FAO resource book for the compilation of food balance sheets

Contents

Tables

Figures

About this document

This document is the first edition of a resource book to guide the compilation of food balance sheets (FBS) at the country level. It is based on the new FBS methodology developed by the Statistics Division of FAO (ESS) and allows FBS compilers at the country level to build up their own FBS by following a step-by-step procedure. To facilitate this capacity development process, the manual will be accompanied by a set of materials (PowerPoint presentations, e-learning modules, online tests, etc.) aimed at facilitating access to and implementation of the new methods. It will also come with a separate software package that allows countries to: i) review the basic data used by FAO in the past; ii) revise and possibly replace the principal parameters used by FAO in the new system; iii) reproduce FBS for any year in the past; and iv) provide updated FBS to FAO as new basic data become available. The software package will enable FAO and FBS compilers at the country level to share the same input and output files. It is written in "R" and thus allows all centrally designed procedures to be rolled out smoothly.

The various elements of the revised FAO methodology have been presented to the FAO Interdepartmental Working Group on Statistics whose members provided comments that have been incorporated into the methodology and into this document. In addition, elements of the new methodology have been exposed to external reviews, notably the sections on trade, trade endorsement calculations, food losses and waste, and feed use. Some approaches have already been adopted by other users; for instance, the feed use methodology has been adopted by the Agricultural Market Information System (AMIS), which now offers its members an online calculator for feed use, based on the new FBS feed methodology. Last but not least, all methods have been developed in close cooperation with technical divisions at FAO. Specifically, collaboration with the following FAO divisions helped improve the supply utilization accounts (SUA)/FBS:

- AMIS helped to improve stock data, sources and methodology.
- The Rural Infrastructure and Agro-Industry Division (AGS) of FAO helped to improve sections on post-harvest losses and waste and explored new sources of data on industrial use of food commodities.
- The Animal Production and Health Division of FAO (AGA) helped to improve feed estimates, using expert knowledge and information from various databases, notably the Global Livestock Environmental Accounting Model database.
- The Plant Production and Protection Division of FAO (AGP) helped to improve the paragraphs on seed rates and the method for imputing seed use.
- The African Commission on Agricultural Statistics (AFCAS), the Asia and Pacific Commission on Agricultural Statistics (APCAS) and the Inter-American Institute for Cooperation in Agriculture (IICA) of FAO provided feedback, inputs, and updates on methodological innovations and capacity development.

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

Overarching findings and principles guiding the FBS revisions

The following are the most important findings and principles that guided the FBS revisions:

- All assumptions are made explicit and are documented. The same commitment will apply to future changes and new assumptions.
- FBS are analytical data sets. They will always have to combine measured with imputed information. Imputation methods cannot replace data collection efforts, no matter how sophisticated they are.
- Every effort must be made to collect more and better-quality data at the country level; the quality of the results of any imputation depends critically on the quality of the measured information.
- Poor imputation methods can sometimes create vastly inaccurate results even where they are based on solid data. Every effort has been made to identify sound imputation methods and base them, to the extent possible, on solid data.
- Imputation methods seek to harness links between the various FBS variables and elements and information from outside the FBS. This allows triangulation of information and ensures overall consistency among FBS variables. The new imputation method for feed use is probably the best example of how this has been implemented in practice.
- Analytical data sets are always associated with large inaccuracies, stemming from differences in data definitions and classifications, measurement errors, imputation problems, etc. To reflect these issues in the FBS results, all estimates have expected values and an explicit measurement error. No claim is made that the estimates are point estimates. The overall philosophy guiding the revisions is to be "roughly right rather than precisely wrong".[1]
- The new methodologies seek to harness innovations in both statistical approaches and new information and communications technology (ICT) to the maximum extent possible. However, they are not intended to replace manual inputs and quality checks. On the contrary, time saved through automatic procedures provides more time for quality assurance and quality control.
- Quality control/assurance procedures are built into the system at various stages. Full compliance with FAO's new Statistical Quality Assurance Framework will be achieved as it and the new procedures mature.

The rest of this document is organized as follows. The first part provides a basic introduction with a quick overview of the history of FBS, their basic uses and usefulness, and the limits to these uses and usefulness. Chapter 1 and 2 introduce the main sources of data, new imputation methods and the new balancing mechanism. Chapters 3 and 4 provide a step-by-step introduction to the new methods based on worked examples. The document concludes with an overview of the changes introduced by the shift to new commodity classification schemes, and of how these changes relate to the existing system (chapters 5 and 6). Chapter 7 provides concepts and definitions used in the FAO FBS.

[1] To quote John Maynard Keynes.

Introduction

The FAO Statistics Division (ESS) regularly reviews and revises the methodological approaches for all of its products. Such revisions include all the databases maintained at ESS, their underlying and accompanying metadata, and the approaches for imputing missing data or creating analytical databases such as greenhouse gas inventories, food balance sheets (FBS) or, most recently, the System of Economic Environmental Accounts for Agriculture, Fisheries and Forestry.

All analytical databases, by their very nature, include a large number of imputed data or analytically derived data. FBS require many, and often particularly complex, transformations of primary data. In undertaking these transformations, FAO always emphasizes how important it is for countries to undertake actual data collection, and encourages all countries to improve and increase data collection efforts through the Global Strategy to Improve Agricultural and Rural Statistics. Notwithstanding efforts to obtain measured data in as many cases as possible, the process of establishing FBS often starts with a rather limited set of hard statistics. For many countries and many commodities, actual measurement of the constituting variables is entirely absent or, where available, is associated with large implicit or explicit measurement errors.

In addition to the need to impute an often large number of variables in a balance, setting up a complete set of food balances also requires a multitude of conversion coefficients, extraction rates and nutritive factors. These too can change, albeit typically far more slowly. For this round of FBS revisions, changes in conversion rates and factors were more important than on previous occasions. The main reason for additional changes is that the underlying commodity classification system has been revised as part of the overall reform effort. The commodity list traditionally used in the system, the FAOSTAT commodity list (FCL), has been replaced by a system of international classifications. In particular, the Harmonized System (HS) will be used for all trade variables, and the United Nations Central Product Classification (CPC), expanded for FAO purposes, will be used for all other variables in the balance. Every effort has been made to ensure consistency between old and new systems and across the new systems. However, in spite of all of these efforts, some conversion factors have had to be adjusted to reflect the product definitions of the newly adopted systems.

The revisions focused on updating the various (imputation) methods of the FBS components and – of particular importance – the overall approach for setting up and solving the balance between all variables of supply and utilization. The motivations for these changes and the differences to existing approaches are laid out in the different sections of this document. An important change relates to the new approach to solving the overall balance. In essence, it constitutes a move from a deterministic approach towards a process that takes into account not only the expected values but also the accuracy of the various variables (“elements”) of the balance that are being measured. The approach eventually selects a combination of values for the various variables that provides the most likely outcome while taking into account the boundaries of confidence/measurement for each variable.

History and background of food balance sheets

As defined by FAO, an FBS provides a comprehensive picture of the pattern of a country's food supply during a specified reference period. The preparation of FBS has a long history, extending back to 1934–1938. Over time, FBS have undergone methodological change, but since 1984, when they were first published in a standardized format, FBS methods have been left largely unaltered, despite periodic review.

Overview[2]

For each food item (usually expressed in primary commodity equivalents, with a few expressed as processed products), the FBS shows the amounts of supply and utilization of foodstuffs.[3] The total quantity of foodstuffs produced in a country, added to the total quantity imported and adjusted to any change in stocks that may have occurred since the beginning of the reference period gives the supply available during that period. On the utilization side, distinctions are made between the quantities exported, fed to livestock, used for seed, used in the manufacture of food and non-food items, lost during storage and transportation, and available for human consumption. The per capita supply of each food item available for human consumption is obtained by dividing the respective quantity by the related data on the population partaking of it. Data on per capita food supplies are expressed in terms of quantity, and – by applying appropriate food composition factors for all primary and processed products – caloric value and protein and fat content.

Annual FBS that are tabulated regularly over several years will show trends in the overall national food supply, disclose changes that may have taken place in the types of food consumed, i.e. the pattern of the diet, and reveal the extent to which the country's overall food supply is adequate in relation to nutritional requirements.

It is important to note that the quantities of food available for human consumption as estimated in the FBS, relate simply to the quantities of food reaching the consumer. The amount of food actually consumed will be lower than the quantity shown in the FBS, depending on the degree of losses of edible food and nutrients at the retail level and in the household, such as during storage, in preparation and cooking (which affect vitamins and minerals to a greater extent than they do calories, protein and fat), as plate-waste, or as quantities fed to domestic animals and pets or thrown away.

Waste on the farm and during distribution and processing is taken into consideration as an element in the FBS. Technical losses occurring during the transformation of primary commodities into processed products are taken into account in the assessment of extraction/conversion rates. There are very few surveys on which to base sound figures for waste, and the few that are currently available are sometimes subject to significant margins of error. In most cases, the assumptions for waste used in FBS are based on expert opinion obtained in the countries.

---

[2] Drawn from: http://www.fao.org/waicent/faostat/agricult/fbs-e.htm

[3] Not all items currently presented in the FBS are food commodities. For example, the FBS include non-food alcohol.

The system involves the compilation and maintenance of data on more than 500 primary and processed commodities, using the supply utilization accounts (SUAs) compiled by country, every year, and which are then standardized into more than 90 FBS commodities and respective commodity aggregates for dissemination. The architecture is organized in "commodity trees" linking primary and derived commodities. The links are formalized through extraction rates and a clear hierarchy spanning up to four levels of processing.

"Vertical standardization" is the conversion of processed commodities back to their primary equivalents. The extraction rates or technical coefficients that were used in building up the database are used to carry out this conversion back to the primary level, by multiplication of the reciprocal of the technical conversion coefficient. There are several reasons for this standardization process. First, it reduces excess information on many different and heterogeneous processed and primary products to a set of products that allow analysts and policy-makers to obtain a quick and meaningful overview of the food economy of a country. Second, preparing standardized FBS at an aggregate level makes the food supply and use data more comparable among countries and over time. Third, many practical tools and applications, such as projection models, require a high level of product aggregation; in fact, many require a level of aggregation that is beyond that offered in the FAO FBS. For instance, FAO's long-term outlook is based on 34 commodity groups derived from the more than 60 individual commodities presented in the FBS.

Use and usefulness of FBS

FBS have been compiled for nearly 80 years and have been used for numerous different purposes ever since. They provide a wealth of information on food and nutrient availability and, when available over longer timeframes and compiled with consistent methodologies, provide the basis for identifying trends and developments in the food economy of a country, a region or worldwide:

1. One of the most important uses of the FBS arises from the fact that they provide a measure of the overall average calorie supply in a country. The dietary energy supply (DES) is not only a standard output of the FBS, but also the key input into the FAO indicator of undernourishment, i.e. the number of undernourished people and the prevalence of undernourishment. The DES enters the prevalence of undernourishment measurement as the mean of a distribution, which, together with a cut-off point – the minimum dietary energy requirement – allows FBS users to calculate the number and percentage of people in a population without access to sufficient calories.
2. FBS have also been used to examine changes in dietary patterns. While this is possible in principle, it is important to note that the observed changes refer only to the average diet and therefore do not allow inferences to be drawn on whether the dietary quality in a country has improved or deteriorated. In fact, it is impossible to say much about the quality of the average diet. The estimates provided may simply be the average of unhealthy overnutrition and unhealthy undernutrition, while the average diet may appear to be about right in terms of volume and composition although most people in the country are consuming an unhealthy diet.

3. FBS are the starting and end points of many, if not most, partial equilibrium models.[4] Models such as the Aglink model of the Organisation for Economic Co-operation and Development (OECD) and FAO, the International Food Policy Research Institute's IMPACT model, and partial equilibrium models for food and agriculture generally use commodity balances as the basis for their projections. They typically use every element/variable of the balance on the left-hand side of the equation and project the starting values of the FBS (commodity balances) into the future. As most of these models use zero global net trade as their closure rules, imports and exports are not projected separately but are collapsed into one variable as net trade. FBS are to be seen as a subset of the general family of commodity balances, i.e. commodity balances for food items.
4. FBS also provide a rich basis for calculating numerous other and simpler indicators. The most straightforward indicators are simple ratios such as self-sufficiency and import-dependency ratios. But FBS also provide inputs into policy measures such as the producer subsidy and consumer subsidy equivalents, which are regularly updated by OECD, and the trade policy analogue of FBS provides inputs into the aggregate measure of support used by the World Trade Organization.
5. Probably the most common use of FBS data in the published literature is the citation of DES and fat and protein intake (e.g. by the United States Department of Agriculture's [USDA's] Economic Research Service [ERS], Grigg, Hopper, Pinstrup-Andersen, Svedberg, Trueblood, Smil). Estimates of intakes of other nutrients including vitamins, minerals and amino acids, are also based on FBS data on food availability. FBS information has also been used to spot shortfalls and surpluses in a nation's energy and nutrient intake (e.g. by the Southern African Development Community, USDA/ERS) or to examine the availability of a particular commodity or class of commodities (e.g. by el Obeid, Hopper, Helsing).
6. Finally, the medical community has also made use of FBS. Researchers have used food and protein availability from the FBS to study the availability and importance of various amino acids and different sources of proteins. They have examined relationships among caloric intake, protein types and amino acids in the diet (Hopper, Young, Kazuo). Medical research has used FBS data to investigate connections between diet and health, especially cardiac health and cancers (Sasaki, Helsing). Medical researchers have also evaluated the degree to which FBS data are useable and relevant. For example, Sasaki and Kesteloot examined correlations between FAO data and data from multiple surveys in 19 countries and deemed the FBS data useable and valuable. It should be noted, however, that the majority of these studies were for developed countries, which have more reliable data and clearer methodological approaches.

Limitations

FBS are the most comprehensive collection of data related to food commodity supply and utilization that is available for a very large set of countries. The data are regularly revised, continually improving and becoming more consistent. In their totality, FBS provide a wealth of information and serve numerous uses. However, there are also strict limits to their applicability and usefulness. FBS compilers and users must be aware of these limitations and of the potential errors that exist in estimates: for example, availability is not the same as intake, and FBS do not say anything about the

[4] Strictly speaking, they are just a subset of these starting and end points. Commodity balances for food and non-food items provide the complete basis for these models.

distribution of food and nutrient access within a country. Some major limitations are summarized in the following:

1. FBS provide only estimates of *average* national food or nutrient availability. They do not offer any insights into the distributional aspects of food and nutrient availability. This means that they allow their users to draw inferences on the average diet, but not on the diet of food-insecure or poor people. FBS also do not provide information about the regional distribution within a country, the access of particular groups of households, or dietary habits. With these caveats in mind, FBS can and are being used as either a contributing factor or the sole basis for analysis of food demand and supply. They provide an approximate picture of the overall food situation in each country.
2. The imputation of missing data cannot replace proper data collection. This FBS resource book offers methods for detecting inconsistencies, imputing missing data and filling data gaps. However, no model-based approach can substitute for, let alone produce more accurate information than, actual data collection. It is therefore imperative that countries creating their own FBS first take stock of food supply and utilization data, juxtapose available data with data needs, and make a final decision on whether FBS can be built with confidence contingent on remaining data gaps. Where insufficient data are available for many commodities and variables/elements, it may be necessary first to improve the domestic database, focus on data collection for missing elements and resume the FBS compilation once sufficient basic data are available. The Global Strategy to Improve Agricultural and Rural Statistics provides the basis for such an endeavour by offering cost-efficient methods for data collection and providing capacity development support to facilitate access to these methods.
3. Even if most or all of the data needed to compile FBS are available, there is a considerable challenge in assembling FBS data from different sources whose coverage and quality vary greatly. More often than not, the underlying accounting system uses a mixture of data from different sources of disparate quality with unknown sizes of error. In practice, primary production and trade data are often the only official data regularly collected, while data on the production and utilization of processed commodities are virtually non-existent, or sparse at best. In addition to the collection of basic data, the compilation of FBS requires a multitude of coefficients. These coefficients are used, for instance, in the calculation of product flows in a commodity tree and in standardization, and were collected or estimated years ago. They are also needed in the assessment of feed use, the calculation of nutrient contents of foods, the derivation of food consumption by tourists, and many more calculations. Where food is used as a balancing variable/element, as it was in the FAO FBS system of the past, extra caution in checking the results is warranted. The balancing item in a supply/utilization system always assumes all the measurement errors of all other elements in the balance. As there is no reason to assume that these measurement errors cancel each other out, the balancing element is the least reliable – or at least the most variable – item in the system. This resource book therefore proposes a new approach that explicitly takes into account the fact that all variables/elements are measured with a certain degree of inaccuracy, and avoids relegating measurement errors to one single element.
4. FBS results require careful interpretation. For instance, food availability assessed through the FBS cannot be directly compared with food consumption data from household surveys. The differences are manifold, and include the different coverage of food consumption – household surveys do not include "collective" consumption in hospitals, schools, the military or prisons, nor do they completely cover out-of-home consumption in restaurants, street food, etc. They present food consumption net of retail waste, which is included in the FBS. Household

surveys may also lack representative coverage over the full reference period of an FBS – a calendar year. In this resource book, no systematic attempt is made to tally FBS results with food consumption or expenditures from household surveys. FAO working papers are available that systematically compare FBS estimates with results from household surveys, and these should be consulted before using household survey results to guide FBS food estimates.[5]

Reference

Grünberger, K. 2014. *Estimating food consumption patterns by reconciling food balance sheets and household budget surveys*. FAO Statistics Division Working Paper Series No. ESS/14.08. Rome, FAO. http://www.fao.org/3/a-i4315e.pdf.

[5] For details on the methodology for comparisons between FBS and household surveys, their scope and limitations see Grünberger, 2014.

1. FBS and SUAs, underlying data, sources, measurement and imputation

1.1 FBS and SUAs

This resource book provides descriptions of the methodologies, processes and practical steps for constructing an FBS. All the steps for compiling a final FBS are based on more disaggregated balances of primary and processed products, which are referred to as supply utilization accounts (SUAs). It is only at the final step of aggregation – the so-called "standardization" (see section 2.2 for the methodology and chapter 4 for practical examples) – that the more aggregate and standardized template known as the FBS is created. While almost all compilation steps refer to the finely disaggregated SUAs, the many hundreds of individual balances they contain are of little use to policy-makers or even analysts. They are therefore not disseminated by FAO. National FBS compilers may need to weigh carefully the pros and cons of making these detailed SUA balances available.[6]

Overview and basic description of FBS/SUA variables

FBS and SUAs consist of a set of variables describing the various forms of supply and utilization of food. All variables used and presented in the FBS are also included in the SUAs. In addition, the compilation of SUA information requires a multitude of auxiliary variables such as conversion factors, shares and ratios and particularly information on how the various processed products relate to their primary product equivalents. The relationships between processed and primary products are encapsulated in "commodity trees", which provide the basic roadmap for consistent commodity aggregation, or standardization. The conversion factors, ratios and shares are presented later in this section, along with the trees and the logic of standardization. Before delving into these details, it is important to define and delineate the basic SUA/FBS variables.

Both FBS and SUAs present simple balances between supply and utilization. They include production, imports and stock withdrawals on the supply side, and exports, food (of the resident population), tourist consumption, feed, seed, waste and losses, industrial use and stock additions on the utilization side. These variables are defined as follows.

- Production

Production refers to the total amount of food produced in a country. As FBS aim to provide food security information, production includes not only commercial and marketed products, but also non-commercial food production and subsistence farming, including foods from orchards and home gardens. Production of primary crops is reported at the farm level, excluding harvesting losses. Output of all crop products is expressed in metric tonnes. The same holds for livestock products, with meat production defined in terms of carcass weight. For processed and derived commodities, production refers to the total amount of output at the manufacturing level.

- Trade

[6] FAO does not publish the detailed SUAs not least because their proper use requires full knowledge of the various compilation steps, which is seldom available to policy-makers and analysts. Past experience with published SUAs suggests that many analysts misinterpreted the information in the underlying SUA balances and drew incorrect (policy) inferences.

Trade refers to all transboundary flows of food items destined for, or originating from, a given country. During data collection, exports and imports of processed and primary products are often reported in their initial units of measurement. These can be heads, pieces, or measures of volume such as litres, hectolitres, bushels, barrels, bales and many other units. The FBS/SUA system requires that all variables be expressed in metric tonnes. FBS compilers must therefore ensure that all trade data are converted into metric tonnes. Compilers may want to refer to relevant conversion factors provided by the United Nations Statistics Division (UNSD), or use FAO trade data instead.

- Food

Food in the FBS definition refers to the amounts of edible product available for human consumption during the reference period of a calendar year. This variable presents the amounts of food reaching the retail level and is therefore inclusive of all waste and losses that occur at or after retail. FBS food amounts are therefore typically higher than the amounts of food reported by other measurement methods, such as household income and expenditure surveys or, even more so, nutrition surveys. Both of these methods are exclusive of certain amounts of waste. In contrast to this, FBS food includes all waste of edible products occurring in shops, supermarkets or households, such as during storage, in preparation and cooking, as plate-waste, fed to domestic animals and pets, or simply thrown away. It also includes food consumed in hospitals, schools, restaurants, military establishments, prisons, etc.

- Feed

Feed represents the quantity of edible products – both domestically produced and imported – that is available for livestock production. It therefore excludes all feeds that are not foodstuffs or not destined for human consumption, including roughages and by-products, such as oilcakes, distiller's dried grains with solubles (DDGS), dregs, etc. These feedstuffs are taken into account in the calculation of total feed use, but are excluded from the FBS calculations and presentation.

- Seed

Seed includes amounts of a commodity set aside for sowing or planting (or for reproduction purposes in general) during the year. The average amount of seed needed per hectare planted for a given crop in a given country does not usually vary greatly from year to year.

- Food losses and waste (FLW)

FLW are the amounts of all food products that are lost at any stage of the supply chain, from the farm up to but excluding the retail level. Also excluded are all losses during the pre-harvest and harvesting stages, and technical losses occurring during transformation of the primary commodities into processed products. The latter are taken into account separately in the standardization process through appropriate extraction rates/conversion factors reflecting processing losses.

- Industrial use

Industrial use covers the quantities of the commodity used during the year for non-food purposes such as biofuels, paints, detergents and cosmetics.

- Tourist consumption

Tourist consumption comprises the food consumed by non-resident visitors during their stay in the country.

- Stocks

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

- Nutrient estimates

Based on SUA quantities, population estimates and nutrient conversion factors, the FBS also feature country-wide averages of calorie, protein and fat availability. The sum of all calories available for human consumption – the DES – is arguably the most important output of the FBS table. In addition, FBS tables also provide fat and protein availability, by product, and from all products. Nutrient estimates are calculated during the SUA process and, apart from nutrient conversion factors, do not require any separate inputs.

Auxiliary variables

To produce additional indicators such as per capita use, and to facilitate standardization, additional variables need to be collected by FBS compilers. They include the following.

- Population

Estimates of the resident population are available from the United Nations Population Division (UNPD), and it is highly recommended that FBS compilers use these estimates, even if alternative national sources are available. UNPD estimates are derived through a globally consistent methodology and serve as the basis for all per capita estimates for United Nations (UN) agencies. They have been used as the common denominator for all per capita Millennium Development Goal indicators and will be the common denominator for all Sustainable Development Goal indicators and targets.

- Activity and productivity variables

FBS tables display only quantities of supply and use, but FBS compilers need to collect information on other variables to arrive at the variables displayed in the FBS. Typically, these other variables include underlying activity and productivity indicators. For primary products, underlying activity indicators include area harvested and area sown in the case of crops, and number of animals[7] in the case of livestock. Underlying productivity indicators are yield and cropping intensity for crops, and slaughter weight and take-off rate for livestock.

Calculation of the amounts of processed products requires the corresponding extraction rates (flour to wheat, etc.). The nutrient conversion factors, which provide information about the nutrient content (calories, protein and fat) per kilogram of food, are also required. All relevant FBS variables are summarized in Table 1.

[7] Including any animals kept for draft proposes, for meat and dairy production, or for breeding. The unit of measurement is expressed in number of heads, or 1 000 heads (for poultry, rabbits and other rodents).

Table 1: Basic and auxiliary variables for producing FBS and SUAs

Production	Basic variables for SUAs/FBS
Imports	
Stock changes	
Food manufacturing	
Feed	
Seed	
Losses and waste	
Industrial utilization	
Tourist consumption	
Other uses	
Food: total calorie equivalent	Derived variables and indicators
Calories/capita/day	
Food: total protein equivalent	
Proteins/capita/day	
Food: total fat equivalent	
Fats/capita/day	
Population	Auxiliary variables
Area sown	
Area harvested	
Yield	
Processed	
Stocks	
Slaughtered	
Carcass weights	
Inputs	
Extraction rates	

1.2 Sources and collection of FBS data

FBS constitute an integrated data set. Their need to combine and reconcile data from a variety of sources poses challenges, but also offers benefits. A main advantage is that FBS provide a full picture of the food and agricultural economy of a country, spanning the different domains of agricultural production, international trade, nutrition and the various subsectors of food and agriculture – crops, livestock and fisheries. Coherent collation of data on production, trade and various forms of food utilization allows FBS users to derive a multitude of indicators, including overall supplies of nutrients, and import-dependency and self-sufficiency ratios. FBS also provide a framework for checking and ensuring data consistency. As each set of data is confronted with evidence from other sources, FBS can help identify problems in the compilation of source data, thereby contributing to enhanced accuracy. However, the coherence achieved within the FBS can come at the cost of potential incoherence between the FBS and the underlying source data. Explaining the reasons for these inconsistencies is important for maintaining credibility and may sometimes constitute a challenge.

Compiling complete FBS requires a considerable amount of data. Ideally, these are collected from reliable sources, based on solid measurement methods and grounded in concepts and definitions that are compliant with international standards. FBS compilers should use imputed or estimated data only when measured data are not available. National statistical systems around the world have disparate capacities for collecting and measuring the required source data, usually reflecting the different strengths of national agricultural statistical systems.[8]

Providing primary data on losses, stock changes and food appears to be particularly difficult for national statistical systems in all parts of the world, as can be seen from Table 2, which provides an overview of data availability by region and variable. For instance, the FAO Statistics Division (ESS) has only been able to collect about 1 percent of the data required from government agencies for these variables. Slightly more official data on other forms of utilization, such as feed and seed, are available. Obtaining official data for agricultural production, imports and exports is much less difficult, as systems for compilation of these variables are usually in place for the production of national accounts, among other purposes. This means that the supply part of the FBS can generally be compiled with data of higher quality than the utilization side. Data from official sources are particularly scarce in Asia and Oceania and in Africa.

[8] A global assessment of agricultural data carried out by FAO and the World Bank concluded that many developing countries do not have the capacity to collect and disseminate the basic agricultural data required to monitor national trends or inform the international development discussion, and that the quantity and quality of agricultural statistics worldwide has even decreased over time (World Bank, 2011).

Table 2: Proportion of data for FBS compilation collected from official sources, 2011–2013 (percentages)

	Productio n	***Import s***	***Stock change s***	***Export s***	***Fee d***	***See d***	***Losse s***	***Foo d***
Developed countries	54.8	77.4	0.4	72.5	5.5	15.5	1.6	0.4
Africa	27.1	41.8	0.3	35.1	2.3	3.2	0.2	0.0
Latin America and the Caribbean	40.5	45.6	3.4	44.2	2.6	4.7	1.2	0.3
Asia and Oceania	37.8	54.7	0.0	48.2	1.1	1.3	0.1	0.0

World	**39.0**	**54.9**	**0.8**	**51.5**	**3.3**	**5.9**	**0.8**	**0.1**

Coverage: All data cells for primary products, excluding fishery, with valid non-zero values.

Source: FAOSTAT internal working system, extracted on 14 August 2015.

1.3 Data collection for FBS compilation

The first step in setting up an FBS compilation system is to take stock and review the quality of possible data sources. Based on this initial review, data gaps can be identified and a needs assessment for additional data collection efforts can be conducted. The data gaps can then be reviewed against available data collection methods – including new, cost-efficient methods available through the Global Strategy to Improve Agricultural and Rural Statistics – which should involve both public and private sources. Provisions should be made to enable efficient transfer of data. Ideally, all data will be transferred electronically based on standard templates.

In a second step, efficient working arrangements should be established with the data providers and other actors involved in FBS compilation. This step is of paramount importance in ensuring smooth collaboration and information flow during FBS compilation.[9] Coordination of activities can be facilitated through a national working group that serves as a governing body and in which the main data providers from both the public and private sectors are represented. In the design phase, the working group will serve as a forum for discussing the commodity coverage, concepts and methods to be applied, and the development of infrastructure for data exchange. During the production phase, the working group may validate all data provided by its members, and review the compiled FBS before dissemination. The working group should also pursue the standardization of data formats and concepts and suggest new ways of filling data gaps.

The following subsections provide an overview of the most common – in some cases the most highly recommended – data sources and methods for measuring and collecting data for the main variables included in the FBS.

Data sources and collection

There are several sources of reliable information on agricultural production in a given country. Regular and representative production surveys are the tried and tested way of collecting and improving information on crop and livestock production. To ease the relevant data collection processes and – of particular importance – to lower their costs, the Global Strategy to Improve Agricultural and Rural Statistics proposes a number of innovative survey designs and sampling frames, which should also enable poorer developing countries to embark on regular and cost-efficient surveys. For instance, the Global Strategy has identified ways of enhancing the accuracy of estimates while lowering data collection costs, such as by combining different sampling frames into a multiple-frame sampling exercise (Vogel, 2015). Some countries already apply two-stage sampling, mostly

[9] Lack of integration within national agricultural systems has been identified as a main cause of decline in the coverage and quality of agricultural data in developing countries (World Bank, 2011). Overcoming this deficit is crucial for setting up an efficient FBS compilation system, which requires the involvement of a variety of different organizations, because FBS provide a cross-domain data set.

with cluster-sampling, in which administrative areas or geographic territories are usually used as the primary sampling units.

In addition to regular surveys, useful auxiliary information (e.g. number and type of animals, hectares of area harvested) can be gleaned from agricultural censuses; however, censuses are typically conducted infrequently, often at ten-year intervals.[10] In addition, most of the variables surveyed are of a structural nature, which means that census data leave important data needs (e.g. on productivity) largely unaddressed.

Some countries keep useful administrative records of agricultural production; for example, local government officials need to track the numbers of slaughtered animals in their districts, and so instruct the relevant agencies to keep records (Pica-Ciamarra *et al.,* 2014: 35 ff). These administrative data can constitute a valuable source of information for the compilation of FBS production figures. When production is concentrated in a small number of large production or processing units, or when farmers sell their produce to a small number of agro-industrial companies, as is the case for cash crops, the records kept by private companies can provide another excellent source of production estimates, particularly when these records are highly accurate and frequently updated. The total amount of production can then be derived from the sales or purchase records of several companies. These records may be obtained either directly from the companies concerned, or indirectly through tax authorities or branch-specific industry organizations (flour millers' associations, etc.).

In developing countries where smallholders account for a large proportion of agricultural production, integrating dedicated modules on agricultural production into existing household surveys has rendered very reliable and FBS-relevant results. An example of these dedicated modules is the integrated survey on agriculture that has been created by including a questionnaire on agricultural production in the questionnaire and sampling design of the World Bank's living standards study. As a first step in implementing the survey instruments of the Global Strategy to Improve Agricultural and Rural Statistics, FAO is currently developing a multipurpose survey based on combined samples of agricultural holdings in the household (smallholdings) and non-household (larger farms and cooperatives) sectors. This Agricultural and Rural Integrated Survey (AGRIS) will be implemented in priority countries targeted by the Global Strategy to support development of their national agricultural systems. Its themes will cover the range of topics addressed by the Global Strategy, including production of the main crops and livestock products (ESS, 2015).

Whenever data are collected through surveys, the production of each unit surveyed needs to be measured. The following sections review the variety of methods developed for doing this, first for crop production, then for the production of livestock products.[11]

---

[10] FAO provides basic guidance on how to collect census information, which variables to cover, how to prepare for a census and how to process the data collected. The FAO guidelines for the 2020 round of the World Census of Agriculture have just been released (FAO, 2015b).

[11] This review is based on two unpublished reports recently prepared in the context of the Global Strategy, on crop production by Sud *et al.* (2015) and on production of livestock products by Moss *et al.* (2015).

Data on primary crop production

Production of primary crops can be measured through either a direct approach, which relies on counting the harvests obtained during the reference period, or an indirect approach, which relies on the assessment of harvested areas and average yields and multiplication of the two.

For *direct measurement* of crop production, the whole-plot harvest method is considered the most accurate, albeit time-consuming, method. In this method, the enumerator observes the harvest of the whole plot and records the result. To save time, she/he may count the number of units, such as sacks, baskets or bundles, harvested by the farmer, and multiply this number by the average weight of one unit, assessed from a sample of units. When harvests occur throughout the year, enumerators may visit farmers regularly, ideally every harvesting day, to record the amount of crop harvested since the previous visit. Alternatively, the size of the harvest can be reported by the farmer, ideally in combination with checks by the enumerator – such as inspections of storage facilities – to prevent false reporting. To reduce recall errors, particularly with frequent or extended harvests, "crop cards" are used, on which farmers record the results of each harvest. As only one variable is measured, the direct approach can generally be expected to lead to smaller measurement errors in production figures than the indirect approach, which involves measurement of two variables – area harvested and yield. The direct approach is therefore particularly suited for FBS compilation, although direct measurement of crop production is often time-consuming, unless it is based on farmers' self-assessment or the purchase records of large companies.

Indirect measurement of crop production first requires measurement of the harvested area. This measurement can be based on maps on which segments are identified by physical boundaries, cells of a grid or individual sample points, and the crop(s) cultivated in these area units are based on visual observation shortly before the harvest. Alternatively, the areas under different types of crop can be identified by means of remote sensing, particularly by establishing the proportions of the different coloured pixels assigned to different crop types that are included in a segment of a satellite image, and extrapolating the total area of each crop's share from the known area of the entire segment. A more traditional method involves visiting the sampled pieces of land and assessing their sizes, for example by fitting the area into an equivalent set of rectangles, triangles or other polygons, and measuring the side lengths of these polygons, applying standard geometric formulae to calculate the area of each, and summing the results. Another method, suggested by FAO (1982), relies on measuring the perimeter of the area and dividing by a number between 4 and 5, depending on how close it is to being a square (divided by 4) or a complex polygon (divided by 5), and then squaring the result. Based on field experience, fairly accurate results can be obtained from Global Positioning Systems (GPS). Equipped with a GPS device, the enumerator walks around the perimeter of the field while the GPS automatically calculates the surrounded area, based on the coordinates of points registered during the walk. Assessment of the areas harvested of individual crops is complicated when more than one crop is grown simultaneously on the same field. In these cases, imputations are required to assign a proportion of the area to each crop. An overview of different methods is provided by Fermont and Benson (2011: 29–34).

The second step in the indirect measurement of crop production requires assessment of the average yield. The most common method for this assessment, recommended by FAO (1982) as standard, is the crop-cutting method, in which a small sub-plot is randomly located within each field prior to harvest. Instead of harvesting the entire field, only the sub-plot is harvested by the enumerator or the farmer, and the yield is measured. Using more than one sub-plot per field for crop-cutting enhances the accuracy of the yield estimate and enables estimation of within-field variance. Alternatively, the yields can be predicted before the harvest takes place, either by the farmers themselves, accompanied by the enumerator, or by experts such as extension staff or field technicians. The expected crop yield can be determined by means of visual assessment or empirical formulae, taking into account measurable characteristics or morphological attributes of the growing crop, such as the number of open cotton balls in a certain row length, the grains per head, plant height or ear length. A technically more advanced, and not yet fully developed, method is to analyse spectral data obtained from the spectral reflection of different bands in satellite images.

Data on livestock production

Measurement of livestock production, such as meat, milk or eggs, does not require measurement of area. As in the case of crop production, physical assessment is generally considered to be more reliable than self-reporting by farmers. Physical assessments are usually carried out by an enumerator, extension officer or market agent who observes and records the number of slaughtered animals and their average carcass weight, or the amount of milk, eggs or wool produced. Carcass weights may be predicted before slaughter, based on experts' visual assessment of the animals' characteristics, particularly their live weight and fat content, probably complemented by the use of technical devices such as scanners. Self-reporting by farmers, which generally requires fewer resources, is a more common method, particularly in developing countries. Self-reporting can be conducted through face-to-face or phone interviews, based on questionnaires. Alternatively, respondents can complete questionnaires and return them by surface mail or via the Internet.

During survey design, there is need to pay particular attention to the difficulties that may arise in collecting data on livestock production. First, the owner of the livestock is often not the same person as the owner of the ground on which the livestock is kept, and it is generally the latter who is selected as the survey respondent. While the owner of the livestock may not hold any land at all, the survey respondent, i.e. the land owner, is not usually capable of providing the necessary information on the livestock on his/her farm. Second, in production systems where livestock are moved from one place to another in search of fresh pastures for grazing (nomadic pastoralism), dedicated data collection methods are required – such as aerial flyover surveys and electronic tagging of animals, combined with remote collection of production data – to guarantee complete coverage and avoid double-counting. A final difficulty is that holdings with intensive animal farming, while including large numbers of animals and producing high output, usually cover small areas. The probability of these holdings being included in area frame-based sample surveys is therefore exceptionally small (Vogel, 2015: 35).

Data on imports and exports

National legislation usually obliges the importers and exporters of goods to provide information on their cross-border transactions to the customs office, in the form of tax declarations. Among other information, these tax declarations include a description of the product, the applicable commodity code based on the HS classification (see chapter 6 for details), the weight, quantity or number of the traded good, the country of origin (for imports) or destination (for exports), and the date of the transaction. As all transactions of cross-border merchandize trade should be reported to customs offices through tax declarations, this data collection method ensures large coverage, a low additional reporting burden and low collection costs. UNSD therefore recommends that customs records be treated as the "most prevalent source" of data on international trade, and that "statisticians take advantage" of that source (UNSD, 2004: 11).

Nevertheless, customs records are not free of errors. For example, errors emerge from inaccuracies in customs declarations, such as incorrect designation of product codes, application of incorrect units of measurement, transcription errors or simple typographical mistakes. In addition, traders sometimes fail to report trade in goods to avoid customs duties or to circumvent trade restrictions. Errors may also slip in during the processing of tax declarations by the tax administration or other institutions involved (Eurostat, 2009: Annex I). Less frequently, certain transactions are not reported for confidentiality reasons, as is often the case for military equipment, or for commercial reasons (trade secrets). To clean the trade data from such inaccuracies, additional quality assurance and validation steps are inevitable, both at the country level and for FAO.

The quality of the trade data can be enhanced by merging customs records with data from other sources, such as:

- records of the financial flows accompanying transactions between residents and non-residents kept by the country's central bank and collected from the financial institutions in charge of settling these transactions, in the framework of the international transactions reporting system;
- other countries' shipping manifests prepared for ships, aircraft and vehicles crossing the border, which list details of their cargoes;
- records of port administrations prepared on the basis of shipping manifests;
- records of the parcel and letter post, which need to be kept in most countries as stipulated by the acts of the Universal Postal Union;
- aircraft and ship registers, particularly those recording exports and imports via ship and aircraft, which do not always imply border-crossing of the traded objects;
- reports of commodity boards;
- records compiled for the collection of value-added tax;
- enterprise surveys (UNSD, 2004: chapter 4).

In some countries, such as the Democratic Republic of the Congo, records on cross-border trade in agricultural products are compiled by the national veterinary service, which is represented at the main customs points to verify the compliance of animal- and plant-based products with hygienic standards.[12]

The second step in enhancing the quality of trade data involves measuring and including all the food aid obtained from relief agencies; if this is not possible, additional imputations, such as of data from the Food Aid Information System (WFP, 2015), should be conducted. The data may also be reconciled with values recorded for the same trade flows by the trading partner country (mirrored data). Comparison of exports with mirrored data is considered particularly helpful, as imports are usually documented more exhaustively and checked more diligently than are exports (UNSD, 2004; chapter 13).

National trade data are forwarded to the UN, checked for quality and eventually brought into the public domain. ESS and UNSD regularly collect the official trade files of about 180 countries. ESS makes efforts to enhance the quality of trade records for agricultural products, subjects them to additional quality checks and reconciles them with statistics from other sources. Data gaps are filled

[12] Information gathered during a country visit to the Democratic Republic of the Congo by ESS in 2011.

through various imputation approaches, such as dividing the recorded monetary value of individual flows by appropriate unit values obtained from elsewhere, imputing mirrored data, or adding import and export figures from multinational industry organizations. Eventually, the trade flows are aggregated and the aggregate flows subjected to further validation checks, notably commodity-specific global consistency checks (minimal net trade). Once all validation and imputation steps are completed, the final data are published on the FAOSTAT Website (FAO, 2015a).[13]

A particular challenge in compiling trade statistics is the inclusion of shadow trade, especially in developing countries. Shadow trade comprises all trade in goods that has not been subjected to statutory border formalities such as customs clearance (Afrika and Ajumbo, 2012: 2). It is therefore not covered by the statistics compiled from customs records. Although it is estimated that in Africa shadow trade is an important source of income for a large share of the population, and even provides a noticeable contribution to food security (Afrika and Ajumbo, 2012: 1), only few developing countries regularly compile data on shadow trade to enhance their official trade statistics.[14] The national statistical offices of Cameroon and Rwanda estimate shadow trade on the basis of a survey of people carrying goods when they cross the main border points of the country (Habinshuti, 2014; Nguingnang, 2014). Unofficial data on shadow trade in East and Southern Africa are compiled by the Famine Early Warning Systems Network each quarter using a similar approach (FSNWG, 2012: 2015).

Data on stocks

Stocks of agricultural products are held by a variety of actors, including farms, processing or trading companies, government agencies and international relief organizations such as the World Food Programme (WFP). A number of factors determine the amount of stocks they hold. Stocks are generally held for the purpose of smoothing consumption; in remote locations in developing countries they are also kept as a reserve for insurance against food shortage, or to store value where a fully developed rural finance system is lacking.

There are several ways of measuring on-farm stocks. One option is through agricultural censuses, but stocks are rarely an integral part of these censuses. Of 196 national censuses, only 30 collect information on storage infrastructure and capacities, and only seven include stock levels directly (Fonteneau, 2014). In addition, agricultural censuses are carried out at long time intervals, so do not really constitute a useful tool for stock measurement. The usefulness of censuses is particularly limited by the fact that stock levels usually exhibit significant short-term variations, limiting the usefulness of censuses for the assessment of stocks and making interpolations difficult.

Data on the stocks held by government agencies and private-sector companies, such as warehouses, retail traders and wholesalers, can be collected from the respective offices or companies. Regarding the stocks of agricultural holdings, agricultural surveys are usually the most apt data collection method. The Agricultural Market Information System (AMIS) and the Global Strategy to Improve

[13] UNSD disseminates the collected trade files in raw form on its commodity trade database (UN, 2015). The Global Trade Information System, a private-sector company, also collects official trade files from about 80 countries and makes them accessible on its Web site (GTIS, 2015).

[14] A recent survey of informal cross-border trade in Africa, carried out by UNSD with 22 compilers of national statistics on international merchandise trade, revealed that only two African countries – Cameroon and Rwanda – regularly compile statistics on informal trade (Muryavan and Iversen, 2014).

Agricultural and Rural Statistics are currently making efforts to develop surveys covering stocks in all countries.

The combination of farm survey data and data collected from institutional stock holders provides a fairly accurate estimate of stocks at the national level. USDA operates such an approach for its biannual farm surveys. The East African Community has adopted this methodology for the compilation of monthly FBS.

Data on feed use

Farmers feed their animals with agricultural products that they purchase or source from their own agricultural production. Both components are important, and their relative weights in overall feed use vary across countries, depending on the prevailing livestock production systems and their degree of specialization, commercialization and intensification. FBS compilers need to create a comprehensive picture of food items used as feed supply. This requires coverage of feeds produced by both farmers and feed compounders. Information is needed on the amounts and types of product that farmers and compounders use to produce the feeds. As such surveys are costly, only few representative and comprehensive surveys are currently available (Westcott and Norton, 2012: 5).[15]

In some cases, exhaustive surveys of feed use have been added to the regular agricultural survey.[16] This is a particularly appealing option in developing countries where a large proportion of the livestock is reared by smallholders and the proportion of purchased feed is comparatively low. Compound feeders can be contacted at the same time, either directly or through industry associations. In some countries, such as Germany and Hungary, feed producers are obliged to report regularly to the ministry of agriculture on their sales of fodder products, based on a government regulation (Marktordnungswaren-Meldeverordnung, 1999; BLE, 2011: 1–4; KSH, 2014). Combined with statistics on external trade, these data may provide a substitute for a dedicated data collection process.

Data on food

For the measurement of food consumption, two principal sources can provide estimates. The first source is production estimates from the food processing industry, which can provide excellent, FBS-consistent estimates of food availability, particularly: i) where the processing industry accounts for a large part of food use, i.e. where home/subsistence production and consumption play a small role; and ii) where a processing industry has a high degree of coverage/organization, such as where all flour is produced and reported by commercial flour millers. This is typically the case in developed countries, but is also increasingly frequent in emerging economies. Estimates of production from flour millers, sugar refiners, oilseed crushers or abattoirs can be used as direct inputs into the SUA system and, through the process of standardization, be converted back into primary equivalents for inclusion in the FBS. While a high degree of coverage/industry organization may not always be possible in

[15] An exception is the compulsory collection of “Annual Data of Agricultural Enterprises” carried out by the Central Statistical Office of Hungary among all organizations carrying out agricultural activities other than agricultural surveys (KSH, 2014).

