Standardization & Balancing for Food Balance Sheet Calculation

Cristina Muschitiello
Food and Agriculture Organization of the United Nations

9 August 2018

Abstract

This vignette provides a description 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 the overall Food Balance Sheet workflow and some, more detailed, technical aspect.

Contents

List of Tables

List of Figures

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 August 9, 2018 and is subject to changes and updates.

The Food Balance Sheet Framework

Definitions

A food balance sheet can be defined as an aggregated and analytic 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***) System[2]. 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:

[1] For 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.

[2] The 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*

- P=Production
- I=Imports
- 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[3]:

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 are[4]:

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*.

[3]For more details on FBS variables please see the latest version of the *Resource Book*, the documentation on *Food Balance Sheet workflow in the Statistical Working System* to be found on *GitHub* and *Bitbucket* and specific documentation available about imputation of all variables

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

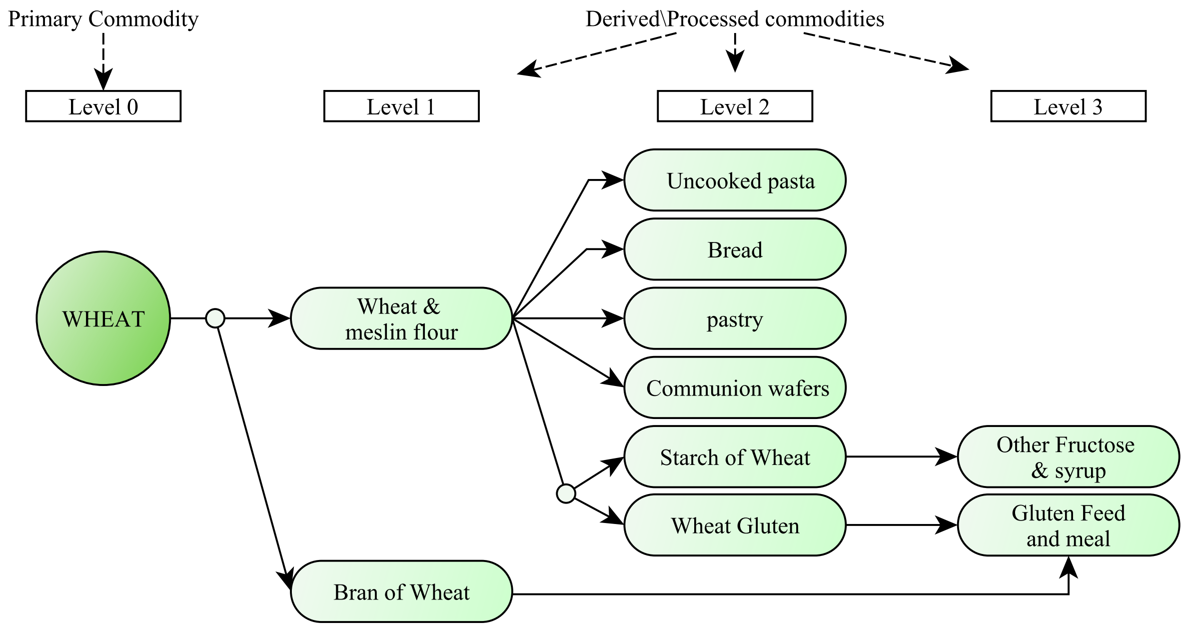

Figure 1: Commodity Tree for Wheat in China mainland 2014

2. Not all the countries have the same production processes: countries have different technologies and primary products availability. 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 that is produced from the specified parent and are expressed as a ratio. Shares are generically defined as:

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

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.

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.

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

- ***No-Tree*** commodities. These are *zero-level* commodities. They are either primary commodities that are never processed (as lettuce) or commodities that are not processed in a specific country/commodity combination. 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. This happens if no production processes have been activated for that specific Country or year.

Standardization and Balancing

The *Standardiation & 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*.

