

Standardization & Balancing for Food Balance Sheet Calculation

Cristina Muschitiello

Food and Agriculture Organization of the United Nations

24 May 2018

Abstract

This vignette provides a description of the Workflow of the Standardization and Balancing Procedure: This represents the process of aggregating and balancing all the accounts of individual products to their primary equivalents. Other documents will cover technical aspects.

Contents

Disclaimer	3
The Food Balance Sheet Framework	3
The Balancing equation and its variables	3
Commodity Tree	6
Standardization and Balancing	7
1 Data Pull	8
the Initial Sua Unbalanced	9
2 The Sua Filling	11
Main rules and the rationale	11
Step 1: Filling of production figures	13
Step 2: Filling of food processing	16
Step 3: Filling of other missing Utilizations	19
$Imbalance = 0$ (Scenario A)	21
$Imbalance < 0$ (Scenario B)	21
$Imbalance > 0$ (Scenario C)	24
-> The Utilization Table	24
-> The Multiple Filler Approach and the Inverse Ranking Rule	24
The “Sua Balanced” table	27
Nutritive Values and DES calculation	29
3 Standardization	30
The “Fbs Standardized” table	32
4 The FBS Balancing	34
Nutritive Values updates	35
5 Final aggregation:FBS items and FBS aggregates	36
Conclusions and future work	39
Related documentation	39

List of Tables

1	Commodity Tree - China/Wheat/2014 example	7
2	Unbalanced Sua table - China/Wheat/2014 example	10
3	Sua table with Imb1 - China/Wheat/2014 example	14
4	Sua table with Production filled/incremented - China/Wheat/2014 example	15
5	Commodity Tree with weights - China/Wheat/2014 example	17
6	Sua table with Food Processing filled - China/Wheat/2014 example	18
7	Sua table with Imb2 - China/Wheat/2014 example	20
8	Sua table with Utilizations filled - China/Wheat/2014 example	23
9	<i>Utilization Table</i> Wheat and meslin Flour - China/Wheat/2014 example	24
10	<i>Utilization Table</i> Bran of Wheat - China/Wheat/2014 example	25
11	Rank Weight of the elements of FBS	26
12	Sua of Cocoa Beans Madagascar 2014 - before <i>Sua Filling</i>	27
13	<i>Utilization Table</i> for Cocoa Beans Madagascar	27
14	Sua line of Cocoa Beans Madagascar 2014 - after <i>Sua Filling</i>	27
15	Sua Balanced - China/Wheat/2014 example	28
16	Sua Unbalanced DES - China/Wheat/2014 example	29
17	Generic SUA for Standardization	30
18	Availabilities and shares of parent/child for Standardization	31
19	DES of <i>Sua Standardized</i> - China/Wheat/2014 example	32
20	Sua Standardized	33
21	Measurement Errors percentages	34
22	Final Balancing - China/Wheat/2014 - Wheat	35
23	Final Balancing - China/Wheat/2014 - Food Preparations	36
24	Fbs Tree [extraction of the main table]	37
25	FBS Final Aggregations - China/Wheat/2014	38

List of Figures

1	Commodity Tree for Wheat in China mainland 2014	6
2	Standardization and Balancing Overall Workflow	8
3	Data Pull from datasets containing data for each separate variable	9
4	The Sua Filling simplified workflow	11
5	The Sua Filling diversified workflow	13
6	Scenarios of the *Sua Filling*	19
7	Workflow of Sua Filling when there is a negative imbalance	22
8	Workflow of Sua Filling when there is a positive imbalance	25
9	Sua Balanced and the calculation of DES	30

Disclaimer

This Working Paper should not be reported as representing the official view of the FAO. The views expressed in this Working Paper are those of the author and do not necessarily represent those of the FAO or FAO policy. Working Papers describe research in progress by the authors and are published to elicit comments and to further discussion.

This paper is dynamically generated on May 24, 2018 and is subject to changes and updates.

The Food Balance Sheet Framework

A food balance sheet can be defined as an aggregated and analytical data set that “presents a comprehensive picture of the pattern of a country’s food supply during a specified reference period.”¹ FBS are presented as products accounts, where the quantities allocated to all the sources of total supply must be equal to the quantities allocated to all the sources of utilization. This balance is compiled for every food item consumed within a country at primary commodity equivalent basis, and all of the primary commodity equivalent balances are then combined into a single overall FBS. FBSs are, then, expressed in terms of per capita supply for each food item by dividing by the country’s population, with the per capita supplies being expressed both in terms of quantity and, through the application of food conversion factors, in terms of caloric value, protein, and fat content. These per capita estimates of caloric value for individual food products are then summed to obtain the total daily per capita Dietary Energy Supply (DES) of a country.

While FBS are typically only published at the primary commodity equivalent level to facilitate interpretation, they are created for all the commodities, accordingly to the *Central Product Classification (CPC)* System². Indeed, a balance for wheat alone would in most cases include little or no food use, because wheat is commonly processed into flour before it is consumed by humans, and flour is then used to produce various other derived products such as bread, pastries and pasta. Because there is both supply and demand both for the primary commodity and the derived commodities, individual accounts are kept for all of products. Individual accounts for commodities are called *Supply-Utilization Accounts (SUA)* and the process of obtaining FBS starting from SUAs is the *Standardization & Balancing* process.

The Balancing equation and its variables

At the most basic level, Food Balance Sheets are, like all commodity balances, simple identities. In these identities, the sum of all supply variables is equal to the sum of all demand variables; the two most common identities set domestic supply equal to domestic demand (first equation) or total supply equal to total demand (second equation).

$$P_{ijt} + I_{ijt} - X_{ijt} - \Delta St_{ijt} = FP_{ijt} + Fo_{ijt} + Fe_{ijt} + Lo_{ijt} + Se_{ijt} + IU_{ijt} + T_{ijt} + ROU_{ijt} \quad (1)$$

$$P_{ijt} + I_{ijt} - \Delta St_{ijt} = X_{ijt} + FP_{ijt} + Fo_{ijt} + Fe_{ijt} + Lo_{ijt} + Se_{ijt} + IU_{ijt} + T_{ijt} + ROU_{ijt} \quad (2)$$

where the i index runs over all countries, the j index over all commodities, and t over years and where, dropping indices for brevity:

- P =Production
- I =Imports

¹For this definition and a more extended description of the motivation behind the development of FBS, see FAO, 2001, *Food Balance Sheets: A Handbook*, available at: <http://www.fao.org/docrep/003/X9892E/X9892E00.HTM>. Accessed on 19 January 2017.

²The CPC represents a comprehensive classification of products into a system of categories that are both exhaustive and mutually exclusive. It is based on a set of internationally agreed concepts, definitions, principles and classification rules. The custodian of this classification is the UNSD. For more information, see the *Guidelines on International Classifications for Agricultural Statistics* and the *UNSD official document on CPC Version 2.1*

- X =Exports
- S =Stock level
- ΔSt_t = Stock Variation = $St_t - St_{t-1}$
- FP_{ijt} = Food Processing
- Fo =Food availability
- Fe =Feed
- Lo =Losses
- Se =Seed
- IU =industrial use
- T =Tourist consumption
- ROU =Residual Other Use
- $TS = Totalsupply = P_{ijt} + I_{ijt} - \Delta St_{ijt}$

All variables are expressed in the same measurement unit: **metric tonnes**. At international level, the primary data source that FAO uses to compile the the Supply Utilization Accounts/Food Balance Sheets are the data as collected through the annual *Agriculture Production Questionnaires*. Unfortunately, measured values are mostly limited to variables on the supply side (production, imports and exports), while, on the demand side, most values are imputed data. In particular³:

- *Production (P)*: Data on production are data at farmgate level. As data on production are very important for countries, these data are very often survey-based data. Nevertheless, not all countries collect data of production for all commodities. Therefore other data collection methods are used, like records of private firms and commodity organization. When no other data are available, Production figures are imputed or estimated. Imputation and estimation procedures depends on the specific commodity. There are different procedures for crops and livestock but all based on an *ensemble approach*. Production data in the FBS framework are collected, imputed or estimated for all the primary commodities and for a set of derived commodities⁴.
- *Import(I) and Export(X)*: Data on Trade are, mainly, official from international trade databases, like UNSD and EUROSTAT, at HS6 commodity level⁵. Official data are integrated with supplementary data having the main aim of filling in all the information hidden by the unrecorded trade and coming, mainly, from trading partners. HS6 classification is more detailed than the used CPC classification, therefore, these data are aggregated in CPC commodities before being used in the standardization process⁶.
- *Stock Variation (ΔSt_t)*: In the FBS framework, stocks are considered as *changes in stocks* from one time period to the next. Moreover they are considered as a component of supply. Therefore, the $-$ sign indicates that the stock is decreased, which means that the stocks are available as a supply, while the $+$ sign indicates that the stocks have increased and they are, therefore, considered as a utilization for that commodity. Changes in stocks are typically limited to a short number of commodities, mainly grains, pulses and sugar and, because they are very rarely measured by country, figures are very often imputed or estimated⁷. Estimation of *changes in stock* is based on opening stocks figures through an approach that maintain time consistency of available data and official data⁸.
- *Food availability (Fo)*: Food availability is defined as the quantity of any substance that is available for human consumption at the retail level by the country's resident population during a given reference period. Official data of food availability come from questionnaires, industrial output surveys and

³For more details on FBS variables please see the latest version of the *Resource Book*

⁴See *Production module* documentation for more details on procedures and list of commodities

⁵Harmonized Commodity Description and Coding Systems (HS) is an international classification of products held by UNSD. Is made of six-digit level codes and used worldwide for trading data classifications. See official *HS6 UNSD webpage* for more details

⁶See *Trade module* documentation for more details

⁷For a complete list of stock commodities and for details about the imputation methodology, please refer to specific documentation for stock

⁸All data are marked as *official*, *semi-official* or *unofficial*, depending on the source they come from, through *flags*. Flag management is one of the core responsibilities of the *Office of the Chief Statistician* Department in FAO. Flags are used from all the estimation procedure for distinguishing between different level of reliability in the data. The most reliable data are used to estimate missing or less reliable data.[this has to be better specified]

household consumption or expenditure surveys. When these sources of data are not available, food availability data are imputed or estimated. Not all CPC commodities are Food commodities, as not all commodities are used for human consumption. Food commodities are divided in two main groups: *Food Estimates* and *Food Residual* and estimated differently depending on the pertaining group, respectively as linear or logarithmic function of income elasticity of demand, GDP per capita, and population, or as residual quantity of production and net trade quantities⁹.

