdrawals were observed in the past, because stocks cannot be continuously withdrawn and vice versa. The results of the imputation are shown in Table 55 in bold.

The $feed$ variable is then imputed based on the animal numbers and feed intensification factors, resulting in calculated feed requirements as explained in Section 9. The resulting figures are also bolded in Table 55.

Table 55: SUA for wheat with imputed values for stocks and feed

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	**1 208 297**	22 874 184	**2 189 300**	-	-	-	/	-	/

Source: Authors' own elaboration.

Losses (Section 10) are assumed to occur only at the primary level, therefore the value reported in Table 56 accounts for all losses, as they are assigned at the primary level and not to derived products. The value reported refers to losses from the post-harvest stage up to, but not including, the retail level. Processing losses are taken into account in the standardization process.

Table 56: SUA for wheat with imputed values for loss

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	-	**20 758**	-	/	-	/

Source: Authors' own elaboration.

Seed values, when missing, are imputed on the basis of the hierarchical linear model described in Section 11. Seed is only allotted to the primary commodity (Table 57).

Table 57: SUA for wheat with imputed values for seed

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	**991 600**	20 758	-	/	-	/

Source: Authors' own elaboration.

Table 58 reports the values for the example as collected from FBS experts for industrial uses.

Table 58: SUA for wheat with imputed values for industrial use

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	-	/	**780 323**	/

Source: Authors' own elaboration.

Food processing now has to be estimated using the processing shares and availabilities of data on the production of derived commodities, depending on whether these data are official or unofficial. In the example, let us suppose that the value for the processed product wheat and meslin flour is available and official at the first processing level. Therefore, the $food\ processing$ quantity for wheat is determined by using **Equation 14** and **Table 59**.

Table 59: Calculation of wheat food processing – level 1

PARENT	CHILD	EXTRACTION RATE	PRODUCTION OF CHILD	BOTTOM-UP SHARE
Wheat	Wheat and meslin flour	0.7686	3 005 000	100%

Source: Authors' own elaboration.

$$food\ processing_{wheat} = \frac{production_{flour}}{eR_{flour-wheat}} \times bottom\text{-}top\ share_{flour-wheat} = \frac{3\ 005\ 000}{0.7686} \times 1 = 3\ 909\ 706$$

The resulting, and unbalanced SUA for wheat is reported in **Table 60**.

Table 60: Unbalanced SUA for wheat

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	**3 909 706**	/	780 323	/

Source: Authors' own elaboration.

17.2.2 Balancing procedure

Once all the variables have been added to the SUA after data collection or imputation, the imbalance is computed (**Table 61**) and the balancing mechanism is applied to solve the imbalance and to ensure that $supply = utilization$.

Table 61: SUA for wheat with imbalance

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	**IMBALANCE**
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	**3 909 706**	/	780 323	/	**309 732**

Source: Authors' own elaboration.

The balanced SUA is obtained through the application of the methodology described in Section **14**. The balancing of wheat, in particular, is obtained after the first step of the balancing flow, by allocating the overall imbalance to $\Delta stocks$. In this case, indeed, $\Delta stocks^{IMB}$ is given by:

$$\Delta stocks^{IMB} = 1\ 208\ 297 + 309\ 732 = 1\ 518\ 029$$

Table 62: Balancing for wheat

WHEAT SUA	OPENING STOCK	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	IMBALANCE	SUPPLY
UNBALANCED	14 056 704	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	3 909 706	780 323	**309 732**	**32 283 900**
BALANCED	14 056 704	32 201 100	82 800	**1 518 029**	22 874 184	2 189 300	991 600	20 758	3 909 706	780 323	**0**	**32 283 900**

Source: Authors' own elaboration.

The following table shows the final SUA for wheat.

Table 63: Balanced SUA for wheat

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	**IMBALANCE**
Wheat	32 201 100	82 800	**1 518 029**	22 874 184	2 189 300	991 600	20 758	3 909 706	-	780 323	-	**0**

Source: Authors' own elaboration.

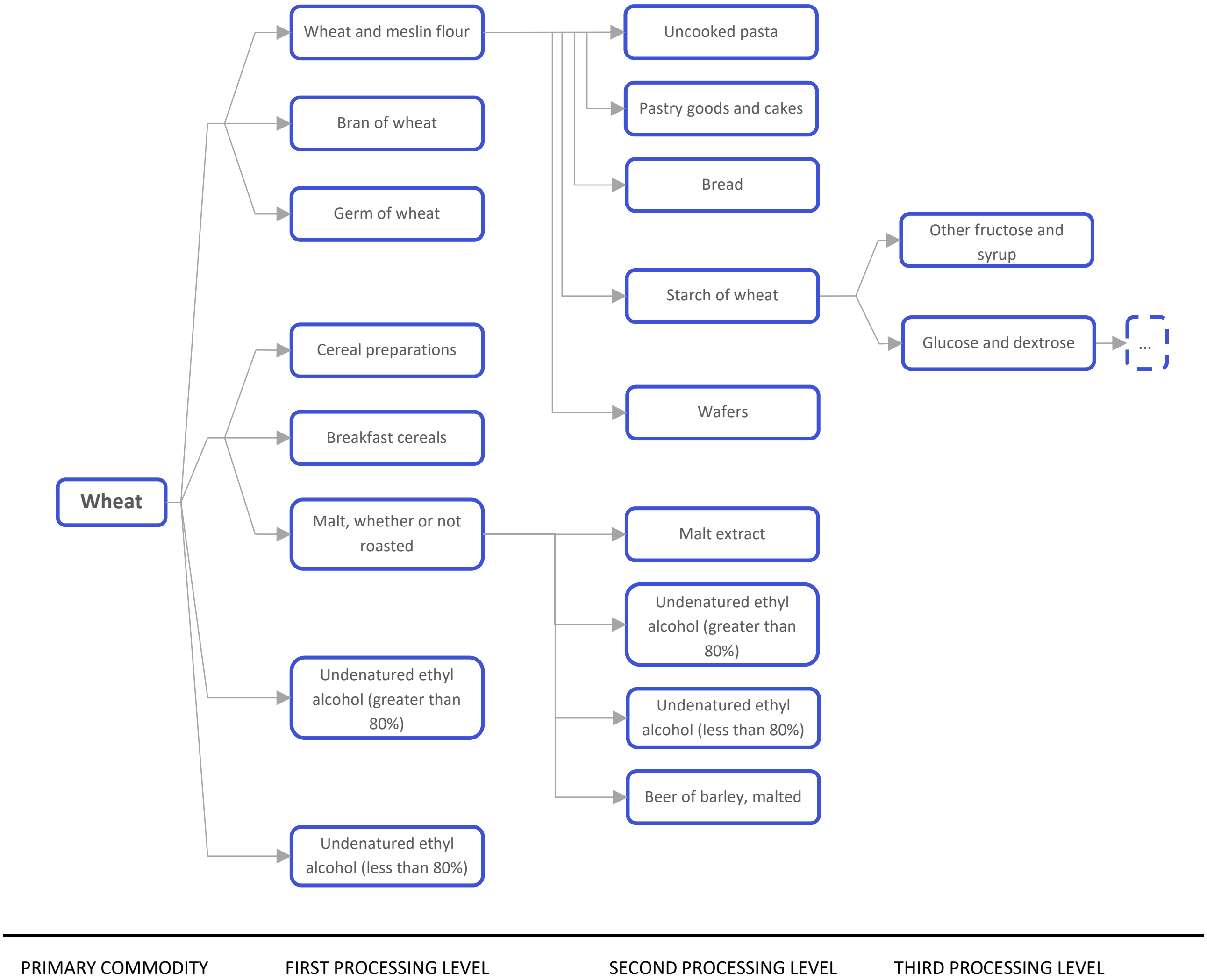
This SUA will be further examined in Section **17.3.1**, where the other SUAs for wheat products are introduced. The standardization procedure will then be explained in Section **17.5**.