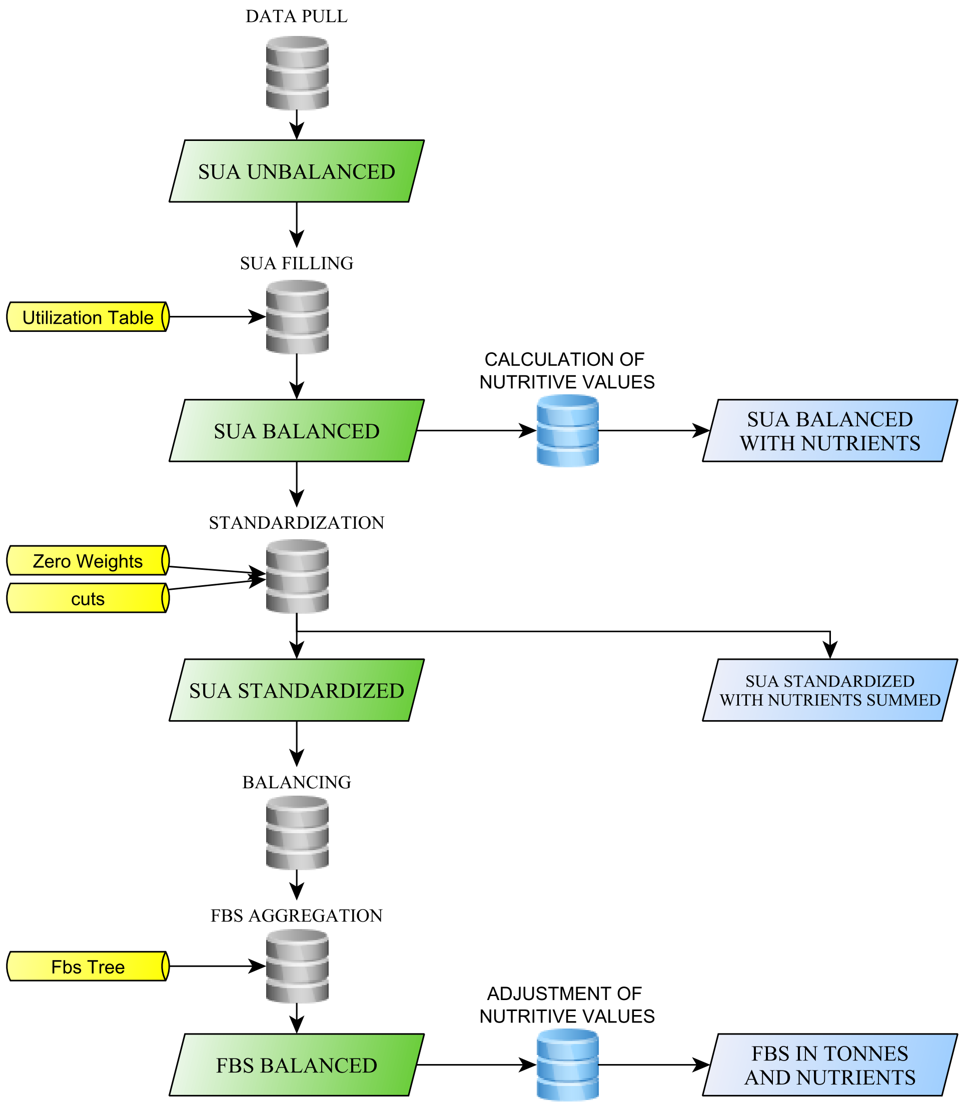

Figure 2: Standardization and Balancing Overall Workflow

1 Data Pull

The process of creating FBSs starts by considering the initial commodity balance for each CPC commodity, either primary and derived. In the balance the different variables of the equation (as listed and briefly described in the previous section) are 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)***[5].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 is given in a separate document.[6].

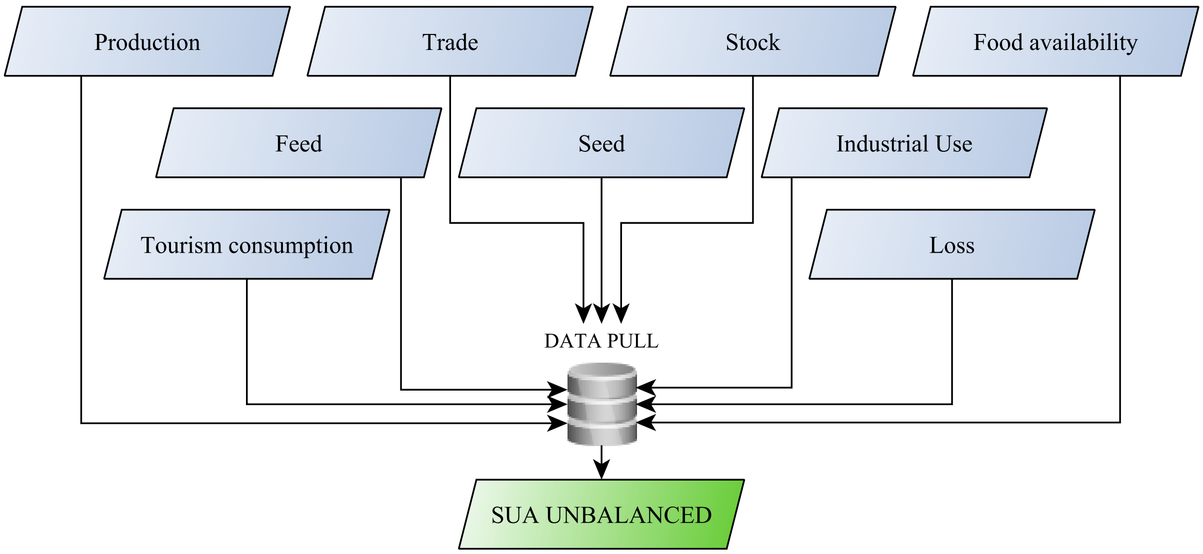

Figure 3: Data Pull from datasets containing data for each separate variable

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 [7]. Third, modules might, sometimes, fail in the imputation, because of the strong complexity and structural diversity of the input data.

[5]SWS 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.

[6][report the document when it will be available]

[7]this opening the strong possibility not to have values where they are supposed to exist and, also, not guaranteeing consistency of data over time

Table 2: Unbalanced Sua table - 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*		-	*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	-			-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.