[16] In the Republic of Moldova, for example, a feed use questionnaire was added as a one-off module of the Agricultural Survey of Smallholders in 2012. In Rwanda, questions on feed use are regularly included in the new Seasonal Agricultural Survey (NISR, 2014).

developing countries, and certainly not for all processing industries, "bottleneck" industries such as sugar refineries provide useful estimates in all countries.

Data compiled from household surveys provide a second source of estimates for food consumption data. However, there are a number of caveats and adjustments to be made before such data can be used to inform estimates of food consumption in FBS. First, such surveys are usually carried out rather infrequently, once every three or five years, and even less frequently in many developing countries. Second, although these surveys are designed to capture a representative sample of a country's population, some population subgroups are typically underrepresented, particularly homeless and marginalized people, and this inevitably introduces a bias in overall food consumption data. Third, these surveys practically always exclude "collective consumption", i.e. food consumed in institutionalized households such as the military, prisons, canteens, schools, universities, hospitals or senior citizen centres. In addition, not all household surveys fully account for food eaten outside the home in restaurants, as street food, etc.

To obtain FBS-consistent food supply information, food consumption estimates from household surveys need to be adjusted in line with the scope and definitions of FBS. For instance, underrepresentation of particular population groups can be adjusted for by re-calculating the extrapolation weights that are meant to reflect households' inverse sampling probability and are used to scale up the values surveyed at the household level to the total population. However, as only one vector of weights is often applied, this adjustment requires selection of one variable for which underrepresentation is to be eliminated, such as employment status, total income or total expenditure. As "food consumption in quantities" is rarely the chosen variable, there is likely to be bias in the estimate for the total population.

The food consumption measured in household surveys needs to be brought into line with food availability as defined in the FBS (see definitions in section 1.1), so all food consumed in collective households and outside the home needs to be estimated, and household survey results adjusted accordingly.

To gauge the extent of differences and facilitate the necessary adjustments, FAO has systematically compared the food figures derived from household surveys with those derived from FBS. The results suggest that the two are only modestly correlated across countries. Throughout a sample of 64 countries, data on food consumption compiled from FBS were on average higher than those obtained from household surveys. In some extreme cases, the difference amounted to two-thirds of the value obtained from household surveys, while in others the value derived from FBS were up to one-fifth smaller than the household survey result. In contrast to the limited correlation in levels, the shares of individual types of food in total consumption appear to match more closely between the two sources of data (ESS, 2015b).

Given the principal and practical discrepancies between household surveys and FBS, FBS compilers are encouraged to: i) ensure that they understand the scope and definitions of household surveys; and ii) identify a way to reconcile the results of household surveys with, or scale them to, the FBS definitions. FAO has devised such a method, the details of which are laid out in an ESS working paper (Grünberger, 2014).

Data on food losses and waste

Measurement of FLW should cover all stages of the food supply chain at which losses and waste can occur (see section 3.8). Data collection could be organized along the food supply chain and should cover the entire chain from production, storage, processing and retailing to the final household – “from farm to fork”.[17] Before setting up the data collection system, it is advisable to conduct a general baseline study at the level at which the most relevant food supply chains are analysed and the involved processes and actors identified (FAO, 2015b: 13; O’Connor *et al.*, 2015).

Whenever FLW data cannot be obtained from records kept by public or private entities, studies based on samples are considered to be the most efficient way of collecting these data. The sampling design should take into account the main characteristics and structure of the subsector of the food industry being surveyed. Objective measurement based on weighing, counting or assessing volume is commonly the preferred data collection instrument, even though it may require more time than other methods. Questionnaires and oral interviews with product owners can be useful for cross-checking the results of the measurement, but they are not recommended as stand-alone instruments as non-response rates and reporting errors are usually high. Product owners are often reluctant to admit losses of large amounts.[18]

Objective measurement can be carried out in a number of ways, depending on the type of food loss prevalent in the particular process of the food supply chain. Food losses can take the form of *disappearance*, when pieces of the product drop out of the food chain, or *biodegradation*, when pieces lose their adequacy for use as food. Disappearance occurs mainly as a result of actions to which the product is exposed along the food supply chain, such as during stacking, stooking, threshing, shelling, cleaning, winnowing and drying of crops on the farm, during transport, during processing and during the commissioning of food at the wholesale and retail levels. Biodegradation occurs mainly over time as a result of biological or chemical transformations or infestation by insects. These transformations take place mainly during storage and, to a certain extent, during drying and long-distance transport.[19]

Measurement of *disappearance* can be carried out by the following methods:

- *Collecting the dropouts caused by a certain activity*, such as by laying out plastic sheets on the floor during the stacking, stooking and threshing of cereals on the farm to catch the scattered grains; analysing the grain that has remained on the straw after threshing; collecting the food left on counters at the end of the serving period in catering companies; or analysing the business’s waste container or – when scanning-based systems are in place, for example in wholesale and retail trade – extracting the records of food items classified as losses from the database.
- *Following up food products and comparing their state before and after an activity*, which is the predominant method of estimating losses during transport and which can also be easily

[17] For FBS purposes, coverage from farm to retail suffices.

[18] Detailed guidelines on the design and implementation of food loss surveys can be found in FLWP (2015: chapters 5, 6, 13) and, tailored to the crop sector in African countries, in Hodges (2013: chapter 2.4).

[19] See Harris and Lindblad (1978: chapter 5) for a more detailed description of the various types of grain loss, and O’Connor *et al.* (2015: chapters 6–8) for a description of food losses caused by wholesalers, retail, markets and food services.

applied, for example, to measure losses during wholesale and retail trading when scanning-systems are in place.[20]

- *Comparing the outcomes of an activity when particular care is taken to prevent losses and when it is not*, which has been recommended, for example, for measuring the grain damage caused by threshing.[21]

A traditional and relatively precise way of measuring *biodegradation* is the count-and-weigh method, initially recommended for measuring grain loss during storage. In this method, a sample of the damaged grain is analysed in a laboratory to assess the weight loss per grain, under the perspective of adequacy for consumption as food. The weight loss per grain is then extrapolated to the total weight loss in the sample by taking into account the proportion of damaged grain (Harris and Lindblad, 1978: 90 ff). A main disadvantage of this method is that it requires large amounts of time and work. In addition, the removal of parts of the produce from the food chain for analysis in the laboratory and the lack of involvement of product owners in the analysis may considerably reduce owners' willingness to participate in such loss measurement surveys. To overcome these disadvantages, the African Postharvest Losses Information System suggests a simplified method based on a visual scale (Hodges, 2013), which provides a set of images depicting grain samples accompanied by numbers representing the percentage weight loss associated with the depicted grain, as established from laboratory analysis. To measure loss, the enumerator, jointly with the grain owner, compares the sampled grain with the pictures on the visual scale and applies the appropriate percentage weight loss to calculate the total loss in the sample.

In later stages of the food supply chain, where agricultural products often occur mixed with other products, before any of these methods are applied, waste composition analysis may be required to distinguish among losses of different food products and to separate food products from other materials (FLWP, 2015: chapter 13.6).

Data on seed use

The most common method of measuring seed use is to include questions on seed use in the questionnaire for an agricultural survey. These questions usually cover only the use of farmers' own production of seed. Farmers may also buy seeds on the market, but these purchases consist mainly of "improved" seed. Data on the use of purchased seed may be obtained from the sales records of commercial seed companies.

As the amount of seed for a specific type of crop is strongly determined by the area sown, it may be sufficient to measure the size of that area, based on some of the techniques outlined for the measurement of area harvested (see the subsection on "Data on primary crop production"), and to multiply this area by the seed rate defined as the quantity of grain sown per hectare. This information may be gathered from farmers or other sources, including from past periods, given that the seed rate does not usually change much from year to year. If the area sown is not known, it can be approximated from the area harvested, although area sown is almost always larger than area

[20] This method can involve more than one activity; for example, in the "load tracking" method, a sample of the primary product is followed throughout the supply chain (FLWP, 2015: chapter 13.3).

[21] Detailed descriptions of these methods can be found in Harris and Lindblad (1978: chapter 5); Møller *et al.* (2014; chapter 5) and FLWP (2015; chapter 13).

harvested. The differences are caused by damage to the crop while it grows, resulting from natural disasters, pests, poor germination, animal grazing or floods. Farmers may also decide not to harvest because of lack of labour, lack of demand or, in the case of cassava, to keep a reserve for times of drought or food shortage (Sud *et al.* 2015: 9 ff).

Data on tourist consumption

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

Surveys collecting detailed information on the different types of food consumed by respondent tourists offer an efficient way of measuring food consumption by tourists. Other valuable data sources include the guest lists of hotels and records of the food purchased by the hotels, as the quantities of food purchased, multiplied by the proportion of days spent by guests from other countries may be used as a second-best estimator of tourist consumption. Information on the number of days spent by international tourists may also be obtained from travel agencies, airlines, railways and shipping companies.

The United Nations World Tourism Organization (UNWTO) collects data on the number of tourists entering a country, including their origins and the durations of their stays, and gathers these data globally (UNWTO, 2015). The UNWTO data set therefore provides another useful starting point for the compilation of data on tourist food consumption.

References

Afrika, J.K. & Ajumbo, G. 2012. *Informal cross border trade in Africa: Implications and policy recommendations*. Africa Economic Brief No. 3/10. Abidjan, African Development Bank.

BLE. 2011. Aufkommen an Futtermitteln in den Wirtschaftsjahren 2005/06, 2007/08, 2008/09 und 2009/10 vorläufig. Regionale Versorgungsbilanz Futtermittel Wirtschaftsjahr 2007/08 endgültig. Bonn, Bundesanstalt für Landwirtschaft und Ernährung (BLE).

ESS. 2015a. Agricultural and Rural Integrated Survey (AGRIS). Rome, FAO Statistics Division (ESS). (unpublished)

ESS. 2015b. Comparing and reconciling food consumption from household surveys and food balance sheets. Presentation at the Expert Workshop on Stock and Utilization Measurement in China, 6–7 July 2015, Beijing, organized by AMIS. FAO Statistics Division (ESS).

Eurostat. 2009. *MEDSTAT II: Asymmetry in foreign trade statistics in Mediterranean partner countries*. Eurostat Methodologies and Working Papers. Luxembourg, Office for Official Publications of the European Communities.

FLWP. 2015. *FLW Protocol Accounting and Reporting Standard*. Food Loss and Waste Protocol (FLWP). (draft)

FAO. 1982. *Estimation of crop areas and yields in agricultural statistics*. FAO Economic and Social Development Paper No. 22. Rome.

FAO. 2015a. FAOSTAT data available from: http://faostat3.fao.org/home/E

FAO. 2015b. *World Programme for the Census of Agriculture 2020, Volume 1 – Programme, definitions and concepts.* Rome.**Fermont, A. & Benson, T.** 2011. *Estimating yield of food crops grown by smallholder farmers. A review in the Uganda context*. Discussion Paper No. 1097. Washington, DC, International Food Policy Research Institute (IFPRI).

FEWS NET. 2012. *Informal cross border trade in Southern Africa*, Issue No. 78. Famine Early Warning Systems Network (FEWS NET)/WFP.

FIO & CFIA. 2012. *China feed industry almanac*. Beijing, China Agricultural Press for the Feed Industry Office (FIO) and the China Feed Industry Association (CFIA).

Fonteneau, F. 2014. Grain stocks direct measurement: Contribution from AMIS/BMGF Project. Presentation at the 2nd Joint Workshop on Rice Data for Thailand and the Philippines, 3–4 December 2014, Manila, organized by AMIS.

FSNWG. 2015. *East Africa Crossborder Trade Bulletin*, volume 9. Food Security and Nutrition Working Group (FSNWG), Famine Early Warning Systems Network (FEWS NET)/FAO/WFP.

Grünberger, K. 2014. *Estimating food consumption patterns by reconciling food balance sheets and household budget surveys*. FAO Statistics Division Working Paper Series No. ESS/14.08. Rome, FAO. Available at: http://www.fao.org/3/a-i4315e.pdf

GTIS. 2015. *Global trade atlas*. Available at: http://www.gtis.com/english/GTIS_GTA.html

Habinshuti, V. 2014. Informal cross border trade statistics in Rwanda. Presentation at the Regional Seminar on International Trade Statistics, 12–16 May 2014, Addis Ababa, organized by the United Nations and the African Union.

Harris, K.L. & Lindblad, C.J., eds. 1978. *Postharvest grain loss assessment methods: a manual of methods for the evaluation of postharvest losses.* Washington, DC, National Academy of Sciences for the American Association of Cereal Chemists.

Hodges, R. 2013. *How to assess postharvest losses and their impact on grain supply: Rapid weight loss estimation and the calculation of cumulative cereal losses with the support of APHLIS*. African Postharvest Losses Information System (APHLIS).

Klapp, C. 2010. Überarbeitung des Getreide- und Vieheinheitenschlüssels. Internationaler Vergleich und Konsequenzen der Anwendung alternativer Viehbewertungen im Kontext rechtlicher Rahmenbedingungen, Endbericht. Georg August Universität Göttingen, Fakultät für Agrarwissenschaften.

KSH. 2014. Metainformation. Koezponti Statisztikai Hivatal (KSH). Available at: http://www.ksh.hu/apps/meta.main?p_lang=EN

Moss, J., Baker, D., Morley, P., Al-Moedhen, H. & Downie, R. 2015. Improving methods for estimating livestock production and productivity. Literature review. Rome, FAO. (unpublished draft)

Møller, H., Hanssen, O.J., Gustavsson, J., Östergren, K., Stenmarck, Å. & Dekhtyar, P. 2014. *Report on review of (food) waste reporting methodology and practice.* European Commission, FUSIONS Reducing Food Waste through Social Innovation.

Muryawan, M & Iversen, K. 2014. Informal cross border trade. Preliminary analysis of results. Presentation at the Regional Seminar on International Trade Statistics, 12–16 May 2014, Addis Ababa, organized by the United Nations and the African Union.

Nguingnang, B. 2014. Méthodologie et outils de collecte des données sur le commerce transfrontalier de marchandises. Presentation at the Regional Seminar on International Trade Statistics, 12–16 May 2014, Addis Ababa, organized by the United Nations and the African Union.

NISR. 2014. *Seasonal agricultural survey report 2013*. Kigali, National Institute of Statistics of Rwanda (NISR).

O'Connor, C., Östergren, K., Quested, T., Soethoud, H., Stenmarck, Å., Svanes, E. & Tostivint, C. 2015. Food waste quantification manual. FUSIONS Reducing Food Waste through Social Innovation, Deliverable No. D1.7. (working document, draft)

Pica-Ciamarra, U., Baker, D., Morgan, N., Zezza, A., Azzari, C., Ly, C., Nsiima, L., Nouala, S., Okello, P. & Sserugga, J. 2014. *Investing in the livestock sector. Why good numbers matter. A sourcebook for decision makers on how to improve livestock data*. World Bank Report No. 85732-GLB. Washington, DC, World Bank.

Sud, U.C., Ahmad, T., Gupta, V.K., Chandra, H., Prachi, M.S., Kaustav, A., Biswas, A. & Singh, A. 2015. Synthesis of literature and framework under the project "Research on Improving Methods for Estimating Crop Area, Yield and Production under Mixed, Repeated and Continuous

Cropping”. New Delhi, International Committee for Animal Recording (ICAR) and Indian Agricultural Statistics Research Institute. (unpublished)

UN. 2014. UN Comtrade database. Available at: http://comtrade.un.org/

UNSD. 2004. *International merchandise trade statistics. Compilers' manual*. New York, United Nations Statistics Division (UNSD).

UNWTO. 2015. Tourism statistics. United Nations World Tourism Organization (UNWTO). Available at: http://www.e-unwto.org/toc/unwtotfb/current

Vogel, F. 2015. Defining the Master Sampling Frame for Agricultural Statistics – Basic principles. *In* FAO. *Handbook on Master Sampling Frame for Agriculture*. (unpublished draft)

Westcott, P.C. & Norton, J.D. 2012. *Implications of an early corn crop harvest for feed and residual use estimates.* Economic Research Service No. FDS-12f-01. Washington, DC, United States Department of Agriculture.

WFP. 2015. Food Aid Information System. Rome, World Food Programme (WFP). Available at: http://www.wfp.org/fais/

World Bank. 2011. *Global Strategy to Improve Agricultural and Rural Statistics*. Report No. 56719-GLB. Washington, DC.

2. Balancing and standardization

The previous chapter provided an overview of possible data sources for the various FBS/SUA variables, product shares, ratios and conversion factors. It made a strong plea for using actual measurements and data (from surveys, censuses, administrative sources) and provided options for the reconciliation of data from different sources. It also emphasized that there is no alternative to data collection and measurement in the long run. However, while every effort should be made to collect data, it is still necessary to deal with missing data, ratios, shares and conversion factors, i.e. to carry out imputation.

This chapter presents the general approaches and specific methods developed for data imputation at FAO. FBS compilers at the country level may want to use the same methods to impute missing values. This chapter presents the principles of the various methods, lays out basic assumptions and presents details of the approaches taken. All concepts are explored again in chapter 3, and illustrated with worked out and practical examples in chapter 4. As well as imputation methods, this chapter also introduces an innovative balancing mechanism and a process for consistently aggregating commodity detail into aggregates – the process generally referred to as “standardization”.

2.1 The basic balance and the balancing mechanism

Introduction

At the most basic level, FBS are simple identities (as are all commodity balances). 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 and total supply equal to total demand.

Domestic supply = domestic utilization:

$$P_{ij} + I_{ij} - X_{ij} - dSt_{ij} = Fo_{ij} + Fe_{ij} + Lo_{ij} + Se_{ij} + IU_{ij} + T_{ij} + ROU_{ij}$$[22] (equation 1)

Total supply = total utilization:

$$P_{ij} + I_{ij} - dSt_{ij} = X_{ij} + Fo_{ij} + Fe_{ij} + Lo_{ij} + Se_{ij} + IU_{ij} + T_{ij} + ROU_{ij}$$ (equation 2)

where $dSt_t = St_t - St_{t-1}$, P = production, I = imports, X = exports, S = stock level, Fo = food, Fe = feed, Lo = losses and waste, Se = seed, IU = industrial use,[23] T = tourist consumption, and ROI = residual other use.[24]

Ideally, as many variables as possible should be measured, and measurement should be made with a maximum degree of accuracy. When empirically measured, measurement should include both an

[22] The ideal equation would replace the stock change variable with two variables: “from stocks” on the left side of the equation, and “to stocks” on the right side. However, such variables are rarely measured, and it is much more reasonable simply to model stock changes.

[23] Industrial use does not include food manufactured into other products. Such quantities are accounted for in the food variable.

[24] All variables/elements are described in detail in their respective sections in chapter 3.

estimate of the expected value of every variable and its measurement error. In reality, a number of problems complicate the process of populating the identity.

First, measured values are usually limited to variables on the supply side (production, imports and exports), but measurement is typically available only for the expected values and not for the measurement errors. On the demand side, most estimates are imputed data, and again estimates are often limited to the expected values without their respective measurement errors.[25]

Second, it is not possible to include values for all variables in the balance, at least not if they enter the identity as point estimates. When this is the case, the balance will not have a solution unless one variable is left as a "balancing item". Inevitably, this balancing item assumes all the measurement errors implicit in all the other variables. If all the estimates for the other elements are unbiased, i.e. have measurement errors with expected value 0, then the expected value for the error of the balancing item/residual will be 0. However, the variance of the error for the balancing item/residual is the sum of the variance of the measurement errors for all the other elements, and this inevitably causes a large variability in the estimate of the balancing item/residual.

Third, as the balancing item assumes the measurement errors of all the other variables, it could be argued that not all variables are equally suitable for functioning as balancing items. Intuitively, variables with higher degrees of annual variability would be more suitable as balancing items. The underlying rationale for such a choice is that the residual/balancing item assumes the sum of all measurement errors, and is therefore more likely to exhibit greater year-to-year variability.

In practice, however, the choice of a variable for the balancing item often reflects the availability of data (or the lack thereof), rather than a clear economic rationale and empirical evidence. It is therefore not surprising that different SUA compilers/approaches have chosen different variables as their balancing items. For instance, USDA uses feed (and residual use) as the balancing item, while the FBS have often used food to balance supply and demand. Conveniently, the XCBS approach often chooses whatever variable is not explicitly available. Clearly, none of these approaches overcomes the problem of accumulating measurement errors in the balancing item. No matter which variable is used, the balancing item will be fraught with the measurement errors of all the other variables. In the context of the FBS, this means that using food as the balancing item is therefore the least suitable solution.

The approach in detail

The new balancing approach seeks to overcome the fundamental problem of identifying a single variable as the balancing item. To this end, all variables enter the balance with an expected value. Obviously, this is tantamount to creating an over-identified[26] equation, at least if the estimates enter the equation as point estimates, with only their expected value. To overcome this problem, the new approach uses an expected value for every variable, and its measurement error (e_i) or an approximation of its measurement error:

[25] In addition, many estimates are the result of complex imputation methods involving many steps and many variables for the balance, making empirical estimation of the measurement error difficult.

[26] Strictly speaking it is not an overidentification, which would require more than one equation. However, the problem is the same in principle.

$P_{ij} + e_{ij}(P_{ij}) + I_{ij} + e_{ij}(I_{ij}) - dSt_{ij} + e_{ij}(dSt_{ij}) = X_{ij} + e_{ij}(X_{ij}) + Fo_{ij} + e_{ij}(Fo_{ij}) + Fe_{ij} + e_{ij}(Fe_{ij}) + Lo_{ij} + e_{ij}(Lo_{ij}) + Se_{ij} + e_{ij}(Se_{ij}) + IU_{ij} + e_{ij}(IU_{ij}) + T_{ij} + e_{ij}(T_{ij}) + ROU_{ij} + e_{ij}(ROU_{ij})$ (equation 3)

The details of how these measurement errors can be approximated are outlined in the section on "Measurement errors and confidence intervals". At this stage it suffices to mention that the measurement errors used by FAO are based on information inherent to the metadata of each variable. Overall, the term "measurement error" refers to the degree of confidence in a particular element of the FBS. It should also be noted that this approach does not change official data as official data are assumed to have a measurement error of zero.

The availability of expected values and measurement errors does not guarantee a balanced identity of supply and demand. If measurement errors are too small, and if over stringent bounds are placed on too many elements, the identity may not be solvable. Conversely, if the bounds are large, the identity will have multiple or even infinitely many solutions. If many solutions are possible, the next step is to single out the best possible solution.

In the new balancing mechanism, the best solution is defined as the one that provides the combination of variables with the highest aggregate "probability". In more theoretical terms, this means that an objective function must be established that selects the combination of variables with the highest aggregate likelihood. The objective function needs to maximize the joint density of all distributions. In practice, the approach amounts to maximum likelihood estimation – a process that selects the combination in which the product of all likelihoods has a maximum value:

$$Obj_i(x_1, x_2, \dots, x_n) = \prod_{j=1}^{n} f_j(x_{i,j}) \text{ (equation 4)}$$

where Obj_i is the objective function for the ith country, f_j is the density for the jth element of the balance (production, imports, etc.), and x_j is the chosen value for this jth element. The x_j are chosen to maximize the objective function, and hence can be seen as the most likely values of the elements given the provided distributions.

If we allow the distributions of each of the different variables to be any arbitrary distribution, then optimizing the objective function (equation 4) must be done numerically. Numerical optimization can solve difficult problems such as this one, but it has several drawbacks: it can be time-consuming, it can fail to converge to a solution, it can find a local rather than a global optimum, etc. However, if all of the distributions above are assumed to be normal distributions, optimization becomes very simple:

1. Calculate the total imbalance (Imb_{ij}) by computing:
$$Imb_{ij} = P_{ij} + I_{ij} - X_{ij} - dSt_{ij} - Fo_{ij} - Fe_{ij} - Lo_{ij} - Se_{ij} - IU_{ij} - T_{ij} - ROU_{ij}$$
2. Adjust each element based on its standard deviation:
$$P_{ij}(\text{balanced}) = P_{ij} - \frac{Imb_{ij}}{e_{ij}(P_{ij})}$$
$$I_{ij}(\text{balanced}) = I_{ij} - \frac{Imb_{ij}}{e_{ij}(I_{ij})}$$
$$X_{ij}(\text{balanced}) = X_{ij} + \frac{Imb_{ij}}{e_{ij}(X_{ij})}$$
$$dSt_{ij}(\text{balanced}) = dSt_{ij} + \frac{Imb_{ij}}{e_{ij}(dSt_{ij})}$$

and so on, where the sign is positive for variables that have negative signs in the calculation of the imbalance, and vice versa.

3. These equations do not allow the enforcing of any bounds on the variables. If a bound is not met, the variable should be assigned the value of the bound, given a standard deviation of 0, and steps 1 and 2 should be repeated. For example, if the lower bound for feed is 100 tonnes but the feed variable is assigned 50 tonnes, it is fixed to the value of 100 tonnes, and the difference (50 tonnes) is reallocated to the other variables.

Constraints

The maximization/optimization process takes place under constraints. In principle, three types of constraints are possible. The first and most obvious constraint holds for every line of the balance and simply imposes that supply is equal to utilization for every commodity. It also includes the condition that P + I – dSt > X. A second constraint can be imposed to reflect prior distributions over all rows of the food balance – a column constraint. It has been argued, for instance, that year-to-year swings in the sum of all calories in the FBS (DES) may not exceed the value of 100 kcal; larger swings may appear from a mechanical implementation of the FBS framework, but are unlikely to occur in practice, or do so only under special circumstances (massive economic problems or natural disasters). A third constraint arises when there is the desire to impose a multi-year constraint for an individual variable. For instance, stock changes in one direction are unlikely to occur over many years in a row, at least not for large aggregate quantities. In addition, the feed use estimation process imposes the constraint that the sum of the respective nutritive values (energy and protein content) of any likely combination of commodities satisfies the biological requirements of livestock, poultry and aquaculture as estimated in the feed use module. For details see section 3.4.

Possible problems and infeasibilities

The proposed objective function maximizes the overall probability under multiple constraints. An optimization process that takes place under multiple constraints can lead to infeasibilities; the occurrence of infeasibilities will increase as the number and stringency of the constraints increase. Such infeasibilities occur at different stages, and the stages at which an infeasibility occurs helps to determine the strategy for overcoming it.

Different types of constraint and infeasibility can be distinguished:

- *No solution with only a row constraint of supply equal to demand:* If supply cannot be matched by utilization for any given row/commodity, the imbalance is larger than the aggregate measurement error and no combination of variables within their respective measurement errors would allow a balanced identity to be reached. Obviously, this means that the bounds on the distributions are too stringent to establish the balance. In this case, one or several bounds need to be relaxed to ensure that a combination of the various variables can be found to balance supply and demand. The need to respect official data may mean that only one or a few bounds can be enlarged or lifted.
- *Solution at the row level with infeasibility for column constraints:* Once a solution for every row has been established, a problem with the column constraints may be encountered. For example, the year-to-year change in the DES may be larger than the reasonable threshold, so the elasticities in the food model will need to be adjusted and the optimization procedure re-run to meet column constraints. In the best case, a solution can be found with all constraints imposed. Failing this, the column constraints need to be reviewed and eventually lifted. Alternatively, single-row limits on food availability may need to be tightened.

- *Solution at the row and column levels with infeasibility for multi-year constraints:* The strategy is usually straightforward; it involves lifting the limits of stock changes over time. Eventually all constraints should be met. If not, all other constraints will need to be lifted gradually to eventually render a solution. Lifting these additional constraints requires manual intervention and judgement as to what constraint should be considered most binding.
- *All constraints met:* In this case, the FBS can be considered ready for inspection by a food specialist and refined where and as necessary.

Figure 1 captures the basic set-up of expected values and associated distributions. It is based on the wheat balance for the United States of America.

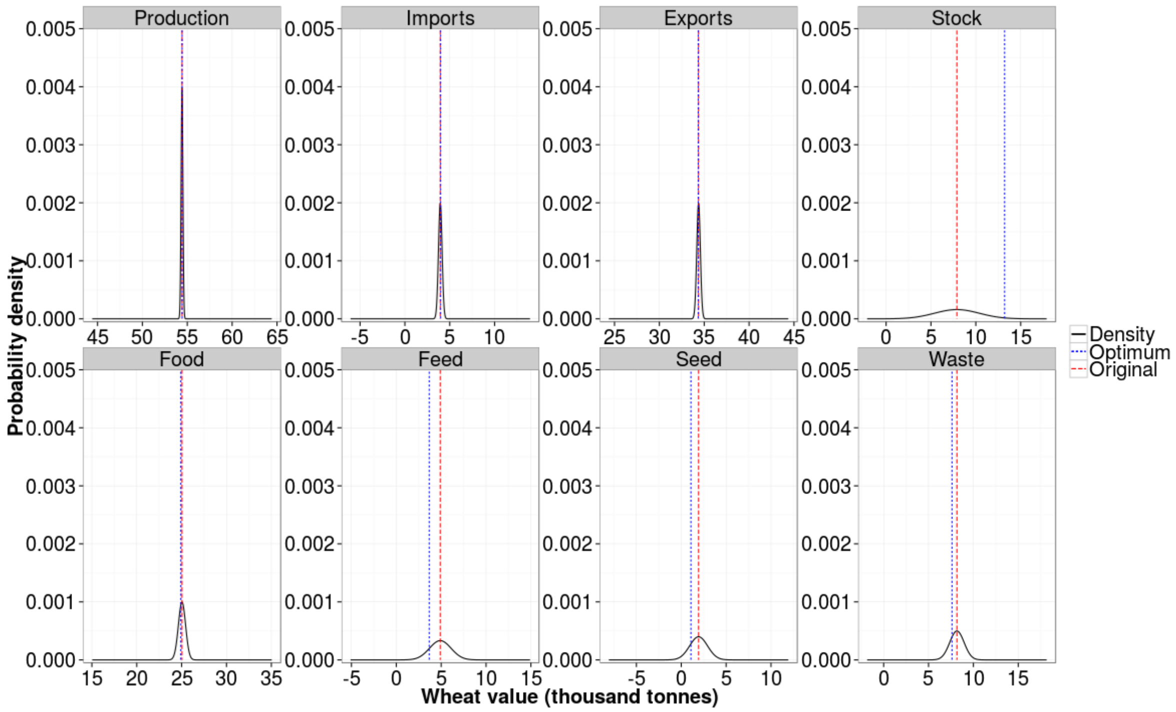

Figure 1: United States wheat, unbalanced (upper level) and balanced (lower level)

In the top pane, the expected value is represented by the dashed line (and the highest point of the distribution). In the bottom pane, the dashed line represents the value after balancing.

Advantages and disadvantages

This radical departure from traditional balancing provides numerous practical advantages and theoretical consistencies. However, it also imposes greater computational burdens. The main advantages and disadvantages are summarized in the following.

Advantages:

- *No residual necessary:* The problem facing all traditional approaches is that the need to identify a residual means not only that this variable is the balancing item of the identity, but also that it takes all measurement errors of the balance. As there is no a priori reason to assume that the sum of the measurement errors is equal to zero, the balancing item is in essence the least reliable estimate of the balance. The inherent year-to-year fluctuations in food consumption are likely to be huge so the reliability and plausibility of this variable as the balancing item would be questionable. Such large fluctuations would not meet the a priori expectations that food consumption moves in a rather smooth manner over time and that other elements, notably stocks, feed use or waste, vary more significantly.

- *All information harnessed:* When available, information about the quality of every variable is retained in the creation of the balance. For instance, if some variables of a balance are measured by high-quality surveys, they will be maintained with their exact values; conversely, variables that come without prior knowledge of accuracy, or a low degree of confidence, will be allowed to vary widely in the balancing process.
- *Rules-based and reproducible process:* The identification of every variable follows a rule-based process, as does the optimization process. This process allows decisions to be traced and results reproduced.
- *Most likely outcome:* The overall solution is a most likely outcome, based on a probability-maximizing algorithm, rather than one of many (notionally infinite) possible solutions of an underidentified equation.
- *Flexibility:* At any stage and for every variable, both the expected value and the confidence interval can be overwritten by expert knowledge. The optimization algorithm then runs within the bounds provided by experts.
- *Multiple constraints:* Apart from the constraint that the balance has to be in equilibrium for every row, the approach allows constraints to be imposed across column sums and time. For instance, a constraint can be introduced to avoid large year-to-year fluctuations in overall calorie availability (DES). Such a constraint reflects prior knowledge about people's desire and ability to maintain a stable calorie intake level, at least within the constraint. Limits on stock changes can also be introduced to reflect multi-year accumulations or drawdowns.

Disadvantages:

- *Need to formalize knowledge of prior information.*
- *Statistically complex procedure; data- and information technology-intensive.*
- *Multiple constraints:* Imposing multiple constraints is both an advantage and a disadvantage. Such constraints can result in infeasibilities in the solution, or lead to border solutions, i.e. solutions along the constraints.
- *Additional effort required for each element*: Expected values and error distributions must be provided for each element, and this requires more advanced modelling than just estimating expected values.

Measurement errors and confidence intervals[27]

The proposed balancing mechanism requires the FBS compiler to provide an estimate for a confidence interval around an initial estimate of a variable. These confidence intervals reflect the degree of reliability of the initial estimate. FBS compilers can refer to various sources and employ strategies to arrive at quantitative estimates for these reliability measures; they may be able to use information about measurement errors that was collected with the primary data (e.g. a household survey), assign confidence ranges based on the imputation method or, in the absence of specific information, simply assign confidence ranges in line with their prior knowledge of how an estimate

[27] The detailed methodology for estimating the mechanics of the balancing method in general and the confidence for individual variables in particular is available as a supplement to this resource book.

was produced, reflecting the approach used, the quality of the statistical system in general or for a particular survey, etc.

Given the need to assign such confidence values in the absence of country-specific information, FAO assigns confidence intervals based on two readily available factors.

The FAO metadata/flags system

ESS collects information from various sources and assigns different reliability levels to these sources. These reliability levels are encapsulated in the FAO "flags" system. The flags are used to approximate confidence ranges as a first step in estimating confidence levels.

Five levels of confidence are distinguished as default values for arriving at confidence intervals around an expected value. The highest confidence is placed on official numbers/estimates, which would be taken as point estimates without a confidence interval; the lowest confidence is placed on estimated data. All confidence estimates are further modified by an allowance for the quality of a country's statistical system, which results in a hierarchy based on the quality of the statistical systems nested in the confidence intervals arising from the general flags. The highest trust/narrowest confidence intervals are assigned to data from countries with excellent statistical systems, while the lowest confidence/widest intervals are assigned to imputed or missing data from countries with very weak statistical systems. An example, categorizing these levels of confidence, is provided in Table 3.

Table 3: *Confidence levels based on FAOSTAT flags*

Source	*Confidence*	*Implied* **measurement** *error*
Official	1.0	0.0%
Semi-official	0.9	10%
Imputed	0.85	15%
Estimated	0.6	40%
Missing	NA	NA

Quality of countries' statistical systems

To capture cross-country differences for the same flag, information from the assessment of statistical systems from the Global Strategy to Improve Agricultural and Rural Statistics was used to fine-tune the levels of confidence and thus determine the final confidence interval for every estimate. Most of the information was gleaned from the country assessment questionnaires processed by the Global Strategy, although these do not provide global coverage.

FBS compilers at the country level may want to use the FAO confidence intervals as a first approximation of the quality of estimates. Country-level FBS compilers are likely to have more, and more reliable, information about the country's statistical system in general and about individual FBS variables in particular. They are therefore strongly encouraged to review the FAO confidence intervals carefully, compare them with specific information available from their statistical system and their own experience, and replace the FAO intervals where appropriate.

2.2 Standardization

FBS and SUAs

In general,[28] FBS present all variables of a supply–utilization balance in their primary equivalents. For instance, in the balance for wheat, all elements are expressed in wheat as a primary product, while the only readily available variables that apply to primary forms are production and, in principle, seed use of wheat. In many cases, information in terms of primary equivalents is not available, or only partially available. For instance, imports or exports of wheat take place in the form of wheat, but also in the forms of different wheat products such as flour (first level of processing), bread or pastry (second level of processing), or even more processed forms.

Some variables of the balance may be available only for the product's processed forms. For example, food of wheat exists only in the form of flour, or flour products such as bread, noodles, pastry or biscuits. Wheat is practically never eaten in its primary form; the same holds for all other cereals and for many other primary products.

[28] A few products are specified as processed products, such as butter and vegetable oils.

As all variables of the balance (apart from production) refer to forms other than the primary form, they all need to be converted back to primary equivalents. This allows comparisons among the various variables, and eventually their union in a balance – only when all elements are expressed in a common denominator can they be added up in the FBS balance. To this end, ESS has developed a process known as "standardization", which is analogous to the process of creating a common denominator that allows processed products to be added up and expressed in their primary product equivalents. The different processing steps in the food chain create many processed products, which all need to be "rolled up" into their primary equivalents. Supply and use of these processed products are also put into balances, referred to as "supply utilization accounts" (SUAs) by ESS. Even if the naming convention reflects tradition rather than statistical or economic rationale, it has been maintained here for the sake of simplicity and continuity.

The new methodology adopts the standardization method of the previous method in principle. It largely maintains existing processing streams and thus the same commodity tree structure. However, the standardization programmes have been rewritten in the R language and problems have been addressed in this process. Important changes have also been made to the parameters that link processing levels, notably the extraction rates. The same holds for other parameters, such as nutrient conversion factors and shares of uses (splitting the utilization side). Chapter 4 describes the standardization process with concrete examples.

Extraction rates: Extraction rates reflect the amount of primary product contained in the next level of processed product; for instance, the extraction rate of wheat flour is about 0.79. This means that 1 tonne of wheat going through a country's flour milling industry renders an average of 790 kg of flour. Flour milling also produces bran, on average 180 kg per tonne of wheat, and wheat germ, about 20 kg. The remaining 10 kg are losses that occur in the milling process. A detailed account of changes in the extraction rates used is provided in a separate document; here, it suffices to mention that all extraction rates are now exogenously assumed. In the past, some extraction rates were implicit (endogenous), calculated from the available information on processed production (e.g. flour) and an allocation of primary products for processing (e.g. wheat). Most scenarios of endogenous extraction rates occur in developed countries and a limited number of products, and this process often created implicit extraction rates with highly unlikely values. For example, the extraction rates for Japan's cereal sector were unduly low, suggesting that the country's flour milling industry was processing wheat inefficiently. In reality, the opposite was true. The low implicit extraction rate arises not from a low estimate of flour produced in Japan, but from a high allocation of wheat going into the country's flour milling industry. The new approach is therefore to maintain flour production as the starting point, apply a reasonable extraction rate and calculate primary food (wheat) based on flour production and the given extraction rate. The added advantage of this process is that it simplifies the standardization method and makes it possible to avoid a complicated *ex-post* reallocation process among processed products.

Nutrient conversion factors: FBS provide an account of not only quantities of food supplied and utilized, but also key nutrients such as calories, protein and fat. Other nutrients can be added. As food products are typically eaten in processed rather than primary form, nutrient conversion factors also apply to and are available for processed products. The new approach typically applies calorie conversion factors at the first level of processing. For wheat, this means that the calorie, fat and protein conversion is made at the flour level. The nutrient conversion for biscuits is at the sugar, flour and vegetable oil levels, after the biscuits have been broken down according to the split shares of these items.

Product specificity: Products of the same name and classification but from different provenances may have different nutrient contents and food values. For instance, grass-fed beef from Argentina is typically less caloric than maize-fed beef from the United States of America. The same holds for many fruits and vegetables, other meats and even cereals from different countries or farming systems. There is growing recognition of the need to account for these differences in the standardization process and to apply more specific calorie conversion factors. Such factors should also be studied over time and adjusted when necessary. There are plans to take these specificities into account in future versions of the SUA/FBS system. In the meantime, while there is awareness of the problem, the specificities are limited to region-specific calorie conversion factors.

Shares for different uses: In the current approach, the utilization side is constructed by applying fixed shares of availability to feed, waste, food manufacture and, in part, other variables. The new approach seeks to determine all forms of utilization by the factors that drive them, rather than applying simple shares of availability. This means that the shares used in the previous methodology (e.g. for feed or waste) are no longer required. The shares of food that enter a manufacturing process outside the normal commodity trees are a notable exception. Such shares can be based on historical allocations in the process, if available, or on global average allocations.

Shares for food manufacture: For some FBS commodity trees, the underlying processing activities are rather diverse, resulting in many links among trees. These commodities are pruned off the main trees and crafted into the FBS as a separate balance. As separate branches, these commodities re-enter the balances as processed rather than primary products. They include products such as sweeteners, beer and other alcoholic beverages. During food manufacturing and processing, these products often receive inputs from several FBS commodities. For example, beer is produced from many different starchy inputs, such as barley, maize, wheat or sorghum, but also from more exotic products such as bananas or plantains. To avoid double counting, beers with different main inputs need to be identified as separate products in the FBS, and are not included in the original commodity tree. All other products are standardized up the commodity tree in which they were produced.

2.3 SUAs and processed products

Every primary product is the basis for a broad variety of processed products. A modern food processing sector is characterized by a vast variety of different products; the further down the processing chain, the more possible combinations and varieties of food products there are. For example, information from the USDA database on food nutrients (ARS SR 27)[29] suggests that from the ten basic commodity groups (cereals, oilseeds, etc.) and the 60 primary FBS commodities, the United States food industry produces more than 8 000 different food products. These products can be consumed in the United States of America, but are also exported to destinations abroad. It is unlikely that all 8 000 processed food products have genuinely different food compositions; many are simply different brand names with the same or similar combinations of first-level food processing products. However, the many different combinations suggest that the variety of food products is large and that many more than 60 products are eaten and traded. Trade information is available for about 3 000 products. About 300 of these are genuinely distinguishable in terms of their composition, and need to be converted into their primary equivalents.

[29] http://www.ars.usda.gov/Services/docs.htm?docid=8964

2.4 SUAs and commodity trees

Standardization – the conversion and subsequent aggregation of processed products into their primary equivalents – aims to bring these products into a hierarchical order that reflects the various food processing chains in which primary products are converted into processed equivalents. The most straightforward way of depicting this hierarchy is in the form of commodity trees. The primary products, such as wheat, represents the stem of such a tree. The main components of first-level processing (flour, bran and germ) form the main branches, which split into ever-finer twigs at higher levels of processing (bread, rolls or dextrose). The hierarchy among various levels of processing reflects the various steps of processing; the different levels and products are connected through the extraction rates (see Figure 2 and other examples in chapters 4 and 6).

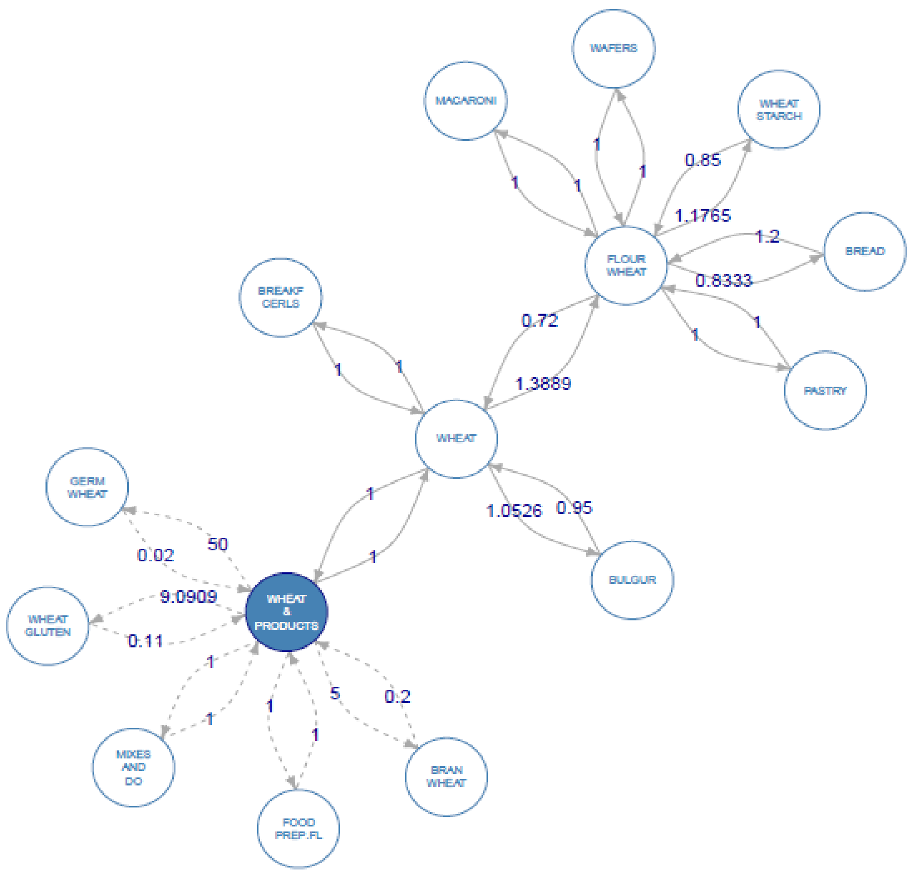

Figure 2: Standardization of wheat processing flows

A single tree cannot reflect the ways in which modern food processing connects trees by combining the processed products of one tree with those of others (e.g. refined sugar with flour and refined vegetable oils), or by combining processed with primary products, such as milk and eggs with flour and vegetable oils for other bakery products. This means that the trees are connected with one another through processed food products. The trees also provide connections at lower levels of processing; for example, wheat or maize germ is connected to the balance of vegetable oils, while bran, the other first-level by-product of flour milling, is connected to the feed balance (see section 3.4).

A review of existing standardization programs revealed a number of issues, including programming errors, breaks in the structure of trees, hard-coded and undocumented exceptions, or – as outlined in section 2.2 – inappropriate conversion factors. These problems become most visible when standardization leads to obviously wrong results, most notably negative utilization. Figure 3 provides an example of how wrong standardization can result in negative utilization. The revision of the FBS methodology identified as many of these cases as possible and addressed them by using appropriate shares, identical commodity trees for different variables/elements of the balance, and more reasonable extraction rates. These rates should be periodically reviewed and updated to ensure they remain valid.