17.3 Wheat commodity tree

Figure 20 represents the wheat commodity tree. It includes all products processed in a given hypothetical country for a given year, from wheat down to the third processing level.

The wheat tree includes multiple parent products (i.e. undenatured ethyl alcohol, which has been already introduced in the maize example in Section 15.2.3), multiple child products (i.e. starch of wheat and malt) and proxy-primary commodities (i.e. beer of barley and wheat-fermented beverages). Indeed, as mentioned earlier, commodity trees can intersect each other, as derived products might have more than one parent in different trees and can also be the parent commodity of more than one processed item (child commodities) . Moreover, some derived products are cut from their tree and treated as if they were primary products (proxy-primary commodities), because of their importance in the food availability of a country.

Figure 20: Wheat commodity tree



Source: Authors' own elaboration.

17.3.1 Compilation of SUAs for wheat products

17.3.1.1 Initial empty SUAs

The following table includes all the SUAs of wheat-related products. In this table the dash ("-") indicates that the value is currently unknown.

Table 64: Empty SUA for wheat products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	-	-	-	-	-	-	-	-	-	-	-
Wheat and meslin flour	-	-	-	-	-	-	-	-	-	-	-
Bran of wheat	-	-	-	-	-	-	-	-	-	-	-
Uncooked pasta	-	-	-	-	-	-	-	-	-	-	-
Germ of wheat	-	-	-	-	-	-	-	-	-	-	-
Bread	-	-	-	-	-	-	-	-	-	-	-
Pastry goods and cakes	-	-	-	-	-	-	-	-	-	-	-
Starch of wheat	-	-	-	-	-	-	-	-	-	-	-
Breakfast cereals	-	-	-	-	-	-	-	-	-	-	-
Malt, whether or not roasted	-	-	-	-	-	-	-	-	-	-	-
Malt extract	-	-	-	-	-	-	-	-	-	-	-
Beer of barley, malted	-	-	-	-	-	-	-	-	-	-	-
Wafers	-	-	-	-	-	-	-	-	-	-	-
Cereal preparations	-	-	-	-	-	-	-	-	-	-	-
Other fructose and syrup	-	-	-	-	-	-	-	-	-	-	-
Glucose and dextrose	-	-	-	-	-	-	-	-	-	-	-
Undenatured ethyl alcohol (> 80%)	-	-	-	-	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	-	-	-	-	-	-	-	-	-	-

Source: Authors' own elaboration.

17.3.1.2 Production of primary products

The first values to be filled in the SUA table are those concerning production.

Table 65: SUA for wheat products filled with official figures for production

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	**32 201 100**	-	-	-	-	-	-	-	-	-	-
Wheat and meslin flour	-	-	-	-	-	-	-	-	-	-	-
Bran of wheat	-	-	-	-	-	-	-	-	-	-	-
Uncooked pasta	-	-	-	-	-	-	-	-	-	-	-
Germ of wheat	-	-	-	-	-	-	-	-	-	-	-
Bread	-	-	-	-	-	-	-	-	-	-	-
Pastry goods and cakes	-	-	-	-	-	-	-	-	-	-	-
Starch of wheat	-	-	-	-	-	-	-	-	-	-	-
Breakfast cereals	-	-	-	-	-	-	-	-	-	-	-
Malt, whether or not roasted	-	-	-	-	-	-	-	-	-	-	-
Malt extract	-	-	-	-	-	-	-	-	-	-	-
Beer of barley, malted	-	-	-	-	-	-	-	-	-	-	-
Wafers	-	-	-	-	-	-	-	-	-	-	-
Cereal preparations	-	-	-	-	-	-	-	-	-	-	-
Other fructose and syrup	-	-	-	-	-	-	-	-	-	-	-
Glucose and dextrose	-	-	-	-	-	-	-	-	-	-	-
Undenatured ethyl alcohol (> 80%)	-	-	-	-	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	-	-	-	-	-	-	-	-	-	-

Source: Authors' own elaboration.

Only the production of primary commodities can be considered in this step. Indeed, all the variables are needed for estimating the production of derived products, which happens after the other variables of the identity are filled or imputed. For primary production data, the table has to be filled first with any available official figures (bold values in **Table 65**). In the current example, only the wheat production quantity is considered as given.

17.3.1.3 Trade

Total imports and exports quantities for wheat are now filled, as well as the other products, into the SUAs table as resulting from the validation steps presented in Section **6.3**.

Table 66: SUA for wheat products filled with trade

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	**82 800**	-	**22 874 184**	-	-	-	-	-	-	-
Wheat and meslin flour	-	**102 099**	-	**276 666**	-	-	-	-	-	-	-
Bran of wheat	-	**79 640**	-	-	-	-	-	-	-	-	-
Uncooked pasta	-	**115 616**	-	**40 300**	-	-	-	-	-	-	-
Germ of wheat	-	-	-	-	-	-	-	-	-	-	-
Bread	-	**12 686**	-	**18 443**	-	-	-	-	-	-	-
Pastry goods and cakes	-	**273 583**	-	**621 318**	-	-	-	-	-	-	-
Starch of wheat	-	**1 954**	-	**3 740**	-	-	-	-	-	-	-
Breakfast cereals	-	**262 717**	-	**127 180**	-	-	-	-	-	-	-
Malt, whether or not roasted	-	**24 401**	-	**571 615**	-	-	-	-	-	-	-
Malt extract	-	**973**	-	-	-	-	-	-	-	-	-
Beer of barley, malted	-	**367 179**	-	**163 564**	-	-	-	-	-	-	-
Wafers	-	**221 853**	-	**70 858**	-	-	-	-	-	-	-
Cereal preparations	-	**8 737**	-	**64 418**	-	-	-	-	-	-	-
Other fructose and syrup	-	**133 329**	-	**60 160**	-	-	-	-	-	-	-
Glucose and dextrose	-	**148 688**	-	**170 046**	-	-	-	-	-	-	-
Undenatured ethyl alcohol (> 80%)	-	**1 357 118**	-	**86 983**	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	**104 707**	-	**183 263**	-	-	-	-	-	-	-

Source: Authors' own elaboration.