- *Feed (Fe)*: Feed demand is increasing because of the increase in income of developing countries. Animal feed may vary among countries due to the difference in livestock and the diversity of commodity used for livestock's rations. Official feed demand data might be available from specific questionnaires. Even when available, these data need to be cross-checked against livestock availability in terms of requirements. When official data, and also other sources of semi-official data, are not available, feed data are estimated as a function of livestock availability and livestock feed demand in terms of energy and protein requirements, in accordance with an inventory of the potential feed supply's products of any country.
- *Seed (Se)*: Official seed data may come from agricultural surveys, while other sources of data might be found in some technical publication. When data are not available these are estimated as a function of a seeding rate and a sown area in the following year.
- *Tourism Consumption (T)*: Tourism consumption is considered here as a separate utilization variable, while in the past it was included in the "other utilization" catchall category. Official data for this variable are rare and may come from tourism offices or collected by tourism boards through surveys. UNWTO is an alternative source of data, but other authorities might also be used. Imputations and estimations of tourist food are made as a function of food figures.
- *Industrial Use (IU)*: This variable refers to utilization of any food items in any non-food industry. Non-food use of food commodities is growing and is highly context and country specific. For this reason there are not, at the moment, suggestions on how to impute and estimate missing figures. As a consequence, Industrial data available for Food Balance Sheets are only those coming from Official or unofficial sources. At the moment the data used comes from USDA and from questionnaires.
- *Loss (L)*: FAO has developed the Global Food Loss Index (GFLI) that focuses on the supply-side aspects of improving the efficiency of global food supply chains. The index is based on a set of primary commodities that are key in agricultural production systems, including crops, livestock, and fisheries. In order to track losses without compounding production variability, losses are expressed as a percentage and are aggregated using fixed quantities and prices. The primary data source that FAO uses for compiling GFLI are loss factors as collected by Questionnaires. Other sources are publications and reports from subnational reports, academic institutions, international organizations and so on. The missing data are imputed using a hierarchical model based on commodity groups¹⁰.
- *Residual and other use (ROU)*: ROU is used to capture categories of products that do not follow in any other category and that might be considered "not important" for the FBS scope. Normally, these residual commodities are different from country to country and for this reason they fall in this variable. ROU are set not to be higher than 5% of total supply and are calculated at-post as absorbing element, in the sense that it absorbs part or all of the imbalance that may exist, at FBS commodity aggregate level, after the standardization process. Any imbalance bigger than 5% of total supply is balanced through a balancing mechanism that will be later specified.
- *Food Processing (FP)*: This variable represents the amount of the availability of a commodity that enters a manufacturing process to be transformed into a derived commodity. Food processing is not officially measured, nor collected via official sources. This variable is entirely calculated during the standardization process by applying extraction rates to the amount of production of the derived commodity. This will be better specified in section 2-Step2 of the present document.

⁹For a complete list of food commodities of the two typologies and for details about the imputation methodology, please refer to specific documentation for Food availability

¹⁰ask for links to a proper documentation

Commodity Tree

The process of combining commodity balances for creating Food Balance Sheets is based on a structured and clear set of relationships between commodity given by the *Commodity tree*. The majority of the commodities are produced from one (or more) commodity (/ies), called *parent* commodity(/ies), and/or are themselves parent of one (or more) *child* (*children*) commodity (/ies). These structure creates an intense and articulated network of relationships at different levels: primary commodities, like crops, are *parent* commodities and, also, *zero-level* commodities from which *children* commodities of *level-1* are produced, which are in turn, used to produce other commodities of a gradually “lower” level. In commodity trees, the bigger the level number, the lower the processing level. There are as many commodity trees as the number of process chains in a country. Fundamental characteristics of commodity trees are¹¹:

1. Each commodity tree is represented as a flowchart of the kind presented in Figure 1 where:
 - **nodes** represent commodities,
 - **edges** represent production processes ,
 - **joints** indicate where a single production process creates more that one commodity. These commodities are, then, called *by-products* or *co-products*.

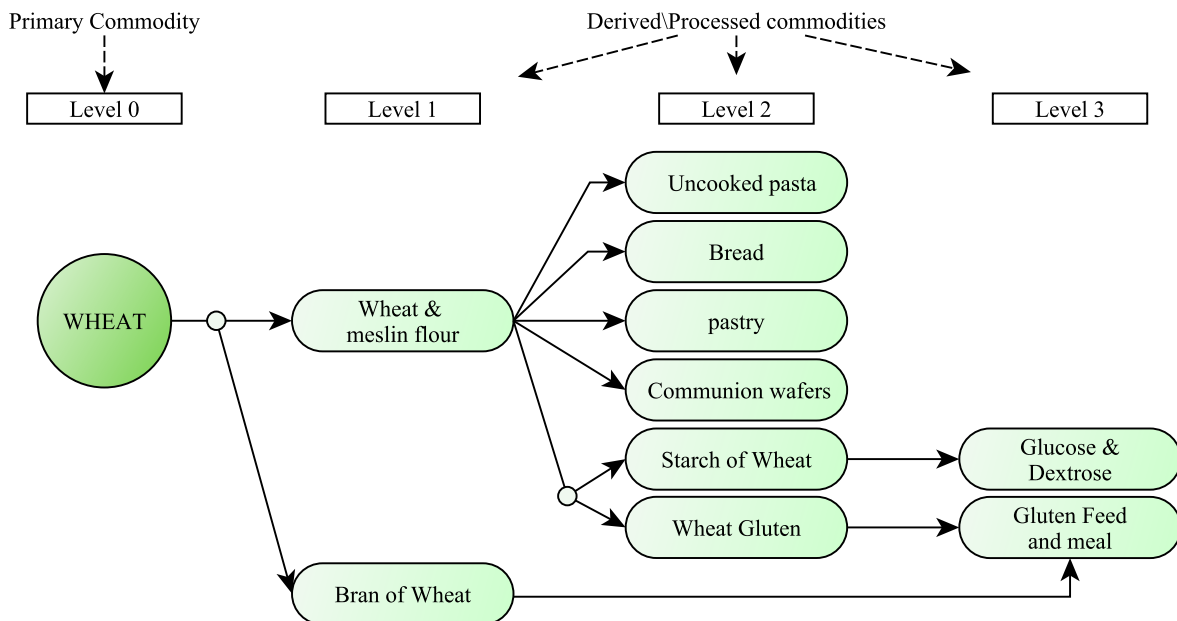


Figure 1: Commodity Tree for Wheat in China mainland 2014

2. Not all the countries have the same production processes, because countries have different technologies and primary products availabilities. Therefore, the commodity trees are not the same across countries.
3. if a production process is active in a country, this is expressed through the existence of a conversion factor called **extraction Rate**. An Extraction rate (eR) represents how much amount of the child commodity is produced from 1 unit of parent commodity. It is expressed as a ratio of the processed product obtained from the processing of the parent/originating product.
4. Some child commodities can be produced starting from more that one parent commodity. A second conversion factor exists representing the amount of a child commodity that is produced from each parent commodity. This conversion factor is called **Share**. Shares represent the amount of the child commodity

¹¹For a more detailed description of commodity trees, please see the specific documentation. The reference document, at the moment is the *technical conversion factor* document available in the documentation folder on *GitHub*

that is produced from the specified parent and are expressed as a ratio. Shares are generically defined as:

$$s_{cp} = \frac{availability_{p(c)}}{\sum_{p=1}^A availability_{p(c)}} \quad (3)$$

where $availability_{p(c)}$ is the availability of each parent p of child c expressed in terms of c (in child equivalent).

Commodity trees are presented in tables like Table 1, which represents the same example of Figure 1. In the table each production process is represented in a separate row.

Table 1: Commodity Tree - China/Wheat/2014 example

Country	Year	ParentName	ChildName	eR	share
China, Mainland	2014	Wheat	Wheat and meslin flo	0.78	1.00
China, Mainland	2014	Wheat	Bran of Wheat	0.22	1.00
China, Mainland	2014	Wheat and meslin flo	Uncooked pasta, not	1.00	1.00
China, Mainland	2014	Wheat and meslin flo	bread	1.00	1.00
China, Mainland	2014	Wheat and meslin flo	pastry	1.00	1.00
China, Mainland	2014	Wheat and meslin flo	Starch of Wheat	0.75	1.00
China, Mainland	2014	Wheat and meslin flo	Wheat Gluten	0.08	1.00
China, Mainland	2014	Wheat and meslin flo	Communion wafers, em	1.00	1.00
China, Mainland	2014	Starch of Wheat	Other Fructose and S	1.00	1.00
China, Mainland	2014	Wheat Gluten	Gluten Feed and Meal	1.00	0.33
China, Mainland	2014	Bran of Maize	Gluten Feed and Meal	1.00	0.67

There are some concepts linked to the *Commodity tree* framework:

- **Proxy-Primary** commodities. These are a set of commodities that are children of other commodities but, because they are important in representing the food availability of a country, are not aggregated to their primary commodities, but are kept separated. These commodities are *cut* from the tree of the primary commodity/ies and, if they can be processed in other products, have their own commodity tree. The name *proxy-primary* is assigned because they are considered as primary-commodities in the *Standardization & Balancing* process.
- **No-Tree** commodities. These are *zero-level* commodities that are primary commodities and are never processed in other products. As they are not involved in any production process, there is no tree associated to them. Notice that, even a commodity that is included in the commodity tree of one Country, might be a No-tree commodity for another country or another year, if no production processes have been activated for that specific Country or year.

Standardization and Balancing

The *Standardization & Balancing* process is presented in Figure 2. It involves 5 main steps and requires a some auxiliary information table.

The 5 steps are:

1. Data Pull,
2. Sua Filling,
3. Standardization,
4. Balancing,

5. FBS aggregation.

The additional information's tables are:

- *Utilization Table*,
- *Zero-Weight* table,
- *cut* table,
- *Fbs Tree*.