A *SUA* also contains other commodities, when existing, that are associated to the primary commodity inside an FBS item. In our example, the *SUA* table includes the commodity *Mixes and dough for the preparation of baker's wares* because it is included in the FBS commodity *WHEAT & PRODUCTS*. That commodity is in the *SUA* of Wheat even if is not processed from wheat directly and is not processed in any commodity. It is a No-tree commodity.

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" here means "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.

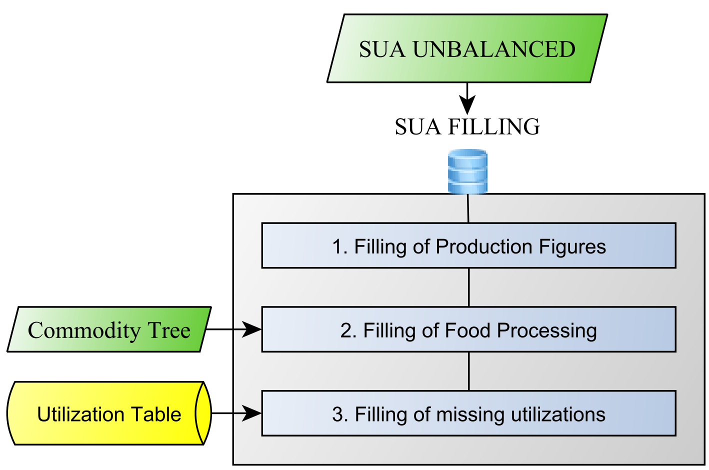

Figure 4: The Sua Filling simplified workflow

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 first, 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. 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 very poor affected 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. Consequently these figures are never created from scratch, but might be sometimes reduced by a small percentage.

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.
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*. These table helps identifying which variables have to be filled, commodity by commodity. Moreover, 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, based on a weight proportional to the size each variable had over the period 2000-2013, this size being represented by the median. As a consequence, the majority of 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, non-tree commodities and derived commodities:

- Primary commodities are the commodities for which more reliable figures are produced. FAO often publishes information about primary commodities and a change of the figures has to be done with care.
- Derived commodities are more affected by modifications during this process, according to the flag of each figure.
- Non-Tree commodities (as previously defined) are treated as derived commodities.

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.

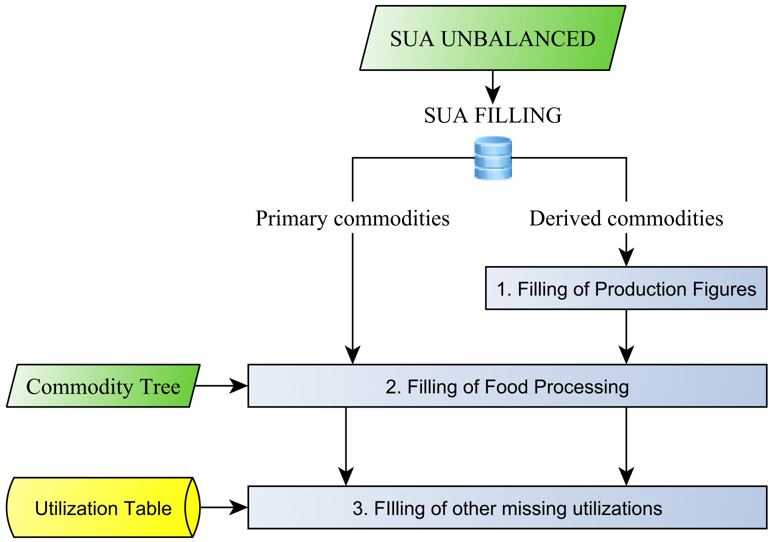

Figure 5: The Sua Filling diversified workflow

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*:

$$Imb1_{ijt} = P_{ijt} + I_{ijt} - X_{ijt} \tag{4}$$

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/increment the Production figure. The new production figure (P^*_{ijt}) is computed as:

$$P^*_{ijt} = P_{ijt} + Imb1_{ijt} \tag{5}$$

- If $Imb1_{ijt} >= 0$ 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, one of which is a no-Tree commodity: *Mixes and dough for the preparation of baker's wares* (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	-	***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	126,277	3,659	162,324		0	-			0			-	***-32,388***
Starch of Wheat	239,816	11,035	40,311			-	172,196			7,919		-	***210,541***
Wheat Gluten	25,580	877	117,373			-	0					-	***-90,916***
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

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		2,985,279	2,713,000	-	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	**31,575**	6,497	38,072		0	-	0		0			-	0
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*), meaning that 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:

$$FP_{pjt} = \sum_{c=1}^{C} \left(\frac{P_{cjt}}{eR_{p \to c}} \right) \times s^1_{cp} \times w_c \tag{6}$$

where:

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

$$s^1_{cp} = \frac{availability1_{p(c)}}{\sum\limits_{p=1}^{A} availability1_{p(c)}} \tag{7}$$

  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 slighlty different availability will be found in a following step:

$$availability1_{p(c)} = (P_{pjt} + I_{pjt} - X_{pjt}) \times eR_{p \to c} \tag{8}$$