17.3.1.4 Stocks

The results of the imputation of stocks variation are shown in **Table 67** (bold values). FAO estimates stocks for primary commodities and a list of "stockable items" (i.e. non-perishable items) that it maintains. For this reason, many values in the table are equal to 0.

Table 67: SUA for wheat products filled with stocks

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	**1 208 297**	22 874 184	-	-	-	-	-	-	-
Wheat and meslin flour	-	102 099	**0**	276 666	-	-	-	-	-	-	-
Bran of wheat	-	79 640	-	-	-	-	-	-	-	-	-
Uncooked pasta	-	115 616	**0**	40 300	-	-	-	-	-	-	-
Germ of wheat	-	-	-	-	-	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	-	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	-	-	-
Starch of wheat	-	1 954	**0**	3 740	-	-	-	-	-	-	-
Breakfast cereals	-	262 717	**0**	127 180	-	-	-	-	-	-	-
Malt, whether or not roasted	-	24 401	**0**	571 615	-	-	-	-	-	-	-
Malt extract	-	973	**0**		-	-	-	-	-	-	-
Beer of barley, malted	-	367 179	**0**	163 564	-	-	-	-	-	-	-
Wafers	-	221 853	**0**	70 858	-	-	-	-	-	-	-
Cereal preparations	-	8 737	**5 414**	64 418	-	-	-	-	-	-	-
Other fructose and syrup	-	133 329	**0**	60 160	-	-	-	-	-	-	-
Glucose and dextrose	-	148 688	**0**	170 046	-	-	-	-	-	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	**0**	86 983	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	104 707	**4 696**	183 263	-	-	-	-	-	-	-

Source: Authors' own elaboration.

17.3.1.5 Food

The $food$ variable is estimated using validated food quantities from the previous year, changes in GDP, and product-related income elasticities (Section 7.2). The allocation to $food$ is considered for any edible item at the SUA level.

Table 68: SUA for wheat products filled with food

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	-	-	-	-	-	-	-
Wheat and meslin flour	-	102 099	0	276 666	-	-	-	-	**2 257 054**	-	-
Bran of wheat	-	79 640	-	-	-	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	**60 248**	-	-
Germ of wheat	-	-	-	-	-	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	**27 766**	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	**56 347**	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	-	-
Breakfast cereals	-	262 717	0	127 180	-	-	-	-	**34 155**	-	-
Malt, whether or not roasted	-	24 401	0	571 615	-	-	-	-	**5 921**	-	-
Malt extract	-	973	0		-	-	-	-	**617**	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	**2 198 300**	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	**0**	-	-
Cereal preparations	-	8 737	5 414	64 418	-	-	-	-	**11 581**	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	**57 733**	-	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	**40 337**	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	104 707	4 696	183 263	-	-	-	-	**2 311**	-	-

Source: Authors' own elaboration.

17.3.1.6 Feed

The next step is to impute the $feed$ variable. The assumption here is that some of the primary-level quantities are used as feed, as well as almost all of the bran (which is a by-product of the flour production process). The feed requirements should cover needs for the entire country.

The calculation of the feed requirements is laid out in Section 9, but is summarized here. The first step is to calculate the feed requirement based on the number of animals, their needs and feeding efficiency. In a second step, the actual amount of compound and concentrate feed is calculated applying country-specific intensification rates. In a third step, individual feedstuffs are allocated by availability. Before this, all feed-only products (e.g. oil cakes and meals, DDGS, dregs, brans) are deducted from the requirements. The remainder of the feed requirements will be satisfied by allocating to the FBS commodities at the primary level (such as cereals and oil crops) according to their availability. Negligible amounts of bran may go into such products as breakfast cereals, but for the sake of simplicity, such quantities will be ignored in this example. The resulting estimates for feed use are shown in bold in Table 69.

Table 69: SUA for wheat products filled with feed

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	**2 189 300**	-	-	-	-	-	-
Wheat and meslin flour	-	102 099	0	276 666	-	-	-	-	2 257 054	-	-
Bran of wheat	-	79 640	-	-	**728 651**	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	-	-	-	-	**62 555**	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	-	-
Breakfast cereals	-	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	-	24 401	0	571 615	-	-	-	-	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	-	8 737	5 414	64 418	**6 425**	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	-	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

17.3.1.7 Loss

Losses are accounted for primary commodities only. The value already seen in Table 56 is added and reported below.

Table 70: SUA for wheat products filled with loss

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	**20 758**	-	-	-	-
Wheat and meslin flour	-	102 099	0	276 666	-	-	-	-	2 257 054	-	-
Bran of wheat	-	79 640	-	-	728 651	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	-	-	-	-	62 555	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	-	-
Breakfast cereals	-	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	-	24 401	0	571 615	-	-	-	-	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	-	8 737	5 414	64 418	6 425	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	-	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

17.3.1.8 Seed

The *seed* values are also imputed for primary commodities only. **Table 71** reports the value already seen in the wheat example.

Table 71: SUA for wheat products filled with seed use

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	**991 600**	20 758	-	-	-	-
Wheat and meslin flour	-	102 099	0	276 666	-	-	-	-	2 257 054	-	-
Bran of wheat	-	79 640	-	-	728 651	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	-	-	-	-	62 555	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	-	-
Breakfast cereals	-	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	-	24 401	0	571 615	-	-	-	-	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	-	8 737	5 414	64 418	6 425	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	-	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	-	-
Undenatured ethyl alcohol (< 80%)	-	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

17.3.1.9 Industrial use

For the wheat commodity tree, the main commodity that has industrial use is wheat starch, but values are imputed also for wheat, other fructose and syrup and undenatured ethyl alcohol (greater that 80 percent). **Table 72** reports the values for the current example as collected from FBS experts.

Table 72: SUA for wheat products filled with industrial use

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	-	-	**780 323**	-
Wheat and meslin flour	-	102 099	0	276 666	-	-	-	-	2 257 054	-	-
Bran of wheat	-	79 640	-	-	728 651	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	-	-	-	-	62 555	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	**7 214**	-
Breakfast cereals	-	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	-	24 401	0	571 615	-	-	-	-	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	-	8 737	5 414	64 418	6 425	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	**6 944**	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	**1 270 135**	-
Undenatured ethyl alcohol (< 80%)	-	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

17.3.1.10 Tourist consumption

The example is based on a country where no tourism data are estimated, therefore no changes are made in the SUA table at this step.

17.3.1.11 Production of derived products

With all other supplies and utilization values now provided in the table, the production of derived products can be computed and estimated using the commodity tree and conversion factors.