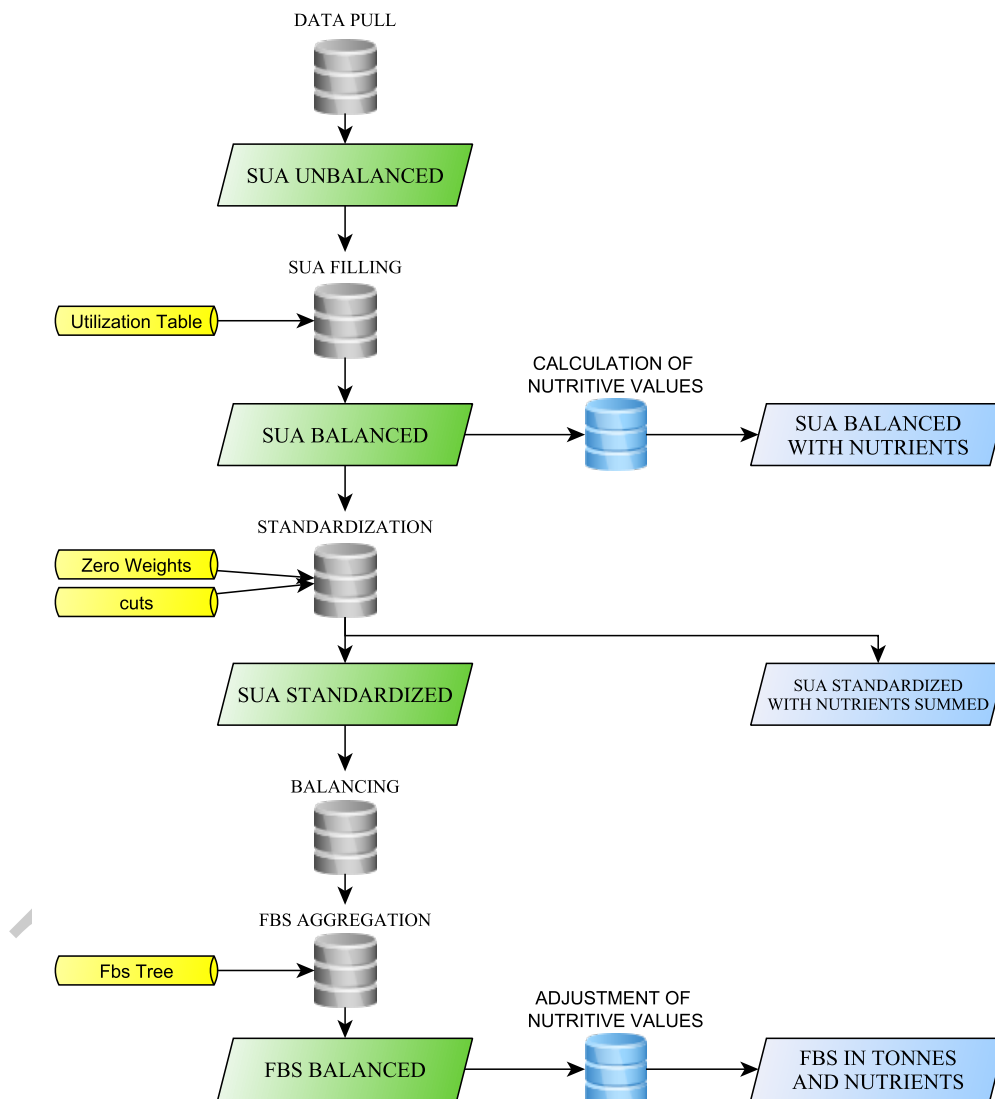


Figure 2: Standardization and Balancing Overall Workflow

1 Data Pull

The process of creating FBSs starts by considering the initial commodity balance for each CPC commodity, either primary and derived, with the different variables of the equation (as listed and briefly described in the previous section) filled with figures as available from official or other sources and from imputation and

estimation methods, when applied. In other words, the process starts by pulling figures inside a so-called *Sua Unbalanced*. In this initial account, food processing and ROU figures are not available (because they, by default, will be measured during the process), whereas the figures for all other variables have been already collected, imputed and estimated through a specific *module* (Figure 3).

A *module*, in the FBS Framework, is an R-script, written by an R-developer, for the execution of a set of operations (either data import, manipulation, imputation or estimation) required for compiling the time series of one variable. There is at least one module (there might be more) for each variable of the FBS. Each module produces figures that are collected in a dataset inside the *Statistical Working System (SWS)*¹². Output data of a module may become input data of another module, this circumstance creating a precise sequence for the execution of a complete FBS. In the present document, we are analyzing the workflow of the Standardization process as starts after all the modules have run and have produced reliable data or each variable. The detailed description of the workflow for the execution of all the modules of the FBS will be given in a separate document.¹³.

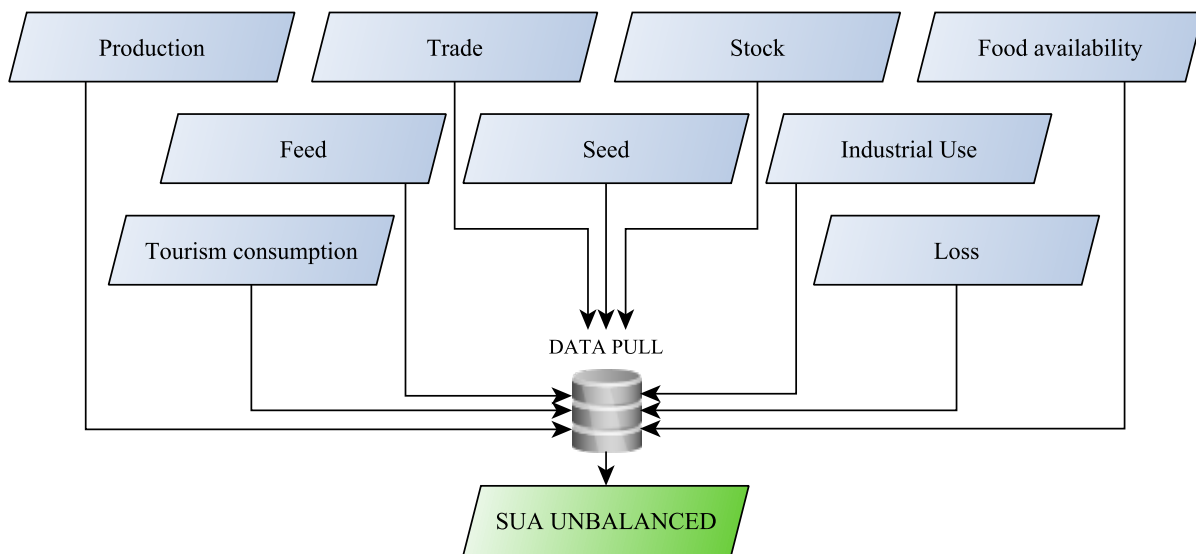


Figure 3: Data Pull from datasets containing data for each separate variable

the Initial Sua Unbalanced

After pulling all data, the process of compiling Food Balance Sheets is a non-complete supply-utilization account. The non-completeness of the SUAs is due to different reasons: first, as already said, some variables are not collected, nor estimated before the process begins, second there is the model for industrial use that does not impute or estimate data, but just collects data from different sources, this opening the strong possibility not to have values where they are supposed to exist and, also, not guaranteeing consistency of data over time. Thirdly, modules might, sometimes, fail in the imputation, because of the strong complexity and structural diversity of the input data.

¹²SWS is an internal Working System providing a platform for statisticians and statistical clerks to collect, collate, validate and correct data. Moreover, the platform supports the possibility of performing imputations of data based on statisticians' knowledge and development.

¹³[report the document when it will be available]

Table 2: Unbalanced Sua table - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU
<i>Wheat</i>	<i>126,208,400</i>	<i>2,971,249</i>	<i>957</i>	<i>1,120,565</i>		-	<i>29,181,617</i>	<i>4,277,567</i>		<i>2,985,279</i>	<i>2,713,000</i>	-
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	-			-17,345			-
Mixes and doughs for		6,497	38,072		0	-	0		0			-
Other Fructose and S	126,277	3,659	162,324		0	-			0			-
Starch of Wheat	239,816	11,035	40,311			-	172,196			7,919		-
Wheat Gluten	25,580	877	117,373			-	0					-
Communion wafers	13,263	8,796	5,822		16,241	-			-4			-
Uncooked pasta	1,415,692	12,520	22,550		1,405,661	-			-362			-
Food Preparations of		69,686	21,977		47,709	-			-12			-
Bran of Wheat	21,414,279	156,359	2,200		16,500,000	-	4,827,244		-4,252			-
Gluten Feed and Meal	793,740	160,231	529,333			-						-
bread	15,485	2,897	4,210		14,175	-			-3			-
pastry	193,950	89,593	117,630		165,914	-			-43			-

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

Supply utilization accounts are typically organized into tables where the SUA for the primary commodity is at the top, and the SUAs for all of the products derived from that commodity follow (Table 2). Commodities in the SUA table are in a parent-child relationship representing a real commodity process. In the example of Table 2 the relationship between commodities is the following:

- Maize is processed in flour, Bran and Breakfast Cereals,
- Flour of maize is processed into Starch and Gluten, Wafers, pastry, bread and pasta
- Starch is processed into Glucose and Dextrose,
- Gluten and Bran are processed into some feed and meal,

All these relationships represent the *Commodity tree* of wheat in China in 2014. In particular, our example represents the commodity tree of Wheat in China in 2014 and is displayed, in its graphical form, in Figure 1.

There are unbalanced SUAs for all combinations of commodity tree/country/year.

2 The *Sua Filling*

Main rules and the rationale

The Standardization, i.e. the conversion of the variables of any child commodity in primary-equivalent commodity, requires all variables of the SUAs to be filled (except for ROU, which is a residual variable and is imputed at the very last step of the process), “all variables” meaning “all variables that are supposed to be filled”. Indeed, not all variables have to exist for all commodities. Consider maize as an example: if there is no figure of *food availability* for maize, that does not mean that there is a missing figure because maize is rarely used for human consumption directly. However, in some country people do eat raw maize and, therefore, for those countries, a food figure is expected and, if missing, that has to be taken into account.

A simplified workflow of *Sua Filling* is reported in Figure 4.

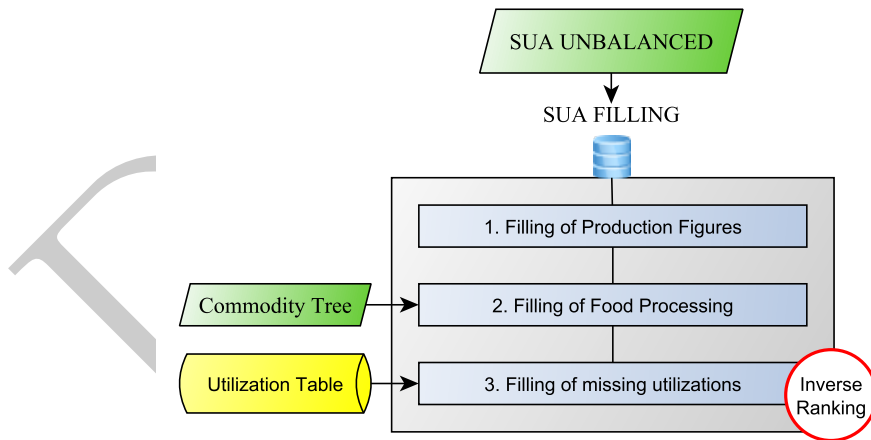


Figure 4: The Sua Filling simplified workflow

Food Processing needs to be computed first and it does not come from any module. For the food processing to be correctly computed, a check on the other variables has to be performed, in order to be sure that the food processing might be correctly imputed. In particular, Production is checked first, as the *Food Processing* figure of any commodity is based on the production one. Then *Food processing* can be computed and, after that, other missing figures, if existing, are filled.

This last step requires two major and not straightforward steps, representing each a separate issue, for solving which, an algorithm has been developed that responds to the following rationale and rules:

1. *Identify which are the missing values that have to be filled.*
2. *Decide how to fill those missing values.*
3. *Preserving reliability of sources of data*

Not all the variables (supplies and Utilizations) are to be filled:

- Trade (import and Export) data are shared with countries and published by FAO. Therefore these figures have a separate process of validation that makes them reliable enough not to be questioned during the standardization process.
- Production figures of primary commodities are also published by FAO. As previously mentioned, production values can have various flags which assign different degrees of reliability to these figures. The *Sua Filling* treats production figures accordingly to flags.
- Stock figures are also not touched during the process of filling. This decision was taken because the standardization process works year by year, while *stock variation* is a time-based variable that has to be always treated looking at the data over time. The cross year approach is incompatible with this characteristic. Any inconsistent or missing value for stock found would require to go back to the module imputing figures and check for possible errors. If no errors can be found, a manual correction would be required.

Information about the remaining variables, which are all Utilizations, is taken from FBSs and SUAs of past years. FBSs published with the previous methodology and their corresponding SUAs can give information about which are the Utilizations supposed to be active for every single commodity in each country.