- 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, 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 commodities 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. 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:

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

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	1
China, Mainland	2014	Bran of Maize	Gluten Feed and Meal	1.00	0.67	1

Shares lower than 1 mean that the child commodity might be produced also from other parents and 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:

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

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

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

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	31,575	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			**264,580**	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. This step is based on the value of the following Imbalance:

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

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

Three different scenarios con be found (Figure 5):

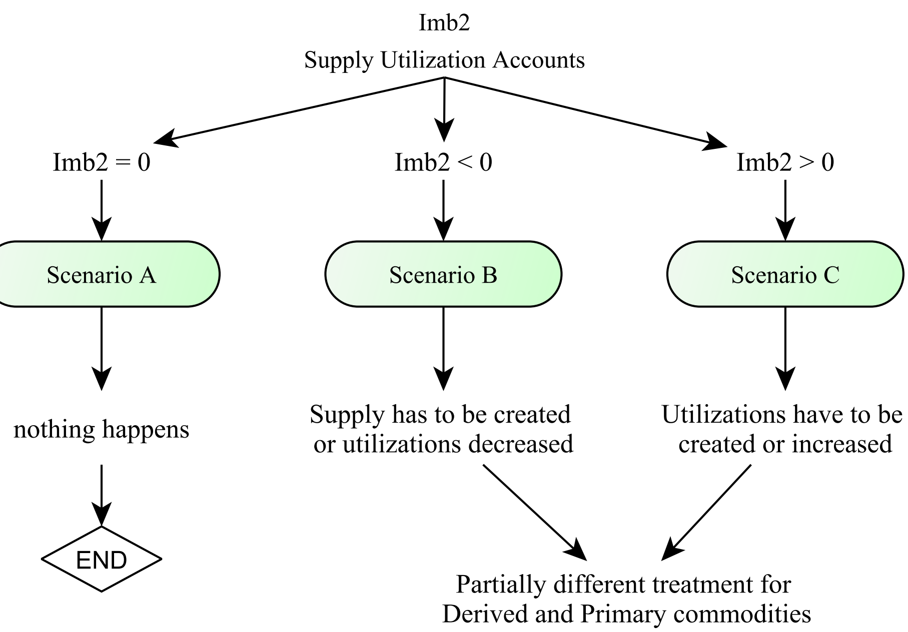

Figure 6: Scenarios of the Sua Filling

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	-	***-1,483,951***
Wheat and meslin flo	70,500,000	33,055	188,674		67,300,000	1,958,146			-17,345			-	***1,103,580***
Mixes and doughs for	31,575	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			158,664	172,196			7,919		-	***-128,239***
Wheat Gluten	116,496	877	117,373			264,580	0					-	***-264,580***
Communion wafers	13,263	8,796	5,822		16,241				-4			-	***0***
Uncooked pasta	1,415,692	12,520	22,550		1,405,661				-362			-	***363***
Food Preparations of		69,686	21,977		47,709				-12			-	***12***
Bran of Wheat	21,414,279	156,359	2,200		16,500,000		4,827,244		-4,252			-	***245,447***
Gluten Feed and Meal	793,740	160,231	529,333									-	***424,638***
bread	15,485	2,897	4,210		14,175				-3			-	***0***
pastry	193,950	89,593	117,630		165,914				-43			-	***42***

[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

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

All commodities enter this step, Primary commodity included.

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, Primary commodities remain unbalanced, while derived commodities can be still changed. In this case 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.This figure is then 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 (in all cases, derived or primary commodities) 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 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:

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

Notice that, after this step, all commodities are balanced.