In the current example, processing shares (the percentage of a given commodity used as input in the process for making a derived product) have to be computed first because no production data are available for child products. Indeed, this example falls under the first case mentioned in Section **5.3.2** concerning the calculation of food processing (the processed quantity of a parent product to be used as input in the production of derived products), where the processing shares are obtained as a moving average of the ratio of the food processing quantity over the availability. Therefore, $food\ processing$ for wheat has to be calculated using the processing share and availability as in **Equation 13**.

$$food\ processing_{wheat} = processing\ share_t \times availability_{wheat} = 0.74 \times 4\ 219\ 438 = 3\ 127\ 767$$

With the value of $food\ processing$ for wheat, the production values of all the products that are derived from wheat can be determined. The values (reported in **Table 74**) are obtained from the values reported in **Table 73** as the multiplication of $food\ processing_{wheat}$, eR and the top-down share (which is the percentage share of the quantity of a primary or derived good used in a transformation process). For example, the production for wheat and meslin flour is obtained as:

$$production_{what\ and\ meslin\ flour} = 3\ 127\ 767 \times 0.7686 \times 1.000 = 2\ 404\ 000$$

Table 73: Wheat values for the calculated production of derived at processing level 1

PARENT	CHILD	EXTRACTION RATE	TOP-DOWN SHARE	PRODUCTION OF CHILD
Wheat	Wheat and meslin flour	0.7686	1.000	2 404 000
Wheat	Bran of wheat	0.2075	1.000	649 011
Wheat	Germ of wheat	0.0200	1.000	62 555
Wheat	Breakfast cereals	0.8500	0.006	15 008
Wheat	Malt, whether or not roasted	0.7300	0.382	872 490
Wheat	Cereal preparations	0.0000	-	-
Wheat	Undenatured ethyl alcohol (> 80%)	0.0000	-	-
Wheat	Undenatured ethyl alcohol (< 80%)	0.6800	0.039	83 664

Source: Authors' own elaboration.

Cereal preparations and undenatured ethyl alcohol (greater than 80 percent) for this country are not derived from wheat, as their eRs are equal to 0.

Table 74: SUA for wheat products filled with the production of derived products at processing level 1

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	**3 127 767**	-	780 323	-
Wheat and meslin flour	**2 404 000**	102 099	0	276 666	-	-	-	-	2 257 054	-	-
Bran of wheat	**649 011**	79 640	-	-	728 651	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	**62 555**	-	-	-	62 555	-	-	-	-	-	-
Bread	-	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	-	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	-	1 954	0	3 740	-	-	-	-	-	7 214	-
Breakfast cereals	**15 009**	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	**872 490**	24 401	0	571 615	-	-	-	-	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	-	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	-	8 737	5 414	64 418	6 425	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	6 944	-
Glucose and dextrose	-	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	1 270 135	-
Undenatured ethyl alcohol (< 80%)	**83 664**	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

The SUA of wheat has a second processing level in this example. At the second level, alcohol, beer of barley and malt extract are obtained from malt; and wafers, starch of wheat, bread, pastry and pasta are obtained from flour. Also at the second level, $food\ processing$ is obtained as the product of the processing share and availability because no production data are available for child products. The processing shares for the example are reported in the **Table 75** together with the availabilities calculated from **Table 74** and the corresponding processed production obtained as the multiplication of the two other values (**Equation 13**).

From Section 5.2.1, the availability used for calculating $food\ processing$ is given by production net of trade. For example, the availability reported in Table 75 for wheat and meslin flour is obtained as:

$$availability_{wheat\ and\ meslin\ flour} = production + imports - exports = 2\ 404\ 000 + 102\ 099 - 276\ 666 = 2\ 229\ 433$$

Table 75: Wheat food processing calculation for processing level 2

COMMODITY	PROCESSING SHARE	AVAILABILITY	FOOD PROCESSING
Wheat and meslin flour	0.16180	2 229 433	360 714
Malt, whether or not roasted	1.09548	325 276	356 333

Source: Authors' own elaboration.

Table 76 is filled with all values for the food processing and production of derived products for the country.

This example highlights some specific points. First, starch of wheat is produced but not used in the country for other products, because it has no food processing. Consequently, glucose and dextrose in this country is not produced from wheat. Indeed, it is linked to wheat through starch of wheat at the third processing level as shown in Figure 20. Since starch of wheat is not processed, it follows that glucose and dextrose, which is a multiple parent commodity, is produced from some other parents.

Table 76: SUA for wheat products filled with production of derived products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	**3 127 767**	-	780 323	-
Wheat and meslin flour	2 404 000	102 099	0	276 666	-	-	-	**360 714**	2 257 054	-	-
Bran of wheat	649 011	79 640	-	-	728 651	-	-	-	-	-	-
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-
Germ of wheat	62 555	-	-	-	62 555	-	-	-	-	-	-
Bread	30 000	12 686	-	18 443	-	-	-	-	27 766	-	-
Pastry goods and cakes	350 000	273 583	-	621 318	-	-	-	-	56 347	-	-
Starch of wheat	9 000	1 954	0	3 740	-	-	-	**0**	-	7 214	-
Breakfast cereals	15 009	262 717	0	127 180	-	-	-	-	34 155	-	-
Malt, whether or not roasted	872 490	24 401	0	571 615	-	-	-	**356 333**	5 921	-	-
Malt extract	-	973	0		-	-	-	-	617	-	-
Beer of barley, malted	2 165 000	367 179	0	163 564	-	-	-	-	2 198 300	-	-
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-
Cereal preparations	81 508	8 737	5 414	64 418	6 425	-	-	-	11 581	-	-
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	6 944	-
Glucose and dextrose	75 000	148 688	0	170 046	-	-	-	-	40 337	-	-
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	1 270 135	-
Undenatured ethyl alcohol (< 80%)	83 664	104 707	4 696	183 263	-	-	-	-	2 311	-	-

Source: Authors' own elaboration.

17.3.2 SUAs balancing for wheat products

Table 77 includes all the unbalanced SUAs for wheat-related products.

Table 77: Unbalanced SUA for wheat products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	**IMBALANCE**
Wheat	32 201 100	82 800	1 208 297	22 874 184	2 189 300	991 600	20 758	3 127 767	-	780 323.41	-	**1 091 672**
Wheat and meslin flour	2 404 000	102 099	0	276 666	-	-	-	360 714	2 257 054	-	-	**-388 335**
Bran of wheat	649 011	79 640	-	-	728 651	-	-	-	-	-	-	**0**
Uncooked pasta	-	115 616	0	40 300	-	-	-	-	60 248	-	-	**15 068**
Germ of wheat	62 555	-	-	-	62 555	-	-	-	-	-	-	**0**
Bread	30 000	12 686	-	18 443	-	-	-	-	27 766	-	-	**-3 522**
Pastry goods and cakes	350 000	273 583	-	621 318	-	-	-	-	56 347	-	-	**-54 081**
Starch of wheat	9 000	1 954	0	3 740	-	-	-	-	-	7 214	-	**0**
Breakfast cereals	15 008	262 718	0	127 180	-	-	-	-	34 155	-	-	**116 391**
Malt, whether or not roasted	872 490	24 401	0	571 615	-	-	-	356 333	5 921	-	-	**-36 977**
Malt extract	-	973	0		-	-	-	-	617	-	-	**356**
Beer of barley, malted	2 165 000	367 179	0	163 564.02	-	-	-	-	2 198 300	-	-	**170 316**
Wafers	-	221 853	0	70 858	-	-	-	-	0	-	-	**150 995**
Cereal preparations	81 508	8 738	5 414	64 418	6 425	-	-	-	11 58	-	-	**2 407**
Other fructose and syrup	-	133 329	0	60 160	-	-	-	-	57 733	6 944	-	**8 492**
Glucose and dextrose	75 000	148 688	0	170 046	-	-	-	-	40 338	-	-	**13 304**
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	1 270 135	-	**0**
Undenatured ethyl alcohol (< 80%)	83 664	104 707	4 696	183 263	-	-	-	-	2 311	-	-	**-1 898**

Source: Authors' own elaboration.