Therefore, these pieces of information have to be given externally, as the process is automated and there is non-human knowledge intervening in training the machine that is performing the calculations. The previous methodology used for compiling FBS made massive use of manual interventions while the new methodology tries to avoid it, but taking advantage of the work and effort of the country experts that compiled the FBSs in the past. Past information is embedded in the new methodology through the use of the *Utilization Table* which helps to identify which variables have to be filled, commodity by commodity, while the amount of them to be assigned to figures is identified through a set of rules, known as *Sua Filling*. This procedure makes use of the imbalance between supply and utilization values for each commodity and fills the figures that have been recognized as missing from the *Utilization Table*. The filling procedure assigns figures using an ad-hoc algorithm, called *Inverse Ranking*, designed for dimensioning the figures in accordance with the importance they had in the past FBSs, as suggested also from the *Utilization Table* through ranks associated to each variable. As a consequence, all the commodities entering the *Sua Filling* procedure are eventually balanced.

Not all the commodities enter the process the same way. In particular, a distinction is made between primary commodities and derived commodities:

- Primary commodities are normally the commodities for which more reliable figures are produced. Moreover, FAO often publishes information about primary commodities, therefore values are not expected to be modified during a computational process.
- Derived commodities are more affected by modifications during this process, according to the flag of each figure.

Therefore, the steps of *Sua Filling* process affect the SUA differently from line to line. Figure 5 shows the general Workflow of the *Sua Filling* highlighting the distinction between primary and derived commodities.

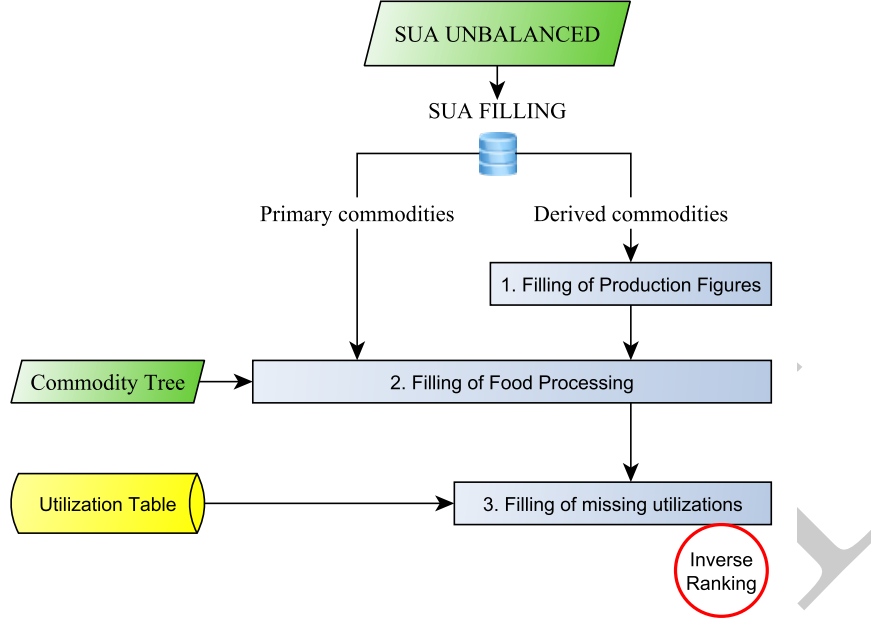


Figure 5: The Sua Filling diversified workflow

Step 1: Filling of production figures

Only Derived commodities are involved in this process. The *Sua Filling* detects the Production figures that need to be filled or modified, by looking at the following *Imbalance*:

$$Imb1_{ijt} = P_{ijt} + I_{ijt} - X_{ijt} \quad (4)$$

where $Imb1_{ijt}$ is the Imbalance and is enumerated because is different from a second Imbalance that will enter into the process later. Here:

- If $Imb1_{ijt} < 0$ there is no supply enough for the export and this is interpreted as the need for creating/incrementing the Production figure. The new production figure (P_{ijt}^*) is computed as:

$$P_{ijt}^* = P_{ijt} + Imb1_{ijt} \quad (5)$$

- If $Imb1_{ijt} \geq 0$ there is enough supply and, therefore, no need for changing the production figure.

The reason for excluding other Variables from the computation of the imbalance here is that, at this step of the process, Trade variables are the most reliable and those for which, to the highest level of probability, there are all the figures filled. In Table 3 *Imb1* of our example is reported. Notice that there are 3 rows with negative values, but one of them is considered a Primary commodity. Indeed, *Mixes and dough for the preparation of baker's wares* is a commodity never processed and, therefore, is considered primary commodity and added to wheat in the FBS item *Wheat and products*. Consequently production of only 2 commodity is changes, as shown in Table 4.

Table 3: Sua table with Imb1 - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb1
Wheat	126,208,400	2,971,249	957	1,120,565		-	29,181,617	4,277,567		2,985,279	2,713,000	-	<i>129,178,692</i>
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	-			-17,345			-	<i>70,344,381</i>
Mixes and doughs for		6,497	38,072		0	-	0		0			-	<i>-31,575</i>
Other Fructose and S	126,277	3,659	162,324		0	-			0			-	<i>-32,388</i>
Starch of Wheat	239,816	11,035	40,311			-	172,196			7,919		-	<i>210,541</i>
Wheat Gluten	25,580	877	117,373			-	0					-	<i>-90,916</i>
Communion wafers	13,263	8,796	5,822		16,241	-			-4			-	<i>16,237</i>
Uncooked pasta	1,415,692	12,520	22,550		1,405,661	-			-362			-	<i>1,405,661</i>
Food Preparations of		69,686	21,977		47,709	-			-12			-	<i>47,709</i>
Bran of Wheat	21,414,279	156,359	2,200		16,500,000	-	4,827,244		-4,252			-	<i>21,568,438</i>
Gluten Feed and Meal	793,740	160,231	529,333			-						-	<i>424,638</i>
bread	15,485	2,897	4,210		14,175	-			-3			-	<i>14,172</i>
pastry	193,950	89,593	117,630		165,914	-			-43			-	<i>165,914</i>

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=Food Processing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=Industrial Use, L=Loss, ROU=Residual and other uses

Table 4: Sua table with Production filled/incremented - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb1
Wheat	126,208,400	2,971,249	957	1,120,565		-	29,181,617	4,277,567				-	129,178,692
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	-			-17,345			-	70,344,381
Mixes and doughs for		6,497	38,072		0	-	0		0			-	-31,575
Other Fructose and S	**158,665**	3,659	162,324		0	-			0			-	0
Starch of Wheat	239,816	11,035	40,311			-	172,196			7,919		-	210,541
Wheat Gluten	**116,496**	877	117,373			-	0					-	0
Communion wafers	13,263	8,796	5,822		16,241	-			-4			-	16,237
Uncooked pasta	1,415,692	12,520	22,550		1,405,661	-			-362			-	1,405,661
Food Preparations of		69,686	21,977		47,709	-			-12			-	47,709
Bran of Wheat	21,414,279	156,359	2,200		16,500,000	-	4,827,244		-4,252			-	21,568,438
Gluten Feed and Meal	793,740	160,231	529,333			-						-	424,638
bread	15,485	2,897	4,210		14,175	-			-3			-	14,172
pastry	193,950	89,593	117,630		165,914	-			-43			-	165,914

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse,

L=Loss, ROU=Residual and other uses

^b Starred figures are those that have changed from the previous table

Step 2: Filling of food processing

When Production has been checked, created or incremented, *Food Processing* can be calculated. Food processing is the amount of the supply of a commodity processed into derived commodities. *Food processing* is calculated for all the parent commodities as the sum of the food processing of all the possible derived commodities. For example, Wheat in China is processed into Flour of Wheat, Bran of Wheat and Breakfast Cereals: Flour and Bran are produced during the same production process (they are called *co-products* or *by-products*), therefore the same amount of wheat is used for producing all of them, Breakfast Cereals comes from a separate production process instead. Therefore, *Food processing* of Wheat is given as the sum of the amount of wheat's supply used for producing Flour and Bran + The amount used for producing Breakfast Cereals. Calculation of Food processing happens by level, then after having calculated the amount for wheat, the amount of supply of flour of wheat used for producing Starch, Gluten, bread and the other derived of wheat flour becomes the *Food processing* of flour of Wheat and so on, for any subsequent level.

The formula for calculating *FP* is based on the production of the child product:

$$FP_{pjt} = \sum_{c=1}^C \left(\frac{P_{cjt}}{eR_{p \rightarrow c}} \right) \times s_{cp}^1 \times w_c \quad (6)$$

where:

- FP_{pjt} is *Food processing* of the generic parent p .
- $c = 1, 2, 3 \dots C$ are the C children c of parent p .
- P_{cjt} is the Production of Child c .
- $eR_{p \rightarrow c}$ is the *extraction Rate* from parent p to child c .
- s_{cp}^1 is the *share* of child c from parent p specific for Food Processing calculation. It is defined as:

$$s_{cp}^1 = \frac{availability1_{p(c)}}{\sum_{p=1}^A availability1_{p(c)}} \quad (7)$$

where $availability1_{p(c)}$ is the availability of each parent p of child c expressed in terms of c (as say in *child equivalent*). Is is enumerated as 1 because a slightly different availability will be found in a following step:

$$availability1_{p(c)} = (P_{pjt} + I_{pjt}) \times eR_{p \rightarrow c} \quad (8)$$

- w_c is the *weight* of child c . This parameter is used for identifying and treating co-products. Indeed, if the production of two or more co-products is used for producing *Food processing* of the parent commodity producing them, the result would be a Food processing twice the size of what it should be. This would happen because the two co-products are produced during the same process, in other words, from the same amount of the parent commodity. For this to be taken into account, a weight is used, which is 1 for the commodity that has to be considered for determining Food processing of parent commodity and 0 for the child commodity that are co-products (equation 9). Please notice that, the co-products are still important in the standardization, because they contribute in the creation of the calories availability of the country, therefore, these commodities (Also called *zero weight* commodities) are multiplied by 0 only when quantities are treated, they will instead be considered when calories will be taken into account:

$$\begin{cases} w_c = 1 & \text{if both quantity and calories have to be standardized} \\ w_c = 0 & \text{if only calories have to be standardized} \end{cases} \quad (9)$$

The following Table reports the commodity tree of the *China/Wheat/2014* example with the extra-column reporting the *weight* of each child-commodity.

Table 5: Commodity Tree with weights - China/Wheat/2014 example

Country	Year	ParentName	ChildName	eR	share	weight
China, Mainland	2014	Wheat	Wheat and meslin flo	0.78	1.00	1
China, Mainland	2014	Wheat	Bran of Wheat	0.22	1.00	0
China, Mainland	2014	Wheat and meslin flo	Uncooked pasta, not	1.00	1.00	1
China, Mainland	2014	Wheat and meslin flo	bread	1.00	1.00	1
China, Mainland	2014	Wheat and meslin flo	pastry	1.00	1.00	1
China, Mainland	2014	Wheat and meslin flo	Starch of Wheat	0.75	1.00	1
China, Mainland	2014	Wheat and meslin flo	Wheat Gluten	0.08	1.00	0
China, Mainland	2014	Wheat and meslin flo	Communion wafers, em	1.00	1.00	1
China, Mainland	2014	Starch of Wheat	Other Fructose and S	1.00	1.00	1
China, Mainland	2014	Wheat Gluten	Gluten Feed and Meal	1.00	0.33	0
China, Mainland	2014	Bran of Maize	Gluten Feed and Meal	1.00	0.67	0

Shares lower than 1 mean that the child commodity might be produced also from other parents and, therefore, the reported ratio represents the ratio of that commodity that is produced from the reported parent. For example, *Gluten Feed and Meal* is produced from *Wheat Gluten* and from *Bran of maize*. The sum of the shares associated to this child (0.33 and 0.67) is equal to 1.