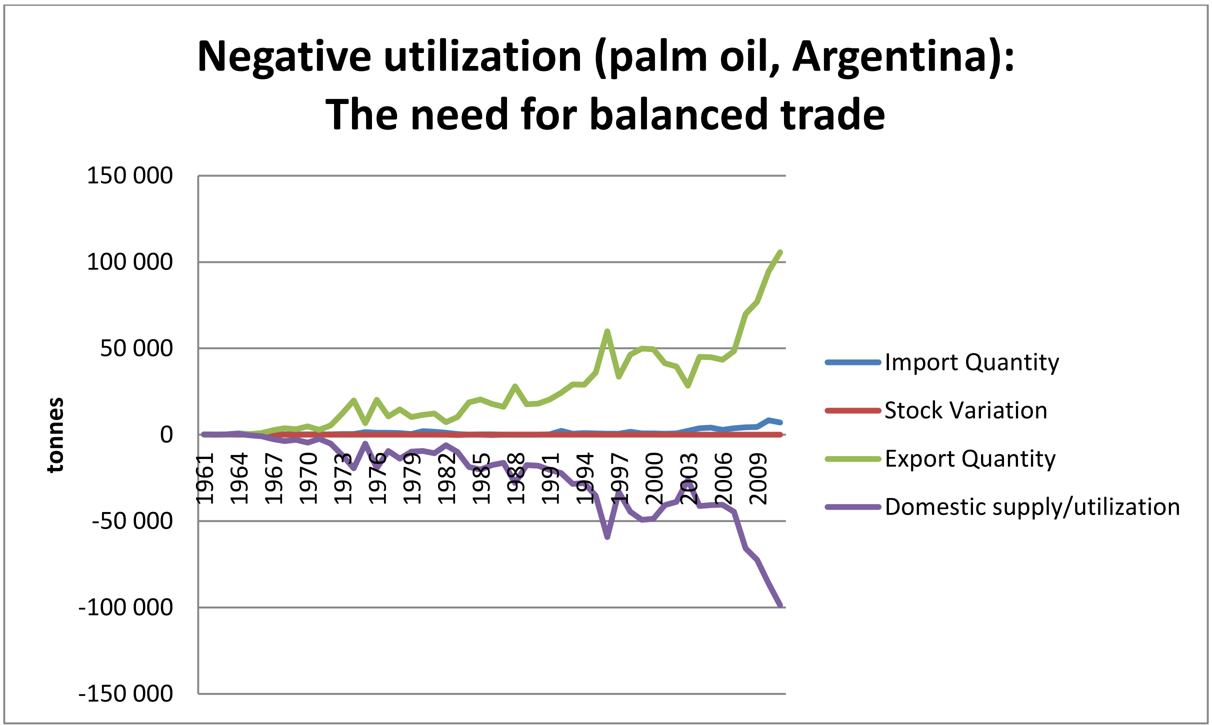

Figure 3: Standardization and negative utilization

2.5 Standardization and trade

Notionally, all commodities in the SUAs are to be standardized, i.e. all underlying processed products are to be converted back into their primary equivalents. For instance, if wheat is stored as flour, the quantities of wheat would need to be converted back into their wheat equivalents for measuring stocks. Where data are available, this is being done, but most information about the use and supply of processed products regards imports and exports.

3. Imputation methods

3.1 Production

The FAO Statistics Division (ESS) collects data on agricultural production via an annual questionnaire, which is dispatched to more than 180 countries, including the Statistical Office of the European Communities (EUROSTAT) and each European Union (EU) member country. Not all countries return the questionnaire. Some countries fail to compile it but make all their data available on the Internet sites of the national statistical office or ministry of agriculture. Others compile the questionnaire, but leave large gaps, usually implying that they have not collected the data.[30] Lack of data provided in a questionnaire reflects a lack of available data at the country level, including for setting up the country's own FBS; such countries will face the same problems as FAO in completing the SUAs and FBS.

It cannot be stressed often enough that the recommended approach for filling data gaps is to undertake surveys and collect conversion factors, extraction rates and other coefficients. FBS compilers should make every effort to obtain empirically measured information. No imputation method, no matter how sophisticated, can replace measurement. This also applies to all the imputation methods presented in this source book: these methods cannot be surrogates for surveys, censuses and measurement in general. Only in the absence of empirically collected/measured information should missing data, conversion factors, etc. be imputed. However, although imputation cannot replace measurement, there are wide differences in the suitability of imputation methods. Good imputation methods are not only necessary to generate missing information, they can also be used to triangulate measured information and discover inconsistencies. For example, the feed use imputation method offers several options for consistency checks and quality control (see subsection on Feed use data in section 1.3).

To improve availability and comparability, data are also collected from sources such as official national publications (general and agricultural yearbooks, monthly bulletins), international databases (EUROSTAT, OECD), or FAO and UN reports of missions to countries. Recently, ESS organized workshops in five continents to present the FAO data collection process and solicit countries' help in enhancing overall data reliability and timeliness and the coherence of data series.

The classification scheme used is the UN's Central Product Classification (CPC) expanded for agricultural statistics.

The imputation procedure

Before embarking on data imputation processes, FBS compilers are encouraged to subject their data to basic quality checks. These checks include a considerable range of operations such as vetting and adjusting units of measurement, weeding out outliers, identifying transcription errors and filling

[30] Most of these countries are characterized by weak statistical systems. The Global Strategy to Improve Agricultural and Rural Statistics provides a detailed account of the state of statistical systems around the world. Surveys suggest that most of the countries with weak agricultural systems are also particularly exposed to food security issues.

obvious gaps. A general concept for improving data quality is laid out in FAO's Statistical Quality Assurance Framework (FAO, 2014), which is also available to all FBS compilers at the country level.

Once the quality assurance/quality control has been undertaken, the outliers weeded out and data gaps filled with information available from alternative sources,[31] the remaining data gaps can be filled through imputation procedures that harness all the information currently available. To impute missing production estimates, a new method has been developed, which combines a robust econometric approach with expert knowledge or the experience of FBS compilers at the country level.

The new imputation method not only harnesses all available data and knowledge, it also incorporates experience with imputation methods that were used previously or that are employed by other institutions. Following examination of these imputation methods, the basic insight is that no single approach/model on its own is able to render satisfactory results across all countries and commodities. Based on this insight, an "ensemble learning model" has been developed, tested and adjusted to the needs of FAO's production domain. The tests showed that a strength of the ensemble model is its successful handling of the wide heterogeneity across countries, and FBS compilers at the country level may also find the same or similar approaches particularly useful for the imputation challenges facing their individual countries.

A detailed description of the methodology is available in a separate document. The presentation here is limited to a few main features. First, imputation of yield in any given country incorporates available information on that commodity from any or all other countries. This enables maximization of information usage and improved stability of the imputation. Second, rather than relying on a single method or model, the imputation process consists of a dozen candidate models, which are averaged in the ensemble, with weights assigned in accordance with the predictability of each model. Third, an ensemble can cope well with sparse data availability. For instance, when only a single observation is available, this value can be carried across many years and be assigned to missing observations. On the other hand, if abundant data are available, more weight will be given to sophisticated models that can capture the complex pattern of the data.

The performance of the new imputation method has been tested and reviewed; the qualitative assessment suggests that the new method is superior to previously employed methods and to the approaches taken by other institutions.

[31] In addition to the Web sites of national statistics offices and ministries of agriculture, ESS also harvests data from authoritative sources, including those that specialize in one commodity, such as Oil World, the International Sugar Organization, the International Coffee Council and the International Cotton Advisory Committee.

The new imputation approach in detail

There are four main steps to the imputation procedure:

1. productivity imputation;
2. imputation by balancing via the productivity = production/activity identity;
3. production imputation;
4. imputation by balancing via the productivity = production/activity identity.

Step 1 is skipped if current productivity data are available. If balancing is possible, imputation will be done via balancing. Production will be imputed if production data are still not available. If one variable remains absent, it will be imputed in the final step.

Step 1: Productivity imputation

Productivity imputation is performed via the ensemble approach alluded to above. Currently ten models are implemented in the ensemble framework. A separate model is fitted to each country and commodity pair, unless otherwise noted:

1. Mean: This model computes a simple mean of all observations and imputes with this value.
2. Linear: A linear regression model is fitted to the available data, and predictions from this model are used to estimate missing values:
$$P_{i,j,t} = \beta_{0,i,j} + \beta_{1,i,j}\, t + \varepsilon_{i,j,t} \text{ (equation 5)}$$
3. Exponential: A regression model is fitted, but the exponential of time is used:
$$P_{i,j,t} = \beta_{0,i,j} + \beta_{1,i,j}\, e^{t} + \varepsilon_{i,j,t} \text{ (equation 6)}$$
4. Logistic: A non-linear regression model is fitted to the data:
$$P_{i,j,t} = A_{i,j} + \frac{B_{i,j}}{1+exp(-C_{i,j}(t-D_{i,j}))} + \varepsilon_{i,j,t} \text{ (equation 7)}$$
If the non-linear regression model fails to converge, then $A_{i,j}$ is assumed to be 0 and a new non-linear model is fitted to the data. If that model also fails to converge to a solution, then $B_{i,j}$ is assumed to be the largest observed value. In this case, a logistic regression (which will always converge) is possible by performing a logit transformation, and this model is then used.
5. Naive: Missing values between two observed values are interpolated linearly. For observations outside the range of the observed data, the nearest observation is carried forward or backward.
6. Autoregressive integrated moving average (ARIMA): Several ARIMA time series models are fitted to the data, and the best one is selected based on the small sample size-corrected Akaike information criterion (AICC). New values are imputed from this model via Kalman filter smoothing.
7. Local regression (LOESS): A local regression model is fitted using linear models as the base models. The model window varies based on the sample size, so more flexible models are fitted when more data are available.
8. Splines: A cubic spline is fitted to the observed values and used for interpolation.
9. Multivariate adaptive regression splines (MARS): A MARS, which is a mathematical model appropriate for piece-wise linear time series, is fitted to the observed values.
10. Mixed model: This model is fitted to all the countries at once, but is restricted to one commodity at a time. The linear mixed model estimates production as a smoothed function of time, and the time trend is considered a random effect by country.

Each of the models is fitted to the available data, and an estimate of the model's ability to explain the data is produced via cross-validation. In other words, each model is fitted to all the available data except for a handful of observations. The performance of an individual model is then evaluated by its ability to estimate these known observations. This process is repeated with different groups of observations so as to provide a good estimate of a model's ability to fit the data, which is captured in the form of an error term, $e_{i,j,k}$, depending on the country i, the commodity j, and the model k. Then, each model is assigned a weight according to its ability to fit the observed data (with better models receiving larger weights), subject to the constraint that the weights sum to one:

$$w_{i,j,k} = (1/e_{i,j,k}) \,/\, \left(\textstyle\sum_{l=1}^{10} 1/e_{ijl}\right) \text{ (equation 8)}$$

A missing observation is imputed using a weighted mean of the estimates of the ensemble models at the missing time. An additional constraint is imposed: models that may not extrapolate well (e.g. an exponential fit) are removed from the weighted average for missing values that lie too far outside the range of observed data. Note that in this case, all weights must be adjusted to ensure that they still sum to one.

Step 2: Imputation by balancing

Once productivity has been imputed, the balancing starts. If both productivity and activity values are available, production is imputed so that it satisfies the production = productivity * activity identity. If productivity and production are available, the activity is imputed by the activity = production/productivity identity.

Step 3: Production imputation

For imputation of production, the procedure outlined in step 1 is followed, but is instead applied to the production data, while values are imputed using a new ensemble fitted to the production data currently available.
Examples on how the imputation process works and the results it renders is available in the example of the FBS for wheat in section 4.1.

3.2 Trade

Country-level data

National FBS compilers may want to start collecting trade data for their countries from their own sources, which will usually be the national customs office. In most countries, trade information will be available in Harmonized System (HS) format and at the level of disaggregation needed to compile the SUAs – the precursors to the FBS. Taking information from the national customs office will also ensure that trade data are available in a timely manner and probably more frequently, such as monthly rather than the trade data from international sources such as UNSD and FAO. However, there are also distinct advantages in using data from international sources, particularly FAO. The main advantage is that FAO subjects trade data to extensive plausibility checks and validation processes, and provides a more complete country-level and global trade data set using information from the trade data of the country (e.g. by calculating trade values from reported trade quantities applying unit values, or vice versa) or its trading partners ("mirroring" flows). Details of how FAO compiles and validates trade data and the variants of trade files that are made available are given later in this section.

International trade sources

All the trade data used and compiled by ESS for all purposes (FBS and others) come from the United Nations Commodity Trade Statistics Database (UN COMTRADE). UN COMTRADE provides access to standardized data from its Web site.[32] It also offers a detailed description of the methodology used to compile trade data.[33] Its guidelines on the compilation of international merchandise trade statistics are particularly important. As details of all methodological aspects are readily available from UN COMTRADE, this resource book does not repeat or summarize any of them. All the aspects presented in this section refer to the work that is undertaken by ESS once all the necessary data have been downloaded from UN COMTRADE database.

UN COMTRADE data are available at different levels of disaggregation, reflecting the internationally recognized standard HS classification, which is maintained by WCO and updated every five years.[34] The latest version is HS 2012; preparatory work for the 2017 is under way. FAO has made numerous contributions to improvement and extension of the HS to include essential data and reflect changing patterns in agricultural and food commodities, and is involved in the development of HS 2017.

The HS is a hierarchically structured classification of products that distinguishes different levels of disaggregation. The first six digits of the HS are common to all countries; lower levels (which can go up to 12 digits) can be developed in national versions. To capture as much detail as possible, trade data are harvested at the lowest level of disaggregation. These data are then mapped to the same classification system that is applied to production data and all forms of utilization (the CPC expanded for agricultural statistics). In this mapping process, a unit conversion must also be performed (see Table 4). The six-digit level is used to perform various consistency and validation checks. For details of compatibility between the HS and the CPC system see chapter 6.

Table 4: Comtrade to CPC conversion factors

Comtrade unit	CPC unit	Conversion factor
1	N/A	N/A
$ value only	N/A	N/A
m^2	mt	1
l	mt	0.001
u	heads	1
u	1000 heads	0.001
u	number	1
kg	mt	0.001

32 http://comtrade.un.org/

33 HTTP://UNSTATS.UN.ORG/UNSD/TRADE/METHODOLOGY%20IMTS.HTM; http://unstats.un.org/unsd/trade/methodology%20imts.htm

34 HTTP://WWW.WCOOMD.ORG/EN/TOPICS/NOMENCLATURE/OVERVIEW/WHAT-IS-THE-HARMONIZED-SYSTEM.ASPX

m^3	mt	1
carat	mt	0.000005

Processing UN COMTRADE data

ESS processes UN COMTRADE data in different ways and for different purposes. The main purpose is to generate specialized trade information (Table 5), but other needs include the FBS/SUAs and many other data sets such as trade of fertilizers and pesticides, machinery, etc.

Data processing involves several steps that lead to different products. The various steps and products are described in detail in the methodology paper for the compilation of trade statistics. Here, it suffices to list the various steps. The resulting products are summarized in Table 5.

Processing steps

i. Direct harvesting of data from UN COMTRADE via an application processing interface. For details see trade documentation.
ii. Data validation and error correction.
iii. Calculation of representative import and export unit values, using information from reporters and/or partners, and – based on these representative values – calculation of missing quantities from available values and missing values from available quantities.
iv. "Mirroring", i.e. completing the trade matrix with information from trading partners/reports wherever available.
v. Calculation of total trade flows (total imports and exports for every commodity at the tariff line level) that are still imbalanced.
vi. Harvesting and/or manual input of data from specialized commodity reports.
vii. Conversion of HS codes into CPC and FCL codes.

Up to this point, all imputation methods keep officially reported data intact (other than weeding out obvious reporting errors and filling in data gaps). All imputations supplement existing data, but do not replace them. However, the trade information that best serves FBS needs is as complete, as balanced (imports versus exports for each commodity) and as consistent as possible at the global level. Obtaining such information requires not only imputing missing data from non-reporting countries, but also adjusting the trade matrix to ensure that global imports are as equal to global exports as realistically/statistically possible. However, balancing the matrix may involve overwriting official data, so requires a separate process. This process involves applying a system of endorsement factors to reflect the reliability in the reporting of trade data. Again, the details of the methodology are described in documentation for the new trade methodology. Here it suffices to list the main steps:

i. Calculation of endorsement factors.
ii. Balancing of trade so that $\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_{ij} = \sum_{i=1}^{n} I_i \sum_{i=1}^{n} I_{ij}$ for every commodity j and all countries i.
iii. Calculation of total trade (total imports and exports for every commodity at the CPC level), of balanced flows.

Figure 4 provides an overview of the new trade processing flow.

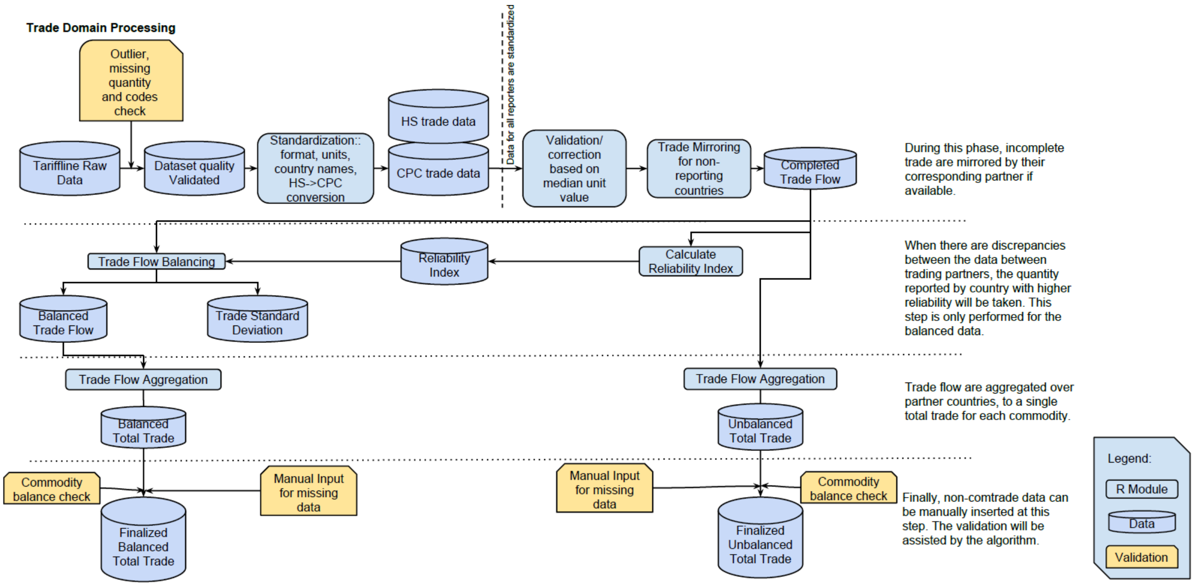

Figure 4: Trade processing work flows

Trade data products

The processing of trade data at FAO ultimately renders six trade data sets, all of which are available through FAOSTAT. They comprise three types of trade matrices, and another three data sets with total trade only. The various products are summarized in Table 5.

Table 5: Trade data products

Products for dissemination	**1. *UN* COMTRADE**	**2. *Mirrored flows***	**3. *Balanced trade***
Matrixes of flows	Cleaned for errors and misreporting	Cleaned for errors and misreporting	Cleaned for errors and misreporting
	No data overwritten	Data gaps filled through mirroring	Data gaps filled through mirroring
	No data added	IUV/EUV calculated and data gaps filled through imputation (quantity/value and value/quantity)	IUV/EUV calculated and data gaps filled through imputation (quantity/value and value/quantity)
		Official data kept intact	Official data overwritten, where necessary
			Balanced trade, endorsement factors
Total trade	**4. Imbalanced official**	**5. Mirrored and imputed, but still imbalanced**	**6. Balanced trade**
	Total exports and imports are reported by country, but	Total exports and imports are reported by country, but	Total exports and imports are reported by country, and
	$\sum_i X_i \neq \sum_i I_i$	$\sum_i X_i \neq \sum_i I_i$	$\sum_i X_i = \sum_i I_i$

EUV = export unit value; IUV = import unit value.

Calculating export and import unit values

The overview provided in Table 5 identifies export and import unit values as the key factors for imputing missing quantities from existing value flows or, much less frequently, missing value flows from existing quantity flows. The imputation process is straightforward; the choice of the most appropriate unit value is not. In theory, there are competing alternatives, but their advantages and disadvantages can only be evaluated empirically/practically and depend on the reliability of reported trade flows, the number of products lumped together into a CPC category, or the number of available trade flows from reporters/partners.

The right choice is therefore selected empirically. Exposing various theoretical alternatives to practical tests resulted in the following procedure. Where the number of trade flows for a given product is low, typically five or fewer, a global average (median) is taken as the basis for calculating the unit values. Conversely, where the number of flows is high, country-specific unit values are identified to calculate unit values. Figure 5 depicts the country-specific case.

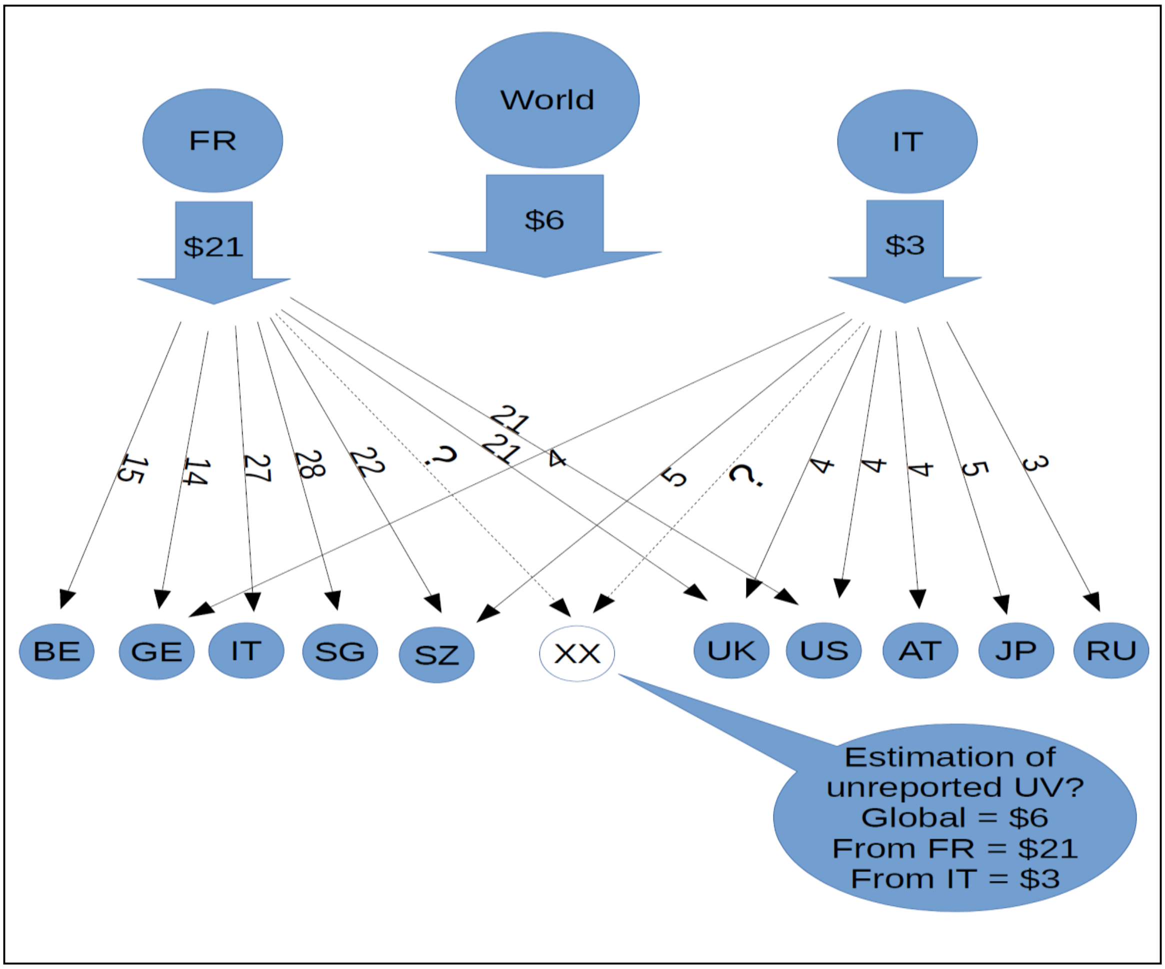

Figure 5: Calculation of unit values based on information from trading partners

In the case captured in Figure 5, country XX imports wine from France but, without quantity or value information, is assumed to pay the same price as all other countries importing wine from France – $21/bottle on average. Likewise, if the same country XX imports wine from Italy, the underlying unit value is based on the one that other countries pay on average when importing wine from Italy – $3/bottle.

When enough flows are available, such a country-specific approach renders a more precise outcome than working with a global average, which would have been $12/bottle. The example provided reflects the fact that France typically exports expensive wine and is therefore assumed also to export expensive wine to country XX, where information is missing. It also captures the fact that Italy typically exports “cheap” wine and is thus assumed to sell the same quality to country XX, where information is missing.

Path dependency of imputation routines

As outlined earlier in this section, the process from “raw” trade data to published trade results encompasses several steps. In general, these steps follow a logical and cogent sequence. For instance, data are cleaned and errors weeded out before other processes such as imputation or balancing can take place. However, there is also a rationale for allowing flexibility in the sequence of the steps, and not all steps follow a cogent and unequivocal logic. There is also no a priori (conceptual) reason for carrying out one step before another in the sequence of implementation, such as undertaking mirroring after the unit value imputations, or vice versa.

Intuitively, FBS compilers may be inclined first to mirror missing data to provide a broader basis for the unit value calculations. As outlined in the subsection on unit values, such a decision would have direct implications on the choice of unit value, with more observations leading to a decision as to whether country-specific or global unit values are employed in the imputation process. This intuition is likely to be confirmed when the mirrored data for a country are at about the same level as the average unit value of a reporter/partner. However, if a mirrored flow is far out of the normal range, it will affect and even distort the unit value used in the next step of the process and, as these unit values will be applied to all flows with missing information, many imputed data would be affected.

Without prior knowledge of the correct sequence of steps, the final results become path-dependent on the sequence of imputation steps. To gauge the magnitude of the potential effect, a number of empirical tests have been conducted. The results suggest that in general, the differences that arise from different imputation sequences are negligibly small. However, there are a few commodity-specific exceptions. Particularly where flows are characterized by large differences in product quality across reporting countries, poor product differentiation in reporting, etc., the path selected can influence the results substantially. Overall, however, such commodity-specific differences were not deemed sufficiently important to deviate from the proposed sequence of data processing steps outlined in Table 5.

Reliability index and endorsement factors

Official data cannot be overwritten arbitrarily. Overwriting requires a rules-based and well-informed process. In the literature, this process is based on the reliability of the reported trade data. In principle, the same approach is applied in this resource book, but with important modifications.

Trade flow data can be used to quantify the reliability of a country’s reported trade, particularly where reported values agree with the reported values of trading partners. However, this approach can be complicated. For example, if country A reports different values from those of two of its trading partners B and C, should it be penalized equally for both disagreements? If B is a highly reliable reporter while C is a highly unreliable reporter, disagreement with B should be penalized more heavily than disagreement with C.

The proposed solution to this problem is as follows:

1. Compute the level of agreement between each pair of countries as the proportion of trade observations on which they agree (within some tolerance).
2. Initialize all countries to the same reliability score.
3. Each country will now transfer its reliability score to all its partners in proportion to the level of agreement with these partners. For example, suppose country A has a reliability score of 1 and trades with countries B and C only. Suppose also that A and B agree 90 percent of the time while A and C agree only 30 percent of the time. A will transfer 90/(30 + 90) percent = 75 percent of its score to country B and 25 percent to country C.
4. Step 3 is repeated until the change in the reliability scores becomes negligible.

This algorithm assigns larger reliability scores to countries that tend to agree with their partners, and the score also accounts for the reliability of the partner country. This algorithm is known to converge to a solution rapidly, and is also used in many other applications for measuring quality within a network (such as Google's PageRank algorithm for determining the quality of Web sites).

Example of reliability scores:

Reliability index calculations are depicted in **Error! Reference source not found.**. The nodes circles) represent different countries, and the thicknesses of the edges indicate the levels of agreement between corresponding countries (with thicker edges indicating stronger agreement). In the first step of the algorithm (left-side plot) all countries have an equal reliability score. After one iteration of the algorithm (middle plot), the top nodes/countries have higher than average reliability scores and the bottom two countries/nodes have below-average scores. After the second iteration (right-side plot), the algorithm approximately converges: the top three countries all have high reliability scores (as they all agree with each other). Country 4 has a slightly higher score than country 5 because countries 4 and 2 agree on some trades. However, countries 4 and 5 both receive low reliability scores, even though they agree between themselves.

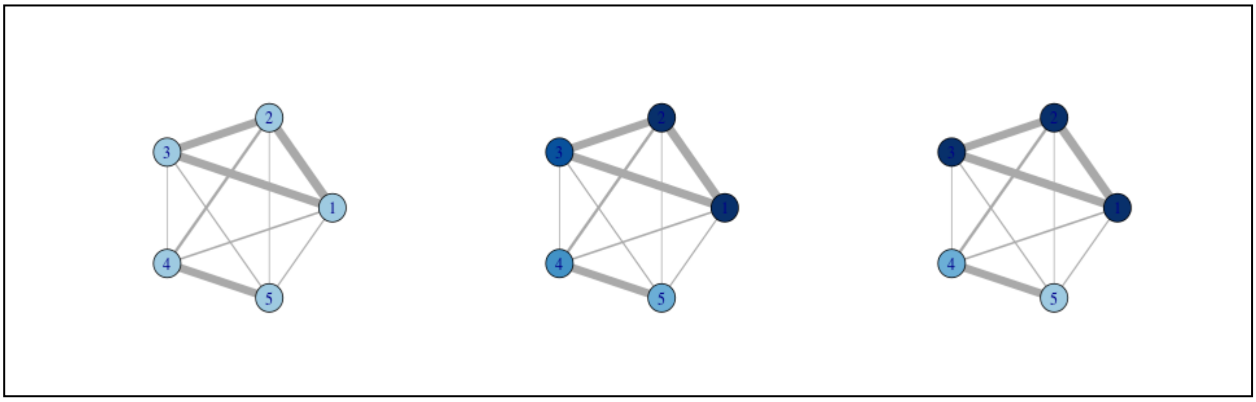

Figure 6: Reliability index calculation over two iterations

3.3 Stocks and stock changes

It has been emphasized that there is no alternative to systematic measurement. In particular, given the special importance of stocks for food security in general and for price stability in particular, there is no alternative to accurately measured stocks, at least in the long run. Several initiatives are under way to improve measured stock estimates, including under the Agricultural Market Information System (AMIS) and the Global Strategy to Improve Agricultural and Rural Statistics. Under the auspices of this strategy, the ten-year Agricultural and Rural Integrated Survey (AGRIS) programme has been developed, which not only lays the foundations for creating an efficient, overall agricultural statistical system, but also offers explicit options for estimating levels of stocks for different agricultural commodities. As AGRIS estimates can provide stocks in conjunction with several other variables of the FBS (production, feed, seed, etc.), it promises to provide particularly relevant estimates of stocks. Countries are therefore encouraged to make all efforts to measure stocks, at least of their most important food staples, either through a specialized survey or, preferably, as part of an integrated survey system such as AGRIS.

The low level of measurement at the country level is directly reflected in the availability of stock data for FAO. The return rates for stock estimates are so low that these estimates are currently not included in any FAOSTAT domains. No separate domain brings such estimates into one place (e.g. analogous to production or trade), neither is there the systematic inclusion of stocks in any other data domain. The only domain where stocks appear is in the commodity balances, of which the FBS are a subset. However, commodity balances do not include levels of stocks, but only year-to-year changes in stocks.

As stocks, or at least stock changes, are an integral part of every supply–demand balance, estimates of stock changes have been included in the FBS system. In many instances, stock changes function as a balancing item.

However, using stocks as a balancing item implies (as it does for any other element that is used as a balancing item) that all measurement errors are relegated to stocks (or the other element chosen as a balancing item). It also means that stock changes no longer only capture changes in stocks, but also function as a catch-all for all measurement errors, and would therefore better be referred to as stock changes and residual uses. This is an undesirable outcome for any element of the balance as it seriously diminishes the element's value as a statistical indicator; given the importance of stock for price volatility analysis, etc., it would be particularly serious. In addition, stock changes would "inherit" errors from previous years, resulting in steadily increasing distortions over time.

Imputation/estimation

The need to move away from a residual approach poses the challenge of identifying an alternative method for generating an estimate, or rather an expected value and a distribution. If empirical stock estimates are available (e.g. the United States of America undertakes a biannual survey of its cereal stocks), they enter the balance as an observed value, ideally with a measured distribution. Clearly, this is the best solution and should be encouraged for as many countries and commodities as possible.

If no information about stock changes is available, a distribution needs to be assumed. However, even assuming a uniform distribution (shown in **Error! Reference source not found.**) is problematic as it mplies assuming that a stock change near the assumed minimum is just as likely as a stock change close to 0; however, a stock change just outside the boundaries has zero probability. Moreover, uniform distributions can be problematic at the balancing stage, as such distributions essentially become residuals. Thus, it is very important to construct some distribution for stocks, even if it is a very wide distribution to indicate that there is much variability in the estimate.

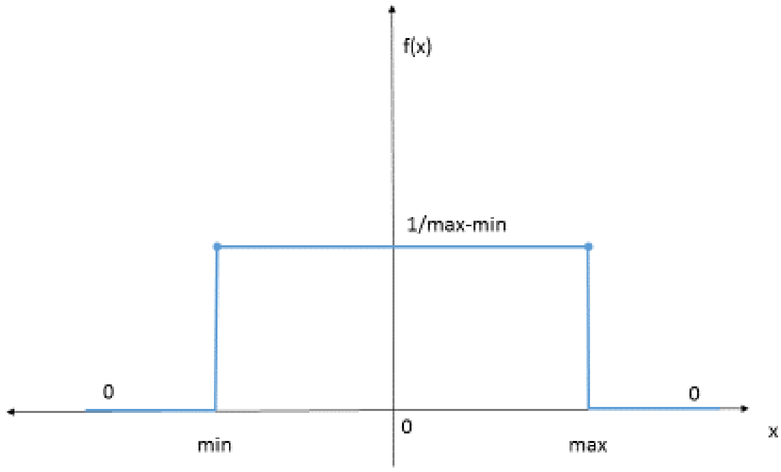

Figure 7: Bounded uniform distribution with a zero mean

The strategy here is to harness additional and readily available information inherent in stock holding practice and economics. Such information is available from, for example, knowledge about the costs of stock accumulation and reduction over time. This prior information can be harnessed to enable a move away from a uniform distribution in which every stock level has the same probability between a maximum (max) and a minimum (min), to an approach that makes some levels of stocks more likely than others, still within the maximum and minimum limits.

Prior information about the economics and dynamics of storage can be used to derive both an expected value and a distribution:

a. *Expected value:* The expected value for the mean stock change should be zero in the long run; empirical information fully confirms this a priori expectation, and Figure 8 captures it for wheat stocks in the United States of America. The reasons for the long-run mean reversion to zero are obvious. Any long-run positive deviation from zero would amount to accumulation of stocks, which would rise the longer such positive shifts prevail and the higher the positive values are. Conversely, successive negative deviations would amount to a permanent drawdown of stocks, and thus be tantamount to an eventual stock-out, or imply unlikely (very costly) high initial levels of stocks. The reasons for a non-zero, short-term stock change lie in its ability to smooth fluctuations in consumption. The desire to keep consumption stable means that stocks function as a short-term buffer to smooth surpluses and deficits, i.e. demand for stocks is – at least at high levels of stocks – very price-elastic while consumption is not.
b. *Distribution:* The above implies that positive and negative stock changes are likely to be symmetrically distributed around the zero mean. The analysis for United States wheat and maize suggests that the empirical distribution could be approximated by a normal distribution with a zero mean (see Figure 10 and Figure 11) in which Shapiro-Wilk metrics are highly

significant). This is already an improvement beyond assuming a uniform distribution for stock changes.

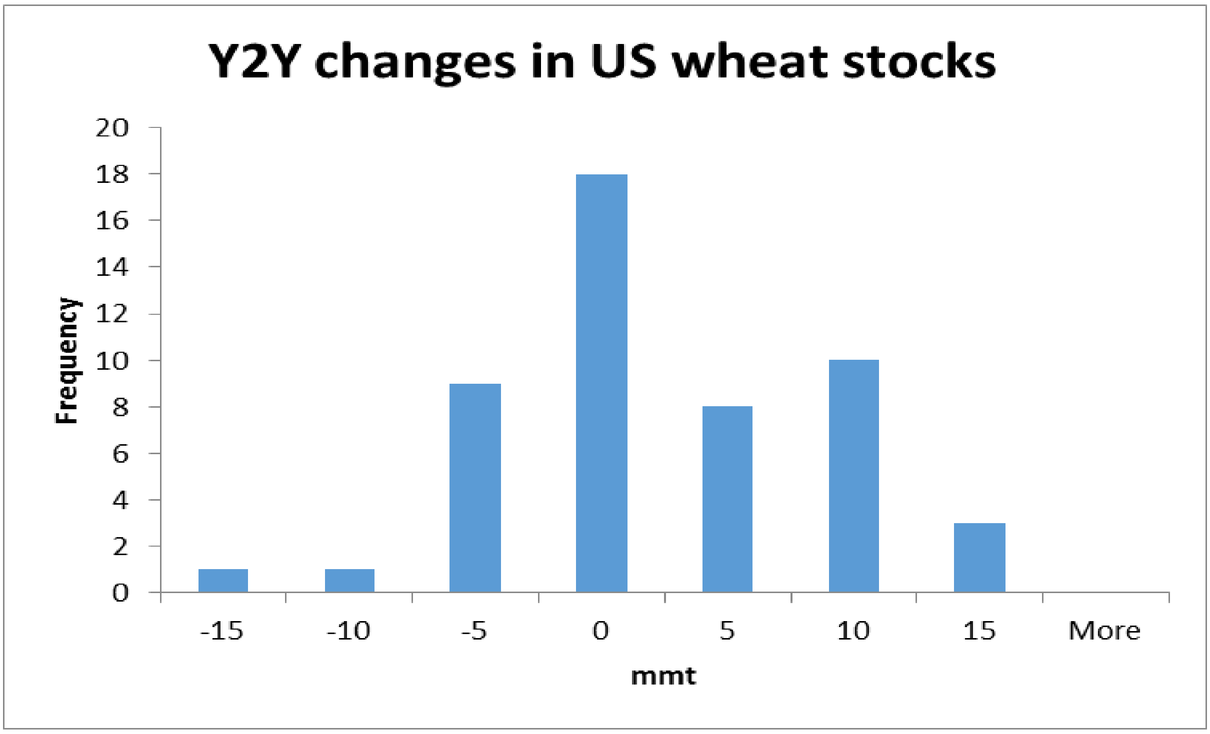

Figure 8: Distribution of year-to-year changes in United States wheat stocks

United States

Estimating stock changes in time t

The analysis presented so far suggests that stock changes are likely to be normally distributed around a zero mean. However, stock changes are unlikely to be independent over time: stock changes in previous years are likely to influence stock change in the current year. Thus, the stock change in the current year may not have an expected value of zero, but rather some amount that depends on previous stock changes.

Information can be harnessed to gauge both the likely direction and the likely amount of stock changes in time t. In fact, a positive stock change in t becomes increasingly likely the longer the preceding period of stock drawdowns and the higher the amounts of drawdowns – the larger the cumulative drawdown over time. If cereal stocks are drawn down for 15 years in a row, the likelihood of drawdowns in successive years becomes increasingly small.[35]

Conversely, if a country accumulates stocks for many years in a row, the cumulative amount of stocks would become so high that storage capacity dwindles, losses loom and the costs of holding stocks become prohibitively high. This means that the probability of a positive change in t increases the longer the history and larger the cumulative amounts of negative changes in the past, and vice versa.

[35] Such a development may lead to (near) stock-out and would be accompanied by a price spike (at least at the global level).

Thus, there is need to construct a model for stock changes that is informed by knowledge of historical changes. One such model would be:

$$\Delta S_t = \beta\left(\sum_{i=1}^{k} \Delta S_{t-i}\right) + \varepsilon_t \text{ (equation 9)}$$

In other words, it is assumed that the stock changes at time t depend on the sum of the previous k stock changes. An error term is added to indicate that the change in stock at time t is not exactly equal to this value; rather, the distribution of the stock change at time t is adjusted given the knowledge of previous stock changes, and still has variability. Alternatively, this could be expressed as:

$$\Delta S_t \sim N\left(\beta \sum_{i=1}^{k} \Delta S_{t-i}, \sigma^2\right) \text{ (equation 10)}$$

This provides a further improvement on the distribution of stock changes. A normal distribution can be assumed and the mean can be adjusted based on previous stock changes. This has the effect of reducing the spread of the distribution, as more is known about what the stock change at time t should be.

Returning to United States cereal data enables examination of whether this hypothesis holds true. Figure 9 shows the relationship between the sum of previous stock changes over 15 years and the current stock change. It shows that there is a generally negative trend, thus validating the original hypothesis. The trends may not be extremely strong or significant, but they allow a distribution for the stock change to be inferred given previous cumulative changes. For example, if it is known that over the previous 15 years the net stock change for "Barley and products" was 2 500, it can be expected that the next stock change should be normally distributed with a mean of -400 and a variability that allows values between -1 200 and 400.