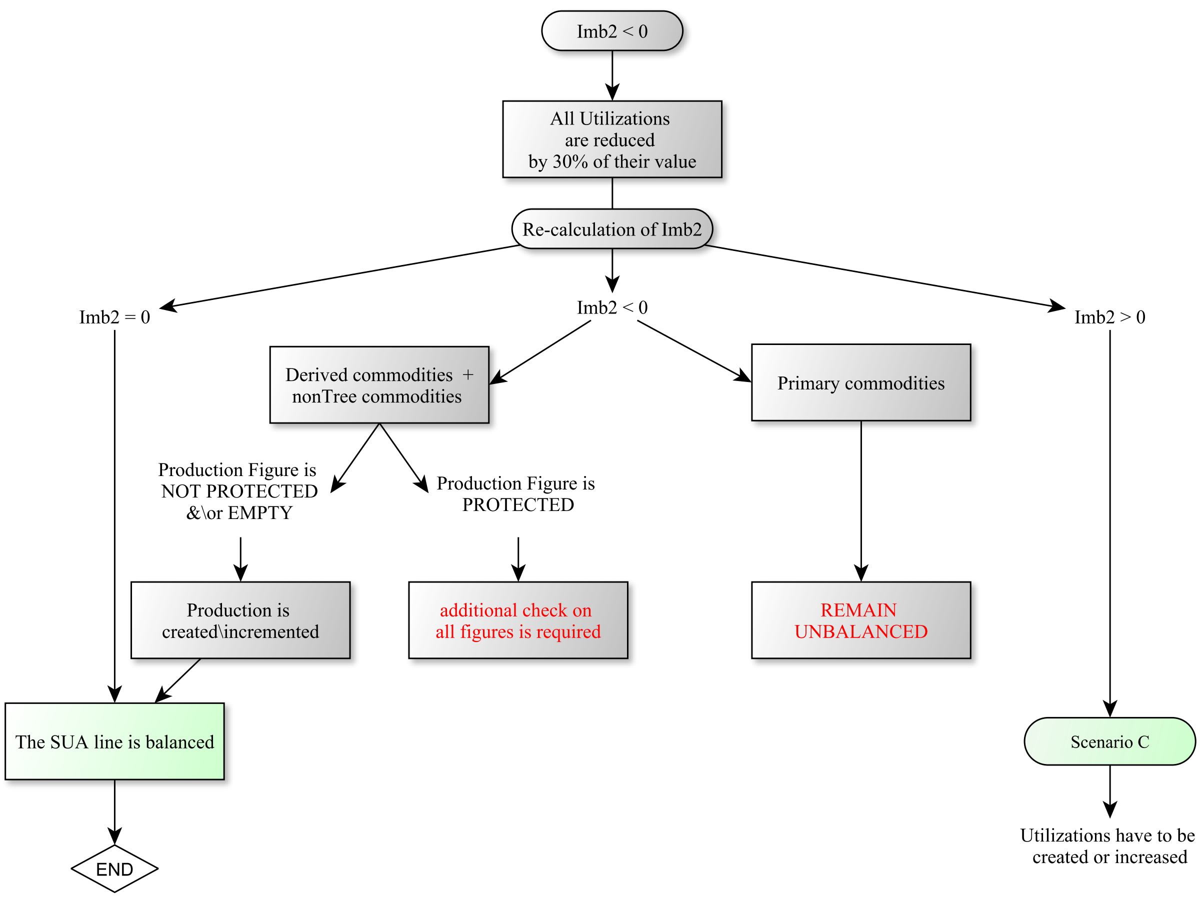

Figure 7: Workflow of Sua Filling when there is a negative imbalance

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	**1,079,280**		90,384,615	**28,106,487**	**4,119,970**		**2,875,293**	**2,613,046**	-	*0*
Wheat and meslin flo	70,500,000	33,055	188,674		**68,372,646**	1,958,146			**-17,621**			-	*0*
Mixes and doughs for	31,575	6,497	38,072		0		0		0			-	*0*
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	**381,076**	877	117,373			264,580	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		**47,721**				-12			-	*0*
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

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

AreaM49	Element	ItemCpc	medianRatio	Wr
1248	FP	23110	0.0303	0
1248	Fo	23110	0.9657	1
1248	DSt	23110	0.0000	0
1248	X	23110	0.0040	0

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

AreaM49	Element	ItemCpc	medianRatio	Wr
1248	Fe	39120.01	0.1366	1
1248	Fo	39120.01	0.8625	1
1248	X	39120.01	0.0009	0

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* that distributes imbalance according to weights proportional to median values of variables over a time range (2000-2013).

–> The *Utilization Table*

Utilization Table is a Country/commodity table telling which Utilizations historically have been active for each commodity and reporting the percentage of the median values of each utilization (over all utilizations). Median values are calculated over the period 2000-2013. Table 9 shows the *Utilization Table* for *Wheat and meslin flour* in China Mainland.

Where 1248 is the m49 code for China Mainland and 23110 is the CPC code of Wheat and Meslin Flour. 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*).

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*

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

As reported in the Figure, Three cases may happen:

1. If all the Utilizations included in the table 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

   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 reamain.