The balancing for wheat has already been explored in Section **17.2.2**. Wheat was balanced using stocks, as all conditions outlined in Section **14.2.1** hold for this commodity. However, the balancing for bread cannot be performed using $\Delta stocks$: **Table 78** shows that bread does not have a value for $\Delta stocks$, therefore none of the conditions for stocks would be fulfilled.

Table 78: Balancing for bread

WHEAT SUA	OPENING STOCK	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	IMBALANCE	SUPPLY
UNBALANCED	0	30 000	12 686	-	18 443	-	-	-	-	27 766	**-3 522**	**42 686**
BALANCED	0	30 000	12 686	-	18 443	-	-	-	-	**24 243**	**0**	**42 686**

Source: Authors' own elaboration.

As bread is a food item and $food$ is its only utilization, all the imbalance will be assigned to $food$, because assigning the whole imbalance to this commodity still keeps a positive value and solves the imbalance.

Other imbalances of the wheat example will be solved in similar ways or by applying the proportional balancing explained in Section **14**. **Table 79** represents the balanced SUA for wheat.

Table 79: Balanced SUA for wheat products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	**IMBALANCE**
Wheat	32 201 100	82 800	**2 299 969**	22 874 184	2 189 300	991 600	20 758	3 127 767	-	780 323	-	**0**
Wheat and meslin flour	2 404 000	102 099	0	276 666	-	-	-	360 714	**1 868 719**	-	-	**0**
Bran of wheat	649 011	180 204	-	-	728 651	-	-	-	-	-	-	**0**
Uncooked pasta	-	115 616	**15 068**	40 300	-	-	-	-	60 248	-	-	**0**
Germ of wheat	62 555	-	-		62 555	-	-	-	-	-	-	**0**
Bread	30 000	12 686	-	18 443	-	-	-	-	**24 243**	-	-	**0**
Pastry goods and cakes	350 000	273 583	-	621 318	-	-	-	-	**2 266**	-	-	**0**
Starch of wheat	9 000	1 954	0	3 740	-	-	-	-	-	7 214	-	**0**
Breakfast cereals	15 009	262 717	0	127 180	-	-	-	-	**150 546**		-	**0**
Malt, whether or not roasted	872 490	24 401	**-36 977**	571 615	-	-	-	356 333	5 921	-	-	**0**
Malt extract	-	973	**356**		-	-	-	-	617	-	-	**0**
Beer of barley, malted	2 165 000	367 179	**170 316**	163 564	-	-	-	-	2 198 300	-	-	**0**
Wafers	-	221 853	0	70 858	-	-	-	-	**150 995**	-	-	**0**
Cereal preparations	81 508	8 737	**1**	64 418	**8 992**	-	-	-	**16 834**	-	-	**0**
Other fructose and syrup	-	133 329	**8 492**	60 160	-	-	-	-	57 733	6 944	-	**0**
Glucose and dextrose	75 000	148 688	**1 322**	170 046	-	-	-	-	**52 320**	-	-	**0**
Undenatured ethyl alcohol (> 80%)	-	1 357 118	0	86 983	-	-	-	-	-	1 270 135	-	**0**
Undenatured ethyl alcohol (< 80%)	83 664	104 707	**0**	183 263	-	-	-	-	**5 109**	-	-	**0**

Source: Authors' own elaboration.

17.4 Balanced SUAs for barley products and sweeteners

This Section provides examples for two other FBS groups, giving their commodity trees and showing the final balanced SUAs. These are barley products (**Table 80** and **Figure 21**) and sweeteners (**Table 81** and **Figure 22**). It is assumed that the SUA compilation process has already been applied for all products in these groups for each country and each year.

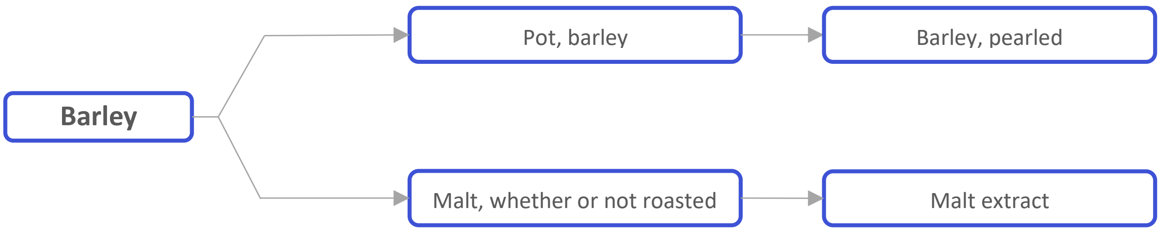
Both these commodity groups share some products with the wheat SUAs as they both contain derived products that also have a parent commodity in the wheat SUAs or commodities that are cut from the wheat tree.

Table 80: Balanced SUA for barley and products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	IMBALANCE
Barley	8 379 700	43 657	27 636	2 238 693	4 351 837	257 800	381 429	1 165 963	-	-	-	0
Barley, pearled	-	1 173	-	-	-	-	-	-	1 173	-	-	0
Malt extract	-	973	356	-	-	-	-	-	617	-	-	0
Malt, whether or not roasted	872 490	24 401	-36 977	571 615	-	-	-	356 333	5 921	-	-	0
Pot barley	-	-	-	-	-	-	-	0	-	-	-	0

Source: Authors' own elaboration.

Figure 21: Barley commodity tree



Source: Authors' own elaboration.