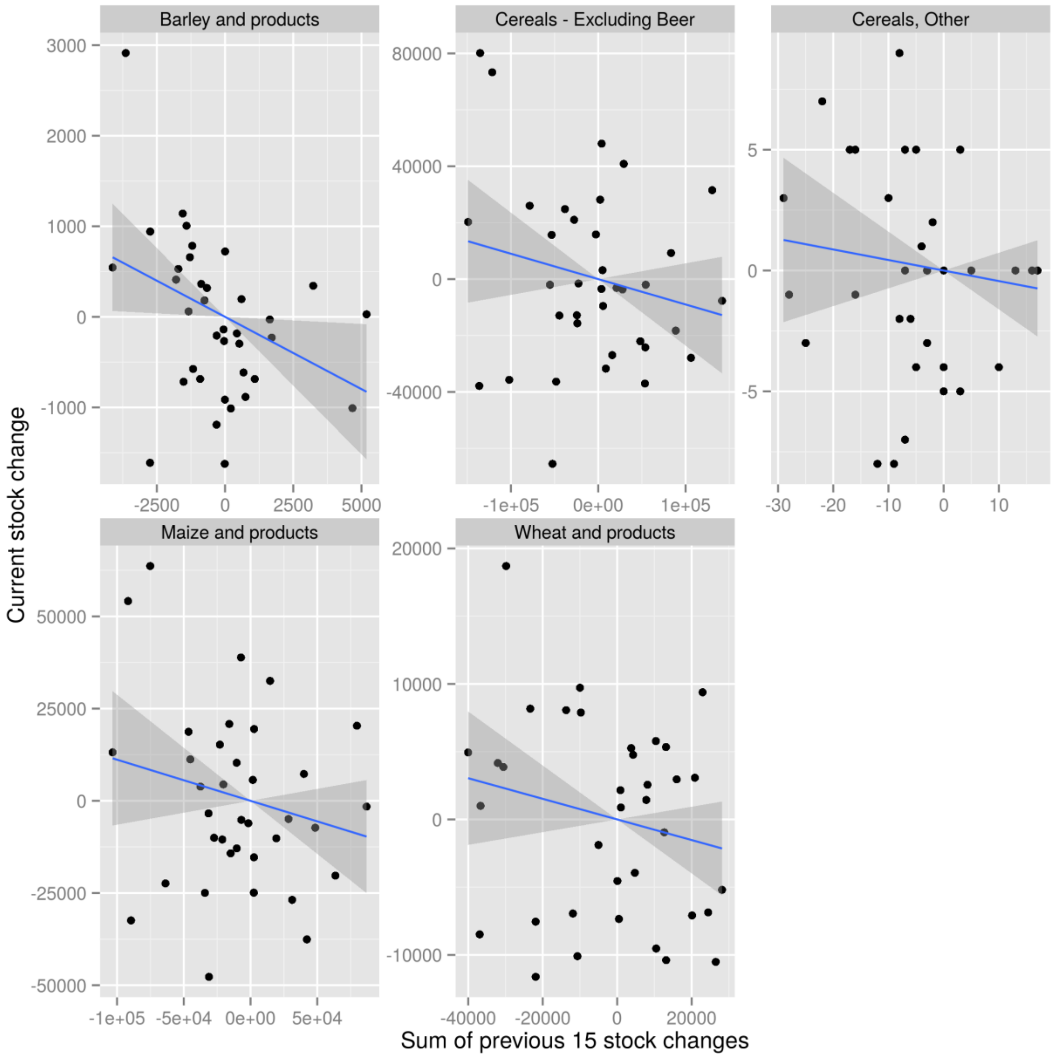

Figure 9: Stock changes in t in relation to cumulative past stock changes

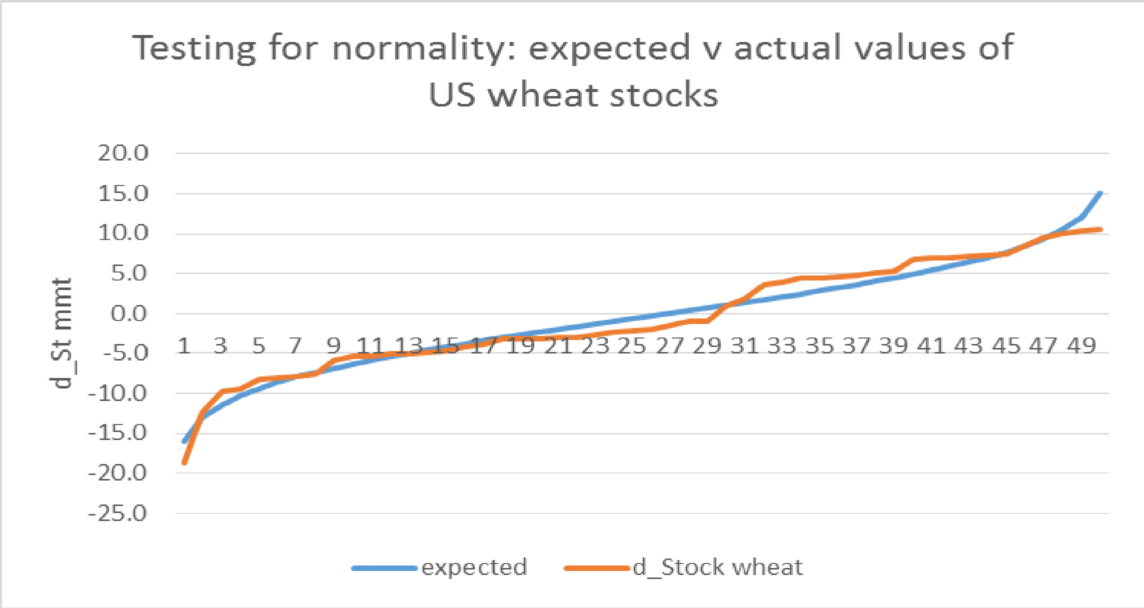

Figure 10: Testing for normality in United States wheat stocks

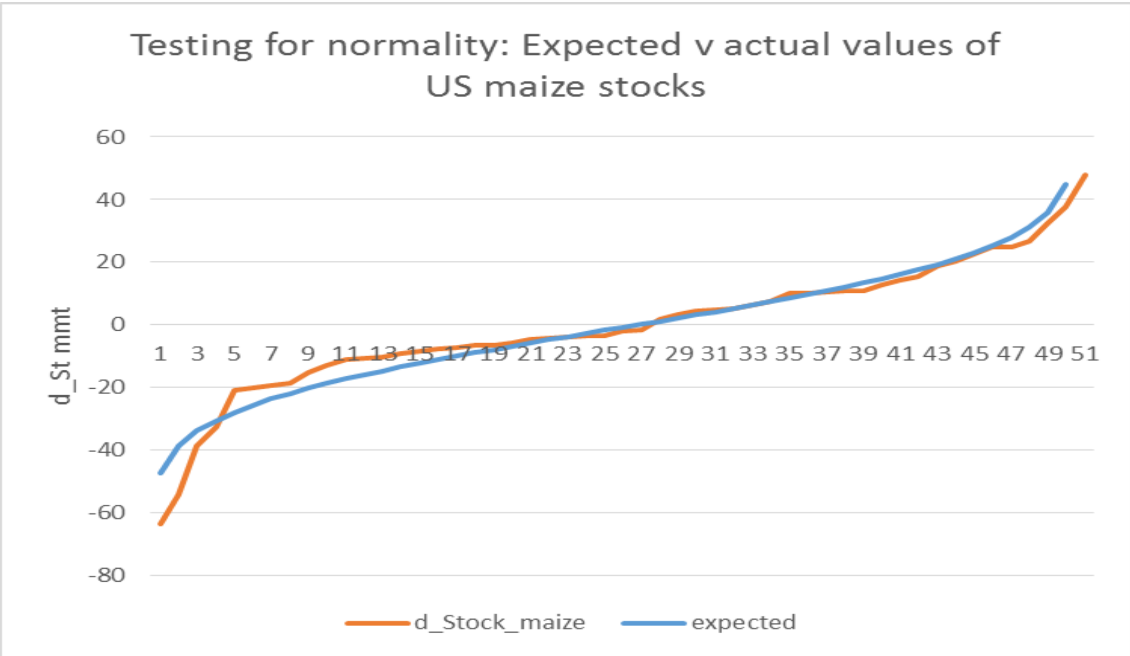

Figure 11: Testing for normality in United States maize stocks

3.4 Feed use

Only a few feed surveys available to FAO are representative at the country level and cover feed production by both farmers and feed compounders. Even in the United States commodity balances, feed serves as the balancing item and is accordingly called "Feed and residual use". When feed demand is measured, feed surveys are often limited to industrial feed sectors and are based on data from commercial feed companies only. For example, China, Germany and Spain obtain their feed estimates through surveys of feed companies. Such data are of limited use to FBS. Very few countries other than Hungary and the Netherlands have surveys that also include farm production of feedstuffs.

This lack of measurement appears even more surprising given that feed use is a major source of disappearance in the FBS and is of rising importance. Growing consumption of meat, milk and other livestock products, and ever-more intensive production systems have resulted in rapidly rising demand for feed products, notably compound and concentrate feeds (CC feeds). Indirect but compelling evidence stems from the rapidly rising production of commodities that are exclusively, or at least mainly, destined for feed use, notably oilmeals and certain coarse grains.

Moreover, shifts to higher consumption levels of livestock products and to more intensive feeding systems have not been limited to developed countries. Available evidence, including from the production of livestock products, suggests that the same trends now prevail in developing countries, particularly in many emerging economies (Brazil, China).

The absence of measured estimates, combined with the rapidly growing importance of feed use means that any methodology for imputing/estimating feed use requires particular attention. Above all, it must capture intensifying livestock production systems and the rising overall importance, volume and diversity of livestock products. In addition, short-term shocks such as epidemic outbreaks of diseases or other mortality-inducing events must be captured in feed use estimates. The imputation method should also capture the rapid technical progress in rearing animals in intensive livestock systems and, as part of this progress, the growing efficiency with which feedstuffs are used in modern systems. Not all of these trends drive feed use in the same direction. For instance, information available from extension services and modern livestock operations suggests an offsetting effect of growing feed intensity, with ever-higher shares of CC in total feed rations on the one hand and steady progress in using feedstuffs more efficiently on the other, thereby reducing the quantities of CC feed needed to produce a given amount of livestock product. The dynamics of these changes should be reflected in an FBS imputation system.

The previous imputation methodology

Before conceiving any new methodology, it is useful to review and appraise the existing approach typically used to estimate feed demand, assessing whether it adequately represents the current and prospective characteristics of feed use and livestock systems. This review can help in deciding whether: i) the old system can be used with minor changes; ii) parts of the old system can be salvaged; or iii) an entirely new system needs to be developed.

The basic approach currently used in the FBS system is easily stated. Feed use is simply calculated as a share of feed availability:

$$Fe_{i,j} = \left(P_{i,j} + I_{i,j} - X_{ij}\right) * \ r_{i,j} \qquad \text{(equation 11)}$$

where the feed use of feedstuff j in country i is determined as a ratio (r) of the availability of the feedstuff in that country. The share is specific to a given country and product, (largely) stable over time, and not specific to the different elements of availability, i.e. it does not distinguish between production and trade.

From an *ex-post* perspective, the rationale behind the choice of this approach is difficult to understand. It may have been motivated by the need to capture the dynamics of feed use in non-intensive feeding systems, in which feed use rises when supplies are ample and shrinks when feedstuffs are in short supply.

Being supply-oriented, the approach could reflect reality insofar as CC feed use rises in a non-intensive system after bumper crops, and contracts – often sharply – after crop failures. However, while possibly reflecting the livestock production systems of several decades ago (when it was conceived), the method is hardly representative of modern feeding systems, and seems to fail to capture the level of feed use even in the least developed countries. It was therefore decided to change the imputation approach radically, from a simple, supply-driven method to a more complex but tractable approach that correctly respects the demand-driven nature of feed use in most countries.

The new feed use estimation procedure

This subsection provides an introduction to the new feed use methodology. Being a radical departure from the previous approach, considerable attention is given to the main elements of the methodology; a more detailed description is available from the specialized documentation of the new approach to feed estimation.

The new feed use estimation process can be divided into three stages. In stage 1, a step-wise and demand-oriented approach is used to determine CC feed use. Total feed requirements ($R_{i,t}$) in country i and year t are calculated as a function of the herd size (N) and the species, country and time-specific requirement index ($r_{s,i,t}$, where the s index runs over all commodities).

Determining total feed requirements

The requirement indices are expressed in terms of an animal unit and incorporate maintenance needs (M) and output requirements (P), such as for meat, milk or wool, and the overall efficiency (E) over the entire herd[36] that is characteristic of the given feeding system:

$$R_{i,t} = \sum_{s=1}^{S} N_{s,i,t} * r_{s,i,t} \text{ (equation 12)}$$

where:

$$r_{s,i,t} = f(M_{s,i,t}, P_{s,i,t}, E_{s,i,t}) \text{(equation 13)}$$

and S is the total number of all commodities. The requirements are calculated for metabolizable energy (MJ ME) and tonnes of crude protein (t CP) via the requirements index, which is constructed

[36] For example, for pigs, the requirements are those of all “types” of pigs along the production chain, from piglets, to hogs and sows, etc. The feeding efficiency also refers to the arc efficiency across all types. This indicator necessarily lies above the better-known feeding efficiency ratios that refer to rearing hogs only – growing them from 30 kg to a slaughter weight of about 100 kg. Lower feeding ratios signify higher efficiency, as fewer kilograms of feedstuff are needed to produce 1 kg of livestock output, in this example, pig meat.

dualistically. Requirements are inclusive of all feedstuffs, not only CC feeds. In a second step, the shares of actual CC feed in total requirements are determined.

Determining actual use of CC feed

The first stage calculated total requirements based on the maintenance and output needs of a herd for all types of animals within the herd and across species. These needs are covered by both CC and other feedstuffs such as roughages, products from pastures or even table scraps. To determine the share of CC feeds, it is necessary to multiply the requirements (R_i) with an intensification factor (I_i). The intensification factor is a simple ratio that determines how much of the total requirements will be covered by CC feeds and how much by other feedstuffs.

This can be expressed as:

$$CC_{Feed} = R_i * I_i \qquad \text{(equation 14)}$$

Again, all calculations distinguish between metabolizable energy (MJ ME) and quantities of crude protein (t CP).

The intensity factors (I_i) vary across feeding systems and animal species (A_i). Across animal species, the main dividing line is between ruminants and monogastric animals (pigs and poultry). For monogastric animals kept in modern feeding systems, the intensity factors have been rising steadily over past decades and have reached values close to unity in developed countries, particularly for poultry, for which values have reached high levels in all countries; only organic agriculture leaves current ratios for poultry below unity. For pigs, intensification factors can still be well below unity where backyard production and non-intensive family farms account for a large share of output. This is also still the case for very large production systems such as pig production in China, but even for these systems, increasing intensification factors now feature prominently in most developing countries. Current data for intensification rates have been taken from the FAO GLEAM database (2005, 2010) data on animal feed baskets in countries, and extrapolated to capture the most recent situations.

The spectrum of intensification rates is much wider for ruminants than for monogastric animals, even in mature feeding systems. In all feeding systems, intensification rates of unity or close to unity cannot be found and are unlikely to occur in the future. Ruminants' physiological need of at least a minimum amount of crude fibre means that even the most intensive feeding systems (e.g. milk production in Israel or meat production in United States feeder lots) will have intensification rates of less than unity. In some systems at the other end of the spectrum, roughages remain the most economically efficient source (e.g. pasture in New Zealand), or the opportunity costs of CC are too high for feed use (e.g. for milk production in India). In these systems, the intensification rate I_i would be only marginally above zero.

Determining the requirements for fed/farmed fish:

The methodology presented so far has not taken into account the feeding of fish, as required in aquaculture. In fact, most available feed statistics do not take fish feeding into account. Given the increasing importance of farmed fish as a source of protein for human consumption, it is expected that the industry absorbs rather significant amounts of agricultural and marine outputs. In absolute terms, the most important users of fish feeds are in East and Southeast Asia, particularly China. As a share of total feed, aquaculture is particularly important in small island states such as Iceland and the Faroe Islands, where fed fish are often the main consumer of feedstuffs. Given the absolute and relative importance of farmed fish, a complete and correct assessment of feed use requires the inclusion of feeds fed to farmed fish in the estimation process.

Estimates of energy and crude protein needs for aquaculture are derived in a similar way to those for livestock. To arrive at a precise assessment, species-specific feed conversion ratios are directly applied to aquaculture production data (tonnes of fish and crustaceans) and converted into energy and crude protein equivalents to yield the total nutrient requirements of the produced amounts of fish.

Fish are also farmed at different intensity levels. On the one hand, salmon and carnivorous trout must be fed entirely on feeds from outside the pond, while other species, such as carp, may retrieve nutrients from microorganisms, aquatic plants, other fish and aquatic wastes that are available in their surrounding waters, depending on the production system. Hence, an intensity factor is applied to the overall biological requirements to circumscribe the actual aquafeed demand.

The parameters of both feed conversion ratio and feeding intensity are retrieved and extrapolated from survey data published by the FAO Fisheries and Aquaculture Department. Eventually, the aquaculture requirements are added to those of livestock to arrive at the total feed energy and crude protein use for every year and country.

Determining use of individual feedstuffs

So far, the new approach has produced only total feed requirements and total feed use in CC feeds in terms of metabolizable energy and crude protein. No distinction has yet been made to identify which of the various feedstuffs contribute to the CC feeds. The next stage in producing this information for the FBS is to break down the energy and protein contents of CC feeds into individual feedstuffs at the commodity level.

The most common approach for carrying out such an allocation is to set up a linear programming system that determines the most price-efficient combination of feedstuffs that meets the constraints of total requirements and the nutrient composition of individual feedstuffs. Such an approach may be very appealing in theory, but it lacks practical relevance for the FBS system. One problem is that it requires exact knowledge and constant updating of feedstuff prices. Collecting this information is challenging in developed countries and infeasible or impossible in developing ones. Another problem is that it would render an economically optimal but not necessarily an empirically likely solution, simply because actual use of CC feeds is determined by many factors beyond relative price quotations, as supply chain constraints and opportunity costs also need to be considered.

An alternative solution is to break down the total requirements into availability shares. This means allocating feedstuffs in accordance with their overall availability: feedstuffs that are available in abundance are used more than those that are scarce.[37] To allocate CC feeds to individual components, a two-stage process is employed. In the first step, all the CC feeds that can be used only (or at least in the vast majority of cases) as feeds are allocated. These feeds are oilmeals, oilcakes and all brans that are not used for human consumption (as breakfast cereals). By-products from biofuel production, particularly distiller's dried grains with solubles (DDGS), are assumed to be fed exclusively to animals. Their energy and protein content is deducted from the estimated overall feed use:

[37] This procedure retains the basic idea of the previous approach used in determining feed use. The decisive difference is that the availability shares are applied after requirements are determined, and intensification is respected.

$$\text{cereals, tubers, pulses etc.} = CC_F - (OM+OC+brans)$$

In a second step, all other feeds from commodities (cereals, tubers, pulses, etc.) in and outside the FBS are allocated as simple proportions of their availabilities.

As requirements are expressed in energy and protein units, two factors are constructed, one for the proportion of energy availability and one for the proportion of protein:

$$a^{n}_{f,i,t} = \frac{n_f q_{f,i,t}}{\sum_{f=1}^{F} q_{f,i,t} n_f} \quad \text{(equation 15)}$$

The proportion of feed availability (a) of feedstuff f in country i and year t for the respective nutritive conversion factor (n), which is either energy units or protein units per unit of feedstuff, is determined by its available nutritive quantity (q * n) divided by the total nutritive value provided by all available feedstuffs.

These factors can now be applied to the requirements and converted into quantities (tonnes). Given the different compositions of feedstuffs in terms of crude protein and energy content, two slightly different allocations are obtained: energy-rich feedstuffs will be more important contributors to overall energy availability than protein-rich feedstuffs, and vice versa.

The final quantities of feedstuffs must meet both the energy and protein requirements. Hence, it is possible that from the two allocations: i) one allocation meets both requirements; ii) no allocation meets both requirements; or iii) both allocations meet both requirements.

When one allocation meets both requirements, that allocation is preferred. When no allocation meets both requirements, upper and lower bounds for each feedstuff can be established as each allocation will meet its own demand. In other words, if the protein-based allocation meets the protein demand, and the energy-based allocation meets the energy demand, the combined maximum values of both in each feedstuff will at least satisfy, if not exceed, both requirements, while the sum of the minimum values of both feedstuffs will not satisfy any of the two requirements. An optimal solution that satisfies both demands therefore lies between these two extremes. To avoid allocating excessive feed, the final quantity of each feedstuff is obtained by optimizing the allocation linearly, such that both requirements are satisfied using the least amount of additional feedstuffs in accordance with their respective nutritive values. This means that preference is given to energy- and protein-rich feedstuffs after the minimum quantities have been allocated. When both allocations meet both requirements the same procedure is applied, but in the opposite way – quantities are subtracted to arrive at the minimum solution that meets both requirements. For instance, the final allocation of wheat for feed use is determined as a share of total availability and is retrieved as:

$$W_{Feed} = a_W R_{Fi} c_n \quad \text{(equation 16)}$$

where c is the respective conversion factor from either protein or energy into wheat quantities and depends on the derivation of a.

Results and consistency checks

Figure 12 depicts global energy and protein feed demands from 1992, grouped in the World Bank income level classification. The shift towards pig and poultry production in emerging economies becomes apparent when the protein demand of these countries exceeds the levels in high-income countries before the energy demand does. This reflects the fact that monogastric animals require relatively more protein than ruminants.

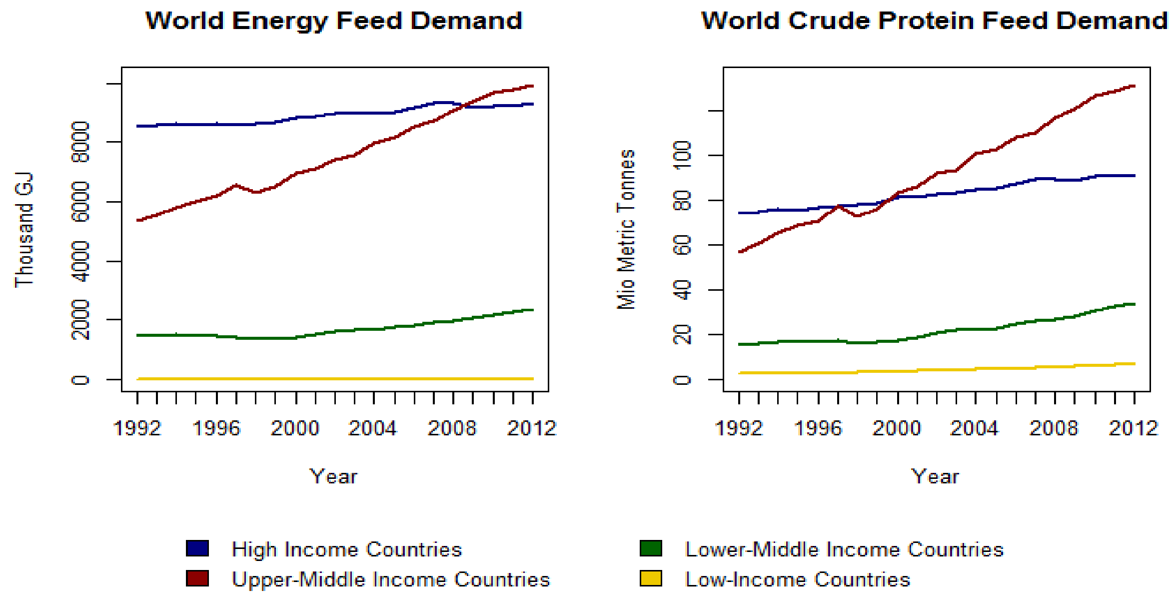

Figure 12: Feed estimates at the global level

Validation of feed estimates is a challenging task, not least because reliable feed statistics are seldom available. The first stage of results – feed requirements expressed in crude protein and metabolizable energy – are compared with reported feed figures. As some countries report on feed in the annual FAO questionnaire, these figures have been taken as the starting point for comparisons with the results of the new method. Figure 13 depicts this comparison.

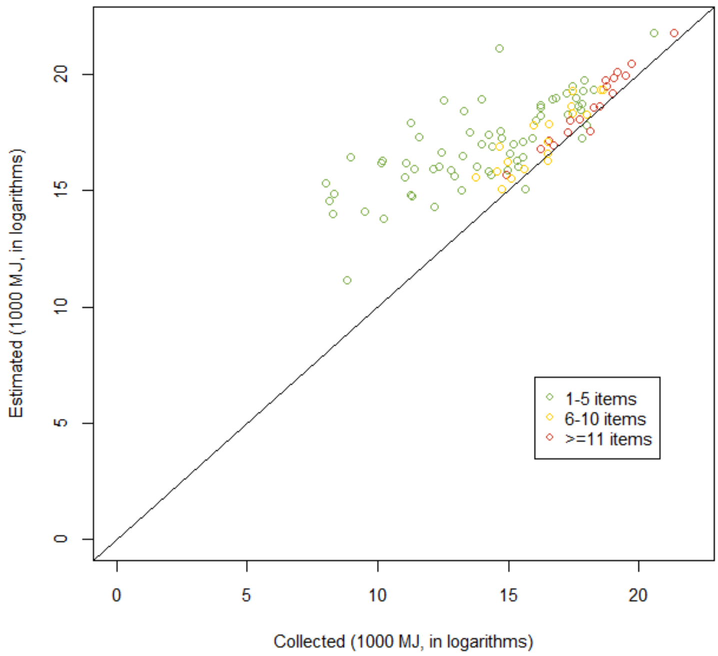

Figure 13: Estimated versus collected values for metabolizable energy

In Figure 13 each dot represents a pair of estimated feed demand and collected feed use in logarithms; the straight line represents scenarios in which collected energy feed fully coincides with estimated energy feed in a given country. It is worth noting that most countries report only sporadically on individual feeds, which means that the whole range of feedstuffs is seldom covered. As the information on feed use is not complete, reported feed should not exceed estimated feed. This implies that the dots in Figure 13 should never fall below the diagonal line, as occurs in only five cases (Azerbaijan, Cyprus, Egypt, Kazakhstan and Saudi Arabia). These cases may reflect either overreporting of feed or underestimation of demand.

The difference between reported feed and estimated feed can reflect feeds that have not been reported. From Figure 13 it is apparent that the more extensively a country reports on feed, the closer the estimate is to the collected aggregate. Assuming that countries covering more items in their reporting of feed are also more confident about their real feed usage levels, this finding provides strong evidence that the estimated requirements are not far from reality. The correlation coefficients between collected feed use and estimated demand increase steadily as the item coverage threshold is raised. Starting with 0.93 for an item coverage of greater than six, for which 36 observations are available, and reaching 0.99 for an item coverage of 20+.[38] In addition, the discrepancy between absolute feed required and feed used in these cases has shrunk to 13 percent, implying that the methodology for estimating feed demand is effective.

The next step is to check how the final feed estimates, after allocating particular feedstuffs, compare with other prominent data sources. Figure 14 depicts global cereal feed data from USDA and FAOSTAT (estimates derived under the old methodology) for 1992–2011. Apart from the newly

[38] In this round of estimation camels and horses have not yet been taken into account. Given the importance of these animals in these countries, this could well have led to underestimation in these particular cases.

modelled feed estimates, all other data sources use entirely supply-driven estimates, except for FAOSTAT figures that are officially reported by countries. However, reported figures are not modified by the new feed estimation procedure.

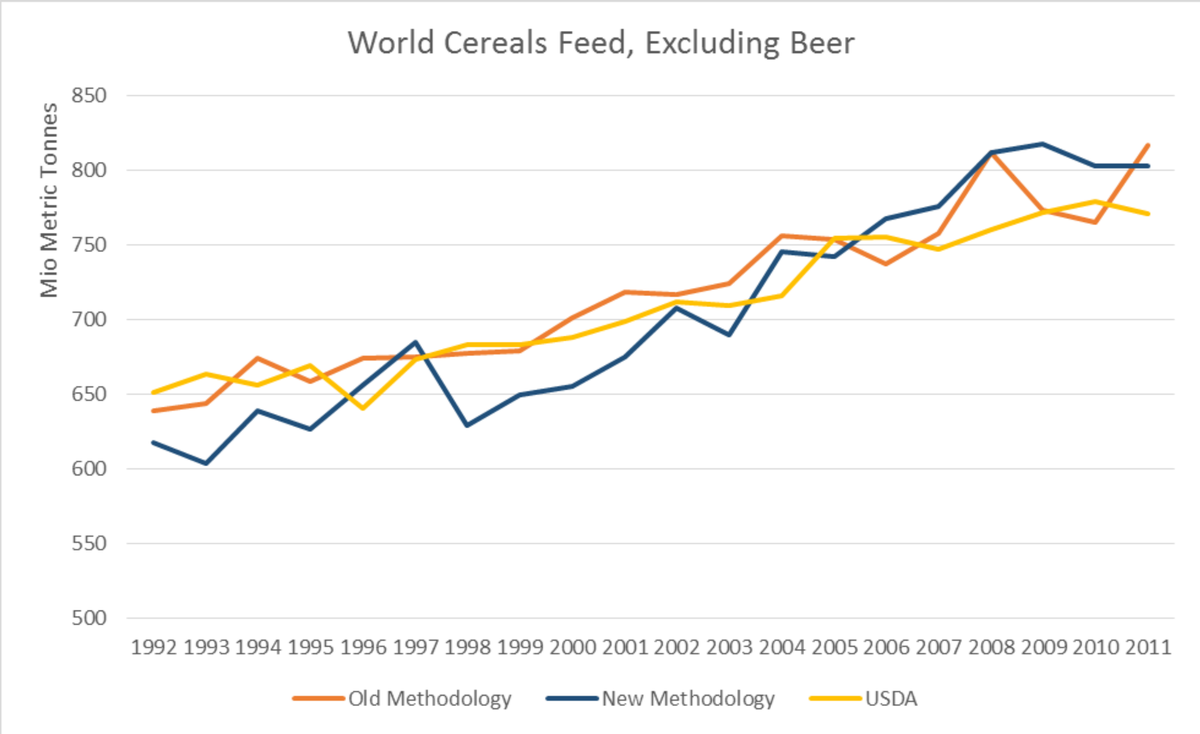

Figure 14: World cereal feeds, comparing results at the global level

Review of the empirical results

Overall, the new methodology suggests a slightly steeper increase of cereal feed use than other data sources/estimates do. There are several explanations for this difference. From the demand side, the sharp increases in herd sizes and animal production over past decades can lead to higher feed use than supply-oriented allocations suggest. In addition, the growing intensification of animal husbandry has favoured the use of energy-rich grains, particularly maize. Combined with the shift towards pork and poultry, these factors support the accelerated increase of grains estimated in the new methodology (Figure 14). In addition, feed use in aquaculture, which was not taken into account in the old methodology, has exhibited extraordinary growth and also supports the faster increase produced by the new methodology. Perhaps of even greater importance is that these findings show how the results of the two approaches are much more different from each other than Figure 14 suggests. Had aquaculture been taken into account in the old approach, the existing FAOSTAT estimates would have been higher by the equivalent of 47.3 million tonnes of aquafeed equivalents[39] in 2012, while in 1992 only 5.9 million extra tonnes of aquafeed equivalents are not considered.

[39] Similar to compound feed for livestock, aquafeed is composed of feedstuffs from outside fish farms and may include a variety of ingredients such as grains, meat or fishmeals and oils: 47.3 million tonnes of aquafeed is equivalent to about 140 million tonnes of maize; 5.9 tonnes of aquafeed corresponds to about 17.3 million tonnes of maize.

Another explanation of the faster increase is that the new methodology suggests more variation of feed use over time. It seems counterintuitive to believe that cereal feed use is steady over time. Feed is at least the second largest factor in disappearance for most cereals. As food use tends to remain stable when commodity availability varies, feed is expected to absorb parts of the variation in availability.

A third explanation is that while the global use estimates may not be entirely dissimilar, even after accounting for aquafeed and feedstuffs for camels, horses, etc. there are considerable differences across countries, which cancel out in the global aggregate.

Consistency and quality checks of feed allocations

Overall, while the new method is more data- and knowledge-intensive, it allows consistency checks of feed allocation against the amounts of nutrients that are required to produce the animal products reported by a country.

Negative protein balances:

It can be argued that some of the most reliable estimates of feed requirements in FBS come from outside the FBS system; these are estimates of the availability of oilmeals and -cakes. The reliability of these estimates is partly the result of oilmeal and -cakes being destined exclusively for feeding purposes[40] and being by-products of oilseed crushing. They are therefore industry products, and data on both the production and imports of such variables are more reliable than data on primary products, not to mention use data. A similar rationale can be put forward for brans. In addition, both oilmeals/-cakes and brans are protein-rich commodities that are often used as the main ingredients to cover protein needs in a feed ration. The high confidence that can be put in these estimates is not only the basis for identifying them separately in the feed allocation process (step 1), but also offers possibilities for undertaking consistency checks.

The main approach to consistency checks is to juxtapose the availability of these feedstuffs with the needs for covering protein requirements, or rather the total use of protein. The most obvious case of an inconsistency is when the protein available from oilcakes/-meals and brans exceeds the calculated needs/total use of protein (requirement times intensification rate). Figure 15 shows 5 256 data points, each representing the logarithm of the amount of protein required and the amount of protein available from oilmeals/-cakes and brans in a given country and year. In 526 cases (about 10 percent), the protein availability exceeds demand. In Figure 15, these cases are shown as dots below the 45-degree line.

[40] Some oilcakes are used as biofuels (e.g. of olives), but these are generally used only in negligible quantities and would not be considered feedstuffs under normal circumstances.

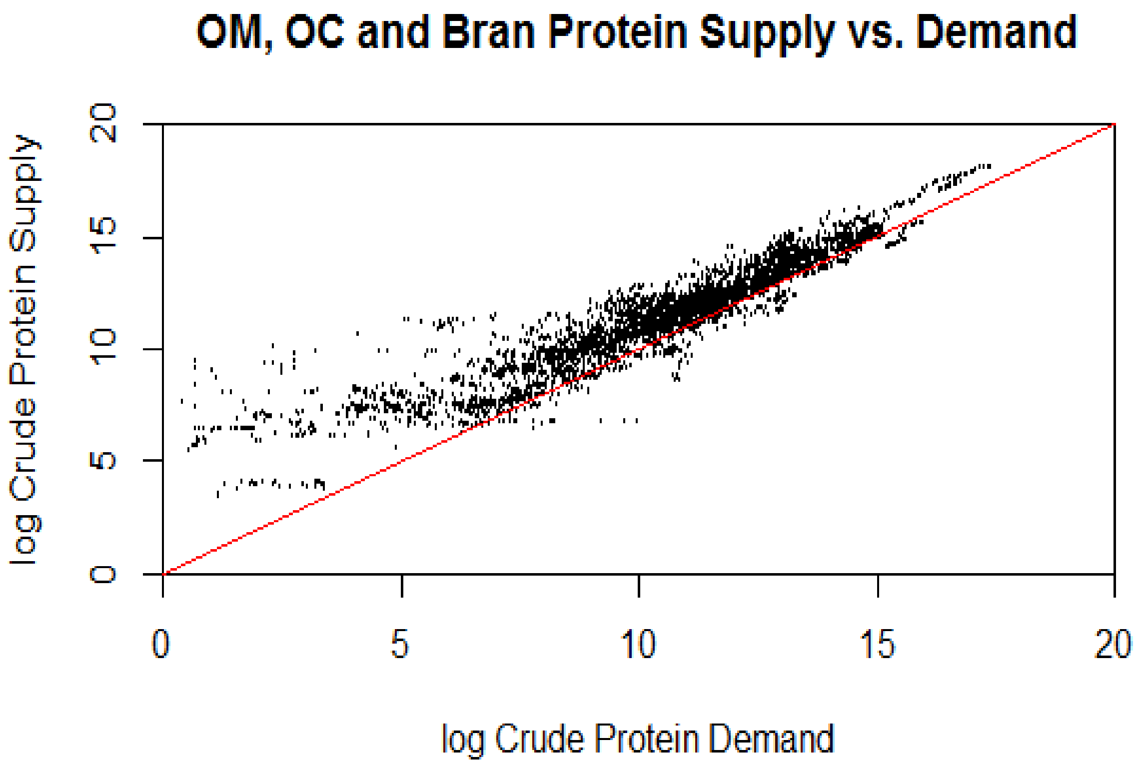

Figure 15: Oilcake, oilmeal and bran supply versus demand

Such an imbalance can have two principal reasons. First, the intensification rate is too low and thus underestimates the actual needs. Second, either the number of animals or the feeding efficiency ratio is too low, which again leads to understatement of actual needs. Such negative protein balances seem unlikely to occur, as they would suggest a gross underestimation of needs.[41] However, in practice, they have occurred in many instances after systematic triangulation of protein needs with availability. Case-by-case review of the results (country by country, year by year) has resulted in many adjustments of animal numbers and livestock productivity, in line with the newly obtained information from consistency checks. Where negative balances are based on official estimates, the inconsistencies are brought to the attention of the data suppliers. A separate document has been compiled providing a complete overview of the results.

Insufficient availability:

Just as there are negative balances, in some countries, vastly positive balances have occurred, at least in some years. While such large positive balances can occur without violating the feed approach in principle, they can absorb so many cereals and pulses that they would render the results for food and other uses unreasonably low.

[41] A negative balance would not only not leave room for any other CC feed, but also suggest that availability from oilmeals and -cakes already exceeds the requirements inherent in the needs calculations.

3.5 Food

The basic idea behind the FBS has always been to derive food as the balancing item between supply and all other forms of utilization. This is still a valid approach, and FBS compilers are encouraged to continue using it, as long as certain conditions are met. Initial estimates[42] of food as a balancing item can still be produced for products where food use is the only, or at least the only major, form of utilization. For example, for meats, eggs and, to a lesser extent, milk and dairy products. Even where products such as skim or whole milk powder are used for feed, FBS compilers will often have access to reliable information on feed use of these products, allowing them to arrive at a reliable food use estimate as a balancing item. In addition, where all other variables of a balance are measured (rather than imputed) with great accuracy, food as a balancing item would also provide a good basis for a reliable food estimate.

However, where non-food uses play important roles, and particularly where many variables are imputed, estimating food via a balancing approach is unlikely to render reliable results. When food is used as a balancing item that has to absorb all measurement errors, the resulting time series is likely to exhibit considerable fluctuations and would thus run counter to economic theory, which suggests that consumers will try to smooth consumption, at least for the total food basket, and will adjust other forms of utilization, such as stocks and feed use, to achieve a smooth consumption path.

Defining FBS food use

The definition of food use in the FAO FBS covers food *availability* up to the retail level for a given reference period, typically one calendar year. The FBS food variable is therefore equivalent to the amount of food available for consumption at the retail level, rather than the amount consumed. It is typically considerably higher than actual consumption, as it includes food waste and losses at the retail and household levels – food that ends up as table scraps or pet food or is simply thrown away.

The FBS food consumption variable refers to the food available to the resident population of a country. The estimates of population come from the United Nations Population Division (UNPD) and also include, to the extent known, migrants, guest workers and refugees, but not tourists. Food consumed by tourists – or, rather, available to them – should therefore not be included in food consumption. In the past, an ad hoc allowance for food consumed by tourists was included in other uses. The new approach makes food available to tourists more transparent, as it estimates it directly and presents it as a separate variable. The methodology applied to estimate food available to tourists is presented, in the following section of this document.

Estimates of food use

Harvesting and harnessing of data

As for all FBS variables, estimates of food use should come from measured information. In the case of food, estimates based on actual data are not only difficult to come by, but these empirically measured observations also typically refer to a different definition. Most empirical estimates come from household surveys, principally household income and expenditure surveys. The information contained therein requires many steps of adjustment, but even a comprehensive set of adjustments will not render a fully compatible definition. As the FBS cover all food disappearance in a country, they

[42] Final estimates will be derived via the balancing mechanism presented in section 2.2 of this document.

also include food consumption/disappearance in public households such as hospitals, prisons and military establishments. Allocating allowances for consumption in these entities is challenging, particularly when the adjustments go beyond the overall average levels, by breaking consumption down into the more than 60 individual commodities contained in the FBS. If adjustments are made, they should include adjustments:

- of the reporting period to a calendar year;
- of expenditures to quantities (applying appropriate prices);
- for intra-household distribution of consumption;
- for food consumed in public/collective households such as hospitals, prisons, military establishments and schools;
- where necessary, for food consumed by non-collective households outside the home, in restaurants, in cafeterias, as street food, etc.;
- for waste at the retail level.

Information from food processors

FBS-compatible food consumption information may also be available from food processors, at least in certain circumstances. For instance, if the flour milling industry of a country keeps good records of total flour produced in a year in the country (as is the case in many developed countries) and if the flour milling industry is the main source of flour production, these production amounts can serve as a basis for estimating the total flour produced in the country. By adding flour imports, deducting flour exports and adjusting for stock changes, it is then possible to arrive at flour availability for food. As flour is destined for food use only, applying an appropriate extraction rate (which is also available from the flour milling industry) enables generation of an accurate and FBS-consistent estimate of primary wheat food. As emphasized, such estimates require the presence of an industrialized flour milling sector. Where flour milling takes place at the household level or in artisanal milling environments with low reporting levels, this method would inevitably underestimate consumption levels of the primary product. However, where first-level processors (flour millers, oilseed crushers, abattoirs, sugar refineries/factories, dairies, etc.) represent a large share of domestic processing (a “bottleneck”), processing statistics can serve as an excellent basis for generating FBS-compatible food estimates.[43]

Creating expected values for time t

The basic idea of the previous FBS methodology was to derive food availability from the intersection of supply and utilization, i.e. food as a residual/balancing item. As outlined previously, this method will not, or at least not always, render plausible results in practice. In fact, it may lead to unintuitive and implausible results, which are visible through large year-to-year fluctuations in per capita food consumption levels. At the same time, deriving FBS-compatible food estimates from household surveys is not always straightforward, while estimates from the food processing industry may not always be reliable or may be available only sporadically. In these cases, it is necessary to generate estimates through an imputation method.

[43] Of cereals from flour millers, vegetable oils from oilseed crushers and meat from abattoirs.

The approach proposed here for imputing values for food consumption is straightforward. It rolls out food consumption in year t as a function of a known level of food consumption in the past, adjusted for changes due to income and other factors (captured by a trend factor), such as changing preferences or known, sudden supply disruptions. In a stylized form, imputed food consumption is:

$$Food_t = f(Food_{t-1}, \Delta Y, t) \text{ (equation 17)}$$

where ΔY represents an average change in per capita income, and t is the estimate for a simple time trend factor. Four different functional forms have been distinguished: a linear specification, a log-log specification, a semi-log specification and a log-inverse specification. The functional form was chosen in line with the functional form that was used for estimation of the underlying income elasticities. All equations for all commodities and all countries have been parameterized with an income elasticity (ε_j) for every commodity j and country i, and a trend factor (t_j):

Linear specification:

$$FoodPC_{t,ij} = FoodPC_{t-1,ij} + t_i * FoodPC_{t-1,ij} \text{ (equation 18)}$$

Log-log specification:

$$FoodPC_{t,ij} = FoodPC_{t-1,ij} + \exp\left(\varepsilon_{ij}\ log\left(\Delta Y_{i,j}\right)\right) + t_i * FoodPC_{t-1,ij} \text{ (equation 19)}$$

Semi-log specification:

$$FoodPC_{t,ij} = FoodPC_{t-1,ij}\left(1 + \varepsilon_i\ log\left(\Delta Y_{i,j}\right)\right) + t_i * FoodPC_{t-1,ij} \text{ (equation 20)}$$

Log-inverse specification:

$$FoodPC_{t,ij} = FoodPC_{t-1,ij} \exp\left(\varepsilon_{i,j}\left(1 - {}^{1}/_{\Delta Y_{i,j}}\right)\right) + t_i * FoodPC_{t-1,ij} \text{ (equation 21)}$$

where $\Delta Y = \frac{GDP\_PC_t}{GDP\_PC_{t-1}}$.

Rationale

Creating sound FBS is a particularly challenging task for countries where a large amount of the food available is not consumed by the resident population. As available food is assumed to be available to all population groups, including refugees and migrants, food available to migrants and refugees is assumed to be captured by UNPD estimates. This approach cannot be used for countries where non-resident populations are not included in the UNPD estimates, as is the case for tourists and consumption by tourists, which are particularly challenging where tourists account for large shares of the population and food consumption. In addition, the consumption patterns of tourists are not necessarily the same as those of the resident population. FBS compilers at the country level are likely to have detailed information about the main parameters necessary for gauging the exact extent of tourist consumption – the number of tourists visiting the country, their average stay, their consumption patterns, etc. Based on this information, it is straightforward for a national FBS compiler to make an allowance for tourist consumption in the country. In the absence of this information, tourist consumption could be estimated through the approach outlined in the following.

Data availability

The basic data for estimating consumption by tourists are available from the United Nations World Tourism Organization (UNWTO), which provides tourist flows between all pairs of countries. Unfortunately, no data are available regarding the consumption patterns of these people while they travel.

The proposed methodology

The methodology at FAO is straightforward, and will be refined as more and better information becomes available at the country level. The methodology is based on the number of day visitors (N_D) and overnight visitors (N_O) in and out of each country, and on information on the average number of nights stayed within each country ($\overline{D}$). With this information, the first step is to compute the total number of "tourist days" (N_T) in and out of each country by adding the day visitor counts to the product of the overnight visitor counts and the average nights per visitor:

$$N_T = N_D + N_O\overline{D} \text{ (equation 22)}$$

In the absence of better information, it is assumed that tourists follow the same consumption levels (calories) abroad as they do at home, while they follow the local consumption pattern in terms of preference. Thus, the total number of tourist days is multiplied by the average daily consumption within the country of origin, and this amount is allocated to tourist consumption in the destination country in accordance with the destination country patterns. This total is also deducted from the food consumption in the country of origin, as the tourists will not be at home to consume these calories. Thus, the change in the amount of food availability for commodity i in country j is:

$$\Delta TC_{ij} = -\sum_{k=1,k\neq j}^{m} N_{jk}f_{ij} + \sum_{l=1,l\neq j}^{m} N_{lj}f_{ij} * \frac{\sum_i f_{il}}{\sum_i f_{ij}} \text{(equation 23)}$$

where N_{lj} represents the number of tourists travelling from country l to country j, and f_{il} represents the historic amount of daily nutrients consumed from commodity i and in country l. This equation can be simplified:

$$\varDelta TC_{ij} = -N_T f_{ij} + \sum_{l=1,l\neq j}^{m} N_{lj} f_{ij} * \frac{\sum_i f_{il}}{\sum_i f_{ij}} \text{ (equation 24)}$$

Average daily consumption is computed based on historic consumption patterns, thus tourist consumption data can be provided at the full FBS level. The measurement error can likewise be derived from the measurement error of food consumption of the resident population. It should be noted that the tourist consumption element for a particular country can be negative if many residents of the country leave as tourists and few tourists from other countries visit.

3.7 Industrial use

The twenty-first century has seen important changes in the global commodity landscape. Many of these changes have originated outside food and agriculture, but many have had a direct and sometimes massive impact on food and agricultural markets. Rising energy prices in particular have made an increasing number of agricultural products become competitive inputs in non-food markets in growing volumes (Figure 16). One of the most important examples is the biofuels market, where rising energy prices in conjunction with policies to promote the use of agricultural feed stocks to generate energy have diverted a growing amount of (food) commodities into the fuel market. Maize for bioethanol in the United States of America and rapeseed oil for biodiesel in the EU are the most important examples, but not the only ones. With rising energy prices, many countries have embarked on similar programmes, albeit to a smaller extent.