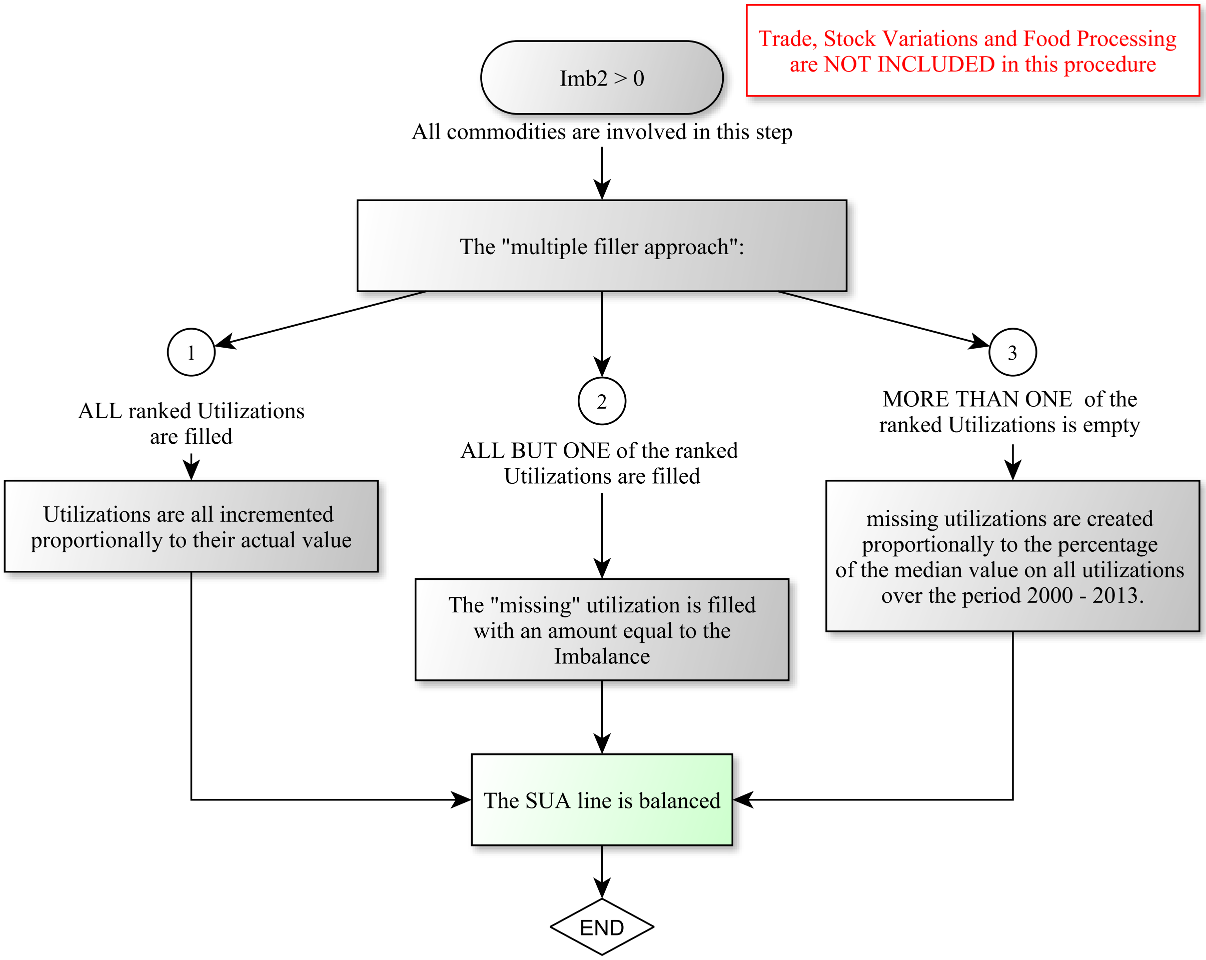

Figure 8: Workflow of Sua Filling when there is a positive imbalance

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 activable 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 8).

3. Finally, if more that one utilization is empty, a weighted approach is used. This rule assigns part of the imbalance to the Utilizations proportionally to the percentage of the median of that utilization on all other utilizations, as included in the Utilization Table. The rule works independently from the number of Utilizations that have to be excluded.

   Given the *Utilization Table* and considering the weights of the elements Wr, as reported in the following table:

element (Ut_h)	median ratio (MeR_h)	Utilization weight (Wr_h)	
Ut_1	MeR_1		Wr_1
...	...	...	...
Ut_h	MeR_h	Wr_h	
...	...	...	...
Ut_n	MeR_n	Wr_n	

Table 11: Weight of the elements in the Multiple filler approach

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

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
Cocoa beans	10,865	36	8,326	-	-	-	-	-	-	-	743	-	***1,832***

where:

- Ut_h is the generic utilization, with $h = 1, 2, ...n$
- MeR_h is the Median ratio of utilization Ut_h, with $h = 1, 2, ...n$
- Wr_h is the weight of utilization Ut_h and is equal to 0 or 1 depending on the element being, respectively, exluded or not from the *Sua Filling* (Table 11).

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

$$V_h = Imb2 \times pR_h \tag{15}$$

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

$$\begin{cases} pR_h = 0 & \text{when} \quad \sum_{h=1}^{n} Wr_h = 0 \\ pR_h = \frac{MeR_h \times Wr_h}{\sum_{i=1}^{n}(MeR_h \times Wr_h)} & \text{when} \quad \sum_{h=1}^{n} Wr_h > 0 \end{cases} \tag{16}$$

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)

Besides the excluded elements, there are 2 Utilizations empty: *food* and *Industrial Use* (even if this last is equal to zero. The Values V_{Fo} and V_{IU} of these two elements are calculated as follows

Table 13: Utilization Table for Cocoa Beans Madagascar

Ut	medianRatio	Wr
FP	0.00	0
Fo	0.20	1
IU	0.06	1
DSt	0.00	0
X	0.74	0

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
Cocoa beans	10,865	36	8326	-	**1,409.23**	-	-	-	-	**422.78**	743	-	***0***

and the result shown in Table 14:

$$\begin{aligned} pR_{Fo} &= \frac{0.2 \times 1}{0+0+0.2+0.06+0+0} = 0.77 \\ V_{Fo} &= 1,832 \times 0.77 = 1,409.23 \end{aligned} \tag{17}$$

$$\begin{aligned} pR_{IU} &= \frac{0.06 \times 1}{0+0+0.2+0.06+0+0} = 0.23 \\ V_{IU} &= 1,832 \times 0.23 = 422.78 \end{aligned} \tag{18}$$

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.