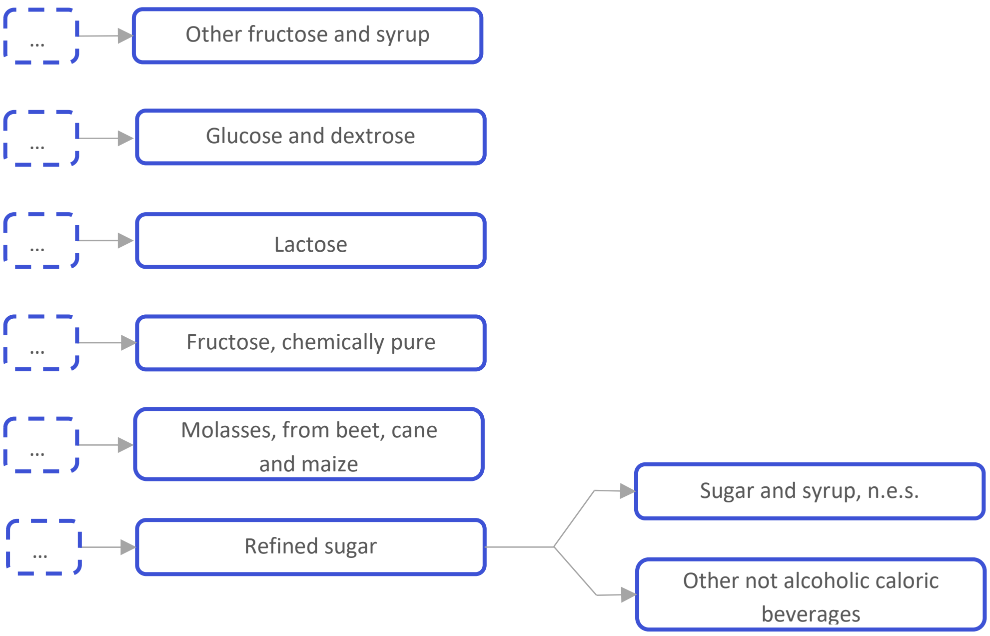
The SUAs for the sweetener products, in particular, are special as they are all composed of products that are cut from the tree of the primary commodities from which they are produced, and considered separately as if they were primary, due to the importance they have in the food composition of the country (*cut commodities* or *proxy-primaries* as defined in Section 15.2.1).

Table 81: Balanced SUA for sweeteners and products

COMMODITY	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST	**IMBALANCE**
Fructose, chemically pure	-	2 850	0	58	-	-	-	-	-	2 791	-	0
Glucose and dextrose	75 000	148 688	1 322	170 046	-	-	-	-	52 320	-	-	0
Lactose	685	8 231	5 426	1 873	-	-	-	-	1 617	-	-	0
Molasses (from beet, cane and maize)	22 960	142 077	-	12 171	54 745	-	-	98 121	-	-	-	0
Other fructose and syrup	-	133 329	8 492	60 160	-	-	-	-	57 733	6 944	-	0
Other non-alcoholic caloric beverages n.e.c.	-	874 700	0	227 473	-	-	-	-	647 226	-	-	0
Refined cane or beet sugar, solid; maple sugar and maple syrup	55 000	2 506	0	48 373	-	-	-	-	9 134	-	-	0
Sugar and syrups n.e.s.	-	46 926	4 163	10 195	-	-	-	-	32 568	-	-	0

Source: Authors' own elaboration.

Figure 22: Commodity tree of sweeteners products



Source: Authors' own elaboration.

17.5 Standardization example

After SUAs have been produced and balanced, standardization can be performed using **Equation 68** to **Equation 70**, using extraction rates and weights. For the current example, **Table 82** shows the correspondence between the SUAs and the corresponding FBS group. The table shows the correspondence at the item level, mapping CPC items to the FBS group. All items in the commodity trees are shown.

Products relating to one single commodity tree can fall into different FBS groups. The table also shows the SUA products having a weight equal to 0, which are not standardized in terms of quantities, but only in terms of macronutrients for reasons discussed in Section **15.3.2**.

For co-products and zero-weight products reported in **Table 82**, the extraction rates and shares are used to standardize quantities for the FBS groups of the example. **Table 83** shows the resulting FBS after standardization and aggregation for the three groups. The example does not show all possible FBS groups in the commodity tree but only the three selected for this example for illustrative purposes.

Table 82: CPC–FBS commodity mapping with weights

FBS GROUP	CPC ITEM (as in SUA)	ZERO
BARLEY & PRODUCTS	Malt, whether or not roasted	
BARLEY & PRODUCTS	Malt extract	X
BARLEY & PRODUCTS	Barley	
BARLEY & PRODUCTS	Barley flour and grits	X
BARLEY & PRODUCTS	Barley, pearled	
BARLEY & PRODUCTS	Bran of barley	X
BARLEY & PRODUCTS	Pot barley	
BARLEY, BEER	Beer of barley, malted	
BEVERAGES, ALCOHOLIC	Undenatured ethyl alcohol (> 80%)	
BEVERAGES, ALCOHOLIC	Undenatured ethyl alcohol (< 80%)	
CEREALS, OTHERS &PRODUCTS	Cereal preparations	X
SWEETENERS, OTHER & PRODUCTS	Other fructose and syrup	
SWEETENERS, OTHER & PRODUCTS	Glucose and dextrose	
SWEETENERS, OTHER & PRODUCTS	Fructose, chemically pure	
SWEETENERS, OTHER & PRODUCTS	Lactose	
SWEETENERS, OTHER & PRODUCTS	Molasses (from beet, cane and maize)	X
SWEETENERS, OTHER & PRODUCTS	Other non-alcoholic caloric beverages n.e.c.	
SWEETENERS, OTHER & PRODUCTS	Refined cane or beet sugar, solid; maple sugar and maple	
SWEETENERS, OTHER & PRODUCTS	Sugar and syrups n.e.s.	
WHEAT & PRODUCTS	Wheat	
WHEAT & PRODUCTS	Wheat and meslin flour	
WHEAT & PRODUCTS	Bran of wheat	X
WHEAT & PRODUCTS	Uncooked pasta	
WHEAT & PRODUCTS	Germ of wheat	X
WHEAT & PRODUCTS	Bread	
WHEAT & PRODUCTS	Pastry goods and cakes	
WHEAT & PRODUCTS	Starch of wheat	
WHEAT & PRODUCTS	Breakfast cereals	X
WHEAT & PRODUCTS	Wafers	

Source: Authors' own elaboration.

Table 83: FBS for wheat, barley and sweeteners

FBS GROUP	PRODUCTION	IMPORTS	DSTOCK	EXPORTS	FEED	SEED	LOSS	FOOD PROCESSING	FOOD	INDUSTRIAL	TOURIST
SWEETENERS, OTHER & PRODUCTS	130 685	1 217 230	19 403	518 180	0	-	-	0	800 598	0	56
BARLEY & PRODUCTS	8 379 700	81 702	-21 184	3 002 577	4 351 837	257 800	381 429	476 189	12 754	0	2
WHEAT & PRODUCTS	32 201 100	1 009 609	2 319 574	25 606 330	2 189 300	991 600	20 758	3 127 767	2 814 205	0	780 323

Source: Authors' own elaboration.

17.6 Standardization of nutrients example

As a final step of the FBS calculation, nutrients are computed using the food consumption values at the SUA level (i.e. for each commodity) as reported in **Table 79**, **Table 80** and **Table 81**. Nutrients are calculated by applying the calorie/fat/protein content factors to all SUA items with a non-zero food quantity. The nutrient values for barley and products of the current example are reported in **Table 84**.