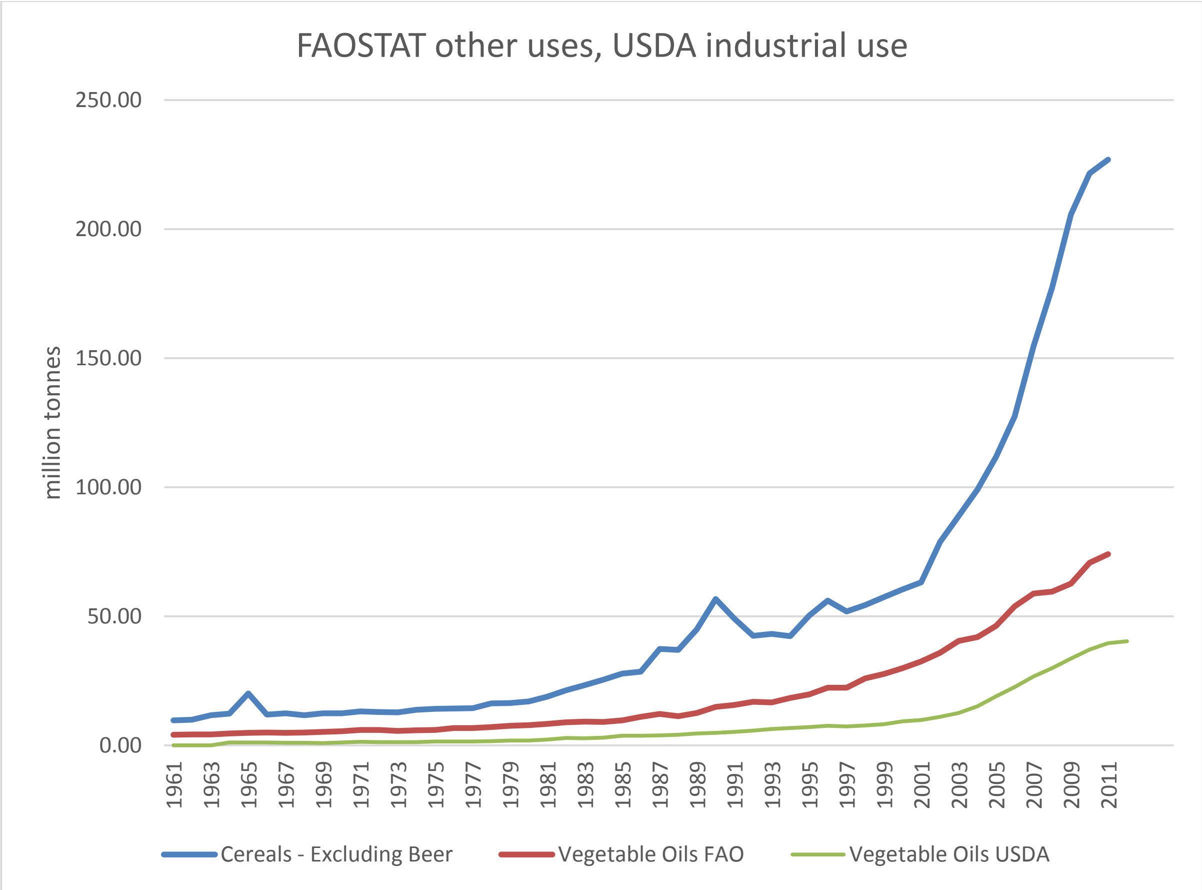

Figure 16: Vegetable oils: other uses from FAOSTAT versus industrial use from USDA's PS&D database

Much less prominent, but equally important, is a similar trend for industrial use outside the energy sector. Similar to the biofuels market, this market also affects both starch-rich and oil-rich food commodities. For example, rapidly rising shares of coconut oil and palm kernel oil are now diverted into the production of cosmetics, while many other vegetable oils have become core ingredients for the production of paints, soaps and other detergents. Starch-based products have likewise become increasingly important as construction materials and are being used for other, non-food purposes. The non-food use of vegetable oils is particularly important for the FBS. These uses can easily be underestimated in quantity terms and, given their high caloric content (9 kcals/g), the impacts on food availability and DES are particularly great.

Unfortunately, the wide variety of different uses (energy, paints, detergents, cosmetics, etc.) makes it difficult to gather actual information. In addition, there is no straightforward way of imputing information. Questions on industrial uses are included in FAO questionnaires, but very few countries provide this information. The only area where information is readily available is in the use of agricultural feedstuffs for biofuel production. This information is collected by FAO for its medium-term outlook work and is taken into account in the new version of the FBS. In the interim, USDA's Production, Supply and Distribution (PS&D) database provides the necessary information on industrial use of vegetable oils. The amounts of calories implicit in these uses are considerably higher than the estimates of the FBS system, and will therefore affect food availability and the DES.

Imputation and data sources

Several imputation methods have been tried, but none have rendered satisfactory results. The main reason for this lack of good results is that there is no straightforward model that explains the amounts of agricultural products used for industrial purposes; even energy prices have only limited explanatory power. The conclusion is that estimates of industrial use must be based on collected data. At the international level, data are collected by United States attachés and made available through the PS&D database, at least for industrial use of vegetable oils. Data are also collected on both vegetable oils and biofuels in the context of the OECD/FAO medium-term outlook. FBS compilers at the country level are encouraged to collect information on industrial use by carrying out specific surveys. Auxiliary information could be derived from biofuel mandates and other policy stipulations governing the use of agricultural feed stocks for industrial uses.

3.8 Food losses and waste

Definition of FLW in the FBS system

The FAO FBS system covers food availability up to the point of food purchases, typically at the retail level. In line with the definition of food availability, food losses cover all losses up to the same level (Figure 17). Post-harvest losses and losses that occur during storage, transportation and processing are therefore included, while losses at the retail or even the household level are *not* included in the FBS system. Processing losses are estimated in the standardization process (see section 2.2), in which a small allowance is made when extraction rates are applied to primary products. Figure 17 provides an overview of losses along the entire value chain and distinguishes losses covered in the FBS system from those outside the FAO food balances.

Data availability

The FAO production questionnaire (see section 3.1) includes a section where countries report estimates of waste and losses. While the vast majority of FAO member countries report production and/or productivity data, estimates of losses are notoriously scarce, and the few that are available are unlikely to be based on reliable measurement efforts. The Global Strategy to Improve Agricultural and Rural Statistics found that there is not even a readily available and reliable method for measuring losses. As a result, the few estimates available may not be fully reliable and comparable across countries. The next step is therefore to devise an experimental design that could ultimately be used to measure FLW accurately.

The data on FLW reported by countries can differ considerably, even among similar countries and even as percentage loss rates – shares of production and trade. Cross-country comparisons have found such differences in estimates of both individual commodities and overall losses at the country level. Although these estimates are not an optimal basis for developing an imputation approach, in the absence of better information, they are the only basis available. Before detailing the new method, a quick review of the previous approach is provided in the following subsection.

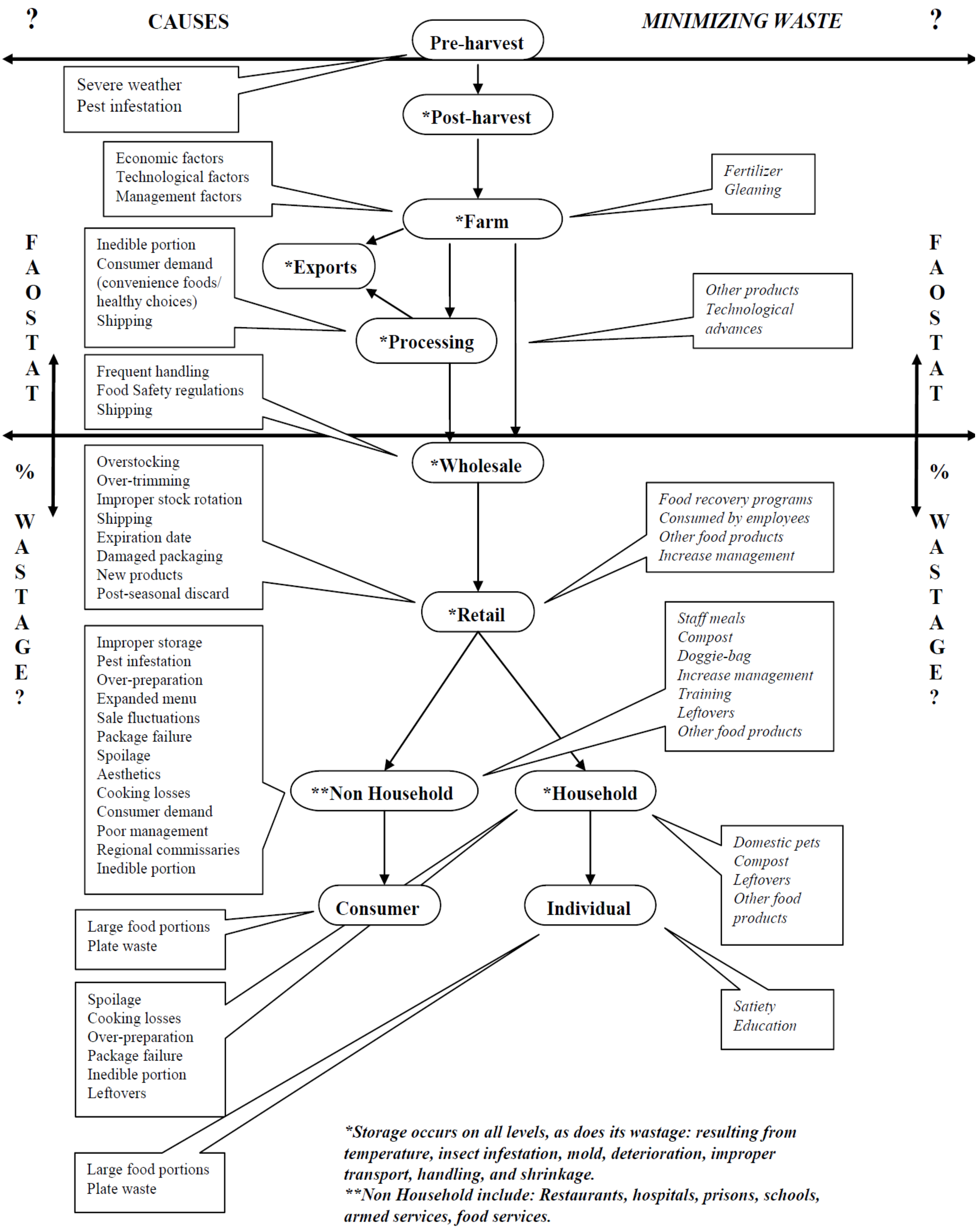

Figure 17: Flow chart of food losses and waste

Source: ESS, 2004.

The previous imputation approach

Similar to feed use, FLW was also estimated as a simple ratio (share) of availability in the previous approach;[44] the amount of food lost for commodity i in country j ($Lo_{i,j}$) was:

$$Lo_{i,j} = (P_{i,j} + I_{i,j}) * r_{i,j} \text{ (equation 25)}$$

Again, similar to the previous approach for estimating feed use, no differentiation was made between trade and domestic production in estimating FLW, i.e. a uniform loss rate was applied to all elements of availability.

The new imputation approach

Given the possible shortcomings of the old imputation method, and despite the data and methodological limitations, it was decided to develop a new method for estimating FLW based on those data points that are officially reported. The scarcity of official data made it necessary to pool available data into a panel. It was also deemed necessary to distinguish losses of domestic products from those of trade products, as there are strong a priori reasons for assuming that the latter are generally lower. The main reasons for lower trade losses are that post-harvest losses do not apply to trade, and traded commodities are likely to be handled and processed in more advanced systems, which are typically less prone to losses.

FBS compilers at the country level are encouraged to use all country-specific information to gauge the amounts of losses in their home countries. They can probably draw on country- and commodity-specific knowledge about farm production (harvesting machinery, storage facilities, etc.) and food processing systems that is not readily available to FAO. For all variables, FAO appreciates receiving such information, ideally through its production questionnaire.

The model/estimation procedure

In the absence of empirical estimates, countries may want to consider using the FAO method of estimating FLW. For this model, a hierarchical linear model was developed with loss quantity as the response. The strength of the approach is its ability to pool different levels of information to derive the best possible inference and optimal prediction.

Modelling is performed hierarchically, with country/commodity-specific estimates at the lowest hierarchical level, followed by commodity-specific estimates, food group estimates, and finally perishable food group estimates. The coefficients in the hierarchy are estimated simultaneously to ensure that they are consistent within the hierarchy and among hierarchies. The specific model employed is:

$$\log(\mathrm{Loss}_{\mathrm{ijklm}}) = \alpha_1 \mathrm{t} + \alpha_{2\mathrm{ijkl}} \log(\mathrm{Production}_{\mathrm{ijklm}} + 1) + \mathrm{A}_{\mathrm{ijklm}}$$

[44] Total availability notionally includes stock withdrawals, but as no data on stock levels were available in the past, stock withdrawals were not included in the previous approach.

$$\alpha_{2ijkl} = \beta_{20} + \beta_{2ijk}(\text{Country: Commodity})_{ijkl} + B_{ijkl}$$
$$\beta_{2ijk} = \gamma_{20} + \gamma_{2ij}(\text{Commodity})_{ijk} + C_{ijk}$$
$$\gamma_{2ij} = \delta_{20} + \delta_{2i}(\text{Food Group})_{ij} + D_{ij}$$
$$\delta_{2i} = \zeta_{20} + \zeta_{21}(\text{Food Perishable Group})_{i} + E_{i}$$

(equation 26a–e)

where the i_{th} index represents the effect of perishable food group, the j_{th} index the effect of food group, the k_{th} index the effect of commodity, the l_{th} index the effect of country/commodity, and the m_{th} index the individual observations. The capital letters represent random errors, which are all assumed to follow normal distributions with unknown variance. The Greek letters indicate coefficients to be estimated from the data.

The utility of this model is that if a coefficient is missing owing to non-availability of official data for estimation, the algorithm will impute losses based on the next level of the hierarchy. This means that if loss estimates are missing for a specific country and commodity, the global loss rate for the commodity is applied. If that is not available, the global loss rate for the food group of the commodity is used, and if even that is not available, the loss rate for perishable food is used. However, if large quantities of data for a particular commodity within a particular country are available, these data are used for imputation without moving up the hierarchy.

The advantage of this model is that it will suffice on its own, whereas the common approach usually demands the imputation of several models when a single model fails. In addition, when separate models are estimated individually, they do not incorporate the hierarchical information present in the data.

3.9 Seed use

A brief review of the previous approach

Seed use is defined to include "all amounts of the commodity in question used during the reference period for reproductive purposes, such as seed, sugar cane planted, eggs for hatching and fish for bait, whether domestically produced or imported". This definition includes double or successive sowing or planting, whenever it occurs. Seed use also includes the quantities needed for sowing or planting of crops for use as fresh fodder or food (e.g. green peas, green beans, maize for forage), at least when this information is available. On average, the amount of seed needed per hectare planted in a given country and for a given crop does not vary greatly from year to year.

Seed data are collected through a special section in the FAO production questionnaire; where countries do not provide information directly, seed use information is collected through the Web sites of the relevant national authorities, such as national statistical offices or ministries of agriculture. Where neither of these sources provide official data, seed estimates are calculated or estimated either as a percentage of production (e.g. eggs for hatching) or by multiplying the area sown/harvested with a seed rate. Ideally, seed estimates are based on the area sown rather than the area harvested; in practice, however, sparse data availability make it inevitable that seed use estimates are based on area harvested.

Seed use can be formulated in the simple identity:

$$Se_{i,j} = (AS_{i,j}) * r_{i,j} \text{ (equation 27)}$$

where Se is seed use, AS is area sown and r is the seed rate for commodity i in country j. The seed rates vary by both country (i) and commodity (j) because of the different agricultural demands in different climates/regions of the world. As for all other variables, national FBS compilers probably have the best knowledge of and information about relevant national seed rates. They are encouraged to use their own country-specific seed rates in the FBS balances and to make the estimates available to FAO, preferably via the annual production questionnaire.

Seed rates

The seed rates currently used in the FBS/SUA system are based on an ESS publication entitled *Technical conversion factors for agricultural commodities* (ESS, [date]), which provides seed rates for every country and primary commodity/item included in the FCL classification. It brings together information on seed rates by country and commodity, reflecting current production practices under different growing conditions. The compilation of seed rates benefited from information provided through the questionnaire and from FAO expert advice.

The publication also provides information about the share of eggs that is typically used for hatching.[45] Additional information is provided to improve assessment of the reliability of hatching rates. All hatching and seed rates have been reviewed for the new FBS system, and have been changed where necessary.

[45] Only 11 countries have provided official data on this share in recent years.

The new imputation approach

As previously mentioned, FAO collects data on seed and area sown through the production questionnaire. However, while overall response rates to the questionnaire have been rising, not all countries provide estimates for all commodities. Where no official seed use information is available, seed use can be imputed, including by national FBS compilers. The necessary steps are to:

1. impute area sown, when missing;
2. estimate seed rates using a hierarchical linear model, if questionnaire results are unavailable, imputing seed as the product of the previously imputed area sown and the estimated seed rate.

Imputation of area sown

To impute the actual area sown, the following approach is taken:

- If previous values of the area sown and the area harvested are available, an average ratio of the area sown (in year t) to the area harvested (in year t + 1) is computed. Then, if the area sown is unavailable in one year, it is imputed by multiplying the area harvested in the following year by the average ratio.
- If no prior information on the area sown is available, a ratio of 1 is used.

Estimation of seed rates

Seed rates are estimated via a hierarchical linear model (described in section 3.8). The rationale for this model is that it is capable of capturing and modelling complicated trends when data are available. In addition, the hierarchy of the model allows accurate imputation for countries with very sparse data by pooling together global data. The mathematical model can be written as follows:

$$\log(\mathrm{Se}_{i,j,k}) = \beta_0 + \beta_1 \mathrm{Temp}_i + \beta_2 \mathrm{Time} + \beta_{3,j,k} \log(\mathrm{Area\ Sown}_{i,j,k}) + \varepsilon_{i,j,k}$$

$$\beta_{3,j,k} = \gamma_{30} + \gamma_{31,j,k}(\mathrm{CPC\ Code: Country})_{j,k} + \delta_{j,k}$$

$$\gamma_{31,j} = \kappa_{310} + \kappa_{311}(\mathrm{CPC\ Group})_j + \zeta_j$$

(equation 28 a–c)

where β, γ, κ are coefficients to be estimated from the data, $\varepsilon, \delta, \zeta$ are error estimates, $Temp_i$ is the average annual temperature of country i (provided by the World Bank), and $Time$ is measured in years and is included to capture linear trends over time. The i indices run over all countries, the j indices over all CPC groups, and the k indices over all unique country/CPC code combinations.

Thus, the model estimates seed use proportional to the area sown. The model also accounts for changes over time and differences among countries; the latter are captured by the annual temperature variable, assuming that seed rates need to be higher where production conditions are difficult and there is potential for late and frequent frosts (Russian Federation), and can be lower where production conditions are more favourable (United Kingdom of Great Britain and Northern Ireland).

If data for a particular country and commodity are sparse, then κ_{310} and κ_{311} will likely be estimated as close to 0. Thus, $\gamma_{31,j,k}$ will be close to its mean value, and the model will account only for availability within commodity groups. However, if data are available for a country or commodity, the estimates of κ_{310} and/or κ_{311} will be far from 0, and thus the model enables adaptation to the individual characteristics of a particular scenario.

3.10 Residual/other uses

The overall strategy for FBS compilers at both FAO and the country level should be, as far as possible, to identify all individual uses separately and explicitly. However, this process cannot guarantee that all uses for all commodities and in all countries are adequately captured by the use categories described so far. There will be specific forms of use that cannot be subsumed under any of the traditional forms. FBS compilers at the country level need to identify and delineate these residual other uses, and provide and include adequate estimates for them.

References

ESS. 2004. *Preliminary research report on the assessment of intra-household wastage of food*. Rome, FAO.

ESS. [date] *Technical conversion factors for agricultural commodities*. Rome, FAO.

FAO. 2014. *The FAO Statistics Quality Assurance Framework*. Rome. http://www.fao.org/docrep/019/i3664e/i3664e.pdf

4. Compilation of food balance sheets: a step-by-step introduction

This chapter provides a step-by-step introduction to the practical compilation of FBS, starting with the detailed SUA of a given country in a given year. It exemplifies the calculations through three commodities – wheat, sugar and palm oil. It also shows how the SUAs are aggregated into primary equivalents, as presented in the FBS, via the standardization process.

As laid out earlier, the basic challenge is to populate a simple identity that balances all forms of supply with all forms of utilization. This holds for the highly disaggregated SUAs as well as for the primary product equivalents as shown in the FBS. The supply side is defined by the sum of its constituent variables:

Supply = Production + Imports – Stock changes[46]

While the utilization side is defined as:

Utilization = Exports + Food + Food processing + Feed + Seed + Tourist consumption + Industrial use + Loss + Residual/other uses

4.1 Wheat

The first example considers the full process for creating an FBS for wheat. In the following, all quantities are in tonnes unless another unit is stipulated. The process starts with an empty SUA table (Table 6) showing some of the commodities in the wheat "commodity tree" (wheat is the primary commodity, while flour, starch, etc. are the processed commodities). In Table 6, a dash indicates that a value is currently unknown.

Table 6: Initial, empty SUA table for wheat

[46] Stock changes are here defined as $St_t – St_{t-1}$. In FAOSTAT they are currently defined as $St_{t-1} – St_t$. The new definition ensures that increases in stocks have a positive sign, whereas stock drawdowns have a negative one.

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	-	-	-	-	-	-	-	-	-	-	-
Wheat flour	-	-	-	-	-	-	-	-	-	-	-
Bulgur	-	-	-	-	-	-	-	-	-	-	-
Breakfast cereals	-	-	-	-	-	-	-	-	-	-	-
Wheat starch	-	-	-	-	-	-	-	-	-	-	-
Wheat bran	-	-	-	-	-	-	-	-	-	-	-

Production

For production data, the table is first filled in with any available official figures. To impute any missing production figures, data on yield and area harvested (in the absence of data on area sown) must also be considered, as yield is defined as production divided by area harvested (thus, with any two variables the third is uniquely defined).

Table 7 illustrates a case where only limited official data are available. For the sake of demonstration the flour production quantity is assumed to have been reported, while wheat production is unknown and must therefore be imputed. In reality, however, the reverse is almost always true, with wheat production officially available, and flour less so.

Table 7: Introducing production into the SUA table for wheat

Name	***Area*** **harvested** ***(ha)***	***Yield (tonnes/*****ha)**	***Production (tonnes)***
Wheat	18 500 000	-	-
Wheat flour	-	-	18 650 000

In the case illustrated in Table 7, the production quantity is known only for wheat flour and is missing for wheat, as is the wheat yield value. The following is the procedure for imputing production data:

- If all three variables are available, any two variables in the Yield = Production/Area harvested formula are used to cross-check the third variable. If the formula indicates an error for one of the given variables, a quick time series check should identify the incorrect value.
- If only two variables are available, the third is computed with the same formula.
- If data for only production or area harvested are available, yields are imputed from the historical time series (using the ensemble approach described in section 3.1). The other missing element can then be calculated using the formula.
- If only the yield variable is available, production is imputed from the historical time series (using the ensemble approach), and the formula is used to calculate area harvested.
- If all three variables (area harvested, yield and production) are missing, yield and production data are imputed from the historical time series (using the ensemble approach). The area harvested is then calculated through the Area harvested = Production/Yield formula.

In the Table 7 example, the yield needs to be imputed. In Figure 18, several models are fitted to the historical yield values (models are represented as lines and historical data as points). These models are combined in a weighted average (with weights based on how well the model fits the data) to form a final ensemble of models. This ensemble is then used to predict the yield value in the current year.

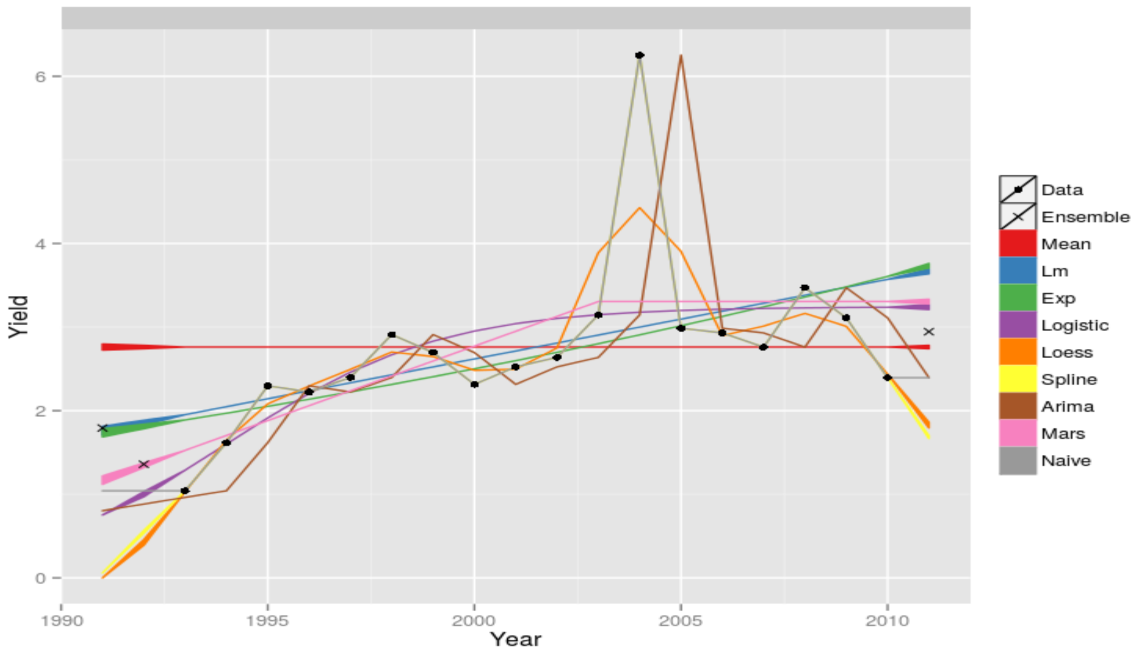

Figure 18: Ensemble modelling of results for yield estimates

The final imputed value for yield in the most recent year (shown in Figure 18 as “x”, representing the ensemble imputation result) is 2.94 tonnes/ha (Table 7). This is a reasonable estimate when compared with the historical time series. Some models (Figure 18) fit the data fairly well, such as the logistic regression, the spline and the autoregressive integrated moving average. Some of these models do not produce good forecasts – particularly local regression, which produces somewhat low forecasts – but averaging the well-performing models results in a good final estimate of yield. See chapter 3 for further details on these models and the ensemble imputation approach.

Table 8: Adding yield estimates to the SUA table for wheat

Name	***Area* harvested *(*ha*)***	***Yield (tonnes/*ha*)***	***Production (tonnes)***
Wheat	18 500 000	**2.9422**	**-**
Wheat flour	-	-	18 650 000

With this information, it is possible to compute the production data (Table 8).

Table 9: Completing the SUA table for wheat production

Name	***Area* harvested (ha)**	***Yield (tonnes/*ha)**	***Production (tonnes)***
Wheat	18 500 000	2.9422	**54 420 000**
Wheat flour	-	-	18 650 000

Next, imputed and official production quantities are added to the table. Production is imputed only for primary products. Official figures are occasionally provided for processed products, as in Table 10, in which case, no additional quantities are filled in other than wheat and flour. However, FBS compilers at the country level are encouraged to add estimates for activity levels (area harvested for crops, herd size for livestock) and productivity (yields and cropping intensity for crops, slaughter weights and take-off rates for livestock) to the FBS and SUA tables. These estimates provide users of FBS tables with a more complete picture of the supply side and additional information for policy decisions (e.g. on promoting activity and/or productivity measures to address a lack of production).

Table 10: Introducing estimates of wheat and flour production into the SUA table for wheat

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Wheat	54 420 000	-	-	-	-	-	-	-	-	-	-
Wheat flour	18 650 000	-	-	-	-	-	-	-	-	-	-
Bulgur	-	-	-	-	-	-	-	-	-	-	-
Breakfast cereals	-	-	-	-	-	-	-	-	-	-	-
Wheat starch	-	-	-	-	-	-	-	-	-	-	-
Wheat bran	-	-	-	-	-	-	-	-	-	-	-

Trade

Trade data are usually recorded in much greater detail than can be presented in the FBS tables and include data on a large number of processed products from a single primary product of the FBS. The national trade data set is usually provided by the customs office and should consist of detailed quantity and value flows for imports and exports, typically classified by commodity using the HS codification. Trade data are available by partner country and are for at least annual, and sometimes also monthly, flows. The level of commodity detail is country-specific, with some countries reporting at the basic standard 6-digit level of the HS, while others go up to 12-digit detail (see subsection on Data on imports and exports in section 1.3 for more detail). Ideally, the country codes used should follow the international standard codes of the UN M49 country classification.

The total imports and exports for each commodity – wheat in this example – are obtained by aggregating the respective trade flows by partner. A typical trade data set for wheat data would look as shown in Table 11 (the data set has been simplified for this example).

Table 11: Trade flow information on wheat and products at HS6 disaggregation

Year	Reporter	Partner	HS-6 code	Flow	Weight (kg)	Value (US$)
2014	950	932	100110	1	3 350 000	502 500 000
2014	950	899	100110	1	1 200 000	264 000 000
2014	950	961	100190	2	870 000	113 100 000

In Table 11, the country codes refer to a specific reporter and three different partners. The HS codes are standard 6-digit codes, indicating wheat in this case. (For more information on the HS classification please see section 3.2.) Flows 1 and 2 indicate imports and exports respectively. The quantity weights are in kilograms and the values are, in this case, in United States dollars. The totals for wheat imports would be obtained by adding all the import flows; total exports would be the sum of all export flows (a typical trade data set would have many more flows than the simple example in Table 11). For the compilation of FBS, only quantities are of interest, and not monetary values.[47] Total import and export quantities for wheat, and the other commodities, can now be inserted into the SUA table (Table 12).

Table 12: Adding trade information to the SUA table for wheat

[47] Monetary values can be used to impute missing quantity data (details in trade discussion in section 1.3).

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Wheat	54 420 000	**1 999 100**	**32 790 000**	-	-	-	-	-	-	-	-
Wheat flour	18 650 000	**341 500**	**572 800**	-	-	-	-	-	-	-	-
Bulgur	-	-	-	-	-	-	-	-	-	-	-
Breakfast cereals	-	**312 500**	**217 300**	-	-	-	-	-	-	-	-
Wheat starch	-	**624 900**	**224 500**	-	-	-	-	-	-	-	-
Wheat bran	-	**2 589 000**	**2 343 700**	-	-	-	-	-	-	-	-

The quantities for bulgur are missing from Table 12. Checks of historical time series and "mirrored" trade data indicate that a quantity should be imputed for bulgur. The historical time series check shows that there has been trade in this commodity in each of the last, say, 20 years. The check of mirrored trade data indicates that there is trade for this country in bulgur in the year in question (see section 3.2 for more details on mirrored trade data). If no estimates of mirrored flows are available for national FBS compilers, the missing trade data (for bulgur in this case) can be obtained from trading partner data available from FAOSTAT.[48] The mirrored imports and exports of bulgur are then inserted into the table, as shown in Table 13.

Table 13: Completing the trade data in the SUA table for wheat

[48] Mirrored data are the total trade flows reported by all other countries (trading partners) for the reporting country. Thus, the mirrored imports are the reporting country's exports; and conversely for mirrored exports.

Name	*Production*	*Imports*	*Exports*	***Stock change***	*Food*	***Food processing***	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-	-	-	-	-	-	-	-
Wheat flour	18 650 000	341 500	572 800	-	-	-	-	-	-	-	-
Bulgur	-	**182 900**	**580 000**	-	-	-	-	-	-	-	-
Breakfast cereals	-	312 500	217 300	-	-	-	-	-	-	-	-
Wheat starch	-	624 900	224 500	-	-	-	-	-	-	-	-
Wheat bran	-	2 589 000	2 343 700	-	-	-	-	-	-	-	-

However, in the example used for this demonstration, data quality validation indicates that there is a quantity error for wheat bran imports, based on analysis of the median unit values in the original trade data set.[49] In this case, the quantity error is simply an extra-digit error, and the actual quantity should therefore have one fewer zeros (to reflect the correct import unit value). The trade data are corrected as shown in Table 14.

Table 14: Improving trade data by correcting obvious reporting errors in the SUA table for wheat

[49] The unit value is the monetary value per quantity (weight, number, etc.) (see subsection on Calculating export and import unit values in section 3.2).

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-	-	-	-	-	-	-	-
Wheat flour	18 650 000	341 500	572 800	-	-	-	-	-	-	-	-
Bulgur	-	182 900	580 000	-	-	-	-	-	-	-	-
Breakfast cereals	-	312 500	217 300	-	-	-	-	-	-	-	-
Wheat starch	-	624 900	224 500	-	-	-	-	-	-	-	-
Wheat bran	-	**258 900**	2 343 700	-	-	-	-	-	-	-	-

So far, the supply side of the SUA table for wheat has been completed. To demonstrate how an SUA table can be produced through imputation, it was assumed that only limited information was available to FBS compilers. To demonstrate other imputation approaches for demand variables, the same basic assumption of very scant information will be made in filling in the SUA table on the utilization side. However, it should be noted that the sole purpose of this assumption is to demonstrate the application of imputation methods. It is not intended to discourage FBS compilers from collecting reliable and measured data. As mentioned repeatedly throughout this book, solid FBS tables require solid data collected from censuses, surveys or at least administrative data.

Stock changes

Stocks are generally held for a select number of primary-level products (such as wheat or rice). The numbers shown in Table 15 represent the estimated stock changes for the example country. The stock imputation methodology described in section 3.3 estimates stock change in the current year as a linear regression on the cumulative stock changes in previous years. The estimate in Table 15 represents a withdrawal (hence the negative sign) in the stocks held. The basic idea behind this assumption is that a drawdown in stocks is more likely to follow a high accumulation of stocks in the past, and vice versa.

Table 15: Adding stock changes to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	**-230 600**	-	-	-	-	-	-	-
Wheat flour	18 650 000	341 500	572 800	**0**	-	-	-	-	-	-	-
Bulgur	-	182 900	580 000	**0**	-	-	-	-	-	-	-
Breakfast cereals	-	312 500	217 300	**0**	-	-	-	-	-	-	-
Wheat starch	-	624 900	224 500	**0**	-	-	-	-	-	-	-
Wheat bran	-	258 900	2 343 700	**0**	-	-	-	-	-	-	-

Food

As explained in section 3.5, there are several options for obtaining estimates of food. In any case, food can be obtained as the balancing item between supply and utilization, which renders good estimates where food is the sole, or at least the main, form of utilization, as is the case of meats, eggs, butter or, albeit to a lesser extent, milk and milk products. Food estimates can also be obtained by gathering data from first-level processors such as flour mills, oilseed crushers, abattoirs and dairies, particularly where these industries represent a bottleneck in the processing chain by covering a large share of primary processing. For the sake of demonstrating imputation methods, it is assumed that no such information is available and that using food as a balancing item is not reasonable. Instead, food use is estimated from previous estimates and changes in consumer income, etc. It should be noted that when information on food is available, the balancing mechanism described here will automatically use food as the residual (as there would be no other element to absorb the imbalance). Thus, only one approach is required.

To this end, food consumption is estimated from the previous year and these estimates are extrapolated forwards using changes in gross domestic product (GDP) and product-related income elasticities. At the SUA level, any edible item can potentially be allocated to food, but the food module estimates variables at the primary level only. The food variable is therefore estimated from the primary level of commodities eaten directly (which is not the case for wheat) and by standardizing/aggregating all the processed consumption quantities to the primary level in the food processing variable. Table 16 shows the imputed food consumption for the example country.

Table 16: Imputing and adding food estimates to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	**0**	**26 720 000**	-	-	-	-	-
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	-	-	-	-
Bulgur	-	182 900	580 000	0	-	-	-	-	-	-	-
Breakfast cereals	-	312 500	217 300	0	-	-	-	-	-	-	-
Wheat starch	-	624 900	224 500	0	-	-	-	-	-	-	-
Wheat bran	-	258 900	2 343 700	0	-	-	-	-	-	-	-

Feed

In imputing the feed variable, it is assumed that some of the primary-level quantities are used as feed, along with most of the bran (which is a by-product of the flour production process). Feed requirements apply at the country level.

Calculation of feed requirements is described in section 3.4. The first step entails calculating feed requirements based on the number of animals, their needs and feeding efficiency. In the second step, the actual amount of compound and concentrate feed is calculated by applying country-specific intensification rates. In the third step, individual feedstuffs are allocated by availability, after all commodities whose sole purpose is as feed for animals (oilcakes and meals, DDGS, dregs, brans, etc.) have been deducted from the requirements. The remainder of feed requirements are then allocated to the FBS primary-level commodities (cereals, oil crops, etc.) according to their availability. Negligible amounts of bran may go into products such as breakfast cereals, but for the sake of simplicity, such quantities are ignored in this example.

The feed amount allocated to bran is computed from the food amount allocated to flour (as flour and bran are produced in the same process). Thus, flour production is converted into wheat by dividing it by the flour extraction rate, and bran production is computed by multiplying by the bran extraction rate. In the example, flour production is provided from official information; if it has to be calculated, the amount that will be processed into other commodities (bulgur, breakfast cereals, etc.) has first to be deducted from the wheat food quantity, to satisfy trade imbalances. The remainder is processed into flour, with bran as a by-product. The resulting estimates of feed use are shown in Table 17.

Table 17: Adding feed estimates to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	-	-	-	-
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	-	-	-	-
Bulgur	-	182 900	580 000	0	-	-	-	-	-	-	-
Breakfast cereals	-	312 500	217 300	0	-	-	-	-	-	-	-
Wheat starch	-	624 900	224 500	0	-	-	-	-	-	-	-
Wheat bran	**5 699 300**	258 900	2 343 700	0	-	-	3 614 500	-	-	-	-

Food losses and waste

FLW are losses from the post-harvest stage up to, but not including, the retail level. Neither retail nor household losses/waste are included in the FBS/SUA system. The methodology for calculating agricultural and food losses is being continuously revised and improved. Currently, the imputation methodology described in section 3.8 uses information about the perishable category of a commodity and the country/region to estimate a hierarchical linear regression model. It should be noted that losses are assumed to occur at only the primary level; processing losses are taken into account in the standardization process.

Table 18: Adding FLW to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	-	-	-	**560 300**
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	-	-	-	**0**
Bulgur	-	182 900	580 000	0	-	-	-	-	-	-	**0**
Breakfast cereals	-	312 500	217 300	0	-	-	-	-	-	-	**0**
Wheat starch	-	624 900	224 500	0	-	-	-	-	-	-	**0**
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	-	-	-	**0**

Seed

FBS compilers typically try to collect estimates of seed use from surveys or administrative data, or they simply impute seed use by multiplying standard seed rates with estimates of area harvested/sown. If neither of these sources is available, seed quantities can be imputed using the methodology described in section 3.9. The seed module fits a hierarchical linear model to seed data from previous years and uses global data. Seed is allotted only to the primary commodity (Table 19).

Table 19: Adding seed use to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	**1 904 200**	-	-	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	**0**	-	-	0
Bulgur	-	182 900	580 000	0	-	-	-	**0**	-	-	0
Breakfast cereals	-	312 500	217 300	0	-	-	-	**0**	-	-	0
Wheat starch	-	624 900	224 500	0	-	-	-	**0**	-	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	**0**	-	-	0

Industrial use

For most commodities, there is no industrial use so the quantity of this variable will be zero (Table 20). Estimates of industrial use are often taken from external sources (the methodology is described in section 3.7). This variable can be important when considering commodities related to biofuels (such as maize) and vegetable oils (such as palm oil). In the wheat commodity tree, the main commodity with industrial use is wheat starch. Considerable efforts have been made to devise a useful method of imputing industrial use, but none has been deemed fit for this purpose. FBS compilers at the country level are encouraged to collect the relevant data from their industry associations, as available.

Table 20: Adding industrial use to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock* **change**	*Food*	*Food* **processing**	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	1 904 200	-	**0**	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	0	-	-	0
Bulgur	-	182 900	580 000	0	-	-	-	0	-	-	0
Breakfast cereals	-	312 500	217 300	0	-	-	-	0	-	-	0
Wheat starch	-	624 900	224 500	0	-	-	-	0	-	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	0	-	-	0

Tourist consumption

National FBS compilers may have access to accurate data on food consumption by tourists and are encouraged to use these estimates. In the absence of such data, compilers may resort to the imputation process used here and described in section 3.6. To estimate tourist consumption, the approach uses tourist data from UNWTO to compute tourist flows, and consumption patterns for the previous year in the tourists' country of origin. Tourist consumption can be negative, as when many nationals travel abroad but no tourists enter the country. In this case, the country will have negative tourist consumption because more calories will be consumed abroad than locally. Table 21 shows the estimates of tourist consumption added to the SUA table for wheat.

Table 21: Adding tourist consumption to the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	1 904 200	**-39 800**	0	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	0	**0**	-	0
Bulgur	-	182 900	580 000	0	-	-	-	0	**0**	-	0
Breakfast cereals	-	312 500	217 300	0	-	-	-	0	**0**	-	0
Wheat starch	-	624 900	224 500	0	-	-	-	0	**0**	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	0	**0**	-	0

Standardization and balancing

Figure 19 shows the commodity tree for wheat.

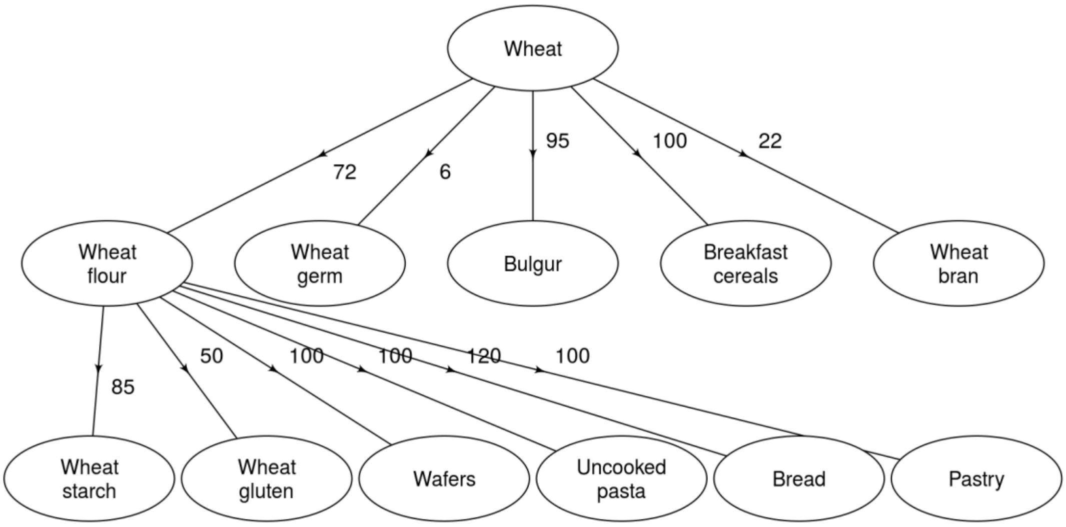

Figure 19: Commodity tree for wheat and products

Table 22 shows the pre-standardized SUA table for wheat.

Table 22: Initial, but complete pre-standardization table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock* **change**	*Food*	*Food* **processing**	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	1 904 200	-39 800	0	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	0	0	-	0
Bulgur	-	182 900	580 000	0	-	-	-	0	0	-	0
Breakfast cereals	-	312 500	217 300	0	-	-	-	0	0	-	0
Wheat starch	-	624 900	224 500	0	-	-	-	0	0	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	0	0	-	0

The initial food processing estimate is based on the module outlined earlier in this section; however, other information may need to be considered, such as trade imbalances (e.g. when exports are higher than production + imports) or official production quantities of processed commodities. From this additional information, it is possible to calculate the production quantities of each processed/derived commodity for which no estimate has yet been made (Table 23).

Table 23: Wheat SUA table after the first standardization steps

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 720 000	4 898 000	1 904 200	-39 800	0	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	-	-	0	0	-	0
Bulgur	**397 100**	182 900	580 000	0	-	-	-	0	0	-	0
Breakfast cereals	**0**	312 500	217 300	0	-	-	-	0	0	-	0
Wheat starch	**0**	624 900	224 500	0	-	-	-	0	0	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	-	3 614 500	0	0	-	0

As wheat starch is a by-product derived from wheat flour (as is wheat bran), it is often necessary first to ensure that the wheat flour in food processing covers any deficits of wheat starch. However, this is not the case in this example, as wheat starch imports exceed exports. All the processed product quantities can now therefore be standardized back to the food processing variable of wheat. The standardized quantities will, of course, be in primary commodity (wheat) equivalents. For example, suppose that 100 tonnes of a primary commodity produces 50 tonnes of the processed product (a 50 percent extraction rate), 50 tonnes of the processed product would be standardized back as 100 tonnes of wheat equivalent (Table 24).

Table 24: Accounting for derived products in the SUA table for wheat

Name	*Production (processed)*	*Wheat equivalent*
Wheat flour	18 650 000	25 910 000
Bulgur	397 100	418 000
Breakfast cereals	0	0
Wheat bran	5 699 300	25 910 000

The main requirement is in wheat flour and bran; these two commodities represent just one requirement (as bran production is consistent with wheat production). In the example, as flour production is an official estimate (and accounts for the vast majority of wheat utilization), the food processing variable is fixed for wheat and set at 26.3 million tonnes with a standard deviation of zero.

The next stage is to ensure that all of the appropriate by-products from processing of the various commodities are included (Table 25). For example, when processing wheat into flour, the by-products bran and germ must be accounted for, and the production numbers for these processed products must be in agreement.

Table 25: Complete, but still imbalanced SUA table for wheat

Name	***Production***	***Imports***	***Exports***	***Stock change***	***Food***	***Food processing***	***Feed***	***Seed***	***Tourist***	***Industrial***	***Loss***
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 330 000	4 898 000	1 904 200	-39 800	0	560 300
Wheat flour	18 650 000	341 500	572 800	0	-	**0**	-	0	0	-	0
Wheat germ	**1 554 300**	**0**	**0**	**0**	**-**	**0**	**-**	**0**	**0**	-	**0**
Bulgur	397 100	182 900	580 000	0	-	**0**	-	0	0	-	0
Breakfast cereals	0	312 500	217 300	0	-	**0**	-	0	0	-	0
Wheat starch	0	624 900	224 500	0	-	**0**	-	0	0	-	0
Wheat bran	5 699 300	258 900	2 343 700	0	-	**0**	3 355 500	0	0	-	0

Some of the SUA lines are not balanced because uses have not been allocated to excess supply. For these commodities, excess trade amounts should be allocated according to the variable that is most appropriate for the commodity concerned (or to several variables if the split shares at which a commodity is used are known).

Table 26: Allocating excess trade in the SUA table for wheat

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Food processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Wheat	54 420 000	1 999 100	32 790 000	-230 600	0	26 330 000	4 898 000	1 904 200	-39 800	0	560 300
Wheat flour	18 650 000	341 500	572 800	0	**18 420 000**	0	0	0	0	0	0
Wheat germ	1 554 300	0	0	0	0	0	**1 554 300**	0	0	0	0
Bulgur	397 100	182 900	580 000	0	0	0	0	0	0	0	0
Breakfast cereals	0	312 500	217 300	0	**95 200**	0	0	0	0	0	0
Wheat starch	0	624 900	224 500	0	0	0	0	0	0	**400 400**	0
Wheat bran	5 699 300	258 900	2 343 700	0	0	0	3 614 500	0	0	0	0

The next step is to aggregate this full table back into the primary commodity equivalent (wheat). The final quantity for wheat-equivalent production is simply the current quantity for wheat production. This is because “production” of bulgur (or any other processed product) is really the conversion of wheat into bulgur and not the actual production of bulgur. Thus, the reported quantity for production will always be production at the primary product level.