Table 6 shows the SUA table after the two steps just described. Notice that, in the specific example, there was no need for the creation/increase of Production on any derived commodity because *Imb1* is positive for all commodities.

From Table 5 and Table 6 here is, as example, the computation of *Food processing* of Wheat, Flour of wheat and Starch of wheat:

$$FP_{wheatChina2014} = (70,500,000/0.78) \times 1 \times 1 + (21,414,729/0.8) \times 1 \times 0 = 90,384,615 \quad (10)$$

$$FP_{flourChina2014} = (1,415,692/1) \times 1 \times 1 + (15,486/1) \times 1 \times 1 + (193,950/1) \times 1 \times 1 + (239,816/0.75) \times 1 \times 1 + (116,496/1) \times 1 \times 0 + (13,263/1) \times 1 \times 1 = 1,958,146 \quad (11)$$

$$FP_{starchChina2014} = (158,665/1) \times 1 \times 1 + (21,414,729/0.8) \times 1 \times 0 = 158,665 \quad (12)$$

Table 6: Sua table with Food Processing filled - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU
Wheat	126,208,400	2,971,249	957	1,120,565		**90,384,615**	29,181,617	4,277,567		2,985,279	2,713,000	-
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	**1,958,146**			-17,345			-
Mixes and doughs for		6,497	38,072		0		0		0			-
Other Fructose and S	158,665	3,659	162,324		0				0			-
Starch of Wheat	239,816	11,035	40,311			**158,664**	172,196			7,919		-
Wheat Gluten	116,496	877	117,373			**0**	0					-
Communion wafers	13,263	8,796	5,822		16,241				-4			-
Uncooked pasta	1,415,692	12,520	22,550		1,405,661				-362			-
Food Preparations of		69,686	21,977		47,709				-12			-
Bran of Wheat	21,414,279	156,359	2,200		16,500,000		4,827,244		-4,252			-
Gluten Feed and Meal	793,740	160,231	529,333									-
bread	15,485	2,897	4,210		14,175				-3			-
pastry	193,950	89,593	117,630		165,914				-43			-

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

^b Starred figures are those that have changed from the previous table

Step 3: Filling of other missing Utilizations

After *Food processing* figures have been computed, the other variables have to be checked and either filled, increased or decreased. Only derived commodities enter this step. Primary commodities are excluded because, for these commodities, the figures are normally more reliable and complete and, very often protected. Only the final balancing may affect Primary commodity figures. This step is based on the value of the following Imbalance:

$$Imb2_{ijt} = P_{ijt} + I_{ijt} - X_{ijt} - \Delta St_{ijt} - FP_{ijt} - Fo_{ijt} - Fe_{ijt} - Lo_{ijt} - Se_{ijt} - IU_{ijt} - T_{ijt} \quad (13)$$

Table 7 reports $Imb2$ for all the lines of the SUAs of our example.

Three different scenarios can be found (Figure 5):

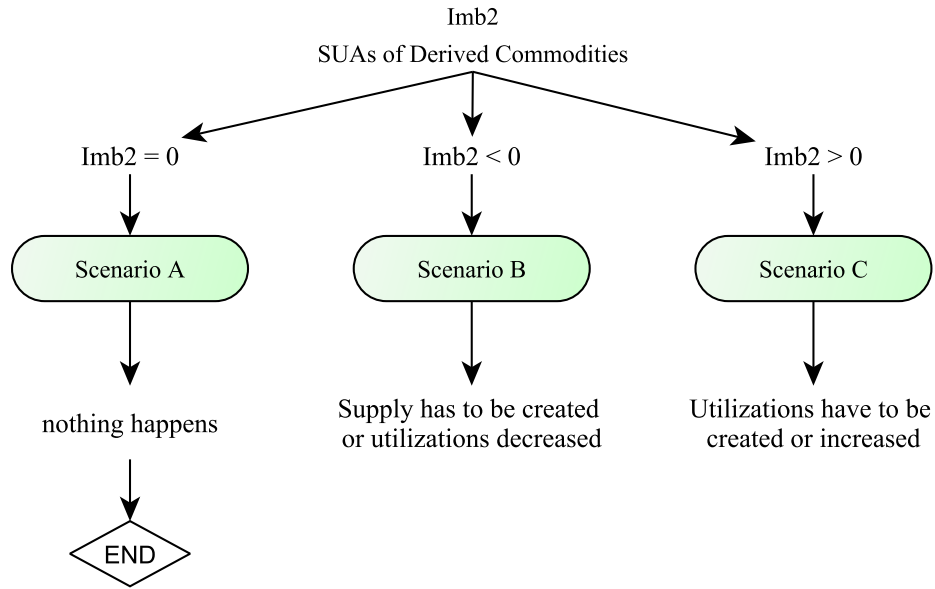


Figure 6: Scenarios of the *Sua Filling*

Table 7: Sua table with Imb2 - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb2
Wheat	126,208,400	2,971,249	957	1,120,565		90,384,615	29,181,617	4,277,567		2,985,279	2,713,000	-	<i>-1,483,951</i>
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	1,958,146			-17,345			-	<i>1,103,580</i>
Mixes and doughs for		6,497	38,072		0		0		0			-	<i>-31,575</i>
Other Fructose and S	158,665	3,659	162,324		0				0			-	<i>0</i>
Starch of Wheat	239,816	11,035	40,311			158,664	172,196			7,919		-	<i>-128,239</i>
Wheat Gluten	116,496	877	117,373			0	0					-	<i>0</i>
Communion wafers	13,263	8,796	5,822		16,241				-4			-	<i>0</i>
Uncooked pasta	1,415,692	12,520	22,550		1,405,661				-362			-	<i>363</i>
Food Preparations of		69,686	21,977		47,709				-12			-	<i>12</i>
Bran of Wheat	21,414,279	156,359	2,200		16,500,000		4,827,244		-4,252			-	<i>245,447</i>
Gluten Feed and Meal	793,740	160,231	529,333									-	<i>424,638</i>
bread	15,485	2,897	4,210		14,175				-3			-	<i>0</i>
pastry	193,950	89,593	117,630		165,914				-43			-	<i>42</i>

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

Please notice that, besides export and stock Variations, that are excluded from this step for the reasons previously explained, also Food Processing is excluded from the next steps, because *FP* is related to the production figures of the children and to modify it, would detach this very important parent-children link.

Imbalance = 0 (Scenario A)

In this scenario, the Imbalance is null, this meaning that there is enough supply for the existing utilization. In this case, the SUA line is balanced and nothing happens for the commodity. In Table 7 this is the case of *Other Fructose and syrup* and *wheat Gluten*.

Imbalance < 0 (Scenario B)

If this happens, it means that there is not enough supply for the commodity to be used in the way the SUA figures suggest. In this case, as shown in Figure 7 the algorithm tries to solve the excess of utilization by reducing all the existing Utilizations by 30% (except Exports¹⁴). If this is enough to cover the Utilization, the balancing line for that commodity is balanced and the process can go to the next step. Instead, after this reduction, there might be 2 different scenarios:

1. If this imbalance is still negative, it is checked if the production figure is “*protected*” or not:
 - If Production figure is not protected, it is interpreted as if there was not enough information for producing a correct production number, therefore, this figure is changed, i.e. created, if missing, or incremented of an amount equal to *Imb2*
 - If Production figure is protected, a warning is given that tells that there is an unsolvable problem and that all figures have to be checked because either the production protected figure is wrong, or the Trade or some utilization have to be changed.
2. If the imbalance becomes positive, the line falls into the Scenario C that will be explained shortly.

Starch of Wheat belongs to this very last case. Indeed, it has a negative Imbalance, therefore, its Utilizations (except those Utilizations that are completely excluded from this process) are reduced by 30% of their values (see Table 8). After that happens, the *Imb2* is recalculated and it is still negative, equal to 74,204 (this value is not reported in the table). In the following step, as the production is not protected, the new value of production is equal to:

$$P_{StarchChina2014} = 239,816 + 74,204 = 314,020 \quad (14)$$

Remember that the commodity *Mixes and dough for the preparation of baker's wares* is considered here as a primary commodity, therefore it will not be balanced even if it has a negative imbalance.

¹⁴Stock is changed in this case, because this is just a proportional reduction of values and does not need a time series analysis