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

$$\text{Nutritive value} = \text{Nutritive factor} \times \text{Food availability (tonnes)} \times 10,000 \tag{19}$$

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

$$DES_{ijt} = \frac{\frac{Kcal_{ijt}}{Population_{jt}}}{365} \tag{20}$$

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 15: Sua Balanced - 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	1,079,280		90,384,615	28,106,487	84,119,970		2,875,293	2,613,046	-	***0***
Wheat and meslin flo	70,500,000	33,055	188,674		68,372,646	1,958,146			-17,621			-	***0***
Mixes and doughs for	31,575	6,497	38,072		0		0		0			-	***0***
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	381,076	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		47,721				-12			-	***0***
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

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*:

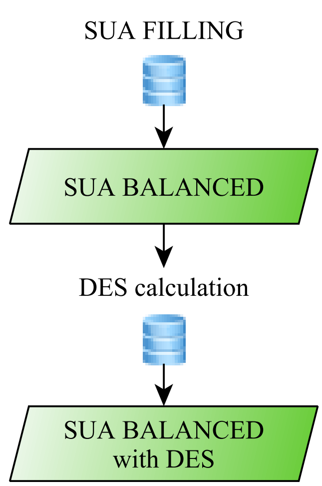

Figure 9: Sua Balanced and the calculation of DES

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. This produces 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 analytic 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.

Table 17: Generic SUA for Standardization for country j and time t

Item	el_1^*	...	el_h^*	...	el_n^*	$Kcal$
$parent_p$	V_{p1}	...	V_{ph}	...	V_{pn}	$Kcal_p$
$child_1$	V_{11}	...	V_{1h}	...	V_{1n}	$Kcal_1$
...	...	...	...	...	...	...
$child_c$	V_{c1}	...	V_{ch}	...	V_{cn}	$Kcal_c$
...	...	...	...	...	...	...
$child_C$	V_{C1}	...	V_{Ch}	...	V_{Cn}	$Kcal_C$

The standardization starts from the SUA table reported, in its generic form, in Table 17,where country and time indexes are omitted and where:

- $c = 1, 2, 3...C$ are the C children c of parent p,
- V_{ch} is Value of element h for child c,
- V_{ph} 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:

  $el^* \in E^* = \{I, X, \Delta St, , Fo, Fe, Lo, Se, IU, T, ROU\}$

- $Kcal$ 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:

a. expressing the vales of all elements V_{ch} of the children of a parent commodity, in terms of their parent commodity,
b. Adding up all the calories of the food commodities and calculate DES.

Elements are transformed in their parent commodity following the equation:

$$\begin{cases} V_{ph} = \sum_{c=1}^{C} \left(\frac{V_{ch}}{eR_{p\to c}} \right) \times s_{cp}^2 \times w_c \\ \\ DES_p = \dfrac{\dfrac{Kcal_p + \sum_{c=1}^{C} Kcal_c}{Population}}{365} \end{cases} \tag{21}$$

where, for country j and time t:

- $c = 1, 2, 3...C$ are the C children c of parent p,
- $eR_{p\to c}$ is the *extraction Rate* from parent p to child c,
- w_c is the *weight* of child c,
- $Kcal_p$ are the Kcal of the parent commodity (if is a food commodity),
- $Kcal_c$ 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:

$$s_{cp}^2 = \frac{availability2_{p(c)}}{\sum_{p=1}^{A} availability2_{p(c)}} \tag{22}$$

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.

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

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

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:

$$availability2_{p(c)} = FP_p \times eR_{p \to c} \tag{23}$$

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

Comparing this table with Table 15:

$$availability2_{wheat(flour)} = 90,384,615 \times 0,78 = 70,500,000 \tag{24}$$

$$availability2_{wheat(bran)} = 90,384,615 \times 0,22 = 19,884,615 \tag{25}$$

Moreover:

$$s^2_{Gluten,wheat} = \frac{5,796,652}{5,796,652 + 1,803,440} = 0.76 \tag{26}$$

$$s^2_{Gluten,Maize} = \frac{1,803,440}{5,796,652 + 1,803,440} = 0.24 \tag{27}$$

By applying equations 21 to 23 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.

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

The "Fbs Standardized" table

The standardized table is called *Fbs Standardized* and is composed of more than 1 line, because, besides the primary commodity there 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 processed, 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

Table 21: Measurement Errors percentages

el	varP_{el}
P	0.10%
I	1%
X	1%
DSt	25%
Fo	5%
FP	25%
Fe	25%
Se	25%
T	25%
IU	25%
L	25%
ROU	0

4 The FBS Balancing

The balancing of Standardized Sua in the new framework is based on the idea that, at this step of the overall process, a *little* imbalance has been cumulated, due to the errors of each variable. The FBS equation can be thought as being composed by each variable and its measurement error (ME).

$$\begin{aligned} P + ME_P + I + ME_I - \Delta St + ME_{\Delta St} = X + ME_X + Fo + ME_{Fo} + Fe + ME_{Fe} + Lo + ME_{Lo} \\ + Se + ME_{Se} + IU + ME_{IU} + T + ME_T + ROU + ME_{ROU} \end{aligned} \tag{28}$$