To standardize imports and exports, the imports and exports of the derived/processed commodities can be aggregated up to their primary equivalents by dividing by the extraction rates. These primary equivalents are added to the current quantity of imports and exports of wheat to derive the final, primary-equivalent import and export quantities of wheat.

Food processing is not usually standardized. In fact, this variable is usually included in the SUA table simply for allocating quantities when one commodity is converted into another. Thus, it is removed from the balance at this point. However, where a processed product is standardized into a different line of the FBS (e.g. barley can be processed into beer, which becomes a separate commodity in the final FBS), production of this processed commodity must be standardized into food processing. In the current FBS, this is reported as food manufacturing.

Quantities of feed commodities such as bran are not standardized back into their primary (wheat) equivalent as they are feed products. Thus, they are not reported in the FBS, but are instead reported in the commodity balances under the appropriate category, such as brans. For the remaining variables, standardization follows the same process as for trade. Table 27 shows the standardized table.

Table 27: Fully standardized, but still imbalanced FBS table for wheat

Name	Production	Imports	Exports	Stock change	Food	Feed	Seed	Tourist	Industrial	Loss
Wheat	54 420 000	3 999 600	34 780 000	-230 600	25 680 000	4 898 000	1 904 200	-39 800	654 300	560 300

It is now necessary to balance to satisfy the FBS equation of supply = utilization. For this balancing, the computed standard deviation of each variable needs to be extracted. Table 28shows the expected value and estimated standard deviation for each of the variables for wheat. The equation is initially not balanced, and will be balanced by adjusting figures according to their standard deviations. For example, a variable with a large standard deviation (low data confidence) can be adjusted substantially, while a variable with zero standard deviation will not be adjusted at all, such as official trade data (see more on this algorithm in section 3.2).

Table 28: Unbalanced FBS table for wheat, including measurement errors

Variable	*Production*	*Imports*	*Exports*	*Stock* change	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Mean	54 420 000	3 999 600	34 780 000	-230 600	25 680 000	4 898 000	1 904 200	-39 800	654 300	560 300
Standard deviation	489 800	0	0	89 900	0	244 900	228 500	-39 800	0	56 000

After balancing, some quantities are updated (and those with a standard deviation of zero remain unchanged). Therefore, the final table (Table 29) now reports on “wheat and products” as it includes wheat and all of its processed products.

Table 29: Accounting for inaccuracies in estimates of the various variables in the FBS/SUA table for wheat and products

Variable	Production	Imports	Exports	Stock change	Food	Feed	Seed	Tourist	Industrial	Loss
Mean	60 850 000	3 999 600	34 780 000	-447 200	25 680 000	3 289 100	503 500	-82 200	654 300	476 100
Standard deviation	489 800	0	0	89 900	0	244 900	228 500	-39 800	0	56 000

Calculating nutrient supplies

The standardization process converts quantities of processed products back into the equivalent quantities of wheat. In fact, standardization is all about aggregating quantities. All nutrients are calculated at the disaggregated SUA level, before standardization.

Returning to the SUA table, the nutritive factors of calories, fat and protein contents are applied to all SUA items for a food quantity. These nutritive factors are obtained from national sources or international standard tables. One GJ is a measure of energy equivalent to 1 billion joules, or roughly 239 000 calories and 1 Mg is 1 million grams. This process renders the nutrient supply levels shown in Table 30.

Table 30: Calculating nutrient supplies (calories, protein, fat)

Name	*Quantity*	*kJ energy/kg wheat*	*g protein/kg wheat*	*g fat/kg wheat*	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Wheat	0	14 200	123.40	18.65	0	0	0
Wheat flour	18 420 000	14 700	110.47	13.39	743 000	5 600	680
Breakfast cereals	95 200	NA	NA	NA	NA	NA	NA

Standardization of nutrients is now a simple last step: all the variables here (calories, fats and proteins) are purely additive, so the standardized calories, fats and proteins are simply the sum of the total calories, fats and proteins for each commodity (Table 31).

Table 31: Total nutrient content of wheat and wheat products

Commodity	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Wheat and products	743 000	5 600	680

To convert these figures into useful indicators, they are divided by the population of the country to obtain per capita data. Table 32 shows the results for a country with 600 million inhabitants.

Table 32: Per capita nutrient content of wheat and wheat products

Commodity	***Calories/person/day***	***g protein/person/day***	***g fat/person/day***
Wheat and products	296	9	1

4.2 Sugar

Establishing sugar balances for FBS/SUAs poses different challenges from those facing wheat balances. Sugar can be produced from two principal crops – cane and beet – and at different levels of final processing (raw, refined), which are associated with different by-products. In addition to the traditional sources of sugar, there are also a growing number of alternative sources for sweeteners such as high fructose corn syrup, which add to the availability of sugar and sweeteners but also to the complexity of building a consistent and standardized FBS for sugar.

As for wheat, the creation of the SUA/FBS balances starts with an empty table (Table 33), comprising all the variables/elements across all sugar crops and sugar products.

Table 33: Initial, empty SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar beet	-	-	-	-	-	-	-	-	-	-	-
Sugar cane	-	-	-	-	-	-	-	-	-	-	-
Sugar and syrups	-	-	-	-	-	-	-	-	-	-	-
Raw sugar	-	-	-	-	-	-	-	-	-	-	-
Refined sugar	-	-	-	-	-	-	-	-	-	-	-
Molasses	-	-	-	-	-	-	-	-	-	-	-

Production

For production data, the table is first filled in with available official figures (Table 34). In the example given in this section, it is assumed that production quantities are known for all primary products (beet and cane) so no imputation is needed. If there were missing quantities, data would be imputed as described in section 3.1 and as shown in the wheat example. Applying appropriate extraction rates for sugar crops allows estimation of sugar production in raw sugar equivalents.

Table 34: Adding production estimates of primary crops to the SUA table for sugar

Name	*Production*	*Imports*	*Exports*	*Stock Change*	*Food*	*Food Processing*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Sugar beet	**26 210 000**	-	-	-	-	-	-	-	-	-	-
Sugar cane	**26 510 000**	-	-	-	-	-	-	-	-	-	-
Sugar and syrups	-	-	-	-	-	-	-	-	-	-	-
Raw sugar	-	-	-	-	-	-	-	-	-	-	-
Refined sugar	-	-	-	-	-	-	-	-	-	-	-
Molasses	-	-	-	-	-	-	-	-	-	-	-

Trade

The detailed process for arriving at trade data for all products, including sugar, is laid out in section 3.2 and demonstrated for wheat in section 4.1. Application of this process renders the import and export estimates shown in Table 35, which features both production and trade and hence almost the entire supply side of the sugar balance.

Table 35: Adding trade estimates to the SUA table for sugar

Name	**Production**	**Imports**	**Exports**	**Stock change**	**Food**	**Food processing**	**Feed**	**Seed**	**Tourist**	**Industrial**	**Loss**
Sugar beet	26 210 000	**194 500**	**300**	-	-	-	-	-	-	-	-
Sugar cane	26 510 000	**9 700**	**860**	-	-	-	-	-	-	-	-
Sugar and syrups	-	**265 400**	**96 200**	-	-	-	-	-	-	-	-
Raw sugar	-	**10**	**194 800**	-	-	-	-	-	-	-	-
Refined sugar	-	**1 275 200**	**111 200**	-	-	-	-	-	-	-	-
Molasses	-	**464 200**	**236 500**	-	-	-	-	-	-	-	-

Stock changes

Stocks are generally held for a select number of primary-level products (such as wheat or rice), but not for sugar cane or beet, as both are highly perishable products. Sugar is therefore stored as raw or refined sugar. The stock change in the SUA table for sugar will ideally be a quantity measured within the country, but such estimates are seldom available. In most cases, at least an initial estimate for stock changes must be imputed using the linear regression model on historical stock change data described in section 3.3. Adding stock changes to the table completes the domestic supply side as shown in Table 36.

Table 36: Adding stock estimates to the SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar beet	26 210 000	194 500	300	**0**	-	-	-	-	-	-	-
Sugar cane	26 510 000	9 700	860	**0**	-	-	-	-	-	-	-
Sugar and syrups	-	265 400	96 200	**0**	-	-	-	-	-	-	-
Raw sugar	-	10	194 800	**0**	-	-	-	-	-	-	-
Refined sugar	-	1 275 200	111 200	**79 500**	-	-	-	-	-	-	-
Molasses	-	464 200	236 500	**0**	-	-	-	-	-	-	-

Food

The module for estimating food allocation uses food consumption estimates from the previous year and extrapolates these estimates forwards using changes in GDP and product-related income elasticities. It should be noted that these are only initial estimates of food consumption and that this is only one of many imputation methods. The final levels of food consumption will be determined in the balancing mechanism, and better estimates of sugar production and eventually availability could come from production statistics on raw and refined sugar from sugar mills.

In addition, sugar is a special case in that the sugar food balance includes not only the crop level (sugar cane and beet) but also raw sugar. Thus, the “primary level” for sugar is the raw sugar commodity. The food and food processing variables for raw sugar are added to the table (Table 37).

Table 37: Adding food and food processing to the SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar beet	26 210 000	194 500	300	0	-	-	-	-	-	-	-
Sugar cane	26 510 000	9 700	860	0	-	-	-	-	-	-	-
Sugar and syrups	-	265 400	96 200	0	-	-	-	-	-	-	-
Raw sugar	-	10	194 800	0	**1 513 800**	**9 795 900**	-	-	-	-	-
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	-	-	-	-
Molasses	-	464 200	236 500	0	-	-	-	-	-	-	-

Food losses and waste

FLW are estimated using the methodology described in section 3.8, unless loss quantities are measured by the country. Losses of only sugar beet and sugar cane are estimated and added to the table (Table 38).

Table 38: Adding food losses and waste to the SUA table for sugar

Name	***Production***	***Imports***	***Exports***	***Stock change***	***Food***	***Food processing***	***Feed***	***Seed***	***Tourist***	***Industrial***	***Loss***
Sugar beet	26 210 000	194 500	300	0	-	-	-	-	-	-	**205 500**
Sugar cane	26 510 000	9 700	860	0	-	-	-	-	-	-	**213 300**
Sugar and syrups	-	265 400	96 200	0	-	-	-	-	-	-	**0**
Raw sugar	-	10	194 800	0	1 513 800	9 795 900	-	-	-	-	**0**
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	-	-	-	**0**
Molasses	-	464 200	236 500	0	-	-	-	-	-	-	**0**

Seed

Seed use of sugar beet is assumed to be zero. However, an allocation for seed from sugar cane should be made (replanting cane every five to six years). In this case, it is not usually actual seed usage, but rather the cutting and replanting of some sugar cane plants, thus resulting in slightly lower availability than would have occurred if all the plants had been harvested. While this is technically not seed use, it simplifies the FBS to allocate it to the seed use variable (Table 39).

Table 39: Adding seed use to the SUA table for sugar

Name	**Production**	**Imports**	**Exports**	**Stock change**	**Food**	**Food processing**	**Feed**	**Seed**	**Tourist**	**Industrial**	**Loss**
Sugar beet	26 210 000	194 500	300	0	-	-	-	**0**	-	-	205 500
Sugar cane	26 510 000	9 700	860	0	-	-	-	**1 572 200**	-	-	213 300
Sugar and syrups	-	265 400	96 200	0	-	-	-	**0**	-	-	0
Raw sugar	-	10	194 800	0	1 513 800	9 795 900	-	**0**	-	-	0
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	**0**	-	-	0
Molasses	-	464 200	236 500	0	-	-	-	**0**	-	-	0

Industrial use

As in the wheat example, there is also industrial use of sugar and its by-products, which is increasing, largely because these products are used for biofuel production. This is particularly the case for sugar cane (and bagasse), but beets and molasses are also used for bioethanol production, notably in Europe. These allocations will be added later, when the sugar cane quantities are converted into processed products. Table 40therefore includes the industrial use of raw sugar only.

Table 40: Adding industrial use to the SUA table for sugar

Name	**Production**	**Imports**	**Exports**	**Stock change**	**Food**	**Food processing**	**Feed**	**Seed**	**Tourist**	**Industrial**	**Loss**
Sugar beet	26 210 000	194 500	300	0	-	-	-	0	-	-	205 500
Sugar cane	26 510 000	9 700	860	0	-	-	-	1 572 200	-	-	213 300
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	-	-	0
Raw sugar	-	10	194 800	0	1 513 800	9 795 900	-	0	-	**0**	0

Refined sugar	-	1 275 200	111 200	79 500	-	-	-	0	-	-	0
Molasses	-	464 200	236 500	0	-	-	-	0	-	-	0

Tourist consumption

As in the wheat example, tourist consumption of sugar is imputed with the methodology outlined in section 3.7. The results are inserted into the SUA table (Table 41). In this example, the imputed amount is negative, indicating that more calories are available in the country (a decrease in utilization is mathematically equivalent to an increase in supply) because inhabitants are consuming calories abroad.

Table 41: Adding tourist consumption to the SUA table for sugar

Name	**Production**	**Imports**	**Exports**	**Stock change**	**Food**	**Food processing**	**Feed**	**Seed**	**Tourist**	**Industrial**	**Loss**
Sugar beet	26 210 000	194 500	300	0	-	-	-	0	**0**	-	205 500
Sugar cane	26 510 000	9 700	860	0	-	-	-	1 572 200	**0**	-	213 300
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	**0**	-	0
Raw sugar	-	10	194 800	0	1 513 800	9 795 900	-	0	**-3 200**	0	0
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	0	**0**	-	0
Molasses	-	464 200	236 500	0	-	-	-	0	**0**	-	0

Feed

To determine feed estimates for raw sugar, all the commodities that are allocated entirely (or assumed to be allocated entirely) to feed must first be deducted from total feed requirements. This means that the molasses produced in the processing of sugar beet and cane must be deducted. It is important to note that when compiling the FBS, the quantities of all commodities allocated to feed must be computed before generating feed estimates for the commodities (wheat, raw sugar, etc.). Thus, FBS cannot be compiled one commodity tree at a time.

To compute the total amount of molasses, it is first necessary to compute the amounts of sugar beet and cane allocated to food processing. This is straightforward: the small amounts allocated to seed and losses are simply deducted from the supply to compute the amounts of sugar beet and cane allocated to food processing (Table 42).

Table 42: Adding food and food processing to the SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar beet	26 210 000	194 500	300	0	-	**26 200 000**	-	0	0	-	205 500
Sugar cane	26 510 000	9 700	860	0	-	**24 730 000**	-	1 572 200	0	-	213 300
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	0	-	0
Raw sugar	-	10	194 800	0	1 513 800	9 795 900	-	0	-3 200	0	0
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	0	0	-	0
Molasses	-	464 200	236 500	0	-	-	-	0	0	-	0

All the sugar beet and sugar cane quantities are now broken down into raw sugar and associated by-products. In this example, the production quantities of sugar crops are officially reported, so after minor allocations to seed and waste, these quantities are allocated to the production of processed commodities (Table 43). These production figures are therefore also considered official.

Table 43: Completing the sugar table for by-products of sugar production/processing

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	0	-	0
Raw sugar	**6 389 200**	10	194 800	0	1 513 800	9 795 900	-	0	-3 200	0	0
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	0	0	-	0
Molasses	**2 284 900**	464 200	236 500	0	-	-	-	0	0	-	0
Beet pulp	**1 834 200**	**0**	**0**	**0**	-	-	-	**0**	**0**	-	**0**
Bagasse	**6 183 700**	**0**	**0**	**0**	-	-	-	**0**	**0**	-	**0**

It is now necessary to determine the quantities of molasses and beet pulp (by-products of sugar processing) feed that are produced (Table 44). These figures inform the amount of feed that is allocated to the commodity being balanced (raw sugar). However, raw sugar is rarely used as a feed product, so only the two processed commodities, beet pulp and molasses, are allocated to feed use.

Table 44: Adding feed use to the SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	0	-	0
Raw sugar	6 389 200	10	194 800	0	1 513 800	9 795 900	**0**	0	-3 200	0	0
Refined sugar	-	1 275 200	111 200	79 500	-	-	-	0	0	-	0
Molasses	2 284 900	464 200	236 500	0	-	-	**2 512 500**	0	0	-	0
Beet pulp	1 834 200	0	0	0	-	-	**1 834 200**	0	0	-	0
Bagasse	6 183 700	0	0	0	-	-	-	0	0	-	0

Standardization and balancing

Figure 20 shows the commodity tree for sugar crops and products.

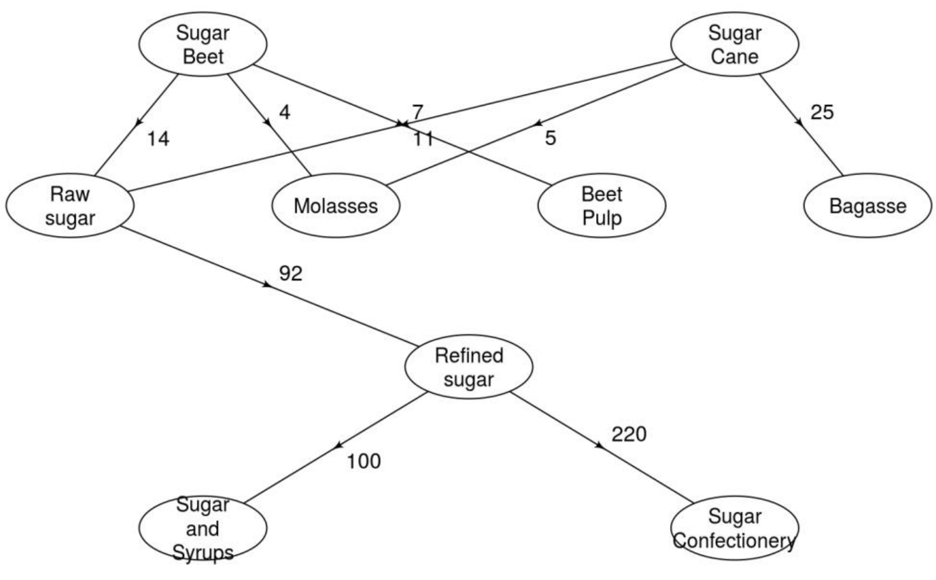

Figure 20: Commodity tree for sugar crops and products

The sugar cane and sugar beet commodities have already been eliminated, so the next task is to roll up processed sugar commodities into raw sugar based on Table 45.

Table 45: Initial pre-standardized SUA table for sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar and syrups	-	265 400	96 200	0	-	-	-	0	0	-	0
Raw sugar	6 389 200	10	194 800	0	1 513 800	9 795 900	0	0	-3 200	0	0
Refined sugar	-	1 275	111 200	79 500	-	-	-	0	0	-	0

		200									
Molasses	2 284 900	464 200	236 500	0	-	-	2 512 500	0	0	-	0
Beet pulp	1 834 200	0	0	0	-	-	1 834 200	0	0	-	0
Bagasse	6 183 700	0	0	0	-	-	-	0	0	-	0

The next step is to balance the processed commodities by creating production quantities. The standardization of the required production quantities would then usually be compared with the estimated quantity of food processing for the raw sugar commodity to ensure that it covers supply/utilization deficits. However, in this example, there are no trade deficits, and the production figure for raw sugar is not official (unlike in the wheat example, which had official data on production of wheat flour). To maintain consistency between the food processing variable and the production of processed commodities, the food processing of raw sugar is therefore allocated to the production of refined sugar (Table 46).

Table 46: Complete, but still unbalanced SUA table for refined sugar

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar and syrups	-	265 400	96 200	0	-	0	-	0	0	-	0
Raw sugar	6 389 200	10	194 800	0	1 513 800	9 795 900	0	0	-3 200	0	0
Refined sugar	**9 012 200**	1 275 200	111 200	79 500	-	0	-	0	0	-	0
Molasses	2 284 900	464 200	236 500	0	-	0	2 512 500	0	0	-	0
Beet pulp	1 834 200	0	0	0	-	0	1 834 200	0	0	-	0
Bagasse	6 183 700	0	0	0	-	0	-	0	0	-	0

Some SUA lines are not balanced, because uses have not been allocated to excess supply. For these commodities, the excess trade amount should be allocated to the variable that makes the most sense for the particular commodity (or to several variables if the split shares at which the commodity is utilized are known) (Table 47).

Table 47: SUA table for refined sugar, balanced at the processing level

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Sugar and syrups	-	265 400	96 200	0	**169 200**	0	0	0	0	0	0
Raw sugar	6 389 200	10	194 800	0	1 513 800	9 795 900	0	0	-3 200	0	0
Refined sugar	9 012 200	1 275 200	111 200	79 500	**10 100 000**	0	0	0	0	0	0
Molasses	2 284 900	464 200	236 500	0	0	0	2 512 500	0	0	0	0
Beet pulp	1 834 200	0	0	0	0	0	1 834 200	0	0	0	0
Bagasse	6 183 700	0	0	0	0	0	0	0	0	**6 183 700**	0

The next step is to aggregate this full table back into the primary commodity equivalent (raw sugar), following the same aggregation/standardization process as is outlined in the wheat example. It should be noted that molasses, beet pulp and bagasse are standardized to a different primary equivalent in the commodity balances and so are not considered here. The results are presented in Table 48.

Table 48: Initial, unbalanced table for raw sugar containing all standardized products

Name	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Raw sugar	6 389 200	1 674 600	420 200	86 400	12 670 000	0	0	-3 200	0	0

It is now necessary to balance to satisfy the FBS equation of supply = utilization. The first step is to extract the computed standard deviation of each variable. The standard deviation is determined by the data source: for official data, a standard deviation of 0 is applied; semi-official data have a higher standard deviation; and estimated quantities an even higher one. In this example, production and trade quantities are official while all other quantities are estimated. These assumptions lead to Table 49.

Table 49: Unbalanced table for raw sugar, including standard deviations

Variable	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Mean	6 389 200	1 674 600	420 200	86 400	12 670 000	0	0	-3 200	0	0
Standard deviation	0	0	0	24 200	1 203 600	0	0	-3 200	0	0

Balancing of Tables 47 and 48 leads to the quantities shown in Table 50. Note that the food variable receives most of the adjustment because it has a substantially higher variability than the other variables.

Table 50: Final balanced table for raw sugar, including all processed commodities

Variable	*Production*	*Imports*	*Exports*	*Stock change*	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Mean	6 389 200	1 674 600	420 200	84 300	7 562 500	0	0	-3 300	0	0
Standard deviation	0	0	0	24 200	1 203 600	0	0	-3 200	0	0

After balancing, some quantities are updated (and those with a standard deviation of zero remain unchanged). In the wheat example, the aggregated table was placed directly into the FBS. In this case, sugar is not the only commodity included in the primary/FBS level, which reports on “Sugar and sweeteners”. Thus, balances must also be performed for the other commodities (honey and artificial sweeteners) and added to this sugar balance to create the final FBS row.

The next step is to calculate the calorie, fat and protein contents. As in the wheat example, these values can be calculated as soon as the food values for SUA-level commodities are available (Table 51). To do this calculation, the nutritive factors of the calorie, fat and protein contents are applied to all SUA items with a non-zero food quantity. However, in this example, the food quantity for the standardized commodity (raw sugar) was adjusted down; to ensure consistency, all the SUA food quantities must be adjusted by the same percentage. Thus, the values reported in Table 51are all roughly 40 percent less than their original values: the pre-balanced food value was roughly 12.7 million and the balanced food value is roughly 7.6 million, a decrease of about 40%. As in the wheat example, 1 GJ is a measure of energy equal to 1 billion joules, or roughly 239 000 calories, and 1 Mg is 1 million grams.

Table 51: Calorie, protein and fat quantities from consumed sugar products

Name	*Quantity*	*kJ energy/kg*	*g protein/kg*	*g fat/kg*	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Sugar and syrups	101 000	NA	NA	NA	NA	NA	NA
Raw sugar	903 600	NA	NA	NA	NA	NA	NA
Refined sugar	6 026 600	17 000	0	0	280 600	0	0
Caloric beverages	142 200	NA	NA	NA	NA	NA	NA

Standardization of nutrients is now a simple last step: all the variables (calories, fats and proteins) are purely additive, so the standardized calories, fats and proteins are simply the sums of the total calories, fats and proteins for each commodity (Table 52).

Table 52: Total calories, proteins and fats from sugar

Commodity	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Sugar	280 600	0	0

To convert these figures into something more meaningful, they are divided by the population of the country to generate daily per capita figures. Assuming that the country has 600 million inhabitants, the per capita estimates shown in Table 53 are obtained.

Table 53: Daily per capita calories, proteins and fats from sugar

Commodity	***Calories/person/day***	***g protein/person/day***	***g fat/person/day***
Sugar	112	0	0

4.3 Palm oil

The step-by-step process described for the wheat and sugar examples provided ample opportunities to demonstrate some of the specificities and intricacies of compiling SUA and FBS tables. While it would be impossible to anticipate all the possible challenges for FBS compilers at the country level, it is worth describing and addressing some of these additional challenges by describing the compilation of SUA/FBS tables for another product, such as palm oil. The basic strategy for building up the SUA and FBS tables follows the same process as already outlined for wheat and sugar, starting with an empty table (Table 54).

Table 54: Initial, empty SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	-	-	-	-	-	-	-	-	-	-	-
Palm oil	-	-	-	-	-	-	-	-	-	-	-
Oil of palm kernels	-	-	-	-	-	-	-	-	-	-	-
Margarine and shortening	-	-	-	-	-	-	-	-	-	-	-
Fatty acids	-	-	-	-	-	-	-	-	-	-	-
Boiled, dehydrated oil	-	-	-	-	-	-	-	-	-	-	-
Fat preparations nes	-	-	-	-	-	-	-	-	-	-	-
Hydrogenated oils	-	-	-	-	-	-	-	-	-	-	-

nes = not elsewhere specified.

Production

For production data, the table is first filled in with any available official figures. In this example, it is assumed that the production quantities are known for all the primary products, and thus no imputation is done. Production data are also available for two processed commodities: margarine and fat preparations (Table 55).

Table 55: Adding production data to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	**0**	-	-	-	-	-	-	-	-	-	-
Palm oil	**0**	-	-	-	-	-	-	-	-	-	-
Oil of palm kernels	-	-	-	-	-	-	-	-	-	-	-
Margarine and shortening	**3 714 000**	-	-	-	-	-	-	-	-	-	-
Fatty acids	-	-	-	-	-	-	-	-	-	-	-
Oil boiled, dehydrated	-	-	-	-	-	-	-	-	-	-	-
Fat preparations nes	**194 100**	-	-	-	-	-	-	-	-	-	-
Hydrogenated oils	-	-	-	-	-	-	-	-	-	-	-

Trade

The relevant steps for collecting and compiling trade data are presented in section 3.2 and are described in detail in the wheat example. In the palm oil example, the available trade figures are inserted into the table (Table 56).

Table 56: Adding trade estimates to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	**0**	**0**	-	-	-	-	-	-	-	-
Palm oil	0	**1 087 600**	**94 900**	-	-	-	-	-	-	-	-
Oil of palm kernels	-	**0**	**0**	-	-	-	-	-	-	-	-
Margarine and shortening	3 714 000	**0**	**0**	-	-	-	-	-	-	-	-
Fatty acids	-	**0**	**0**	-	-	-	-	-	-	-	-
Oil boiled, dehydrated	-	**0**	**0**	-	-	-	-	-	-	-	-
Fat preparations nes	194 100	**0**	**0**	-	-	-	-	-	-	-	-
Hydrogenated oils	-	**0**	**0**	-	-	-	-	-	-	-	-

Stock changes

Stocks are generally held for a select number of primary-level products (such as wheat or rice). Similar to refined sugar, margarine is another exception to this rule: a processed product of which stocks are occasionally held (Table 57).

Table 57: Adding stock changes to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	**0**	-	-	-	-	-	-	-
Palm oil	0	1 087 600	94 900	**0**	-	-	-	-	-	-	-
Oil of palm kernels	-	0	0	**0**	-	-	-	-	-	-	-
Margarine and shortening	3 714 000	0	0	**3 900**	-	-	-	-	-	-	-
Fatty acids	-	0	0	**0**	-	-	-	-	-	-	-
Oil boiled, dehydrated	-	0	0	**0**	-	-	-	-	-	-	-
Fat preparations nes	194 100	0	0	**0**	-	-	-	-	-	-	-
Hydrogenated oils	-	0	0	**0**	-	-	-	-	-	-	-

Food

The module for estimating food allocation uses food consumption estimates from the previous year and extrapolates these estimates forwards using changes in GDP and product-related income elasticities. The food variable is generally reported only at the primary level, as the food processing variable is estimated by standardizing the food quantities for all processed commodities. However, in some cases (such as this example), there may be no available data on the primary product and a large import value for a processed product. It therefore makes sense to standardize food and food processing to the primary level, as the processed product is likely to be created at this level.

In this example, supposing that the food module estimated margarine consumption at 3.7 million tonnes, this estimate would be standardized back to the food processing variable for oil of palm fruit for the balancing. However, as margarine can be created from many different commodities, the margarine food amount must be standardized back to all the possible commodities that can be processed to create margarine. For example, the country may have only oilseeds and oil crops for processing into margarine, as in Table 58.

Table 58: Primary products (oilseeds and oil crops) used to produce margarine

Commodity	*Availability*	***Equivalent* availability for m*argarine***	*Share*
Oil-palm fruit	0	0	0%
Palm oil	993 000	1 122 100	30.1%
Soybeans	3 874 100	788 000	21.1%
Sunflower seed	3 930 500	1 821 000	48.8%

The equivalent availability for margarine is derived from the availability of the primary commodity multiplied by the extraction rate(s) for converting the primary commodity into margarine. The share is then computed from the availability: 30.1 percent of the total availability of primary products comes from palm oil, so it is assumed that 30.1 percent of the margarine is produced from palm oil. This is, of course, just an approximation, but in the absence of country-specific information it is a reasonable assumption. Thus, 30.1 percent of the 3.7 million tonnes of margarine production must come from palm oil, requiring 1.1 million tonnes of palm oil. This translates into 980 000 tonnes of palm oil (dividing by the extraction rate). If there is also a small palm oil requirement for fat preparations, the final food processing estimate for oil-palm fruit may be higher, say about 1.2 million tonnes (Table 59).

Table 59: Adding food processing to the SUA table for palm oil

Name	**Production**	**Imports**	**Exports**	**Stock change**	**Food**	**Food processing**	**Feed**	**Seed**	**Tourist**	**Industrial**	**Loss**
Oil-palm fruit	0	0	0	0	**0**	-	-	-	-	-	-
Palm oil	0	1 087 600	94 900	0	**0**	**1 159 700**	-	-	-	-	-
Oil of palm kernels	-	0	0	0	-	-	-	-	-	-	-
Margarine and shortening	3 714 000	0	0	3 900	-	-	-	-	-	-	-
Fatty acids	-	0	0	0	-	-	-	-	-	-	-
Oil boiled, dehydrated	-	0	0	0	-	-	-	-	-	-	-
Fat preparations nes	194 100	0	0	0	-	-	-	-	-	-	-
Hydrogenated oils	-	0	0	0	-	-	-	-	-	-	-

To avoid later confusion, the production value of margarine and shortening is also adjusted (as only 37.5 percent of this production will be covered by oil-palm fruit). Table 60 provides the updated information, inclusive of production estimates for margarine and shortening.

Table 60: Adding production of processed products to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	-	-	-	-	-
Palm oil	0	1 087 600	94 900	0	0	1 159 700	-	-	-	-	-
Oil of palm kernels	-	0	0	0	-	-	-	-	-	-	-
Margarine and shortening	**1 117 000**	0	0	3 900	-	-	-	-	-	-	-
Fatty acids	-	0	0	0	-	-	-	-	-	-	-
Oil boiled, dehydrated	-	0	0	0	-	-	-	-	-	-	-
Fat preparations nes	194 100	0	0	0	-	-	-	-	-	-	-
Hydrogenated oils	-	0	0	0	-	-	-	-	-	-	-

Food losses and waste

FLW are estimated using the methodology described in section 3.8, unless the country measures loss quantities. The FLW that occur during the processing of palm oil are taken into account in the standardization procedure. FLW for oil-palm fruit are likely to be greater than zero, but only if the country produces palm fruit. As this is not the case for the country in the example, the FLW estimates should be assumed to be zero (Table 61). If the country produced palm fruit, the FLW would need to be imputed or, preferably, measured with appropriate methods, and reported in representative surveys.

Table 61: Adding FLW to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	-	-	-	-	**0**
Palm oil	0	1 087 600	94 900	0	0	1 159 700	-	-	-	-	**0**
Oil of palm kernels	-	0	0	0	-	-	-	-	-	-	**0**
Margarine and shortening	1 117 000	0	0	3 900	-	-	-	-	-	-	**0**
Fatty acids	-	0	0	0	-	-	-	-	-	-	**0**
Oil boiled, dehydrated	-	0	0	0	-	-	-	-	-	-	**0**
Fat preparations nes	194 100	0	0	0	-	-	-	-	-	-	**0**
Hydrogenated oils	-	0	0	0	-	-	-	-	-	-	**0**

Seed

As the country in the example has no production of oil-palm fruit, no amount is allocated to seed. However, even if the country produced palm fruit, the amount would be zero, as propagation takes place through the planting of new trees. The table can therefore be filled in with zeros (Table 62).

Table 62: Adding seed use to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	-	**0**	-	-	0
Palm oil	0	1 087 600	94 900	0	0	1 159 700	-	**0**	-	-	0
Oil of palm kernels	-	0	0	0	-	-	-	**0**	-	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	-	-	**0**	-	-	0
Fatty acids	-	0	0	0	-	-	-	**0**	-	-	0
Oil boiled, dehydrated	-	0	0	0	-	-	-	**0**	-	-	0
Fat preparations nes	194 100	0	0	0	-	-	-	**0**	-	-	0
Hydrogenated oils	-	0	0	0	-	-	-	**0**	-	-	0

Industrial use

Industrial use can account for a major share of palm oil use, and palm oil has become an increasingly important input for many non-food industries. As well as its use in biodiesel and as a replacement for or additive to heating fuel (e.g. in district heating systems), palm oil is also an increasingly popular raw material for the production of soaps, detergents, paints and cosmetics. As already outlined in the wheat and sugar examples, the amounts of raw materials used for industrial purposes are difficult to gauge, and thus to impute; in addition, the amounts are likely to change rapidly with policy incentives and energy prices, both of which are difficult to predict. In short, to estimate the amounts of palm oil used in industry, there is no real alternative to data collection. For the sake of simplicity, industrial use of palm oil has been assumed to be zero in this example (Table 63).

Table 63: Adding industrial use to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	-	0	-	**0**	0
Palm oil	0	1 087 600	94 900	0	0	1 159 700	-	0	-	-	0
Oil of palm kernels	-	0	0	0	-	-	-	0	-	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	-	-	0	-	-	0
Fatty acids	-	0	0	0	-	-	-	0	-	-	0
Oil boiled, dehydrated	-	0	0	0	-	-	-	0	-	-	0
Fat preparations nes	194 100	0	0	0	-	-	-	0	-	-	0
Hydrogenated oils	-	0	0	0	-	-	-	0	-	-	0

Tourist consumption

In this example, it is assumed that small numbers of tourists leave or enter the country, so the allocation of palm oil to tourist consumption is roughly zero (Table 64).

Table 64: Adding tourist consumption to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	-	0	**0**	0	0
Palm oil	0	1 087 600	94 900	0	0	1 159 700	-	0	**0**	-	0
Oil of palm kernels	-	0	0	0	-	-	-	0	**0**	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	-	-	0	**0**	-	0
Fatty acids	-	0	0	0	-	-	-	0	**0**	-	0
Oil boiled, dehydrated	-	0	0	0	-	-	-	0	**0**	-	0
Fat preparations nes	194 100	0	0	0	-	-	-	0	**0**	-	0
Hydrogenated oils	-	0	0	0	-	-	-	0	**0**	-	0

Feed

In the palm oil commodity tree, only one element is allocated to feed (cakes of palm kernels). This commodity is a by-product of the processing of oil of palm kernels. However, in the example, no palm kernels are available, so neither oil of palm kernels nor cakes of palm kernels are produced. Palm oil itself is never fed directly to animals, so no commodity is allocated to feed. Feed is therefore assumed to be zero (Table 65).

Table 65: Adding feed use to the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	**0**	0	0	0	0
Palm oil	0	1 087 600	94 900	0	0	1 159 700	**0**	0	0	-	0
Oil of palm kernels	-	0	0	0	-	-	**0**	0	0	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	-	**0**	0	0	-	0
Fatty acids	-	0	0	0	-	-	**0**	0	0	-	0
Oil boiled, dehydrated	-	0	0	0	-	-	**0**	0	0	-	0
Fat preparations nes	194 100	0	0	0	-	-	**0**	0	0	-	0
Hydrogenated oils	-	0	0	0	-	-	**0**	0	0	-	0

Standardization and balancing

As for all other commodities, producing the detailed entries for the SUA palm oil table is just the first step towards generating balanced, aggregated and standardized data for the FBS. The next important steps are first to aggregate/standardize the SUA tables and then to balance supply and utilization. These steps are laid out in this subsection. To understand the standardization process, it is important to recall the processing structure, which is illustrated as a commodity tree structure (Figure 21).

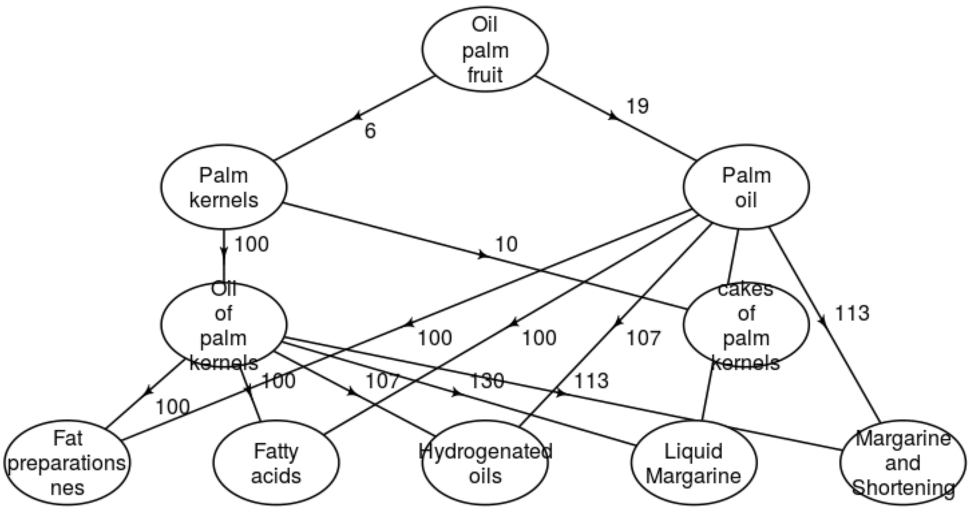

Figure 21: Commodity tree for palm oil and products

The SUA table produced so far (Table 66) serves as the basis for the standardization process and must contain estimates for all the input variables for this process.

Table 66: Initial, pre-standardized SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	-	0	0	0	0	0
Palm oil	0	1 087 600	94 900	0	0	1 159 700	0	0	0	-	0
Oil of palm kernels	-	0	0	0	-	-	0	0	0	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	-	0	0	0	-	0
Fatty acids	-	0	0	0	-	-	0	0	0	-	0
Oil boiled, dehydrated	-	0	0	0	-	-	0	0	0	-	0
Fat preparations nes	194 100	0	0	0	-	-	0	0	0	-	0
Hydrogenated oils	-	0	0	0	-	-	0	0	0	-	0

The next step is to balance the processed commodities by creating production quantities. However, in this example, official production figures are available for the two main uses of oil-palm fruit in the country: margarine and fat preparations. As these figures are official and account for the majority of the end uses of oil-palm fruit, they should be used as the basis for updating the food processing estimate (Table 67).

Table 67: Updating the SUA table for palm oil with food processing information

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	0	0	0	0	0	0
Palm oil	0	1 087 600	94 900	0	0	**1 182 600**	0	0	0	-	0
Oil of palm kernels	-	0	0	0	-	0	0	0	0	-	0
Margarine and shortening	1 117 000	0	0	3 900	-	0	0	0	0	-	0
Fatty acids	-	0	0	0	-	0	0	0	0	-	0
Oil boiled, dehydrated	-	0	0	0	-	0	0	0	0	-	0
Fat preparations nes	194 100	0	0	0	-	0	0	0	0	-	0
Hydrogenated oils	-	0	0	0	-	0	0	0	0	-	0

Some of the SUA lines are not balanced because uses have not been allocated to excess supply. For these commodities, the excess trade amount should be allocated to the variable that makes the most sense for the particular commodity (or to several variables if the split shares at which the commodity is used are known). These operations are reflected in Table 68.

Table 68: Removing remaining imbalances from the SUA table for palm oil

Name	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Tourist	Industrial	Loss
Oil-palm fruit	0	0	0	0	0	0	0	0	0	0	0
Palm oil	0	1 087 600	94 900	0	0	1 182 600	0	0	0	0	0
Oil of palm kernels	-	0	0	0	0	0	0	0	0	0	0
Margarine and shortening	1 117 000	0	0	3 900	**1 113 100**	0	0	0	0	0	0
Fatty acids	-	0	0	0	0	0	0	0	0	0	0
Oil boiled, dehydrated	-	0	0	0	0	0	0	0	0	0	0
Fat preparations nes	194 100	0	0	0	**194 100**	0	0	0	0	0	0
Hydrogenated oils	-	0	0	0	0	0	0	0	0	0	0

The next step is to aggregate this full table back into the primary commodity equivalent (palm oil). The same aggregation/standardization process as outlined in the wheat example is followed to obtain the complete balance for palm oil in primary equivalents (Table 69).

Table 69: Aggregated FBS balance for palm oil

Name	*Production*	*Imports*	*Exports*	*Stock* **change**	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Palm oil	0	1 087 600	94 900	3 400	1 179 200	0	0	0	0	0

It is now necessary to balance to satisfy the FBS equation of supply = utilization. The first step is to extract the computed standard deviation of each variable (Table 70). The standard deviation is determined by the data source: for official data, a standard deviation of 0 is applied; semi-official data have a higher standard deviation; and estimated quantities an even higher one. In this example, production and trade quantities are official, while all other quantities are estimates.

Table 70: FBS balances for palm oil, inclusive of measurement errors

Variable	*Production*	*Imports*	*Exports*	*Stock* **change**	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Mean	0	1 087 600	94 900	3 400	1 179 200	0	0	0	0	0
Standard deviation	0	0	0	1 400	59 000	0	0	0	0	0

Balancing of Table 69 generates the quantities shown in Table 71. The food variable receives most of the adjustment because it has a substantially higher variability than the other variables.

Table 71: Fully balanced FBS table for palm oil

Variable	*Production*	*Imports*	*Exports*	*Stock* **change**	*Food*	*Feed*	*Seed*	*Tourist*	*Industrial*	*Loss*
Mean	0	1 087 600	94 900	3 300	989 400	0	0	0	0	0
Standard deviation	0	0	0	1 400	59 000	0	0	0	0	0

After balancing, some quantities are updated, and those with a standard deviation of zero remain unchanged. As in the wheat example, the aggregated table for palm oil can be placed directly into the FBS. However, if there were any data on oil of palm kernels, it would be necessary to perform a standardization and balance for that commodity and include it as a separate FBS line.

As in the wheat and sugar examples, the calorie, fat and protein contents are calculated using the food consumption values at the SUA level (Table 72). For this calculation, the nutritive factors of the calorie, fat and protein contents are applied to all the SUA items with a non-zero food quantity. However, in this example, the food quantity for the standardized commodity was adjusted down. To ensure consistency, all the SUA food quantities must therefore be adjusted by the same percentage. As in the wheat example, 1 GJ is a measure of energy equal to 1 billion joules, or roughly 239 000 calories, and 1 Mg is 1 million grams.

Table 72: Calculating the nutrient content of the palm oil food supply

Name	*Quantity*	*kJ* **energy**/*kg*	*g* **protein**/*kg*	*g* **fat**/*kg*	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Palm oil	0	37 000	0.00	1000.00	0	0	0
Margarine and shortening	933 900	31 800	1.68	856.79	81 300	5	2 200
Fat preparations nes	162 900	NA	NA	NA	NA	NA	NA

Standardization of nutrients is now a simple last step: all the variables (calories, fats and proteins) are purely additive, so the standardized calories, fats and proteins are simply the sums of the total calorie, fat and protein contents of each commodity (Table 73).

Table 73: Total nutrient supplies from palm oil

Commodity	*Energy (GJ/day)*	*Protein (Mg/day)*	*Fat (Mg/day)*
Palm oil	81 300	5	2 200

To convert these figures into meaningful data on daily per capita consumption, they are divided by the population of the country, which is assumed to be 600 million people in this example (Table 74).

Table 74: Per capita nutrient supplies from palm oil

Commodity	***Calories/person/day***	***g* protein*/person/day***	***g* fat*/person/day***
Palm oil	32	0	4

5. Processing of agricultural trade data

Through its FAOSTAT corporate statistical database, FAO provides access to the detailed agricultural trade data of many countries of the world. These trade data sets have been standardized by the FAO Statistics Division (ESS) and are available by quantity and value, imports and exports, reporter and partner country, and at different disaggregation levels within the commodity classification system of the Harmonized System (HS).[50] National FBS compilers are encouraged to make use of these trade data sets for the production of their FBS. More generally, countries are encouraged to compare their national agricultural trade with information from other countries, to study regional and global patterns, and – particularly – to make use of mirrored data from trading partners for cross-checking national figures or filling in missing data. These data sets can also be used to adjust national data where required, based on the commodity unit-value concept.