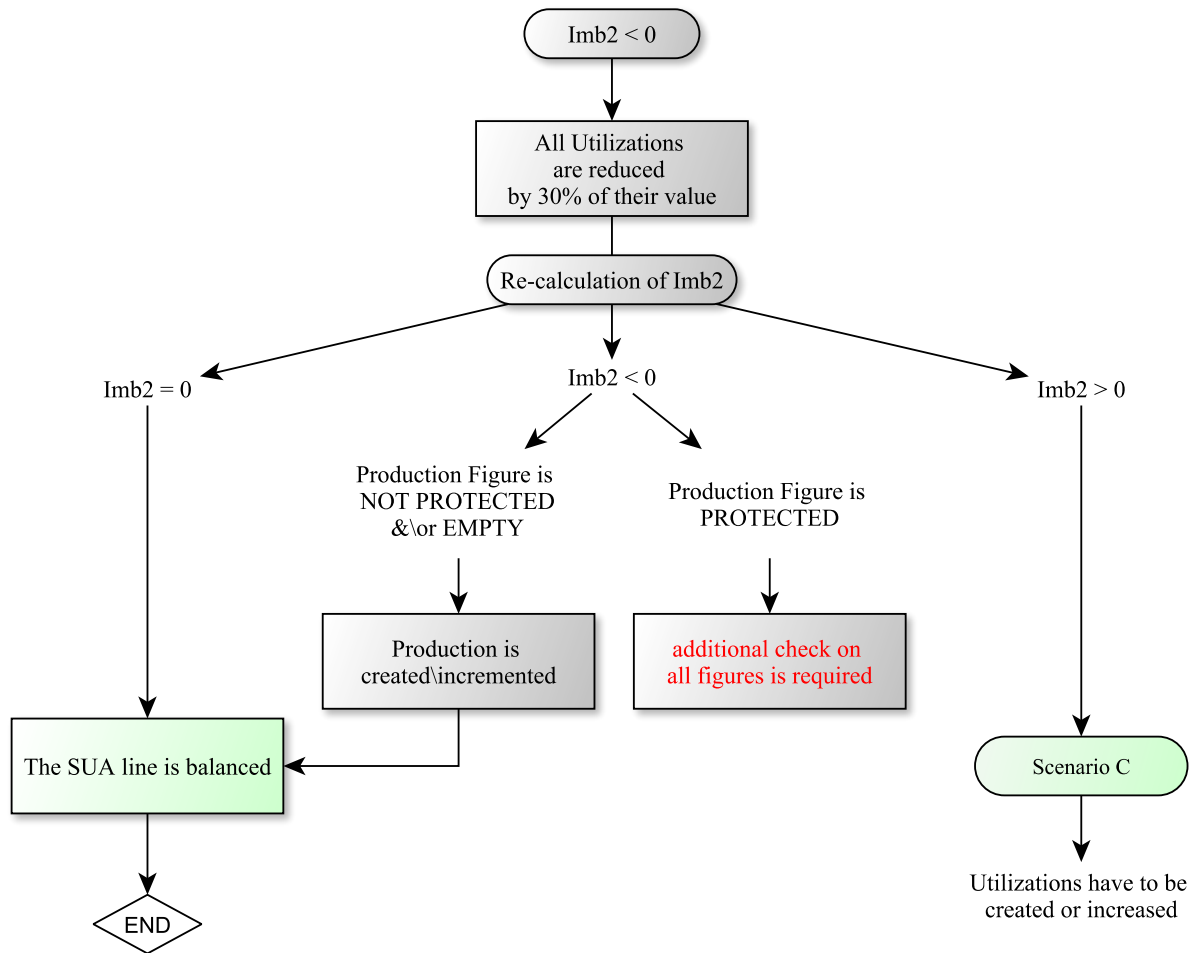


Figure 7: Workflow of Sua Filling when there is a negative imbalance

Table 8: Sua table with Utilizations filled - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb2
Wheat	126,208,400	2,971,249	957	1,120,565		90,384,615	29,181,617	4,277,567		2,985,279	2,713,000	-	-1,483,951
Wheat and meslin flo	70,500,000	33,055	188,674		**68,372,646**	1,958,146			**_17,621**			-	0
Mixes and doughs for		6,497	38,072		0		0		0			-	-31,575
Other Fructose and S	158,665	3,659	162,324		0				0			-	0
Starch of Wheat	**314,020**	11,035	40,311			158,664	**120,537**			**5,543**		-	0
Wheat Gluten	116,496	877	117,373			0	0					-	0
Communion wafers	13,263	8,796	5,822		16,241				-4			-	0
Uncooked pasta	1,415,692	12,520	22,550		**1,406,024**				**_363**			-	0
Food Preparations of		69,686	21,977		47,709				-12			-	12
Bran of Wheat	21,414,279	156,359	2,200		**16,689,929**		**4,882,810**		**_4,301**			-	0
Gluten Feed and Meal	793,740	160,230	529,333				**424,637**					-	0
bread	15,485	2,897	4,210		14,175				-3			-	0
pastry	193,950	89,593	117,630		**165,957**				**_42,75**			-	0

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

^b Starred figures are those that have changed from the previous table

Imbalance > 0 (Scenario C)

When $Imb2 > 0$ there is an excess of supply that has to be distributed through Utilizations. Here:

- A *Utilization Table* tells which figures are supposed to be filled.
- The Imbalance is distributed among variables according to a, so called, *Multiple Filler approach* which make use of the *Inverse Ranking Rule*.

→ The Utilization Table

Utilization Table is a Country/commodity table telling which Utilizations historically have been active for each commodity and assigning a Rank and an inverse rank to each utilization based on its mean value over the period 2000-2013. Table 9 shows the *Utilization Table* for *Wheat and meslin flour* in China Mainland.

Table 9: *Utilization Table* Wheat and meslin Flour - China/Wheat/2014 example

AreaM49	Element	ItemCpc	rank	rankInv
1248	FP	23110	2	3
1248	Fo	23110	1	4
1248	DSt	23110	4	1
1248	X	23110	3	2

Where 1248 is the m49 code for China Mainland and 23110 is the CPC code of Wheat and Meslin Flour. The table tells the following pieces of information:

1. The reported commodity in China, as been historically used:
 - for producing other commodities (*Food Processing*),
 - for human consumption (*Food*),
 - for increasing stocks (*Stock Variation*),
 - for exports (*exports*).
2. The biggest amount of availability of the reported commodity was used for Human consumption, the second big for processing other commodities, then for exports and, in the end, for increasing stocks.

These pieces of information are combined with the amount of $Imb2$ for filling the SUAs of each derived commodity according to a set of rules all falling into the, so called, *Multiple Filler approach*.

→ The Multiple Filler Approach and the Inverse Ranking Rule

This approach distributes the $Imb2$ across Utilizations, by the rules reported in Figure 8.

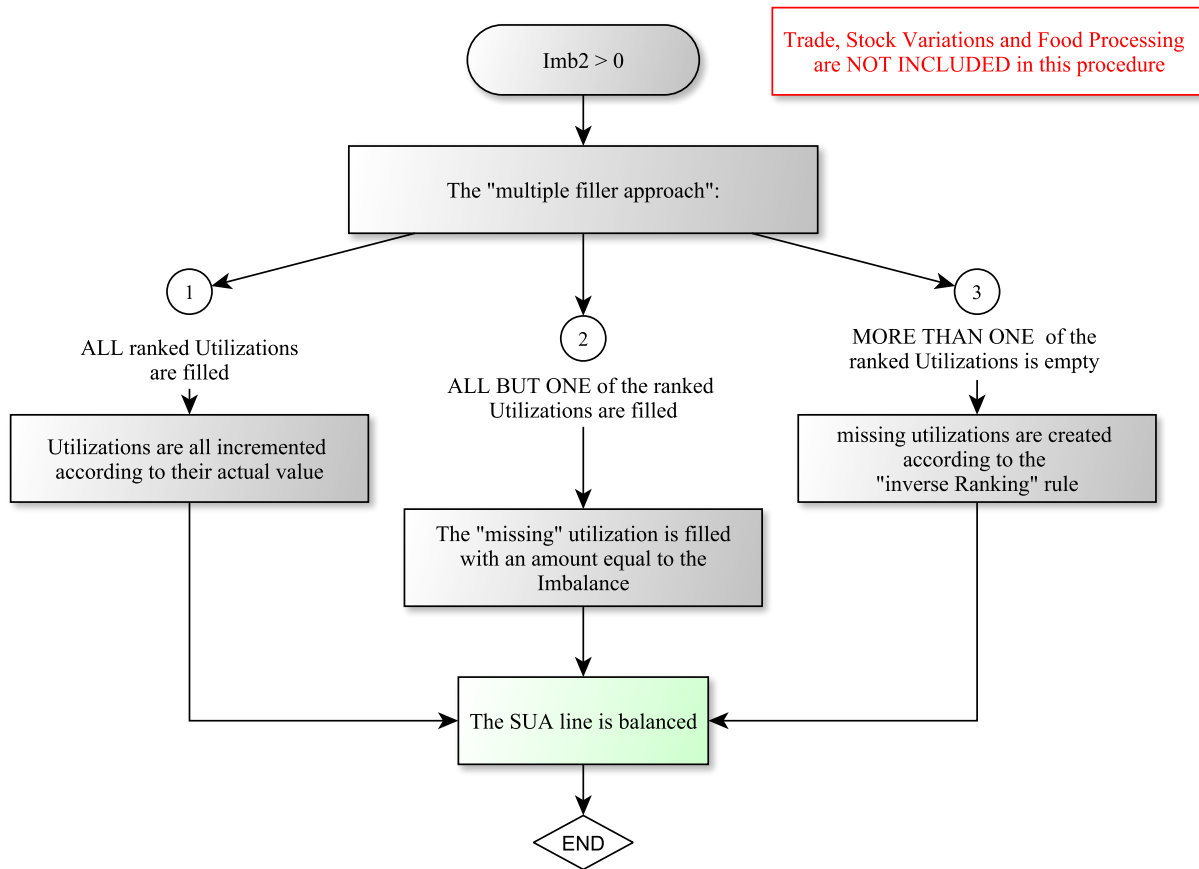


Figure 8: Workflow of Sua Filling when there is a positive imbalance

As reported in the Figure, Three cases may happen:

1. If all the ranked Utilizations are already filled, all of them are increased proportionally to their values. *Bran of Wheat* in our example, belongs to this case. *Utilization Table* for *Bran of Wheat* is reported in Table 10

Table 10: *Utilization Table* Bran of Wheat - China/Wheat/2014 example

AreaM49	Element	ItemCpc	rank	rankInv
1248	Fe	39120.01	2	2
1248	Fo	39120.01	1	3
1248	X	39120.01	3	1

According to this table, there are 3 Utilizations that are suppose to be active for this commodity: *feed*, *food* and *exports*. *Exports* is excluded from this step, therefore, food and feed remain. Both Variables are filled in the SUA line of This Commodity, therefore, all of them are increased proportionally to their values.

2. If all the ranked Utilizations are already filled but one, the all imbalance goes to that commodity. In our Example, this is the case of *Gluten Feed and meals* (Table 10).
3. Finally, if more that one utilization is empty, the so called ***Inverse Ranking rule*** is activated. This rule assigns part of the imbalance to the ranked Utilizations proportionally to the ranks. This happens

by making use of the inverse rank of all the Utilizations that are “activable” and not excluded from the procedure. The rule works for any number of ranks and independently from the number of Utilizations that have to be excluded.

The *Inverse ranking rule* starts from the *Utilization Table* with the addition of a column containing the *rank weighth* Wr , as reported in the following table:

element (Ut_h)	rank (r_h)	Inverse rank (Ir_h)	weight Rank (Wr_h)
Ut_1	r_1	Ir_1	Wr_1
...
Ut_h	r_h	Ir_h	Wr_h
...
Ut_n	r_n	Ir_n	Wr_n

where:

- Ut_h is the generic utilization, with $h = 1, 2, \dots, n$
- r_h is the rank of utilization Ut_h , with $h = 1, 2, \dots, n$
- Ir_h is the Inverse rank of utilization Ut_h , with $h = 1, 2, \dots, n$ and is equal to:

$$Ir_h = \max(r) - r_h + 1 \quad (15)$$

- Wr_h is the rank weight of utilization Ut_h and is equal to 0 or 1 depending on the element being, respectively, excluded or not from the *Sua Filling* (Table 11).

Table 11: Rank Weight of the elements of FBS

Utilizations	Wr
X	0
DSt	0
Fo	1
FP	0
Fe	1
Se	1
T	1
IU	1
L	1
ROU	0

The value V_h assigned to the Utilizations of each commodity is calculated as a rank based proportion pR_h of the total $Imb2$ according to the following equation:

$$V_h = Imb2 \times pR_h \quad (16)$$

with pR_h depending on $\sum_{h=1}^n Wr_h$:

$$\begin{cases} pR_h = 0 & \text{when } \sum_{h=1}^n Wr_h = 0 \\ pR_h = \frac{Ir_h \times we_h}{\sum_{i=1}^n (Ir_h \times Wr_h)} & \text{when } \sum_{h=1}^n Wr_h > 0 \end{cases} \quad (17)$$

As an example consider the SUA line of *Cocoa Beans* in Madagascar in 2014 (Table 12) and the *Utilization Table* of the same country/commodity combination (Table 13)

Table 12: Sua of Cocoa Beans Madagascar 2014 - before *Sua Filling*

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb2
<i>Cocoa beans</i>	10,865	36	8,326	-	-	-	-	-	-	-	743	-	1,832

Table 13: *Utilization Table* for Cocoa Beans Madagascar

Ut	r	Ir	Wr
FP	4	2	0
Fo	2	4	1
IU	3	3	1
DSt	5	1	0
X	1	5	0

Besides the excluded elements, there are 2 Utilizations empty: *food* and *Industrial Use*. The Values V_{Fo} and V_{IU} of these two elements are calculated as follows and the result shown in Table 14:

$$pR_{Fo} = \frac{4 \times 1}{0 + 0 + 4 + 3 + 0 + 0} = 4/7 = 0.57143 \quad (18)$$

$$V_{Fo} = 1,832 \times 0.57143 = 1,047$$

$$pR_{IU} = \frac{3 \times 1}{0 + 0 + 4 + 3 + 0 + 0} = 3/7 = 0.42857 \quad (19)$$

$$V_{IU} = 1,832 \times 0.42857 = 785$$

Table 14: Sua line of Cocoa Beans Madagascar 2014 - after *Sua Filling*

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb2
<i>Cocoa beans</i>	10,865	36	8326	-	1,047	-	-	-	-	785	743	-	0

The “Sua Balanced” table

The table resulting after the *Sua Filling* is called *Sua Balanced*. For the *China/Wheat/2014* example the *Sua Balanced* is reported in Table 15.