Because is not possible to know the ME a-priori, an approach is based that computes adjustment values for each variable proportionally to some *a-priori allowed percentages of variation*

Therefore, equation 1 can be rewritten, in its balanced form, as:

$$P^*_{ijt} + I^*_{ijt} - X^*_{ijt} - \Delta St^*_{ijt} = FP^*_{ijt} + Fo^*_{ijt} + Fe^*_{ijt} + Lo^*_{ijt} + Se^*_{ijt} + IU^*_{ijt} + T^*_{ijt} + ROU^*_{ijt} \tag{29}$$

where (dropping the indices for brevity):

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

and where each *adjustment ($adj_{el \in E}$)* is a balancing adjustment calculated starting from *a-priori allowed percentages of variation* for each element $el \in E = \{P, I, X, \Delta St, FP, Fo, Fe, Lo, Se, IU, T, ROU\}$.

The process is a two-step process. It start with the definition of the *allowed percentages of variation ($varP_{el \in E}$)* for each element as reported in the following Table:

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

Indeed, the *final balancing* happens according to the Residual and Other Uses is set equal to the Imbalance steps (figure 10:

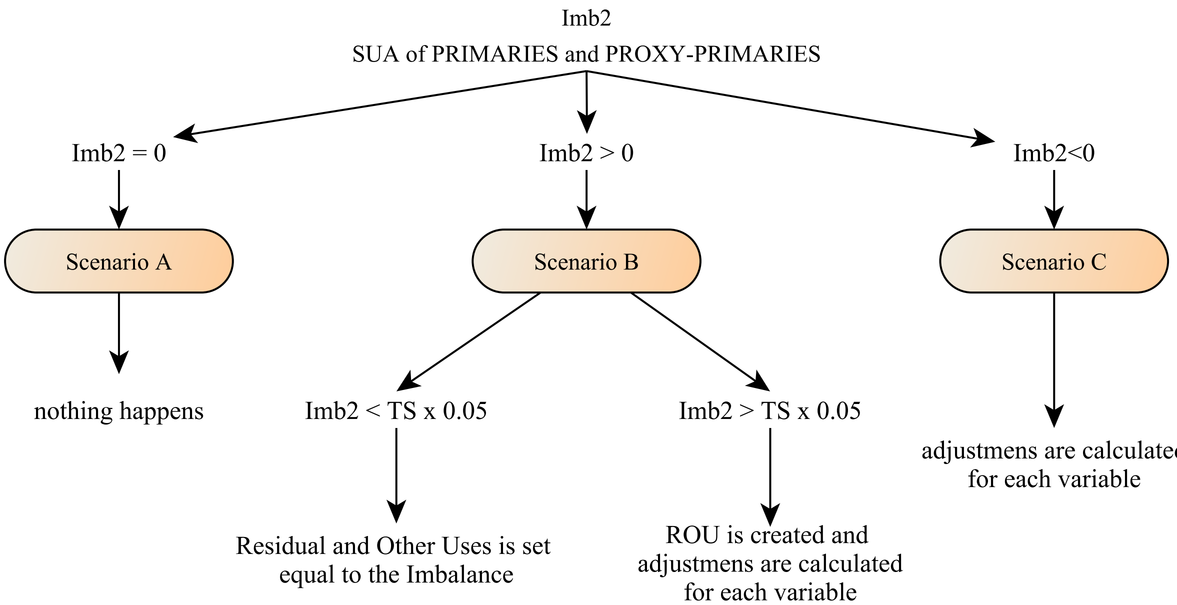

Figure 10: Scenarios of the final balancing

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:

$$adj_{el \in E} = Imb2 \times V_{el} \times p_{adj_{el}} \tag{30}$$

with

$$p_{adj_{el}} = \frac{varP_{el}}{\sum_{el \in E} (V_{el} \times varP_{el \in E})} \tag{31}$$

and

$$\sum_{el \in E} p_{adj_{el}} = 1 \tag{32}$$

adjustments here are considered with Positive sign for utilization and negative for Supply.

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

ROU is not activated and the other elements are balanced accordingly to equations 30 to 32, with the opposite sign.

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:

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

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

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

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 level,
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 23: 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
			2514 MAIZE & PRODUCTS ...	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 24: 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

Room/needs for improvement

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.

Some aspects that might need revision and/or update are:

1. How to manage the balancing of *stock variation* during the sua Filling. As said across the document, figures of stock variation are not touched during the sua filling. While this appeared the only way of treating stocks in the present version of the methodology, some effort might be taken in the future for letting also stock being included in the process.
2. How to manage the imbalance. A check on the amount of imbalance might be added before the balancing happens, so to warn in case of huge imbalance, for the need of revision at input-data level.
3. The user might need to change figures in the *sua unbalanced* table. If it happens, an additional module should exists that synchronize back the figures in the original data-sets. In this way, the new figures will be available the following year, as starting point of new imputations.

This methodology has been used for producing FBS of year 2016.