United Nations Statistics Department tariff line data set

ESS receives trade data for most countries from UNSD, which collects the data from countries and stores them in the UN COMTRADE database.[51] These data sets are usually provided as originally reported by national customs offices and are by detailed flow (as described in the previous paragraph). The level of commodity detail is country-specific, with some countries reporting at the basic 6-digit level of the HS, while others go up to as much as 12-digit HS detail.

The global coverage, by value, of this data set is more than 90 percent in any given year. A typical non-standardized country trade data set, at the 10-digit HS level, is provided in Table 75.

Table 75: Basic trade information by commodity, partner and reporter

[50] While only flows at the 6-digit HS level follow the same classification across all countries, ESS plans also to make flows at a higher level of disaggregation (the tariff line level) available.

[51] HTTP://UNSTATS.UN.ORG/UNSD/COMTRADE_ANNOUNCEMENT.HTM

Year	Reporter	Partner	HS code	Flow	Weight	Quantity	Q-unit	Value
2011	842	484	8501314000	Import		33 498 070	5	366 057 280
2011	842	484	7321190040	Import		242	5	21 321
2011	842	591	7326908588	Export	1 896	1 896	8	13 011
2011	842	792	6104499060	Import	829	232	11	78 839
2011	842	699	2907191000	Import	20 000	20 000	8	128 598

Reporters and trade partners are represented by the three-digit numerical codes used by UNSD, based on the international standard M49 codes. The flows indicate imports or exports, but can also be re-imports or re-exports. Weight is in kilograms, while the quantity indicates other units of measurement such as number of items or volume in litres. The full list of units and their descriptions is in Annex I of UN COMTRADE's quantity and weight data.[52] The commodity codes, as mentioned previously, follow an internationally comparable standard up to the 6-digit level, but become country-specific at lower levels (8 digits and beyond).[53] Table 76 provides an example of the HS classification at the 10-digit level reported by a country.

Table 76: Example of trade flow disaggregation at the 10-digit HS level

[52] HTTP://UNSTATS.UN.ORG/UNSD/TRADEKB/KNOWLEDGEBASE/QUANTITY-AND-WEIGHT-DATA-IN-UN-COMTRADE

[53] Descriptions of country-specific HS codes are provided by the European Commission's Market Access Database and copyrighted by Mendel Verlag, Germany: http://madb.europa.eu

HS	Description
02	Chapter 2 Meat and edible meat offal
0202	Meat of bovine animals, frozen:
0202.10	• Carcases and half-carcases:
0202.10.05	– Described in general note 15 of the tariff schedule:
0202.10.05.10	• Veal
0202.10.05.90	• Other

5.1 Key processing steps

The process outlined in this section describes the key processing steps employed by FAO to prepare trade data sets for dissemination and for use in the SUA/FBS process. FBS compilers at the country level may want to review these steps, evaluate their suitability for the country's domestic data and other purposes, and either develop their own trade data processing system or simply use the data made available by FAO.

FAO's processing of trade data is organized around the steps described in the following subsections. Figure 22 provides an overview of the process.

Pre-validation

In a first step, a series of programs for identifying missing quantity data are run and validations of unit value-based outliers are performed based on the findings. These checks ensure that the data meet the minimum requirements for processing. They also provide a first opportunity for resolving data problems and preventing low-quality information from being used throughout the data processing.

Data set standardization

Standardization of the data set includes preparatory steps for facilitating data processing. These steps include basic formatting (e.g. unifying the two quantity columns into one column), converting units of measurement (e.g. to tonnes, numbers, etc.), linking country codes to country names, and conversion from the HS to the FAOSTAT commodity list (FCL) or the Central Product Classification (CPC).[54] When successfully completed, these steps will produce the FAO standardized trade data set.

These steps are applied to the data from all reporting countries. Once completed, it is also possible to calculate regional and global median unit values for every commodity and for a specific year. The unit value is defined as the price of the commodity divided by the quantity (number of units).

[54] FAO uses CPC version 2.1 expanded: http://www.fao.org/economic/ess/ess-standards/commodity/en/

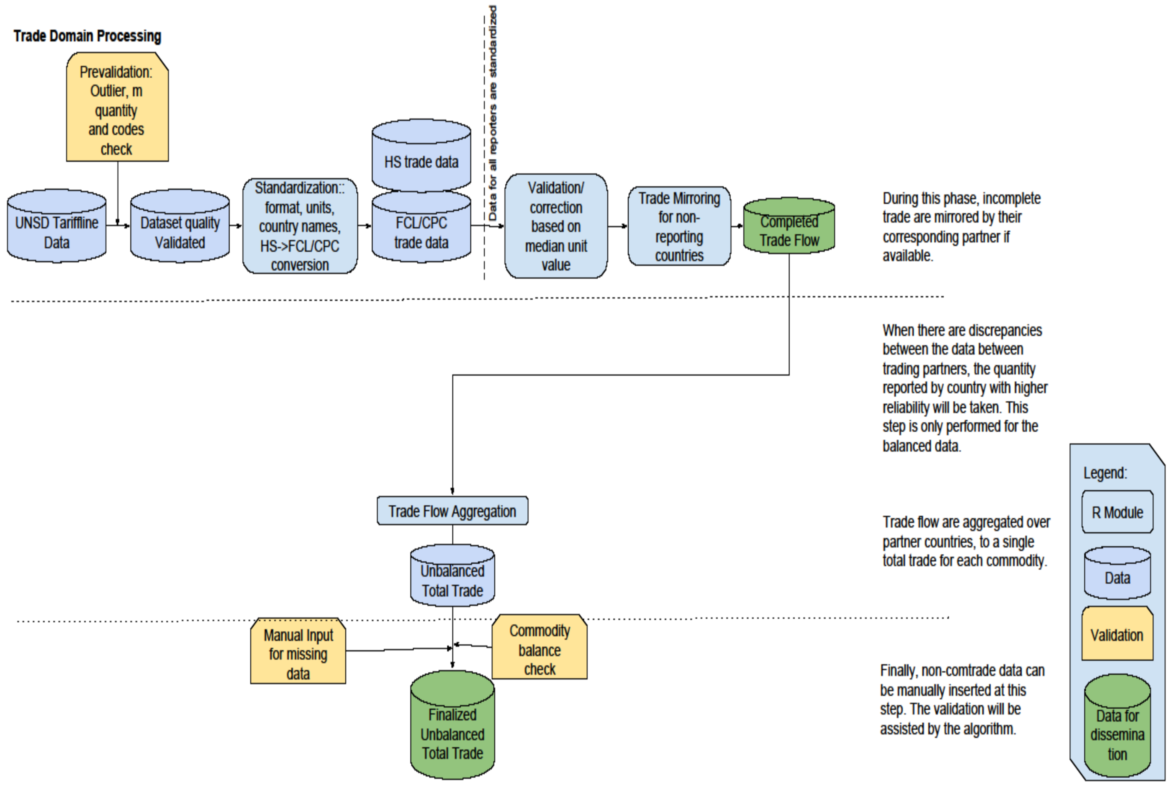

Figure 22: Processing of trade data at FAO

Validation of trade data

Based on the findings of the missing data and outlier report, the following further pre-validation steps can be undertaken:

1. Estimate missing quantities (and very rarely missing monetary values) based on the median unit value of the reporter's trade flow.
2. Adjust quantity outliers (out-of-range values), again based on the median unit value of the reporter's trade flow. Quantity errors are often the result of incorrectly placed decimal points or commas.
3. When a median unit value cannot be calculated because there are insufficient trade flow observations in the reporter's data set, the continental/regional or global median unit values for the commodity are used instead.
4. In some cases, the unit value of the trading partner's mirrored flow may be used. However, as imports are reported on a free on board (FOB) basis, while exports are on a cost, insurance, freight (CIF) basis, an FOB/CIF adjustment factor must be used. This factor, which reflects the difference in FOB and CIF values, is commodity-specific and depends on the mode of transportation and the geographical distance. The mirrored partner data are the corresponding flows of imports or exports declared by the partner country; needless to say they are specific to the commodity and the year.

These four steps warrant a more detailed description so that they can be repeated or modified by national FBS compilers. A more detailed description will also aid in the outlier detection process.

Outlier detection

Unit value (UV = monetary value/quantity) outliers – observations located outside the range – are identified using this formula:

$$[Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$$

where Q_1 and Q_3 are the lower and upper quartiles respectively (25–75 percent, with the median at 50 percent), and k is a non-negative constant. In this case, k = 1.5, as suggested by John Tukey (1977) in his basic approach for detecting potential outliers. Such an approach is more outlier-robust than the arithmetic mean. Figure 23 shows the combination of missing quantities and unit value outliers for reporting countries in a sample year.

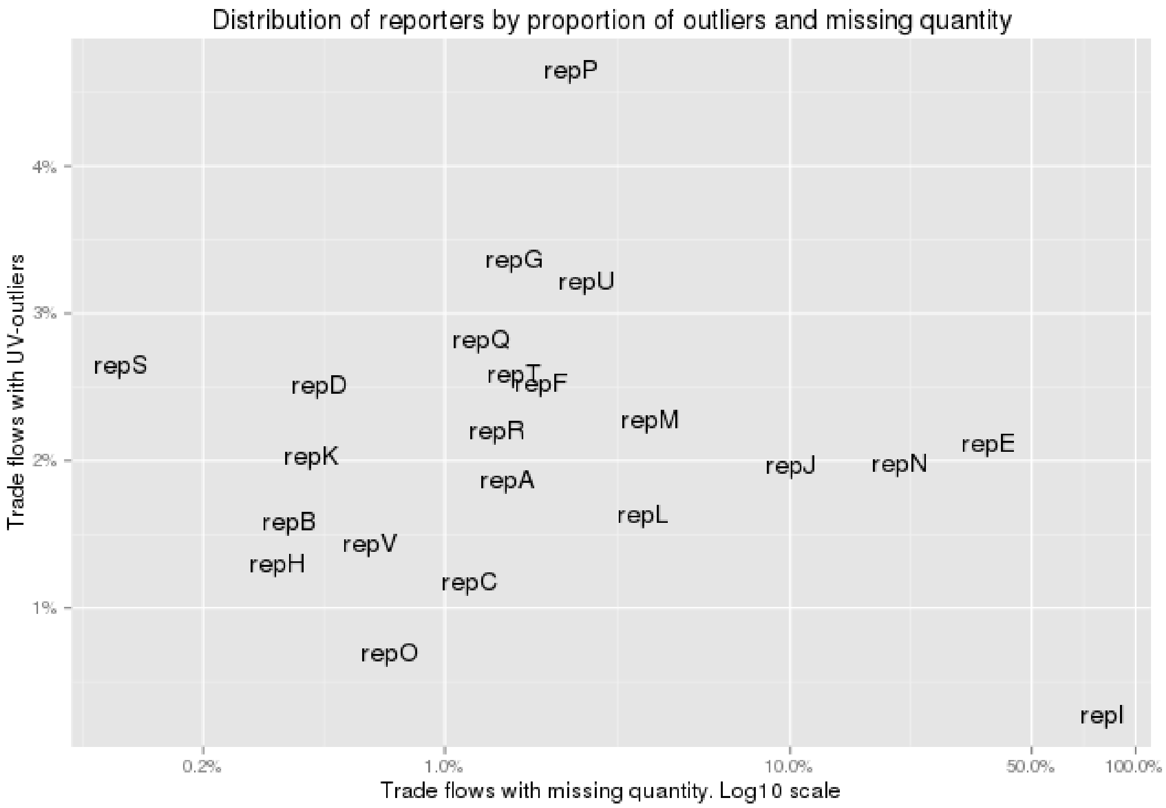

Figure 23: Results of outlier tests

Imputation of missing quantities/values and replacement of outliers

Once the outlier detection process is completed, the next step is to find criteria for removing or keeping outliers and imputing missing quantities.

In the example shown in Table 77, validation checks show that the data reported by country X in a particular year, in HS chapters 2, 10 and 15, have 17 trade flows with missing quantities and 33 trade flows with unit value outliers. The reporting partners are Y1 to Y4.

Table 77: Identifying and mending outliers based on trade flows with missing quantities and outlier unit values

Reporter	Partner	Flow	Commodity (FCL code)	Quantity	Value	Unit value	Median unit value
X	Y1	Import	271	-	9 193	-	3.71
X	Y1	Import	271	89	3 267	36.71	4.91
X	Y2	Import	1035	-	194 271	-	4.76
X	Y3	Import	56	23	2 937	127.7	3.84
X	Y4	Import	870	-	112 727	-	4.49
X	Y1	Import	79	-	2 479	-	1.88
X	Y1	Import	258	192	4 382	22.82	1.69
X	Y1	Import	1243	2 853	24 030	8.42	3.87
X	Y1	Import	89	-	7 833	-	1.24
X	Y1	Import	340	769	22 842	29.7	3.66

For simplicity, not all the trade flows used to calculate the median unit values are shown.

Missing quantities can be imputed and outliers adjusted using the median unit value for the respective commodity:

$$AdjustedQuantity = \frac{ReportedValue}{MedianUV_{reporter}}$$

$$Adjusted\ Quantity = \frac{Reported\ Value}{Median\ UV_{reporter}}$$

The first calculation is 9 193/3.71 = 2 478, and this value is shown Table 78. The other calculations follow the same procedure.

Table 78: Trade flows with adjusted quantities

Reporter	*Partner*	*Flow*	*Commodity (FCL code)*	*Quantity*	*Adjusted quantity*	*Difference*
X	Y1	Import	271	-	2 478	-
X	Y1	Import	271	89	665	-576
X	Y2	Import	1035	-	40 813	-
X	Y3	Import	56	23	765	-742
X	Y4	Import	870	-	25 106	-
X	Y1	Import	79	-	1 319	-
X	Y1	Import	258	192	2 593	-2 401
X	Y1	Import	1243	2 853	6 209	-3 356
X	Y1	Import	89	-	6 317	-
X	Y1	Import	340	769	6 241	-5 472

Imputation using mirrored data from a trading partner

In an additional step, data from a trading partner can be used to impute missing information through mirroring. Before doing so, possible outliers need to be validated or rejected. In the example, the validation algorithm identifies the following trade flows as having outlying unit values (Table 79)

Table 79: Updated trade matrix with validated data

Year	*Reporter*	*Trade partner*	*Flow*	*Commodity (FCL code)*	*Quantity*	*Value*	*Unit value*	*Median unit value*
2011	X	**Y5**	**Import**	1223	**16 057**	4 932 146	307.1	5.83
2011	X	Y6	Import	994	105 784	2 191 650	20.72	7.15
2011	X	Y7	Export	1064	726	11 780	16.23	4.27

In this case, the *import* quantity outlier will be adjusted using the mirrored *export* data from the trading partner. Therefore, the corresponding records in the partner's data must be found, as shown in Table 80.

Table 80: First step in the mirroring process

Year	*Reporter*	*Trade partner*	*Flow*	*Commodity (FCL code)*	*Quantity*	*Value*	*Unit value*	*Median unit value*
2011	Y5	**X**	**Export**	1223	131 206	2 665 550	**20.32**	5.07

The corresponding export unit value is calculated based on this quantity and the value data of the partner:

Partner export value/Partner export quantity = Partner export unit value

2 665 550/131 206 = 20.32

However, as exports are usually reported on an FOB basis (which excludes additional costs such as transport and insurance), while imports are on a CIF basis, the export unit value is expected to be lower than the import unit value. To adjust for this, an FOB/CIF adjustment factor is used. This factor, as previously mentioned, is commodity-specific and depends on the mode of transport and the geographical distance, among other factors. While specificity is always important, a number of studies provide standard factors of adjustment, which are applicable to most commodities and trading partners. For this example, it is assumed that the FOB/CIF factor is 1: 1.15 (15 percent).

Therefore, the final unit value from the above calculation is:

20.32 x 1.15 = 23.34

This adjusted partner export unit value is now applied to estimate the reporter's import quantity (Table 81):

Reporter import total value/Partner adjusted export unit value = Adjusted reporter quantity

4 932 146/23.34 = 279 132.3

Table 81: Introducing CIF/FOB corrections to the mirrored trade data set

Year	*Reporter*	*Trade partner*	*Flow*	*Commodity (FCL code)*	*Mirrored unit value*	*Value*	*Corrected quantity*
2011	X	Y5	Import	1223	23.34	4 932 146	**279 132**

Using mirrored trade data for non-reporting countries

Not all countries provide their trade data to UNSD. Thus, the trade data set obtained by FAO from UNSD does not have full global coverage. For non-reporting countries, other sources and methodologies – mainly mirrored trade data – have to be used to obtain or estimate trade figures. To identify the countries that are not reporters in the fully validated and processed data set, a check using the full FAO country codes list is run.

The global trade flow tables are used to fill in the missing imports and exports of non-reporting countries. Briefly, the steps are as follows:

1. For each non-reporting country, all the reported trade flows that indicate this non-reporter as a trading partner are identified in the validated data set. In the example shown in Table 82, a simplified filter is used for non-reporting partner N, showing only two commodities that are reported as being traded with N globally in that year.

Table 82: Starting table for mirroring flows after validation and imputation

Year	*All reporters*	*Non-reporting trading partner*	*Flow*	*Commodity (FCL code)*	*Quantity*	*Value*
2011	X	**N**	Import	1223	160 571	4 932 146
2011	A	**N**	Import	1223	80 120	2 301 650
2011	C	**N**	Export	1064	726	11 780
2011	D	**N**	Export	1064	2 000	35 780

2. Three distinct operations are then undertaken:

(a) The *exports* of the reporters become the *imports* of the non-reporting country for each commodity, and vice versa.

(b) The quantities are taken as they are, while the monetary values are adjusted for FOB/CIF, as described previously.

(c) The non-reporting country is now listed as a reporter, and the reporters become partners (Table 83).

Table 83: Final trade matrix after mirroring

Year	Mirrored reporter	Partner	Flow	Commodity (FCL code)	Quantity	Value
2011	**N**	**X**	Export	1223	160 571	4 120 350
2011	**N**	**A**	Export	1223	80 120	2 110 443
2011	**N**	**C**	Import	1064	726	13 810
2011	**N**	**D**	Import	1064	2 000	41 455

With this final step, the global trade flow data set is complete (Figure 22) and will be disseminated through the FAOSTAT Web site.

5.2 Trade flow aggregation and commodity balances

All the trade flows are now aggregated by reporter, commodity and flow (import/export) for the year. So, the trade flows in Table 83 would be aggregated as in Table 84.

Table 84: Aggregated total flows for dissemination

Year	Country	Flow	Commodity (FCL code)	Quantity	Value
2011	N	**Export**	1223	240 691	6 230 793
2011	N	**Import**	1064	2 726	55 265

This "totals" data set by commodity, country and region and for the world exists alongside the original validated trade flow data set. It allows for the manual input of commodity/country data from other sources, and shows the global totals by commodity and year (for checking of import versus export commodity balances).

The other sources used to obtain trade data for non-reporters, or to adjust reported data, include official statistical publications (usually online), data from specialized commodity institutes (OilWorld, the International Coffee Organization, etc.), and food aid statistics. Currently, data from such sources are entered manually into the system, but in the near future, they will be "harvested" automatically (similar to the downloading of data from UNSD).

This final data set with global totals (including data from other sources and minimized commodity imbalances) is also disseminated through the FAOSTAT Web site. National FBS compilers can use extractions from these data sets for compilation of their national FBS. Integration of trade into the FBS data set is explained in section 3.2.

References

Tukey, J.W. 1977. *Exploratory data analysis*. Boston, Massachusetts, USA, Addison-Wesley.

6. Commodity classifications

6.1 Overview of classification systems

The FAO Statistics Division (ESS) is revising the classification system used in FAOSTAT and replacing the FAOSTAT commodity list (FCL)[55] with the Harmonized System (HS) for trade variables and with the UN Central Product Classification (CPC), expanded for FAO purposes to cover all the variables in the balance. Commodities in FBS are defined in terms of primary equivalents and follow an ad hoc codification as in the old system.

Changing the classification system was a challenging process that required significant resources and a major collaboration effort among FAO divisions and other international organizations. However, efficiency gains are expected in the long run.

The HS[56] is developed and maintained by the World Customs Organization (WCO), and is the most widely used trade nomenclature in the world: more than 200 countries, territories and customs or economic unions utilize it as the basis for customs tariffs and the compilation of trade statistics. In the HS, commodities are generally classified according to their raw or basic material, degree of processing, use or function, and economic activities. The HS comprises about 5 000 commodity groups identified by 6-digit codes, which are binding for contracting parties. Countries can also expand their national HS by creating additional sub-headings at lower levels of aggregation (generally 8, 10 or 12 digits). Lower levels are country-specific. Maintenance of the HS includes measures for ensuring uniform interpretation of HS data and the system's periodic updating in line with developments in technology and changes in trade patterns. WCO manages this process through the HS Committee and the HS Review Sub-Committee. Each review cycle typically lasts five years.

The CPC is developed and maintained by the United Nations Statistics Division (UNSD). It is a comprehensive classification of products[57] into a system of categories that are exhaustive, mutually exclusive and based on a set of internationally agreed concepts, definitions, principles and classification rules. The CPC has a five-level hierarchical structure in which each digit provides information on the product grouping.

The CPC is a general-scope classification, i.e. it covers products of all economic activities (and is not sector-specific), but it can be customized for sector-specific applications. It is also a general-purpose classification, so potential areas of application range from production, to trade, prices and consumption.

---

[55] The FCL is the commodity classification used in FAOSTAT since the 1960s. It was originally based on the UN Standard International Trade Classification (SITC). It includes 683 commodities, in 20 chapters (or groups) and covers crops, livestock and their derived products. It excludes agricultural inputs (e.g. fertilizers, pesticides and machinery) and fishery and forest products, for which different classifications and lists are used in FAOSTAT.

[56] HTTP://WWW.WCOOMD.ORG/EN/TOPICS/NOMENCLATURE/INSTRUMENT-AND-TOOLS/HS_NOMENCLATURE_2012/HS_NOMENCLATURE_TABLE_2012.ASPX

[57] Products follow the UN System of National Accounts definition: *all output of economic activities* that can be the object of domestic or international transactions or that can be entered into stocks, including transportable goods, non-transportable goods, services and other products.

The following are some of the benefits of using the CPC:

- It is an international classification, which is constantly updated and reviewed by the Expert Group on International Classification, chaired by UNSD and with the participation of representatives of countries and international organizations.
- It is used by other organizations and statistical domains, allowing data comparability across domains.
- It is already used by countries, so its use at FAO will reduce reporting burdens. In addition, the expanded CPC is designed not only for FAO but also for countries engaged in the collection and dissemination of data on agriculture and food products: it provides a flexible tool that allows increased granularity at lower levels, to include local species and varieties, while maintaining comparability across countries at the higher level.
- It aligns with the International Standard Industrial Classification of all Economic Activities (ISIC) and the HS: when these classifications are updated, the CPC is also updated.
- As it is closely aligned to the HS, data conversion for SUAs/FBS is better than in the FCL: 65 percent of correspondences between HS 2007 and CPC Ver. 2 are one-to-one or many-to-one, versus 35 percent of HS–FCL correspondences; the alignment between HS 2012 and CPC Ver. 2.1 is even closer (Figure 24).

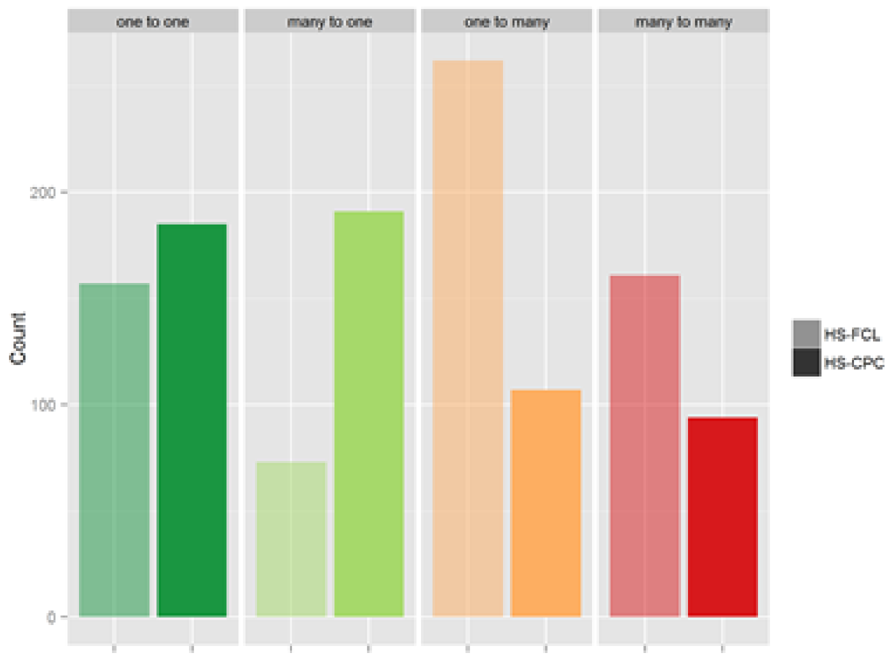

Figure 24: Quality of links between the HS and the FCL (shaded) and between the HS and the CPC

The latest CPC Ver. 2.1[58] is complemented with an official annex developed by FAO to meet the needs of agricultural statistics; this structure is called the "CPC expanded for agricultural statistics". The CPC expanded provides additional detail on agricultural commodities (primary products) and is obtained by adding one level (two digits) to the lower level of the standard CPC (Ramaschiello, 2011; 2013, 2015) (Figure 25).

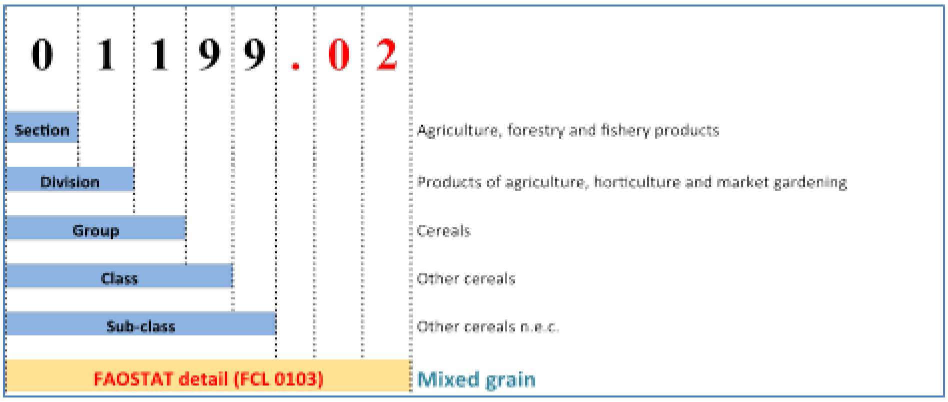

Figure 25: CPC expanded code for "Mixed grain" (0103 in the FCL)

6.2 CPC implementation in the production and trade domains

ESS collects production data through a production questionnaire (PQ) that is sent annually to national statistical offices and ministries of agriculture around the world (supplementary sources are also used for selected commodities and countries). Product lists included in the PQs are country-specific, which means that the number and types of commodities can vary from country to country. The generic template of the PQ includes 209 primary commodities (167 crops and 42 livestock commodities) and 47 processed products (34 vegetable oils and cakes, four dried fruits, six alcoholic beverages, and three sugar products). The classification currently used is the CPC expanded for agricultural statistics, in place of the FCL.

ESS does not collect trade data through a questionnaire: countries previously sent ESS their full trade files, but now most data are received from UNSD (through the UN COMTRADE database). In the past, trade data were converted from HS to FCL format. From now on, trade series will be published in HS format and converted to the CPC to compile SUAs.

A prerequisite for implementation of the CPC in the new system was to increase the detail on agriculture, forest and fishery products in the classification. To this end, over the past ten years, FAO has contributed significantly to CPC Ver. 2 and Ver. 2.1, together with UNSD and the Expert Group on International Classifications.

The CPC expanded for agricultural statistics has been added as an official annex to CPC Ver. 2.1. When detail in CPC Ver. 2.1 expanded is not sufficient, the classification is expanded further in FAOSTAT according to FAO's needs and the available data. This ensures harmonization between the FCL and the CPC.

[58] HTTP://UNSTATS.UN.ORG/UNSD/CR/REGISTRY/CPC-21.ASP

It is planned to use the CPC for future data collection and to apply it to old time series, to allow data comparability over time and to avoid breaks in the series.

Although the basic condition for data back-cast is to have double-coded data for at least one year, it was difficult for FAO to increase its data request to countries: the additional burden on national offices might have lowered the response rate and hampered the data collection process. ESS therefore identified alternative solutions to facilitate the change of classification and data back-cast, while reducing the cost involved. The solution adopted depended on the type of link encountered and allowed full alignment between the FCL and the CPC:

- *One-to-one* cases are resolved easily, as old data are transferred to the new classification by assigning codes and definitions according to the new classification, while the data remain the same (the "key method").[59]

- For *many-to-one* cases data conversion is also straightforward, as data in the FCL are aggregated into the target classification (the CPC). Such aggregation entails a loss of information, as the CPC is less detailed than the FCL. To avoid losing information in FAOSTAT, many-to-one cases have been turned into one-to-one correlations: first the target classification is expanded further according to the detail available in the FCL, then the key method is applied. When detail in the CPC 2.1 expanded is not sufficient, the classification is expanded further for FAOSTAT purposes.

More difficulties are faced with one-to-many and many-to-many links. In these cases, data are converted based on the statistician's best judgement according to the *dominant* correspondence. Coefficients of conversion are not calculated, as there is a lack of information in both formats for at least one year, so there is a risk of compromising data quality in the conversion. The conversion keys used are 1 and 0 exclusively:

- *One-to-many* relations between the FCL and the CPC are managed by identifying the dominant correlation based on the statistician's best judgement, and assigning the conversion key 1.

- In *many-to-many* cases, which represent a minority of FCL–CPC correlations, the target classification is modified and aligned to the source classification.

Details and examples are provided in the Appendix to this chapter.

6.3 CPC implementation for SUAs and FBS

The compilation of SUAs and FBS is based on commodity trees. However, the structure referred to as a "commodity tree" in FAO should not be confused with a classification tree or "hierarchy".

[59] In the key method, a classification at the lowest aggregation level is recoded directly to the revised classification. For example, the old code 12345 is recoded to 56789 and the historical data for 12345 are assigned to 56789. This method assures a straightforward relationship between old and new results, as the old data are simply transferred to the new classification. However, the process and outcomes should be documented and communicated to users. The key method is described in Buiten, Kampen and Vergouw, 2009.

A *commodity tree* is a "symbolic representation of the flow from a primary commodity to various processed products derived from it, together with the conversion factors from one commodity to another" (FAO, 2001).

A *statistical classification* is "a set of categories which may be assigned to one or more variables", where "the categories are defined in terms of one or more characteristics of a particular population of units of observation. A statistical classification may have a flat, linear structure or may be hierarchically structured, such that all categories at lower levels are sub-categories of a category at the next level up" (Hancock, 2013).

The FCL is a flat classification (or list) in which commodities are listed in ascending order (in most cases). The FCL itself does not set the relations among commodities, as all categories are at the same level: to distinguish primary from processed products the printed version of the FCL uses capital letters, which is not a classification feature (Figure 26). It is the commodity tree that sets the links among commodities listed in the FCL by applying extraction rates. Extraction rates "indicate, in percentage terms, the amount of the processed product concerned obtained from the processing of the parent/originating product, in most cases a primary products" (FAO, 2001) (Figure 27).

Group 1: Cereals
0070 MILLET
0080 Flour of millet
0081 Bran of millet
0082 Beer of millet

Figure 26: Classification of millet and its derived products in the FCL

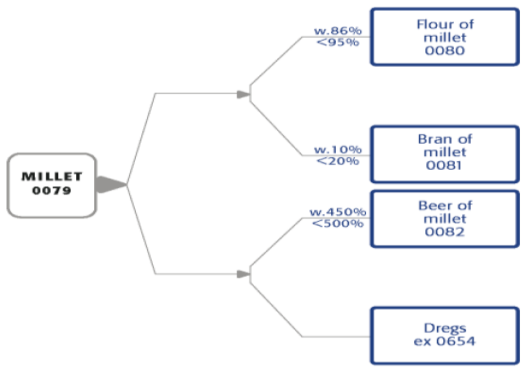

Figure 27: Commodity tree for millet

This is a simple example that includes only first-level processed products.

As long as single commodities are identified in the reference product classification (one-to-one and one-to-many correspondences), commodity trees can be developed. Commodity trees are independent of the statistical classification used: their structure does not depend on the reference classification hierarchy, and the relations set in the trees should not be confused with the classification hierarchy. In a hierarchical classification, items at the lower level can be grouped/aggregated into an item at the higher level. For example, millet, wheat, barley and maize can be grouped into cereals, or seeds and grains of millet can be grouped into millet. This is not true for commodity trees, in which flour, bran and beer cannot be grouped into millet unless quantities are first expressed in terms of primary equivalents by applying extraction rates.

Commodity trees therefore show the *relations* among commodities (Figure 28), while the commodities themselves are listed in the reference product classification (Figure 29). This approach based on commodity trees will be maintained in the new system. The innovation is that SUAs are expressed in the CPC expanded, instead of the FCL. However, commodities in FBS, defined in terms of primary equivalents, will continue to follow the ad hoc codification as in the old system (while links are built to the CPC). For example, millet is coded 2517 in FBS (as in the old system) and is defined as millet (0118) while processed products expressed in terms of primary equivalent are defined as flour of millet including groats, meal and pellets (23120.05), bran of millet (39120.07), etc.

The classification scheme used in the new system is summarized in Figure 30.

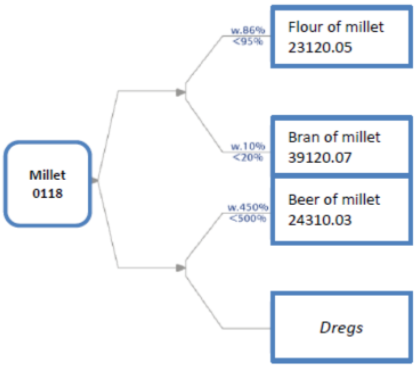

Figure 28: Commodity tree for millet in the CPC

Section 0: Agriculture, forestry and fishery products
Division 01: Products of agriculture, horticulture and market gardening
Group 011: Cereals
Class **0118: MILLET**

Section 2: Food products, beverages and tobacco [...]
Division 23: Grain mill products, starches and starch products; other food products
Group 231: Grain mill products
Class 2312: Other cereals flour
Subclass 23120: Other cereal flours
FAO expansion **23120.05 Flour of millet**
Division 24: Beverages
Group 243: Malt liquors and malt
Class 24310: Beer made from malt
FAO expansion **24310.03 Beer of millet**

Section 3: Other transportable goods
Division 39: Wastes or scraps
Group 391: Wastes from food and tobacco industry
Class 3912: Bran and other residues from the working of cereals or legumes
Subclass 39120: Bran and other residues from the working of cereals or legumes
FAO expansion **39120.07: Bran of millet**

Figure 29: Classification of millet and its derived products in the CPC

The CPC hierarchy reflects the economic activity of origin.

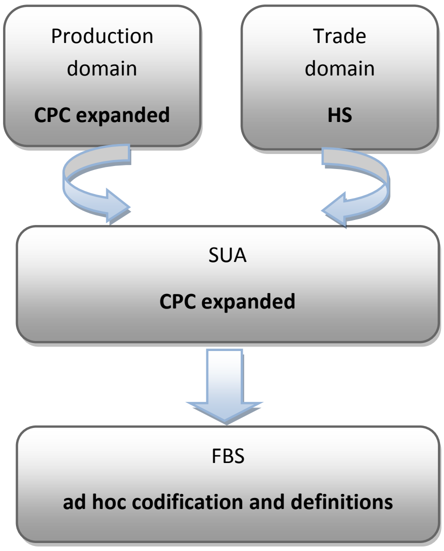

Figure 30: Classification scheme used in the new system

Appendix

The following are examples of solutions adopted to convert FAOSTAT data on agricultural commodities from FCL to CPC format.

In *one-to-one* cases, old data are transferred to the new classification, i.e. codes and definitions are reassigned according to the new classification while the data remain the same (Example 1).

Example 1: Data conversion from the FCL to the CPC for one-to-one link

FCL			FCL→CPC conversion factor	CPC Ver. 2.1 expanded		
Code	**Descriptor**	**Data (old format) production quantity**		**Code**	**Descriptor**	**Data (new format) production quantity**
0125	cassava	**4 082 903 tonnes**	**1**	01520	cassava	**4 082 903 tonnes**

The data in this example refer to cassava production in Cameroon in 2011.
Source: FAOSTAT.

Data conversion is also straightforward for *many-to-one* cases, as data in the FCL are aggregated to the CPC. Such aggregation entails a loss of information, as the target classification is less detailed than the source one (Example 2).

Example 2: Data conversion from the FCL to the CPC for a many-to-one link

FCL			FCL→CPC conversion factor Code	CPC Ver. 2.1 expanded		
Code	**Descriptor**	**Data (old format) production quantity**		**Code**	**Descriptor**	**Data (new format) production quantity**
0430	okra	**5 784 000 tonnes**	**Σ**	01239	other fruit-bearing vegetables	5 784 000+ 27 557 000= **33 341 000 tonnes**
0463	other vegetables	**27 557 000 tonnes**				

The data in this example refer to the production of okra and other fresh vegetables in India in 2011.
Source: FAOSTAT.

To avoid losing information in FAOSTAT, many-to-one cases are turned into one-to-one correlations: first the target classification is expanded according to the detail available in the FCL (new CPC expanded codes 01239.01 and 01239.90 in Example 3), then the key method is applied as in Example 1. When the detail in CPC 2.1 expanded is insufficient, the classification is expanded further for FAOSTAT purposes.

Example 3: FCL–CPC data conversion when many-to-one are turned into one-to-one correlations

FCL			FCL→CPC conversion factor Code	CPC Ver. 2.1 expanded		
Code	**Descriptor**	**Data (old format) production quantity**		**Code**	**Descriptor**	**Data (new format) production quantity**
n/a	n/a	n/a		01239	other fruit-bearing vegetables	33 341 000 tonnes
0430	okra	**5 784 000 tonnes**	**1**	**01239.01**	okra	**5 784 000 tonnes**
0463	other vegetables	**27 557 000 tonnes**	**1**	**01239.90**	other fruit-bearing vegetables NEC	**27 557 000 tonnes**

NEC = not elsewhere classified.
Codes in **bold blue** text are the CPC expanded codes.
The data in this example refer to the production of okra and other fresh vegetables in India in 2011.
Source: FAOSTAT.

More difficulties are faced with one-to-many and many-to-many types of links. In these cases, data have been converted based on the statistician's best judgement according to the *dominant* correspondence. Coefficients of conversion have not been calculated, as there is a lack of information in both formats for at least one year, so there is a risk of compromising data quality in the conversion. The conversion keys assigned are exclusively 1 and 0.

One-to-many relations between the FCL and the CPC concern mainly agricultural (primary) versus industrial (processed) products. For example, fresh and dried fruit are sometimes classified together in the FCL while they are separated in the CPC. This separation is because the CPC is closely linked to the ISIC, and dried fruit is considered an output of the manufacturing industry rather than agriculture (unless sun-dried). In FAOSTAT, when dried fruit is not dedicated to a specific class (as in the case of dates), FCL data are associated to items in only the agricultural section of the CPC, leaving corresponding blanks in the industrial goods section. In Example 4, the one-to-many correlation is converted into a one-to-one correlation, assigning the conversion factor 1 to the class that – in the statistician's judgement – covers the FCL boundaries (dominant correspondence) better. In the metadata it will be noted that 01314 may, in some years for some countries, include information on dates dried on-farm.

Example 4: Data conversion from the FCL to the CPC for a one-to-many link

FCL			FCL→CPC conversion factor Code	CPC Ver. 2.1 expanded		
Code	**Descriptor**	**Data (old format) production quantity**		**Code**	**Descriptor**	**Data (new format) production quantity**
0577	dates (fresh+dried)	**724 894 tonnes**	**1**	01314 (agriculture)	dates, fresh	**724 894 tonnes**
			0	214190.03 (industrial)	dates, dried	**0**

The data in this example refer to date production in Algeria in 2011.
Source: FAOSTAT.

In *many-to-many* cases, which represent the minority of FCL–CPC correlations, the CPC is modified and aligned with the FCL.

In Example 5, the FCL code 0619 combines “subtropical fruit” with “fruit fresh NEC”, while in the CPC subtropical fruit is classified with “other tropical and subtropical fruits NEC.” (01319). This difference generates a mismatch between the two classifications. Given the impossibility of estimating split shares, and to avoid introducing breaks in the series, the CPC is adapted and aligned with the FCL (Example 6): the component “subtropical fruit” of the CPC is moved to “other fruits NEC” as in FAOSTAT (01359.90). Definitions in the metadata are adjusted accordingly.

Example 5: Many-to-many correlations between the FCL and the CPC for tropical, subtropical and other fruit NEC

FCL		**CPC Ver. 2.1 expanded**	
Code	**Descriptor**	**Code**	**Descriptor**
0603	fruit tropical fresh NEC	01319	other tropical and subtropical fruit NEC
0619	fruit fresh NEC (including subtropical)	01359.90	other fruits NEC

Example 6: Data conversion from the FCL to the CPC for a many-to-many link

FCL			**FCL→CPC conversion factor Code Code**	**CPC Ver. 2.1 expanded**		
Code	**Descriptor**	**Data (old format) production quantity**		**Code**	**Descriptor**	**Data (new format) production quantity**
0603	fruit tropical fresh NEC	**52 684 tonnes**	**1**	01319	other tropical **and subtropical** fruit NEC → other tropical fruit NEC (excluding subtropical fruit)	**52 684 tonnes**
0619	fruit fresh, n.e.c. (incl. subtropical)	**193 686(E) tonnes**	**1**	01359.90	other fruit NEC → other fruit NEC **(including** subtropical fruit)	**193 686(E) tonnes**

The data in this example refer to the production of tropical fruit NEC and fruit NEC in Ecuador in 2011.
Sources: FAOSTAT and FAO estimates (E).

References

Buiten, G., Kampen, J. & Vergouw, S. 2009. *Producing historical time series for STS-statistics in NACE Rev.2: Theory with an application in industrial turnover in the Netherlands (1995–2008).* Discussion Paper No. 09001. The Hague, Statistics Netherlands. HTTP://WWW.CBS.NL/NR/RDONLYRES/A8A9AB3B-37F6-480A-BA76-253979DED22D/0/200901X10PUB.PDF

FAO. 2001. *Food balance sheets: A handbook*. Rome. HTTP://WWW.FAO.ORG/DOCREP/003/X9892E/X9892E00.HTM

Hancock, A. 2013. Best practice guidelines for developing international statistical classifications. Presented at the Meeting of the Expert Group on International Classifications, UNSD, New York, 13–15 May 2013. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2013/AC267-5.PDF

Ramaschiello, V. 2011. CPC Ver.2 review and harmonization with food and agriculture statistics in FAO. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 18–20 May 2011. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2011/AC234-15.PDF

Ramaschiello, V. 2013. Aligning classifications for agricultural statistics with other international standards. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 13–15 May 2013. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2013/AC267-21.PDF

Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 19–22 May 2015. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2015/AC289-16.PDF

7. Concepts and definitions used in the FAO food balance sheets

7.1 Commodity coverage

In principle, all potentially edible commodities should be taken into account in preparing FBS, regardless of whether they are actually eaten or are used for non-food purposes.

Generally, FBS are constructed for primary crop, livestock and fish[60] commodities. They include processed products at different stages of processing even though it is difficult to obtain data for all the different forms of processed products, and even more difficult to trace the components of processed composite products accurately.

Commodities are grouped and coded for FBS purposes and defined in terms of primary equivalent. Links to the CPC Ver. 2.1 expanded are provided in brackets. Default compositions can be adjusted by countries according to the availability of the commodities at the national level.

The following is an example of a default composition and commodity codes in FBS.

FBS ad hoc code

FBS descriptor

Default composition in primary equivalent with links to CPC Ver. 2.1 expanded

FBS 2517 - **Millet**

Includes millet (0118) and the following processed products: flour of millet, including groats, meal and pellets (23120.05; 23130*) and bran of millet (39120.07)

Millet includes small-grained cereals of different botanical species and with different local names, whether or not processed, including:
Echinocloa frumentacea (barnyard or Japanese millet)
Eleusine coracana (ragi, finger or African millet)
Panicum miliaceum (common, golden or proso millet)
Paspalum scrobiculatum (koda or ditch millet)
Pennisetum glaucum (pearl or cattail millet)
Setaria italic (foxtail millet)

It does not include Indian millet, large African millet, and teff (*Eragrostis abyssinica*) cf. 01199.01

Description of the primary commodity

[60] The previous version of this publication (FAO, 2001) defines fish and fisheries products to include freshwater fish, demersal fish, pelagic fish, crustaceans, molluscs, aquatic mammals meat, and aquatic plants. Data on these products are not available in the FCL and are not owned by ESS; definitions may be provided by the FAO Fisheries and Aquaculture Department if needed.