Table 84: Nutrients for barley and products

COMMODITY	CALORIES (kcal/100g)	PROTEINS (g/100g)	FATS (g/100g)	FOOD QUANTITY (t)	CALORIES/YEAR (kcal)	PROTEINS/YEAR (t)	FATS/YEAR (t)
Barley	-	-	-	-	-	-	-
Barley flour and grits	-	-	-	-	-	-	-
Barley, pearled	330	10	1	1 173	3 871	117	12
Bran of barley	-	-	-	-			
Malt extract	310	6	0	617	1 913	37	0
Malt, whether or not roasted	357	10	2	5 921	21 138	592	118
Pot barley	-	-	-	-	-	-	-
TOTAL					**26 922**	**746**	**130**

Source: Authors' own elaboration.

The aggregation of nutrients is now a simple last step: all the nutrient variables (i.e. calories, fats, and proteins) are purely additive, so the standardized calories, fats and proteins for the FBS group barley and products are simply the sum of the total calories, fats and proteins, respectively, for each commodity.

Table 85 reports the final values of the nutrients for the current example. In the table, total values are also reported, which represent the overall values for the FBS of this example country and include values from all the other FBS groups.

Table 85: Wheat, barley and sweeteners nutrients

FBS GROUP	CALORIES/YEAR (kcal)	PROTEIN/YEAR (t)	FATS/YEAR (t)
WHEAT & PRODUCTS	8 099 333	247 733	47 924
BARLEY & PRODUCTS	26 922	746	130
SWEETENERS, OTHER & PRODUCTS	756 002	173	0
…	…	…	…
GRAND TOTAL	47 250 118	1 379 487	2 136 555

Source: Authors' own elaboration.

18 REFERENCES

Berthou, A. & Emlinger, C. 2011. The Trade Unit Values database. *Économie internationale, 128*, 97-117. https://doi.org/10.3917/ecoi.128.0097

Dietterich, T.G. 2000. Ensemble Methods in Machine Learning. In: J. Kittler & F. Roli, eds. *First International Workshop on Multiple Classifier Systems, Lecture Notes in Computer Science* (pp. 1-15). New York, Springer Verlag. https://web.engr.oregonstate.edu/~tgd/publications/mcs-ensembles.pdf

Eurostat. 2024. *European business statistics geonomenclature applicable to European statistics on international trade in goods – 2024 edition*. Luxembourg. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/w/KS-GQ-24-008

FAO. 2001. *Food Balance Sheets: A Handbook.* Rome. https://www.fao.org/4/x9892e/x9892e00.htm

FAO. 2017a. *Guidelines for the compilation of Food Balance Sheets.* Rome. https://openknowledge.fao.org/handle/20.500.14283/ca6404en

FAO. 2019. *The State of Food and Agriculture 2019. Moving forward on food loss and waste reduction.* Rome. https://doi.org/10.4060/ca6030en

FAO. 2023a. *Units of Measure – Version 3*. Statistical Standard Series. Rome. https://openknowledge.fao.org/handle/20.500.14283/cc6207en

FAO. 2023b. *FAO Statistics and Data Quality Assurance Framework*. Rome. https://openknowledge.fao.org/handle/20.500.14283/cc6683en

FAO. 2023c. *Observation Status Code List – Version 3*. Statistical Standard Series. Rome. https://openknowledge.fao.org/handle/20.500.14283/cc6208en

FAO. n.d.a. Methods, standards and classifications. In: FAO. [Cited 7 February 2025]. https://www.fao.org/statistics/methods-and-standards/agriculture/en

FAO. n.d.b. Questionnaire on Crop and Livestock Production and Utilization. In: FAO. [Cited 7 February 2025]. https://www.fao.org/statistics/data-collection/agriculture/en

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A. & Tempio, G. 2013. *Tackling Climate Change through Livestock – A Global Assessment of Emissions and Mitigation Opportunities.* Rome, FAO. https://openknowledge.fao.org/handle/20.500.14283/i3437e

Grigg, D. 1993. International Variations in Food Consumption in the 1980s. *Annals of American Geography 78*, pp. 251–266. http://www.jstor.org/stable/40572331

Grünberger, K. 2014. *Estimating food consumption patterns by reconciling food balance sheets and household budget surveys*. FAO Statistic Division Working Paper Series No. 14/08. Rome. https://openknowledge.fao.org/handle/20.500.14283/i4315e

Helsing, E. 1995. Traditional diets and disease patterns of the Mediterranean, circa 1960. *American Journal of Clinical Nutrition, 61*. https://doi.org/10.1093/ajcn/61.6.1329s

HLPE (High Level Panel of Experts on Food Security and Nutrition). 2014. *Food losses and waste in the context of sustainable food systems*. Rome. https://openknowledge.fao.org/handle/20.500.14283/i3901e

Hopper, G. 1999. Changing food production and quality of diet in India, 1947-98. *Population and Development Review* 25, 443–477. https://doi.org/10.1111/j.1728-4457.1999.00443.x

Jacobs, K.A. &Sumner, D.A. 2002. *The Food Balance Sheets of the Food and Agriculture Organization: A Review of Potential Ways to Broaden the Appropriate Uses of the Data.* University of California.

Kazuo, H. 1991. Current and World Supplies of and Requirements for Essential Amino Acids. *Food and Nutrition Bulletin* 13(1):1–12. https://doi.org/10.1177/156482659101300130

Mingione, M., Fabi, C. & Jona Lasinio, G. 2021. Measuring and Modeling Food Losses. *Journal of Official Statistics* 37(1):171–211. http://dx.doi.org/10.2478/JOS-2021-0008

National Research Council. 1994. *Nutrient Requirements of Poultry.* Ninth Revised Edition. Washington, DC, National Academy of Science.

National Research Council. 1998. *Nutrient Requirements of Swine.* 10th Revised Edition. Washington, DC, National Academy of Science.

Perthel, D. 1975. Engel's Law Revisited. *International Statistical Review*, vol. 4, n. 2, pp. 211–2018. https://doi.org/10.2307/1402900

Pinstrup-Andersen. 1993. *World Food Trends and How They May Be Modified.* Presented at the CGIAR International Centers Week, 25–29 October 1993. Washington, DC, CGIAR. https://cgspace.cgiar.org/server/api/core/bitstreams/d408c119-4076-4e4f-b51a-623aa8244601/content

Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Paper presented at the Meeting of the Expert Group on International Statistical Classifications, 19–22 May 2015. New York, United Nations Statistics Division. https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-16.PDF

Salathe, H. 1979. An Empirical Comparison of Functional Forms for Engel Relationships. *Journal of Agricultural Economics Research*, vol. 31. http://dx.doi.org/10.22004/ag.econ.148544

Sasaki, S.A. 1992. Value of Food and Agricultural Organization Data on Food-balance Sheets as a Data Source for Dietary Fat Intake in Epidemiologic Studies,. *American Journal of Clinical Nutrition 56*, 716–723.

Seré, C., Steinfeld, H. & Groenewold, J. 1996. *World Livestock Production Systems: Current Status, Issues and Trends*. Rome, FAO. https://openknowledge.fao.org/handle/20.500.14283/w0027e

Smil, V. 1987. *Energy, Food, Environment: Realities, Myths, Options.* Oxford, Clarendon Press.