The red lines are not balanced lines. Indeed, the *Sua Filling* procedure does not intervene on all the commodities. It leaves unbalanced all the “top level” commodities falling in 2 main categories:

1. The *Primary commodities*. These are all the crop commodities, like cereals, pulses, vegetables, fruits etc, that are often processed in other commodities, like the wheat in flour of wheat.
2. The *no tree* commodities. These commodities are excluded from the list of commodities that are processed into something else, they are often mixes of some other crop or commodity and used for specific purposes, i.e. *Mixes and dough for the preparation of baker’s wares* and *Food preparations of Flour, Meal or Malt Extract* from our example.

Therefore 3 lines are not balanced in Table 15, belonging to these 2 categories of “not balanced” commodities.

Table 15: Sua Balanced - China/Wheat/2014 example

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	Imb2
<i>Wheat</i>	126,208,400	2,971,249	957	1,120,565		90,384,615	29,181,617	4,277,567		2,985,279	2,713,000	-	-1,483,951
Wheat and meslin flo	70,500,000	33,055	188,674		68,372,646	1,958,146			-17,621			-	0
<i>Mixes and doughs for</i>		6,497	38,072		0		0		0			-	-31,575
Other Fructose and S	158,665	3,659	162,324		0				0			-	0
Starch of Wheat	314,020	11,035	40,311			158,664	120,537			5,543		-	0
Wheat Gluten	116,496	877	117,373			0	0					-	0
Communion wafers	13,263	8,796	5,822		16,241				-4			-	0
Uncooked pasta	1,415,692	12,520	22,550		1,406,024				-363			-	0
<i>Food Preparations of</i>		69,686	21,977		47,709				-12			-	12
Bran of Wheat	21,414,279	156,359	2,200		16,689,929		4,882,810		-4,301			-	0
Gluten Feed and Meal	793,740	160,230	529,333				424,637					-	0
bread	15,485	2,897	4,210		14,175				-3			-	0
pastry	193,950	89,593	117,630		165,957				-42.75			-	0

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

The reason for the exclusion of the *zero-level* commodities from the filling stays in the circumstance for which the standardization will affect the figures of the primary commodities, this opening the need for a final *balancing* procedure. In order not to intervene twice on these commodities, it has been decided to keep as smaller as possible the number correction, so to reduce error.

Nutritive Values and DES calculation

After the *sua Balanced* table has been produced, **Nutritive values** are calculated (*Calories*, *Proteins* and *Fats*), based on the Food availability produced during the *Sua Filling*. These Variables are calculated for each Food Product of the SUA, based on official Nutritive Factors stored in the SWS. Values of Nutritive elements are stored in the system as:

- *g/100g*, for Proteins and Fats
- *Kcal/100g* for Calories.

As a consequence, the amount of Nutritive Elements for the number of Tonnes of Food is given by

$$\text{Nutritive value} = \text{Nutritive factor} \times \text{Food availability (tonnes)} \times 10,000 \quad (20)$$

In particular, after all Nutritive Values are calculated, only DES is reported for FBS's purposes, DES being given by:

$$DES_{ijt} = \frac{\frac{Kcal_{ijt}}{Population_{jt}}}{365} \quad (21)$$

where the *i* index runs over all countries, the *j* index over all commodities, and *t* over years.

DES for the food commodities in the *China/Wheat/2014* example are reported in Table 16

Table 16: Sua Unbalanced DES - China/Wheat/2014 example

itemName	Fo	DES
Wheat	-	-
Wheat and meslin flo	68,372,646	478.38
Mixes and doughs for	0	0
Other Fructose and S	0	0
Starch of Wheat	-	-
Wheat Gluten	-	-
Communion wafers, em	16,239.38	0.1405
Uncooked pasta, not	1,406,024	10.1699
Food Preparations of	47,708.53	0.3545
Bran of Wheat	16,689,929	70.0635
Gluten Feed and Meal	-	-
bread	14,174.09	0.0696
pastry	165,957	1.2069

Figure 9 Represents the steps following the *Sua Filling*:

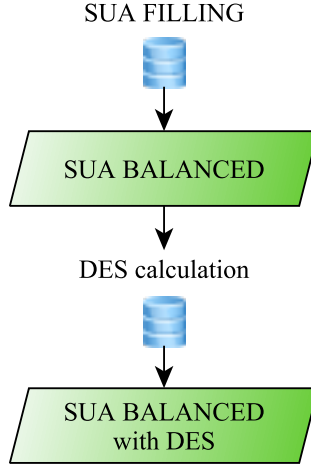


Figure 9: Sua Balanced and the calculation of DES

3 Standardization

The *Standardization* is the process of transforming all the processed commodities in terms of their primary (or *proxy-primary*) commodities and to add up all the calories of the food commodities, so to have all the existing food availability of a country, expressed in terms of the main commodities. Aggregation of data from a detailed product classification to a more comprehensive one is common practice for two reasons. The first reason is to reduce the amount of data, and the number of commodities involved, to a level and size more suited for analytical purposes. The second reason is, that by cancelling out the intermediate production of derived products against the input use of primary products, a clearer view emerges of the domestic availability of a product and its final uses.

The standardization starts from the SUA table (reported, in its generic form, in Table 17)

Table 17: Generic SUA for Standardization

Item	El_1	...	El_h	...	El_n	$Kcal$
$parent_p$	$V_{El_1, pjt}$...	$V_{El_h, pjt}$...	$V_{El_n, pjt}$	$Kcal_{pjt}$
$child_1$	$V_{El_1, jt1}$...	$V_{El_h, jt1}$...	$V_{El_n, jt1}$	$Kcal_{1jt}$
...
$child_c$	$V_{El_1, cjt}$...	$V_{El_h, cjt}$...	$V_{El_n, cjt}$	$Kcal_{cjt}$
...
$child_C$	$V_{El_1, Cjt}$...	$V_{El_h, Cjt}$...	$V_{El_n, Cjt}$	$Kcal_{pjt}$

where:

- $c = 1, 2, 3...C$ are the C children c of parent p ,
- $V_{El_h, cjt}$ is Value of element h for child c ,
- $V_{El_h, pjt}$ is Value of element h for parent p ,
- El_h is the generic Element of the generic child c of Parent p in the range of elements involved in the Standardization. Indeed, P and FP are excluded from this step of the process,
- $Kcal$ is the amount of Calories associated with the commodity. This is positive only if the commodity is a food commodity.

Standardize a SUA is the process of:

- expressing all elements $El_{k,c}$ of the children of a parent commodity, in terms of their parent commodity,
- Adding up all the calories of the food commodities and calculate DES.

Elements are transformed in their parent commodity following the equation:

$$\begin{cases} El_{h,pjt} = \sum_{c=1}^C \left(\frac{V_{El_{h,cjt}}}{eR_{p \rightarrow c}} \right) \times s_{cp}^2 \times w_c \\ DES_{pjt} = \frac{Kcal_{pjt} + \sum_{c=1}^C Kcal_{cjt}}{Population_{jt}} \times 365 \end{cases} \quad (22)$$

where:

- $c = 1, 2, 3 \dots C$ are the C children c of parent p ,
- $eR_{p \rightarrow c}$ is the *extraction Rate* from parent p to child c ,
- w_c is the *weight* of child c ,
- $Kcal_{pjt}$ are the Kcal of the parent commodity (if is a food commodity),
- $Kcal_{cjt}$ are the Kcal of the child commodities,
- s_{cp}^2 is the *share* of child c from parent p and is based on a different availability from the previous one:

$$s_{cp}^2 = \frac{availability2_{p(c)}}{\sum_{p=1}^A availability2_{p(c)}} \quad (23)$$

where $availability2_{p(c)}$ is the availability of each parent p of child c expressed in terms of c (as say in *child equivalent*). Is is enumerated as 2 because is different from the one previously defined. Indeed, here The availability of each parent is the amount of that commodity available to be processed, i.e. the Food processing, expressed in terms of child, therefore:

$$availability2_{p(c)} = FP_{pjt} \times eR_{p \rightarrow c} \quad (24)$$

Table 18 shows the values of *availability2*, *eR* and *shares* and *weights* of the *China/Wheat/2014* example.

Table 18: Availabilities and shares of parent/child for Standardization

Child	Parent	extractionRate	availability	share	weight
Wheat and meslin flo	Wheat	0.78	70,500,000	1.00	1
Bran of Wheat	Wheat	0.22	19,884,615	1.00	0
Uncooked pasta, not	Wheat	0.78	72,458,146	1.00	1
Germ of Wheat	Wheat	0.02	1,920,221	1.00	0
bread	Wheat	0.78	72,458,146	1.00	1
Bulgur	Wheat	0.95	85,865,385	1.00	1
pastry	Wheat	0.78	72,458,146	1.00	1
Starch of Wheat	Wheat	0.58	54,343,610	1.00	1
Wheat Gluten	Wheat	0.06	5,796,652	1.00	0
Communion wafers, em	Wheat	0.78	72,458,146	1.00	1
Other Fructose and S	Wheat	0.58	54,502,275	1.00	1
Gluten Feed and Meal	Wheat	0.06	5,796,652	0.76	0
Gluten Feed and Meal	Maize (corn)	0.09	1,803,440	0.24	0

As you can see, comparing this table with Table 15:

$$availability2_{wheat(flour)} = 90,384,615 \times 0,78 = 70,500,000 \quad (25)$$

$$availability2_{wheat(bran)} = 90,384,615 \times 0,22 = 19,884,615 \quad (26)$$

Moreover:

$$s_{Gluten,wheat}^2 = \frac{5,796,652}{5,796,652 + 1,803,440} = 0.76 \quad (27)$$

$$s_{Gluten,Maize}^2 = \frac{1,803,440}{5,796,652 + 1,803,440} = 0.24 \quad (28)$$

By applying equations 22 to 24 to Table 15, Table 20 is obtained, which contains the so *Standardized* Table, while the DES for the *Sua Standardized* of Table 20 is reported in Table 19.