Svedberg, P. 1999. 841 Million Undernourished? *World Development*, 27(12): 2081–2098. https://doi.org/10.1016/S0305-750X(99)00102-3

Taglioni, C., Rosero Moncayo, J.R. & Fabi, C. 2023. *Food loss estimation: SDG 12.3.1a data and modelling approach.* FAO Statistics Working Paper Series, No. 23-39. Rome, FAO. https://doi.org/10.4060/cc9173en

Trueblood, M.S., Shapouri, S. & Henneberry, S.P. 2001. Policy options to stabilize food supplies: a case study of Southern Africa. Agricultural Information Bulletins 33703, United States Department of Agriculture, Economic Research Service. https://ageconsearch.umn.edu/record/33703/files/ai010764.pdf

United Nations. 2015a. *Central Product Classification (CPC) version 2.1.* New York. https://unstats.un.org/unsd/classifications/unsdclassifications/cpcv21.pdf

United Nations. 2015b. Transforming our world: the 2030 Agenda for Sustainable Development. In: United Nations. [Cited 7 February 2025]. https://sdgs.un.org/2030agenda

United Nations. 2025. UN Comtrade Database. [Accessed 7 February 2025]. https://comtradeplus.un.org/

UNSD. 2013. *International Merchandise Trade Statistics: Compilers Manual, Revision 1*. New York, United Nations. https://unstats.un.org/unsd/trade/EG-IMTS/IMTS2010-CM%20-%20white%20cover%20version.pdf

WCO. 2022. HS Nomenclature 2022 edition. In: WCO. [Cited 7 February 2025]. https://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs-nomenclature-2022-edition/hs-nomenclature-2022-edition.aspx

WCO. n.d. What is the Harmonized System (HS)?. In: WCO. [Cited 7 February 2025]. https://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx

Young, V.R. & Pellett, P.R. 1990. Current concepts concerning indispensable amino acid needs in adults and their implications for international nutrition planning. *Food and Nutrition Bulletin* 12(4):1–13. https://doi.org/10.1177/156482659001200

Zhou, Z.H. 2009. Ensemble Learning. In: S.Z. Li & A. Jain, eds. *Encyclopedia of Biometrics* (pp. 270–273). Boston, Springer. https://doi.org/10.1007/978-0-387-73003-5_293

Table 51 references

FAO/Government of Kenya. 2018. *Kenya Food Composition Tables.* Nairobi. https://openknowledge.fao.org/handle/20.500.14283/i8897en, https://nutritionhealth.or.ke/programmes/healthy-diets-physical/food-composition-tables/

FAO. 2017b. *FAO/INFOODS Global Food Composition Database for Pulses Version 1.0 - uPulses 1.0*. Rome. https://openknowledge.fao.org/handle/20.500.14283/i6832en

Food Standards Australia New Zealand. 2022. Australian Food Composition Database – Release 2.0. [Accessed 7 February 2025]. https://afcd.foodstandards.gov.au/

Haytowitz, D.B., Ahuja, J.K.C., Wu, X., Somanchi, M., Nickle, M. Nguyen, Q.A. *et al.* 2019. USDA National Nutrient Database for Standard Reference, Legacy Release. In: USDA. [Cited 7 February 2025]. https://doi.org/10.15482/USDA.ADC/1529216

Hotz, C., Lubowa, A., Sison, C., Moursi, M. & Loechl, C.U. 2012. *A Food Composition Table for Central and Eastern Uganda, 2012.* HarvestPlus Technical Monograph 9. Washington, DC, IFPRI and Cali, CIAT. https://assets.publishing.service.gov.uk/media/57a08a90e5274a27b200066d/Tech_Mono_9_Web_0.pdf

Longvah, T., Ananthan, R., Bhaskarachary, K. & Venkaiah, K. 2017. *Indian Food composition Tables.* Hyderabad, National Institute of Nutrition. https://www.researchgate.net/publication/313226719_Indian_food_Composition_Tables

Ministry of Education, Culture, Sports, Science and Technology (MEXT). 2015. Standard Tables of Food Composition in Japan. Seventh revised edition. In: MEXT. [Cited 7 February 2025]. https://www.mext.go.jp/en/policy/science_technology/policy/title01/detail01/1374030.htm

National Food Institute, Technical University of Denmark. 2024. Danish Food Composition Database – Frida, version 5.3. [Accessed 7 February 2025]. https://frida.fooddata.dk/?lang=en

The New Zealand Institute for Plant & Food Research Limited & Ministry of Health. 2024. New Zealand FOODfiles™ 2024. In: New Zealand Food Composition Data. [Cited 7 February 2025]. https://www.foodcomposition.co.nz/foodfiles

Public Health England. 2019. Composition of foods integrated dataset (CoFID). In: gov.uk. [Cited 7 February 2025]. https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid

Shaheen, N., Rahim, A., Mohiduzzaman, M., Banue, C., Bari, L., Basak Tukun, A., Mannan, M.A., Bhattacharjee, L. & Stadlmayr, B. 2013. *Food Composition Table for Bangladesh.* Dhaka, University of Dhaka. https://www.researchgate.net/publication/268520007_Food_Composition_Table_for_Bangladesh

University of São Paulo. 2022. Brazilian Food Composition Table (TBCA). Version 7.2. [Accessed 7 February 2025]. http://www.fcf.usp.br/tbca

Vincent, A., Grande, F., Compaoré, E., Amponsah Annor, G., Addy, P.A., Aburime, L.C., Ahmed, D., Bih Loh, A.M., Dahdouh Cabia, S., Deflache, N., Dembélé, F.M., Dieudonné, B., Edwige, O.B., Ene-Obong, H.N., Fanou Fogny, N., Ferreira, M., Omaghomi Jemide, J., Kouebou, P.C., Muller, C., Nájera Espinosa, S., Ouattara, F., Rittenschober, D., Schönfeldt, H., Stadlmayr, B., van Deventer, M., Razikou Yiagnigni, A. & Charrondière, U.R. 2020. *FAO/INFOODS Food Composition Table for Western Africa (2019) User Guide & Condensed Food Composition Table / Table de composition des aliments FAO/INFOODS pour l'Afrique de l'Ouest (2019) Guide d'utilisation & table de composition des aliments condensée*. Rome, FAO. https://openknowledge.fao.org/handle/20.500.14283/ca7779b

Acquiring accurate and timely statistical data is crucial for shaping effective strategies in agrifood systems, directly impacting living standards across nations. Food balance sheets are a structured representation of a country's food availability, presented as an accounting of the supply and use of resources and food during a specified reference period. They are pivotal statistics, which provide comprehensive insights into national food consumption patterns, levels and trends.

This handbook provides an essential description of the methodology used by FAO to develop food balance sheets. It aims to familiarize Member Nations, food budgeting institutions, and other interested parties with the core processes involved in data collection and imputation strategies specifically for agricultural and livestock products.

ISBN 978-92-5-139645-2 ISSN 1014-3378

CD4472EN/1/02.25