Table 19: DES of *Sua Standardized* - China/Wheat/2014 example

itemName	DES
Wheat	560.0300
Mixes and doughs for	0.0000
Food Preparations of	0.3545

Notice that *Production* remains untouched from the Standardization process and *ROU* are still untouched.

The “Fbs Standardized” table

The stadardized table is called *Fbs Standardized* and is composed of more than 1 line, because, besides the primary commodity thare are 2 more typologies of commodities, both :

- *No-tree* commodities, that are also “*zero-level*” commodities, appear here as they were in the previous step. Indeed, as these are not procesed, the standardization does not change any of the element’s values.
- *Proxy-primary* commodities. The Child commodities defined as *proxy-primary* are cut from their trees immediately before the standardization begins. They are treated as primary commodities and appear here together with the primaries.

Table 20 and Table 23

Table 20: Sua Standardized

itemName	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU
Wheat	126,208,400	**3,184,654**	**781,808**	1,120,565	**89,711,591**	**0**	**29,387,664**	4,277,567	**_22,766.5**	**2,994,755**	2,713,000	-
Mixes and doughs for	-	6,496.91	38,071.88	0	0	0	0	0	0	0	0	-
Food Preparations of	-	69,685.96	21,977.43	0	47,708.53	0	0	0	-12.2958	0	0	-

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

^b Starred figures are those that have changed from the previous table

4 The FBS Balancing

The balancing of Standardized Sua in the new framework is based on the idea that each variable of equation 1 includes a measurement error, so that the equation can be rewritten as follows:

$$P_{ijt}^* + I_{ijt}^* - X_{ijt}^* - \Delta St_{ijt}^* = FP_{ijt}^* + Fo_{ijt}^* + Fe_{ijt}^* + Lo_{ijt}^* + Se_{ijt}^* + IU_{ijt}^* + T_{ijt}^* + ROU_{ijt}^* \quad (29)$$

where (dropping the indices for brevity):

- $P^* = P \pm adj_P$
- $I^* = I \pm adj_I$
- $X^* = X \pm adj_X$
- $\Delta St^* = \Delta St \pm adj_{\Delta St}$
- $FP^* = FP \pm adj_{FP}$
- $Fo^* = Fo \pm adj_{Fo}$
- $Fe^* = Fe \pm adj_{Fe}$
- $Lo^* = Lo \pm adj_{Lo}$
- $Se^* = Se \pm adj_{Se}$
- $IU^* = IU \pm adj_{IU}$
- $T^* = T \pm adj_T$

and where each *adjustment* ($adj_{el \in E}$) is an balancing adjustment proportional to the measurement error (*ME*) of each element $el \in E = \{P, I, X, \Delta St, FP, Fo, Fe, Lo, Se, IU, T, ROU\}$

The balancing mechanism is a two step process aimed at modify each value proportionally to the *ME*. It start with the definition of the percentages of *Measurement Errors* (p_{ME}) for each element as reported in the following Table:

Table 21: Measurement Errors percentages

Element	p_{ME}
P	0.1%
I	1%
X	1%
DSt	25%
Fo	5%
FP	25%
Fe	25%
Se	25%
T	25%
IU	25%
L	25%
ROU	0

ROU is not included in this table because *ROU* is created in this step in a slightly different way.

Indeed, the *final balancing* happens according to the following steps:

1. If $Imb2 > 0$:

- **1.a** If $Imb2 < TS \times 0.05$:

Residual and Other Uses is set equal to the Imbalance:

$$ROU = Imb2$$

and the equation is balanced.

- **1.b** If $Imb2 > TS \times 0.05$:
 - $ROU^* = TS \times 0.05$
 - $Imb2 = Imb2 - ROU^*$
 - The $adj_{el \in E}$ s are given by:

$$adj_{el \in E} = Imb2 \times Value_{el} \times p_{adj_{el}} \quad (30)$$

with

$$p_{adj_{el}} = \frac{p_{ME}}{\sum_{el \in E} (Value_{el} \times p_{ME})} \quad (31)$$

and

$$\sum_{el \in E} p_{adj_{el}} = 1 \quad (32)$$

Equation 32 being assured by the two steps approach.

- 2.** If Imbalance is Negative: $Imb2 < 0$

ROU is not activated and the other elements are balanced accordingly to equations 30 to 32

Nutritive Values updates

After the balancing happens, the value of Food availability might have changed, therefore, the Nutritive values have to be adjusted accordingly. The *Balanced Nutritive values* are obtained through a proportion. For the *DES*, this proportion is given by:

$$Fo^* : Fo = DES^* : DES \rightarrow DES^* = DES \times \frac{Fo^*}{Fo}$$

Table 22 reports all the steps of the final balancing in the China/Wheat/2014 example for *Wheat*, including the final adjustment of the *DES*.

Table 22: Final Balancing - China/Wheat/2014 - Wheat

<i>el</i>	$Value_{el \in E}$	p_{ME}	$Value_{el} \times p_{ME}$	$p_{adj_{el}}$	$adj_{el \in E}$	$Value_{el \in E}^*$	<i>DES</i>	<i>DES</i> *
<i>P</i>	126,208,400	0.001	126,208	0.01	13,426	126,221,826	-	-
<i>I</i>	3,184,654	0.01	31,847	0	3,388	3,188,042	-	-
<i>X</i>	781,808	0.01	7,818	0	-832	780,976	-	-
ΔSt	1,120,565	0.25	280,141	0.02	-29,801	1,090,764	-	-
<i>Fo</i>	89,711,591	0.05	4,485,580	0.3	-477,172	89,234,419	560.03	557.05
<i>FP</i>	0	0.01	0	0	0	0	-	-
<i>Fe</i>	29,387,664	0.25	7,346,916	0.5	-781,559	28,606,105	-	-
<i>Se</i>	4,277,567	0.25	1,069,392	0.07	-113,761	4,163,806	-	-
<i>T</i>	-22,766.50	0.25	-5,692	0	605	-22,161	-	-
<i>IU</i>	2,994,755	0.25	748,689	0.05	-79,645	2,915,110	-	-
<i>L</i>	2,713,000	0.25	678,250	0.05	-72,152	2,640,848	-	-
<i>ROU</i>	-	-	-	-	-	-	-	-
<i>Imb2</i>	-1,571,130	-	-	-	-	0	-	-
$\sum_{el \in E}$			14,769,149	-	-	-	-	-

The balancing of the commodity *Food Preparations of* is reported in Table 23. In this case, the imbalance is positive and absorbed by the element *ROU*.

Table 23: Final Balancing - China/Wheat/2014 - Food Preparations

	<i>el</i>	$Value_{el \in E}$	$Value_{el \in E}^*$	<i>DES</i>	<i>DES*</i>
	<i>P</i>			-	-
	<i>I</i>	69,685.96	69,685.96	-	-
	<i>X</i>	21,977.43	21,977.43	-	-
	ΔSt	0	0	-	-
	<i>Fo</i>	47,708.53	47,708.53	0.3545	0.3545
	<i>FP</i>	0	0	-	-
	<i>Fe</i>	0	0	-	-
	<i>Se</i>	0	0	-	-
	<i>T</i>	-12.2958	-12.2958	-	-
	<i>IU</i>	0	0	-	-
	<i>L</i>	0	0	-	-
	<i>ROU</i>	-	12	-	-
	<i>Imb2</i>	12	0	-	-
	<i>TS</i>	47,709	-	-	-
	$TS \times 0.05$	2,385	-	-	-
	$Imb2 < TS \times 0.05$	TRUE	-	-	-

5 Final aggregation:FBS items and FBS aggregates

The Food Balance Sheets have to be expressed in *FBS* items. The conversion from CPC to FBS is a simple aggregation based on an aggregation table, called ***Fbs Tree***, an extract of which is reported in Table 24.

After the Balancing happens, the CPC commodities are aggregated on the basis of the classifications of FBS commodities reported in the table. 3 different level of aggregations are performed:

1. at FBS item L Elevel,
2. At FBS group level
3. at FBS family level
4. at the Total (By country) FBS Level.

Table 25 reports these aggregations for the China/Wheat/2014 example, including the DES.

Table 24: Fbs Tree [extraction of the main table]

FBS				SUA
TOTAL	FAMILY	GROUP	ITEM	P-P & PROC*
2901 GRAND TOTAL	2903 VEGETALE PRODUCTS	2905 CEREALS (EXCLUDING BEER)	2511 WHEAT & PRODUCTS	WHEAT FLOUR WHEAT BRAN WHEAT MACARONI GERM WHEAT BREAD BULGUR PASTRY WHEAT,STARCH WHEAT GLUTEN BREAKF CERLS WAFERS MIXES AND DO FOOD PREP.FL
				MAIZE GERM MAIZE FLOUR MAIZE BRAN MAIZE MAIZE GLUTEN STARCH MAIZE GLUT FEED& ME ...
	
		2909 SWEETENERS	2543 SWEETENERS, OTHER & PROD	FRUCTOSE CHE MALTOSE CHEM MAPLE SUGAR SUGAR CROPS MOLASSES (calories OTHER FRUCTO SUGAR NES GLUC. DEXTR. LACTOSE ISOGLUCOSE BEV NON-ALC ...
				...
		2914 VEGETABLE OILS	2571 SOYABEAN, OIL 2572 GROUNDNUT, OIL 2	OIL SOYABEAN OIL GROUNDNT ...

	2941 ANIMAL PRODUCTS	2943 MEAT (SLAUGHTERED)	2731 MEAT & PRODUCTS BOVINE	BEEF VEAL BEEF BONLESS BEEF DSS MEAT EXTRACT SAUSAGE BEEF BEEF PREP BEEF CANNED HOMOGENIZED BUFFALO MEAT ...
				...

* P-P & PROC stands for "Proxy-primaries and Processed commodities"

Table 25: FBS Final Aggregations - China/Wheat/2014

commodity	P	I	X	DSt	Fo	FP	Fe	Se	T	IU	L	ROU	DES
WHEAT & PROD.	126,221,780	3,272,192	813,228	1,090,866	89,283,771	0	28,608,797	4,164,198	-23,382	2,915,384	2,641,097	12	557
CEREALS & PROD.	557,425,224	20,835,349	3,248,501	10,922,600	266,674,214	9,160,912	199,216,583	13,786,396	-69,280	42,941,839	20,934,770	11,444,037	1,442
VEGETABLE PROD.	1,778,914,939	154,782,065	26,232,702	9,948,701	1,107,614,579	227,982,944	340,218,284	17,403,713	-315,270	78,697,328	97,503,864	28,410,159	2,465
GRAND TOTAL	1,940,733,489	177,486,138	27,427,634	9,948,701	1,272,788,655	229,422,480	347,720,409	18,125,532	-370,776	81,577,706	99,470,600	32,108,686	3,128

^a P=Production, I=Import, X=Export, DSt=Delta Stock, Fo=Food Availability, FP=FoodProcessing, Fe=Feed, Se=Seed, T=Tourism Consumption, IU=IndustrialUse, L=Loss, ROU=Residual and other uses

Conclusions and future work

The methodology presented in this document is the result of the teamwork of the ESS *methodological team* and *data team*. Still some aspects of the methodology have to be fully tested and might require adjustments and/or debug. Nevertheless, this methodology has been used for producing FBS of year 2016.

Related documentation

Here a list of the documents related to the present one:

- 1.

DRAFT