7.2 Cereals and cereal products

CEREALS are generally of the gramineous family and, in the FAO concept, refer to crops harvested for dry grain only. Crops harvested green for forage, silage or grazing are classified as fodder crops. Also excluded are industrial crops, such as broom sorghum (Crude organic materials NEC) and sweet sorghum when grown for syrup (Sugar crops NEC). For international trade classifications, fresh cereals (other than sweet corn), whether or not suitable for use as fresh vegetables, are classified as cereals. Cereals are identified according to their genus. However, when two or more genera are sown and harvested as a mixture they should be classified and reported as "mixed grains".

Production data are reported in terms of clean, dry weight of grains (12–14 percent moisture) in the form usually marketed. However, rice is reported in terms of paddy. Apart from moisture content and inedible substances such as cellulose, cereal grains contain, along with traces of minerals and vitamins, carbohydrates – mainly starches – (comprising 65–75 percent of their total weight), proteins (6–12 percent) and fat (1–5 percent). Cereal products are derived from either the processing of grain through one or more mechanical or chemical operations, or the processing of flour, meal or starch.

FBS 2511 – Wheat

Includes wheat (0111) and the following processed products expressed in terms of primary equivalent: bread and wafers (23410; 23490); breakfast cereals (23140.03); bulgur (23140.02); flour of wheat (23110; 23130.01); macaroni (23710); pastry (23420; 23430); starch of wheat (23220.01); bran of wheat (39120.01); germ of wheat (23140.01); gluten of wheat (23220.02); mixes and doughs and food preparations of flour (23180); meal or malt extract (23999.02).

Wheat, species of *Triticum*, *T. aestivum* (common), *T. durum* (durum) and *T. spelta* (spelt). Common and durum wheat are the main types. The main varieties of common wheat are spring and winter, hard and soft, red and white, whether or not processed.

Wheat includes meslin (a mixture of wheat and rye).

FBS 2805 – Rice (milled equivalent)

Includes rice (0113) and the following processed products expressed in terms of primary equivalent:

- rice husked (23162); starch of rice (23220.03); rice flour (23120.01); rice gluten (39130.01); bran of rice (39120.02);
- rice, semi- or wholly milled (23161):
    - milled (husked) rice (23161.01);
    - rice milled (23161.02);
    - rice broken (23161.03).

Rice, species of *Oryza*, mainly *O sativa*, not husked, also known as rice in the husk and rough rice. Used mainly for human food.

FBS 2513 – Barley

Includes barley (0115) and the following processed products expressed in terms of primary equivalent: pot barley (23140.04); barley pearled (23140.05); malt (24320); malt extract (23999.01); bran of barley (39120.03); barley flour and grits (23120.02).

Barley, species of *Hordeum*, mainly *H. disticum* (two-row barley), *H. hexasticum* (six-row barley) and *H. vulgare* (four-row barley), whether or not processed, including with husk and without (naked).

Barley tolerates poorer soils and lower temperatures than wheat and is used as a livestock feed, for the manufacture of malt, and – when polished or pearled – for preparing foods. The roasted grains are a coffee substitute.

Barley does not include sprouted barley (malt), roasted malt, roasted barley (coffee substitutes), malt sprouts separated from the malted grain during the kilning process and other brewing wastes.

FBS 2514 – Maize

Includes maize (0112) and the following processed products expressed in terms of primary equivalent: flour of maize (23120.03); starch of maize (23220.04); gluten feed and meal (39130.04); germ of maize (23140.06); bran of maize (39120.04); maize gluten (39130.02).

Maize, species of *Zea mays* (corn, Indian corn, mealies), is a grain with a high germ content. At the national level, hybrid and ordinary maize should be reported separately owing to their widely different yields and uses. Maize is used largely for animal feed and commercial starch production.

It includes:

- corn seed; maize, hybrid, seed; maize seed;
- white maize (not considered necessary for popcorn);
- corn, unmilled, golden-yellow or white;
- corn, unmilled, reddish-brown or mottled;
- corn-on-the-cob, fresh (excluding sweet corn);
- grain, maize, not husked or otherwise worked;
- kernels, corn, fresh, suitable for human consumption;
- maize in sheaves, cobs or threshed;
- maize not husked or otherwise worked;
- maize, cut before maturity, complete with husks;
- maize, unmilled;
- popcorn, on or off the cob, not popped.

Maize does not include green corn (01290.01).

FBS 2515 – Rye

Includes rye (0116) and the following processed products expressed in terms of primary equivalent: flour of rye including groats, meal and pellets (23120.04); bran of rye (39120.05).

Rye, species of *Secale cereale*, whether or not processed, including rye denatured and unmilled, is a grain that is tolerant of poor soils, high latitudes and altitudes. Used mainly in making bread, whisky and beer. When fed to livestock, it is generally mixed with other grains.

FBS 2516 – Oats

Includes oats (0117) and the following processed products expressed in terms of primary equivalent: oats rolled (23140.07); bran of oats (39120.06).

Oats, species of *Avena sativa*, grains with their husks as well as those that in their natural state have no husk or hull, whether or not processed, are plants with open, spreading panicle-bearing large spikelets.

There are two main kinds of oats: grey (or black) oats and white (or yellow) oats.

This crop is used primarily in breakfast foods; it makes excellent fodder for horses.

It includes unmilled oats.

FBS 2517 – Millet

Includes millet (0118) and the following processed products expressed in terms of primary equivalent: flour of millet including groats, meal and pellets (23120.05); bran of millet (39120.07).

Millet includes small-grained cereals of different botanical species and with different local names, whether or not processed, including:

Echinocloa frumentacea (barnyard or Japanese millet);
Eleusine coracana (ragi, finger or African millet);
Panicum miliaceum (common, golden or proso millet);
Paspalum scrobiculatum (koda or ditch millet);
Pennisetum glaucum (pearl or cattail millet);
Setaria italic (foxtail millet).

It does not include Indian millet, large African millet, and teff (*Eragrostis abyssinica*) cf. 01199.01.

FBS 2518 – Sorghum

Includes sorghum (0114) and the following processed products expressed in terms of primary equivalent: flour of sorghum including groats, meal and pellets (23120.06); bran of sorghum (39120.08).

Sorghum, species of *Sorghum*, mainly *S. guineense* (guinea corn), *S. vulgare* (common, milo, feterita, kaffir corn) and *S. dura* (durra, jowar, kaoliang), is a cereal that has both food and feed uses. Sorghum is a major food grain in most of Africa, where it is also used in traditional beer brewing. It is desirable to report hybrid and other varieties separately.

Millet includes unmilled doura (durra), federita (feterita), grain sorghum, kafir, kaoliang, milo, sorghum.

Sorghum does not include forage sorghums (which are used for making hay or silage) such as halepensis (*S. halepense*), grass sorghums (which are used for grazing) such as sudanensis (*S. sudanense*) or sweet sorghums (which are used primarily for the manufacture of syrup or molasses) such as saccharatum. It also excludes broomcorn (*S. vulgare var. technicum*).

FBS 2520 – Cereals, other

Includes, inter alia, buckwheat (01192); quinoa (01194); fonio (01193); triticale (01191); canary seed (01195); mixed grain (01199.02); cereals NEC (01199.90); and the following processed products expressed in terms of primary equivalent: flour of buckwheat (23120.07; 23130.08); flour of fonio (23120.08; 23130.09); flour of mixed grain (23120.10; 23130.11); flour of cereals NEC including groats, meal and pellets (23120.90); cereal preparations NEC (23140.08; 23130.90); bran buckwheat (39120.09); bran of fonio (39120.10); bran of triticale (39120.11); bran of mixed grains (39120.12); bran of cereals (39120.13).

Buckwheat, *Fagopyrum esculentum* (Polygonaceae), is a cereal cultivated primarily in northern regions. Buckwheat is considered a cereal, although it does not belong to the gramineous family.

Quinoa, *Chenopodium quinoa* (Chenopodiaceae), is a cereal that tolerates high altitudes and is cultivated primarily in Andean countries. It is used for food and to make *chicha*, a fermented beverage.

Fonio, species of *Digitaria*, mainly *D. exilis* (fonio or findi) and *D. iburua* (black fonio or hungry rice), is a cereal of importance in West Africa where it is eaten in place of rice during famines. The seeds are cooked by steaming the whole grain.

Triticale is a cross between wheat and rye, combining the quality and yield of wheat with the hardiness of rye.

Canary seed is a cereal normally used as bird feed.

Mixed grain is a mixture of cereal species that are sown and harvested together. It does not include meslin (a wheat–rye mixture) that is included under wheat.

Other cereals not elsewhere classified are also included here.

7.3 Roots and tubers and derived products

ROOTS AND TUBERS are plants yielding starchy roots, tubers, rhizomes, corms and stems. They are used mainly for human food (in processed or unprocessed form), for animal feed and for manufacturing starch, alcohol and fermented beverages including beer. The denomination “roots and tubers” excludes crops that are cultivated mainly for feed (mangolds, swedes) or for processing into sugar (sugar beets), and those classified as roots, bulb and tuberous vegetables (onions, garlic and beets). It does include starch and the starchy pith and flour obtained from the trunk of the sago palm and the stem of the Abyssinian banana (*Musa ensete*). Certain root crops, notably bitter cassava, contain toxic substances, particularly in the skins. As a result, certain processes must be undertaken to make the product safe for human consumption.

Apart from their high water content (70–80 percent), these crops contain mainly carbohydrates (largely starches that account for 16–24 percent of their total weight) with very little protein and fat (0–2 percent each). Methods of propagating root crops vary. A live potato tuber or seed must be planted, but only part of the live yam tuber, and only a piece of the stalk (not the root) of the cassava plant. Production data on root crops should be reported in terms of clean weight, i.e. free of earth and mud.

FBS 2520 – Cassava

Includes cassava (01520) and the following processed products expressed in terms of primary equivalent: flour of cassava (23170.01); tapioca of cassava (23230.02); cassava dried (01520.01); cassava starch (23220.06).

Cassava, species of *Manihot esculenta, M. utilissima* (manioc, mandioca, yuca), *M. palmate and M. dulcis* (yuca dulce).

It is a semi-permanent crop grown in tropical and subtropical regions. Sometimes bitter and sweet, cassavas are referred to as separate species – *M. esculenta* being bitter and *M. palmate* sweet – but this differentiation is incorrect as toxicity varies according to location. Cassava is the staple food in many tropical countries. It is not traded internationally in its fresh state because tubers deteriorate very rapidly.

It includes cassava, fresh or dried, whole or sliced; root, manioc, fresh or dried, whole or sliced; and pellets, of manioc, whether or not disintegrated (manioc pellets may be disintegrated, but are classified here provided that they are identifiable by physical characteristics: non-homogeneous particles with broken pieces of manioc pellets, brownish colour with black spots, pieces of fibre visible to the naked eye and a small quantity of sand or silica remaining).

FBS 2531 – Potatoes

Includes potatoes (01510) and the following processed products expressed in terms of primary equivalent: flour of potatoes meal, including powder, flakes, granules and pellets of potatoes (21392); frozen potatoes (21313); tapioca of potatoes (23230.01); starch potatoes (23220.05).

Potatoes, species of *Solanum tuberosum* (Irish potato).

This is a seasonal crop grown in temperate zones all over the world, but primarily in the Northern Hemisphere.

It includes fresh or chilled potatoes of all kinds, seed potatoes intended for sowing, and new potatoes.

It does not include sweet potatoes (01530).

FBS 2533 – Sweet potatoes

Includes sweet potatoes (01530).

Sweet potatoes, species of *Ipomoea batatas*, are a seasonal crop grown in tropical and subtropical regions. Used mainly for human food. Trade data cover fresh and dried tubers, whether or not sliced or in the form or pellets made either from pieces of the roots or tubers of sweet potatoes or from their flours, meals or powders.

FBS 2535 – Yams

Includes yams (01540).

Yams, species of *Dioscorea*, include *D. batatas*, *D. trifida*, *D. alata*, *D. bulbifera*, *D. rotunda*, *D. cayenensis*, *D. exculenta* and *D. dumetorum*.

The principal edible yams are widely grown throughout the tropics. They are a starchy staple foodstuff, normally eaten as a vegetable, boiled, baked or fried. In West Africa, they are consumed mainly as *fufu*, a stiff glutinous dough. Trade data cover both fresh and dried yams.

Yams include fresh, chilled, frozen or dried yams, whether or not sliced or in the form of pellets made either from pieces of the roots or tubers of yams or from their flours, meals or powders.

FBS 2534 – Roots, other

Includes yautia (01591); taro (01550); other edible roots and tubers with high starch or inulin content, NEC (01599); and the following processed products expressed in terms of primary equivalent: flour of roots and tubers (23170.02); roots and tubers dried (01599.10).

- Yautia, species of *Xanthosoma*, mainly *X. sagittifolium* (malanga, new cocoyam, ocumo, tannia). Several plants are included in this group, some with edible tubers and others with edible stems (also called aroids). Yautia is grown mainly in the Caribbean and is used for food.

- Taro, *Colocasia esculenta* (Dasheen, eddoe, taro, old cocoyam), is an aroid cultivated for its edible starchy corms or underground stems. Taro is grown throughout the tropics for food.

It includes other tubers, roots or rhizomes, fresh, that are not identified separately because of their minor relevance at the international level. Because of their limited local importance, some countries use this commodity heading to report roots and tubers that are classified individually by FAO:

- *Arracacoa xanthorrhiza* (arracacha);

- *Maranta arundinacea* (arrowroot);

- *Cyperus esculentus* (chufa);

- *Metroxylon* spp. (sago palm);

- *Oxalis tuberosa* and *Ullucus tuberosus* (oca and ullucu);
- *Pachyrxhizus erosus, Pachyrxhizus angulatus* (yam bean, jicama);
- *Tropaeolum tuberosum* (mashua);
- *Helianthus tuberosus* (Jerusalem artichoke, topinambur).

Products included may be fresh, chilled, frozen or dried, whether or not sliced or in the form of pellets made either from pieces (e.g. chips) of the roots or tubers of the plant or from its flours, meals or powders.

Other roots and tubers not elsewhere classified are also included here.

7.4 Sugar crops and sweeteners and derived products

In addition to providing the source for the manufacture of sugar, SUGAR CROPS are used to produce alcohol and ethanol. In some countries, sugar cane is eaten raw in minor quantities. It is also used in the preparation of juices and for animal feed. There are two major sugar crops – sugar beets and sugar cane – but sugar and syrups are also produced from the sap of certain species of maple trees, from sweet sorghum when cultivated explicitly for making syrup, and from sugar palm. Sugar beets that are cultivated solely as a fodder crop, and red or garden beets that are classified as vegetable crops are excluded from the FAO list of sugar crops. Sugar cane is a perennial grass (replanted at intervals using pieces of the cane stalks) that is cultivated mainly in the tropics. Sugar beet is an annual crop that is propagated by the seeds of the flowers. It is cultivated in cooler climates than sugar cane is, mainly above the 35th parallel of the Northern Hemisphere. Both sugar beets and sugar cane have high water content, accounting for about 75 percent of the total weight of the plants. The sugar content of sugar cane is 10–15 percent of the total weight, while that of sugar beets is 13–18 percent. The protein and fat content of both beets and cane is almost nil.

Under SWEETENERS, FAO includes products used for sweetening that are derived from sugar crops, cereals, fruits or milk, or that are produced by insects. This category includes a wide variety of monosaccharides (glucose and fructose) and disaccharides (sucrose and saccharose). Sweeteners are either in a crystallized state as sugar, or in thick liquid form as syrups. The traditional sources of sugar are sugar cane and sugar beets, but in recent years, ever-larger quantities of cereals (mainly maize) have been used to produce sweeteners derived from starch.

OTHER DERIVED PRODUCTS: In addition to sugar, molasses is also obtained with varying degrees of sugar content. The by-product obtained from extracting sugar from molasses is called bagasse in the case of sugar cane and beet pulp in the case of sugar beets.

FBS 2536 – Sugar cane

Includes sugar cane (01802).

Sugar cane, species of *Saccharum officinarum,* fresh, chilled, frozen or dried, ground or unground; in some producing countries, marginal quantities of sugar cane are consumed, either directly as food or in the form of juice.

Sugar cane does not include bagasse, the fibrous portion of the sugar cane remaining after the juice has been extracted.

FBS 2537 – Sugar beet

Includes: sugar beet (01801).

Sugar beet, species of *Beta vulgaris var. altissima*, fresh, chilled, frozen or dried, ground or unground; in some producing countries, marginal quantities are consumed, either directly as food or in the preparation of jams.

FBS 2541 – Sugar non-centrifugal

Includes sugar non.centrifugal (23511.02).

Generally derived from sugar cane through traditional methods without centrifugation, in the form of brown crystals or other solid forms, the colour depending on the presence of impurities, is generally destined for processing into refined sugar products. However, raw sugar may be of such a high degree of purity that it is suitable for human consumption without refining.

FBS 2827 – Sugar (raw equivalent)

A non-refined, crystallized material derived from the juices of sugar cane stalks and consisting either wholly or mainly of sucrose, and from juices extracted from the root of the sugar beet (raw, in solid form, not containing added flavouring or colouring matter) and consisting either wholly or mainly of sucrose.

FBS 2543 – Sweeteners, other

Includes sugar crops NEC (01809) and the following processed products expressed in terms of primary equivalent: fructose, chemically pure (23210.01); maltose, chemically pure (23210.02); maple sugar and syrups (23530); glucose and dextrose (23210.05); lactose (23210.06); isoglucose (23210.08); molasses (23540); other fructose and syrup (23210.03); beverages, non-alcoholic (24490); sugar NEC (23210.04).

- Fructose, chemically pure, or levulose, monosaccharide, present with glucose in sweet fruits and honey.
- Maltose, chemically pure, produced industrially from starch by hydrolysis with malt diastase. Used in the brewing industry. Includes invert sugar and other sugar and sugar syrup blends containing, in the dry state, 50 percent by weight of fructose.
- Maple sugar and syrups, produced by atmospheric boiling of maple obtained from the sap of varieties of the maple tree, chiefly *Acer saccharum* and *A. nigrum*, in an open-pan evaporator. Continuing the evaporation process until the syrup crystalizes yields maple sugar.
- Sugar crops NEC, including *Sorghum saccharatum* (sugar palm) and *Arenga saccharifera* (sweet sorghum). This subclass does not include sugar cane (01802); sugar beet (01801); sugar beet seeds (01803); and locust beans (carobs) (01356).
- Other fructose and syrup, monosaccharide found in fruits and honey, commercially produced from glucose, sucrose or by hydrolysis of inulin (polysaccharide found mainly in the tubers of the dahlia and the Jerusalem artichoke), containing, in the dry state, more than 50 percent by weight of fructose, excluding invert sugar. Especially suitable for use by diabetics. Both commercial and chemically pure fructose are included.
- Sugar and syrups NEC include invert sugar, caramel, golden syrup, artificial honey, maltose other than chemically pure, sorghum and palm sugars. (See the note in the introduction to this section.) Includes invert sugar and other sugar and sugar syrup blends containing, in the dry state, 50 percent by weight of fructose.
- Glucose and dextrose, monosaccharides produced by hydrolyzing starch with acids and/or enzymes. Dextrose is chemically pure glucose. Used in the food industry, in brewing, in tobacco fermentation and in pharmaceutical products. Includes glucose and glucose syrup, not containing fructose, or containing, in the dry state, less than 20 percent by weight of fructose

or at least 20 percent but less than 50 percent by weight of fructose. Invert sugar is not included.

- Lactose, also known as milk sugar, produced commercially from whey. Such products must contain, by weight, more than 95 percent lactose, expressed as anhydrous lactose, calculated from the dry matter. Includes both commercial and chemically pure lactose.

- Isoglucose, also known as high-fructose corn syrup (HFCS), high-fructose starch syrup (HFSS) or high-fructose glucose syrup (HFGS). A new type of starch syrup in which glucose is isomerized to fructose by using one or more isomerizing enzymes. It is the most important of the sweeteners manufactured from maize starch. Widely used in the production of food and soft drinks.

- Beverages, non-alcoholic, include sweetened or flavoured mineral waters and other non-alcoholic beverages such as lemonade, orangeade and cola. Include waters, including mineral waters and aerated waters, containing added sugar or other sweetening matter or flavoured; and other non-alcoholic beverages, not including the following fruit or vegetable juices: orange, grapefruit (including pomelo); any other single citrus fruit; pineapple; tomato; grape (including grape must); apple; cranberry (*Vaccinium macrocarpon*, *V. oxycoccos*, *V. vitis*‑*idaea*); mixtures of juices; and juice of any other single fruit or vegetable.

Molasses, a by-product of the extraction or refining of beet or cane sugar or the production of fructose from maize. Used for feed, food, industrial alcohol, alcoholic beverages and ethanol.

Other sugar crops not elsewhere classified are also included here.

FBS 2745 – Honey

Includes natural honey (02910).

Honey produced by bees (*Apis mellifera*) or other insects, centrifuged, in the comb, or containing comb chunks, provided that neither sugar nor any other substance has been added.

It excludes artificial honey and mixtures of natural and artificial honey.

7.5 Pulses and derived products

PULSES are annual leguminous crops yielding one to 12 grains or seeds of variable size, shape and colour within a pod. They are used for both food and feed. The term "pulses" is limited to crops harvested solely for dry grain, thereby excluding crops harvested green for food (green peas, green beans, etc.), which are classified as vegetable crops. Also excluded are crops used mainly for oil extraction (soybeans, groundnuts, etc.), and leguminous crops (seeds of clover, alfalfa, etc.) that are used exclusively for sowing purposes. In addition to their food value, pulses also play an important role in cropping systems because of their ability to produce nitrogen and thereby enrich the soil. Pulses contain carbohydrates, mainly starches (55–65 percent of the total weight); proteins, including essential amino acids (18–25 percent – much higher than cereals); and fat (1–4 percent). The remainder consists of water and inedible substances. Production data should be reported in terms of dry clean weight, excluding the weight of the pods. Certain kinds of pulses can be skinned and partially crushed or split to remove the seed coat, but the resulting products are still considered raw for classification purposes.

FBS – 2546 Beans

Includes: beans dry (01701).

Beans dry, species of *Phaseolus* and *Vigna* spp., *P. vulgaris* (kidney bean, white pea bean and haricot bean), *P. lunatus* (lima, butter bean), *P. angularis* or *V. angularis* (adzuki bean), *P. aureus* (mungo bean, golden, green gram), *P. mungo* or *V. mungo* (black gram, urd), *P. coccineus* (scarlet runner bean), *P. calcaratus* (rice bean), *P. aconitifolius* (moth bean), *P. acutifolius* (tepary bean), *V. radiata* and *V. aconitifolia*, dried, shelled, whether or not skinned or split.

The class does not include *V. sinensis* (cowpeas) (01706); *Vigna or Voandzeia subterranea* (Bambara beans) (01708); soybeans (0141); green beans (01241); lentils, green (01241.02); bean shoots and sprouts, (01290); locust beans (carobs) (01356); castor beans (01447); broad beans and horse beans (01243); garbanzo beans (chickpeas) (01703); lentils, dry (01704).

FBS 2547 – Peas

Includes peas dry (01705).

Peas, species of *Pisum sativum* (garden pea) and *P. arvense* (field pea), dried, shelled, whether or not skinned or split.

Includes dried peas, shelled, whether or not skinned or split; peas for fodder, dried, shelled; green peas, dried, shelled, whether or not skinned or split; peas seed, dried

FBS 2549 – Other pulses

Include broad beans (01702); chickpeas (01703); cowpeas (01706); pigeon peas (01707); lentils (01704); Bambara beans (01708); lupins (01709.02); vetches (01709.01); other pulses NEC (01709.90); and the following processed products expressed in terms of primary equivalent: flour of pulses (23170.03); bran of pulses (39120.14).

Includes the following commodities (dried and shelled, of a kind used for human or animal consumption, even if intended for sowing or other purposes):

- broad beans, horse beans, species of *Vicia faba*, mainly *V. faba var. equina* (horse bean), *V. faba var. major* (broad bean) and *V. faba var. minor* (field bean);
- chickpeas, species of *Cicer arietinum* (chickpea, Bengal gram, garbanzos);
- cowpeas, species of *Vigna unguiculata*;
- pigeon peas, species of *Cajanus cajan*;
- lentils and dhal;
- Bambara beans, species of *Vigna subterranea* or *Voandzeia subterranean*;
- lupins, vetches and similar forage products, fresh or dried, whole, cut, chopped or pressed. These products are included under this heading whether or not they have been salted or otherwise treated in a silo to prevent fermentation or deterioration.

Seeds of vetches (other than broad beans and horse beans) are not included.

Other pulses not elsewhere classified are also included here.

7.6 Nuts and derived products

Tree NUTS are dry fruits or kernels enclosed in woody shells or hard husks, which are in turn generally covered by a thick, fleshy/fibrous outer husk that is removed during harvest. Similar products, such as groundnuts, sunflower seeds and melon seeds, although often used for similar purposes, are included with oil-bearing crops. FAO includes only dessert or table nuts in nuts and derived products. Nuts that are used mainly for flavouring beverages, and masticatory and stimulant nuts are excluded, with the exception of Areca nuts and kola nuts, which FAO considers inedible nuts but that are included in the nut and derived products group to be consistent with international trade classifications. Nuts used mainly for the extraction of oil or butter (e.g. shea nuts) and nuts contained in other fruits (e.g. peaches) are excluded. It should be noted that some countries report certain nut crops (chestnuts, pignolia nuts) with forestry products. Production data relate to the weight of nuts in the shell or inner husk, but without the outer husk. The weight of the kernel contained in the nut ranges from as little as 30 percent for cashew nuts to as much as 80 percent for chestnuts. The edible portion of nut kernels is, with the major exception of chestnuts, very rich in fat content, at 50–65 percent. Protein makes up 15–20 percent and carbohydrate 10–15 percent. Starch and saccharose are the main components of dry chestnuts, accounting for about 75 percent.

NUT PRODUCTS include shelled nuts, whole or split; and further processed products, including roasted nuts, meal/flour, paste, oil, etc. Nut oils are not identified separately in the FAO classification; instead they are included under the heading Oil of vegetable origin NEC. The most commonly marketed oils are almond oil and cashew nut oil and its derivative cardol.

FBS 2551 – Nuts

Includes brazil nuts (01377); cashew nuts (01372); almonds (01371); walnuts (01376); pistachios (01375); kola nuts (01379.02); hazelnuts (01374); Areca nuts (01379.01); other shelled nuts NEC (21429.90); other nuts (excluding wild edible nuts and groundnuts), in shell NEC (01379.90); and the following processed products expressed in terms of primary equivalent: nuts, uncooked or cooked, frozen (21493.02); nut puree and nut pastes (21494.02); other prepared nuts (21495.90); nuts, provisionally preserved, not for immediate consumption (21496.02); other prepared and preserved nuts NEC (21499.03).

- Brazil nuts, *Bertholletia excelsa* (Brazil, Para or cream nut).
- Cashew nuts, *Anacardium occidentale*, produced mainly in East Africa, India and Brazil.
- Chestnuts, species of *Castanea, C. vesca, C. vulgaris* and *C. sativa*, produced mainly in Europe and Asia.
- Almonds, species of *Prunus amygdalus*, *P. communis* and *Amygdalus communis*, produced mainly in Mediterranean countries, the United States of America and Asia.
- Walnuts, species of *Jugland, J. regia*, produced in temperate zones of the Northern Hemisphere, particularly the United States of America.
- Pistachios, species of *Pistacia vera*, produced mainly in the Near East and the United States of America.

- Kola nuts, species of *Cola nitida, C. vera,* and *C. acuminata* (kola, cola, Sudan cola nut), produced mainly in Africa. Kola nuts, containing 2.4–2.6 percent caffeine, are commonly chewed by the local population. Much used in Europe and America in the production of beverages.

- Hazelnuts (filberts), species of *Corylus avellana*, produced mainly in Mediterranean countries and the United States of America.

- Areca nuts, species of *Areca catechu* (Areca, betel nut), produced mainly in the Far East. Areca nuts are used mainly as a masticatory. They contain alkaloids (arecoline and arecaidine).

- Nuts NEC, including *Carya illinoensis* (pecan nut); *Caryocar nuciferum* (butter or swarri nut); *Canarium* spp. (pili nut, Java almond, Chinese olives); *Lecythis zabucajo* (paradise or sapucaia nut); *Macadamia ternifolia* (Queensland, macadamia nut); *Pinus pinea* (pignolia nut). Other nuts that are not identified separately because of their minor relevance at the international level. Because of their limited local importance, some countries report nuts under this heading that are classified individually by FAO.

Wild edible nuts and groundnuts are not included.

Other nuts not elsewhere classified are also included here.

7.7 Oil-bearing crops and derived products

OIL-BEARING CROPS OR OIL CROPS include both annual (usually called oilseeds) and perennial plants whose seeds, fruits or mesocarp and nuts are valued mainly for the edible or industrial oils that are extracted from them. Dessert and table nuts, although rich in oil, are listed under nuts. Annual oilseed plants that are either harvested green or are used for grazing and for green manure are included in fodder crops. Some of the crops included as oil-bearing crops group are also fibre crops as both the seeds and the fibre are harvested from the same plant. Such crops include coconuts, yielding coir from the mesocarp; kapok fruit; seed cotton; linseed; and hemp seed. For several other crops, both the pulp of the fruit and the kernels are used for oil. The main crops of this type are oil-palm fruit and tallow tree seeds. Production data are reported in terms of dry products as marketed. Exceptions to this general rule include groundnuts, which are reported as groundnuts in the shell; coconuts, which are reported on the basis of the weight of the nut, including the woody shell, but excluding the fibrous outer husk; and palm oil, which is reported in terms of oil, by weight. Because of the very different natures of the various oil crops, the primary products cannot be aggregated in their natural weights to obtain total oil crops. For this reason, FAO converts these crops into either an oil or an oilcake equivalent before aggregating them.

Only 5–6 percent of the world production of oil crops is used for seed (oilseeds) and animal feed, while about 8 percent is used for food. The remaining 86 percent is processed into oil. The fat content of oil crops varies widely, from as low as 10–15 percent of the weight of coconuts to more than 50 percent of the weight of sesame seeds and palm kernels. Carbohydrates, mainly polysaccharides, account for 15–30 percent of the oilseeds, but are generally lower in other oil-bearing crops. The protein content is very high in soybeans, at up to 40 percent, but is much lower in many other oilseeds, at 15–25 percent, and lower still in some other oil-bearing crops.

PRODUCTS DERIVED FROM OIL CROPS. Edible processed products from oil crops, other than oil; include flour, flakes or grits; groundnut preparations (butter, salted nuts, candy); preserved olives; desiccated coconut; fermented and non-fermented soya products.

FBS 2555 – Soybeans

Includes soybeans (0141) and the following processed products expressed in terms of primary equivalent: soya sauce (23995.01); soya paste (23995.02); soya curd (23999.03).

Soybean, species of *Glycine soja*, is the most important oil crop. It is also widely consumed as a bean and in the form of various derived products (soymilk, soymeat, etc.) because of its high protein content.

Roasted soybeans used as a coffee substitute are not included.

FBS 2556 – Groundnuts (shelled equivalent)

Includes groundnuts (0142) and the following processed products expressed in terms of primary equivalent: prepared groundnuts (21495.01); peanut butter (21495.02).

Groundnuts (shelled equivalent), species of *Arachis hypogaea* (peanuts), whether or not shelled or broken, but not roasted or otherwise cooked. They are used as direct food and for extracting oil.

For trade data, groundnuts in shell are converted at 70 percent and reported on a shelled basis.

FBS 2557 – Sunflower seed

Includes: sunflower seed (01445).

Sunflower seed, species of *Helianthus annuus*, whether broken or unbroken. It is valued mainly for its oil; minor uses include as a human food and as feed for birds.

FBS 2574 – Rape and mustard oil

Includes rapeseed or canola oil, crude (21641.01); mustard seed oil, crude (21641.02); rape, colza and mustard oil, refined (21642).

Rape and mustard oil is obtained by dry pressure extraction of the seeds of several species. It generally contains a high level of erucic acid, and has both food and industrial uses.

- Oil of rapeseed or canola oil is obtained for food use from seeds of *Brassica*, particularly *B. napus* and *B. rapa* (or *B. campestris*). Canola oil is produced from new varieties of rapeseed. Oil recovered with solvent from the residues of the pressure extraction is used for industrial purposes, for salad dressings, to produce margarine, and in other industrial products. The refined oils (generally colza oil) are edible.
- Oil of mustard seed is obtained from seeds of *Sinapsis alba* and *B. hirta* (white mustard), *B. nigra* (black mustard) and *B. juncea* (Indian mustard). It is used in medicines, for cooking, and in industrial products.

FBS 2575 – Cottonseed oil

Includes: cottonseed oil (2168).

Cottonseed oil, species of *Gossypium*, obtained first by pressure extraction from the kernels of cotton seeds. It is used mainly as a food but also in industry. The pure refined oil is of great value as a salad or cooking oil and for making margarine and lard substitutes. The residue from the extraction process is then exposed to a solvent.

FBS 2560 – Coconut (including copra)

Includes coconut (01460); copra (01492); and the following processed products expressed in terms of primary equivalent: coconuts desiccated (21429.07).

Coconut including copra, species of *Cocos nucifera* (husked coconut), in shell, including meat, fresh, whether shredded or unshredded, covered by the endocarp, while the exocarp (smooth outer skin) and mesocarp (fibrous covering) are removed. Immature nuts contain a milky juice that is consumed as a refreshing drink. Mature nuts are consumed unprocessed, or are processed for copra or desiccated coconut. The flesh, from which copra/oil is extracted, constitutes 40–70 percent of the weight of the husked coconut. The oil content is about 36 percent of the flesh.

Copra is the dried flesh of coconut from which the oil is extracted; it is unsuitable for human consumption.

FBS 2561 – Sesame seed

Includes sesame seed (01444).

Sesame seed, species of *Sesamum indicum*, whether broken or unbroken, is valued for its oil, but also as a food, either raw or roasted, and in bakery products and other food preparations.

FBS 2576 – Palm kernel oil

Includes palm kernel oil (21691.14).

Palm kernel oil, mainly from *Elaeis guineensis* (African oil-palm), obtained from the kernel of the nut of the fruit by pressure extraction in two or three stages at different temperatures. Includes oil of babassu kernels. It is used in the margarine and candy industries and the manufacture of glycerol, shampoos, soap and candles.

FBS 2563 – Olives

Includes olives (01450) and the following processed products expressed in terms of primary equivalent: olive oil, crude (21671); olive oil, refined (21672); oil of olive residues (21673).

Olives, species of *Olea europaea*, fresh or chilled, include table olives and olives for oil.

FBS 2570 – Oil crops, other

Includes karité nuts (01499.01); castor beans (01447); tung nuts (01499.02); jojoba seeds (01499.03); safflower seeds (01446); poppy seeds (01448); melon seeds (01449.01); tallow tree seeds (01499.04); kapok fruits (01499.05); linseed (01441); hemp seeds (01449.02), other oilseeds NEC (01449.90), of kinds used for the extraction of edible or industrial oils and fats; and the following processed products expressed in terms of primary equivalent: flour of oilseeds (21920) – flours and meals of oilseeds or oleaginous fruits, except those of mustard.

- Karité nuts (shea nuts), species of *Butyrospermum parkii*. Production data refer only to the nut contained in the fruit, although the pulp around the nut is also edible.
- Castor beans, species of *Ricinus communis*, valued mainly for their oil, which is used in pharmaceutical products. Ground seedcakes are used as fertilizers (castor oil pomace).
- Tung nuts, species of *Aleurites cordata* and *A. fordii*, valued mainly for their oil.
- Jojoba seeds, species of *Simmondsia californica* and *S. chinensis*, from the shrub or small tree of the Buxaceae family.
- Safflower seeds, species of *Carthamus tinctorius*, whether broken or unbroken, valued mainly for their oil. Minor uses include as a human food and as poultry feed.
- Poppy seeds, species of *Papaver somniferum*, the source of opium, poppy seeds are also used in baking and confectionery.
- Melon seeds, species of *Cucumis melo*, include seeds of other Cucurbitaceae.

- Tallow tree seeds, species of *Shorea aptera, S. stenocarpa* (Borneo tallow tree), *Sapium sebiferum,* and *Stillingia sebifera* (Chinese tallow tree), grown wild and cultivated. FAO considers vegetable tallow (21691.09) and stillingia oil (21691.10) to be primary products.

- Kapok fruit, species of *Ceiba pentandra*, the fruit of kapok contains fibre and seeds, which FAO treats as primary crops. The soft shell is approximately 40–50 percent of the total weight of the nut; shelled kapok fruit is used for extracting oil.

- Linseed, species of *Linum usitatissimum* (flaxseed), whether broken or unbroken, is an annual herbaceous that is cultivated for its fibre and oil. It includes the seeds of the flax plant.

- Hemp seed, species of *Cannabis sativa,* is an annual herbaceous that is cultivated for its fibre and oil. In major producing countries, oil is extracted from the seeds.

Other oilseeds not elsewhere classified are also included here.

7.8 Vegetable oils and fats

VEGETABLE OILS AND FATS. Oil extraction by traditional methods often requires preliminary operations, such as cracking, shelling and dehulling, after which the crop is ground to a paste. The paste, or the whole fruit, is then boiled with water and stirred until the oil separates and can be collected. Such traditional methods have a low rate of efficiency, particularly when performed manually. Oil extracted by pressing without heating is the purest method and often produces an edible product without refining. Modern methods of oil recovery include crushing and pressing, and dissolving the crop in a solvent, usually hexane. Extracting oil with a solvent is a more efficient method than pressing. The residue left after removal of the oil (oilcake or -meal) is used as a feedstuff. Crude vegetable oils are obtained without further processing other than degumming or filtering. To make them suitable for human consumption, most edible vegetable oils are refined to remove impurities and toxic substances, a process that involves bleaching, deodorization and cooling (to make the oils stable in cold temperatures). The loss involved in these processes ranges from 4 to 8 percent.

The FAO vegetable oils and fats category includes raw, refined and fractioned oils, but not chemically modified oils. With some exceptions, and in contrast to animal fats, vegetable oils contain predominantly unsaturated (light, liquid) fatty acids of two kinds: monounsaturated (oleic acid – mainly in extra virgin olive oil); and polyunsaturated (linoleic acid and linolenic acid – in oils extracted from oilseeds). Vegetable oils have a wide variety of food uses, including as salad and cooking oils and in the production of margarine, shortening and compound fat. They are also used in many processed products, such as mayonnaise, mustard, potato chips, French fries, salad dressing, sandwich spread and canned fish. Industrial and non-food uses of vegetable oils include the production of soaps, detergents, fatty acids, paints, varnishes, resins, plastics and lubricants.

FBS 2571 – Soybean oil

Includes soybean oil (2161).

Soybean oil, crude and refined, is obtained by hydraulic or expeller presses or solvent extraction from the seeds of the soybean. It is used mainly for food, and also for industrial purposes.

FBS 2572 – Groundnut oil

Includes groundnut oil (2162).

Groundnut oil (or peanut oil), crude and refined, is obtained by pressure or solvent extraction from the seeds of nuts of the common groundnut (*Arachis hypogaea*). It is used mainly for food, and also for making soaps or lubricants.

FBS 2573 – Sunflower seed oil

Includes sunflower seed oil, crude (21631.01); sunflower seed and safflower seed oil, refined (21632).

Sunflower seed oil, whether refined or unrefined, but not chemically modified, is obtained by pressure extraction from the seeds of the sunflower (*Helianthus annuus*). It is used mainly for food, and also for making medicines, alkyd resins, paints and varnishes.

FBS 2577 – Palm oil

Includes palm oil, crude (2165) and the following processed products expressed in terms of primary equivalent: industrial monocarboxylic fatty acids; acid oils from refining (34120); residues of fatty substances (21932.02).

Palm oil, whether refined or unrefined, mainly from the African oil-palm (*Elaeis guineensis*), is obtained from the mesocarp of the fruit of the oil-palm by pressure, and also by solvent from the residues of the pressure extraction, for use in the manufacture industry. Refined palm oil is used as a foodstuff.

Palm kernel oil or babassu oil is not included (2576).

FBS 2578 – Coconut oil

Includes coconut oil (2166).

Coconut oil, whether refined or unrefined, is obtained from the copra of the coconut (*Cocos nucifera*) by pressure, and by solvent from the residue of the pressure extraction. It is used in the manufacturing industry; when refined it is used as food.

FBS 2579 – Sesame seed oil

Includes sesame oil (21691.07).

Sesame seed oil, whether refined or unrefined, but not chemically modified, is obtained from the seeds of an annual herb (*Sesamum indicum*) by pressure extraction in two or three stages at different temperatures. Sometimes the oil is also extracted by solvent from the residue of the pressure extraction. It is used mainly for food, also used for industrial purposes.

FBS 2580 – Olive oil

Includes olive oil crude and refined (21671, 21672); olive oil residues (21673).

Olive oil, whether refined or unrefined, but not chemically modified, is obtained from the fruit of the olive tree (*Olea europaea* L.) by mechanical or other physical means. Olive oil is the only vegetable oil that can be consumed without refining.

It includes oil of olive residues, extracted with solvents from olive residues left after the olives have been pressed to produce olive oil.

FBS 2581 – Rice bran oil

Includes: rice bran oil (21691.01).

Rice bran oil, whether refined or unrefined, but not chemically modified, is extracted from the bran of the seeds of an annual herb (*Sesamum indicum*) by pressure, or – more frequently – solvents.

FBS 2582 – Maize germ oil

Includes maize germ oil (21691.02).

Maize germ oil, whether refined or unrefined, but not chemically modified, is extracted from the germ of maize or Indian corn by pressure or by solvents. The refined oil is edible and is used for cooking, in bakeries and for mixing with other oils.

FBS 2586 – Oilcrop oil, other

Includes butter of karité nuts (21691.03); oil of castor beans (21691.04); oil of tung nuts (21691.05); oil of jojoba (21691.06); safflower oil (21631.01, 21632); poppy oil (21691.08); vegetable tallow (21691.09); stillingia oil (21691.10); oil of kapok (21691.11); linseed oil (21691.12); oil of hemp seed (21691.13); oil of vegetable origin NEC (21691.90); cocoa butter (23620); liquid margarine (21700.01); margarine and shortening (21700.02); castor oil, hydrogenated (21693.02); hydrogenated oils and fats (21693.03); animal or vegetable fats and oils and their fractions, chemically modified, except those hydrogenated, inter-esterified, re-esterified or elaidinized, inedible mixtures or preparations of animal or vegetable fats or oils (34550).

Other oilcrop oil includes the following vegetable fats and oils, whether refined or unrefined, but not chemically modified:

- Butter of karité nuts: a very important vegetable oil in West Africa. Used as a substitute for cocoa butter and in cosmetics.
- Oil of castor beans, from the seeds of *Ricinus communis*, obtained by pressure or solvent. Uses are mainly industrial, in pharmaceuticals and cosmetics.
- Oil of tung nuts, from the seeds of different species of the genus *Aleurites* (*A. fordii*, *A. montana*), obtained by pressure and used exclusively for industrial purposes. The resulting cake contains a toxic protein and thus cannot be used as feed.
- Oil of jojoba, from the seeds of desert shrubs of the genus *Simmondsia* (*S. californica* or *S. chinensis*), obtained by cold pressure. Its peculiar chemical properties make it the only vegetable oil with the same characteristics as spermaceti. Below 15 ˚C it solidifies and assumes the characteristics of wax. It is used as a lubricant, in cosmetics and in pharmaceuticals, and is considered a product with good growth prospects.
- Safflower oil, from the seeds of the safflower (*Carthamus tinctoris*), obtained by either pressure or solvent. Has both food and industrial uses.
- Poppy oil, obtained by pressure extraction. Has both food and industrial uses.
- Vegetable tallow, obtained by pressure extraction or solvent from the kernels of the fruit of the Borneo tallow tree and from the outer coating that surrounds the seeds of the fruit of the Chinese tallow tree. Used as a substitute for cocoa butter, and in soap, candles, medicines and cosmetics.
- Stillingia oil, obtained by solvent from the seeds of *Stillingia sebifera*. Used as a drying agent in paints and varnishes.

- Oil of kapok, obtained from shelled seeds by pressure. Used for food and soap.
- Linseed oil, from the seeds of the flax plant (*Linum usitatissimum*), obtained by pressure extraction. Used mainly in non-food items. Cold-pressed linseed oil is fit for human consumption.
- Oil of hemp seed, obtained either by pressure extraction or by solvent. Used mainly in non-food items.
- Oil of vegetable origin NEC, includes myrtle wax and Japan wax.
- Cocoa butter, obtained by hot pressing either cocoa paste or the whole bean. Includes the fat and oil. Used in chocolate making to enrich cocoa pastes, in confectionery, perfumery, cosmetics and pharmacy.
- Liquid margarine and margarine shortening, made principally from one or more hydrogenated vegetable or animal fats or oils in which is dispersed an aqueous potion containing milk products, salt, flavouring agents and other additives. Shortening is a product similar to margarine, but with a higher animal fat content. Shortening and compound fats are used primarily for baking and frying. The fat content of margarine and shortening varies from 70 to 90 percent.
- Castor oil, hydrogenated, also called “opal wax”. Vegetable oil and its fractions, partly or wholly hydrogenated, inter-esterified, re-esterified or elaidinized. Frequently used in the preparation of edible fat.
- Oils, boiled, dehydrated include oxidized and sulphurized oils. Animal and vegetable fats and oils whose chemical structure has been modified to improve viscosity, drying ability or other properties. Include:

- linoxyn;

- mixtures of animal or vegetable fats or oils or fractions of different fats or oils not elsewhere specified or included, inedible;

- oil, castor, dehydrated;

- oils, animal or vegetable, blown;

- oils, animal or vegetable, boiled;

- oils, animal or vegetable, oxidized;

- oils, animal or vegetable, polymerized by heat in vacuum or in inert gas;

- oils, animal or vegetable, sulphurized (excluding fractions);

- oils, brominated;

- oils, deep-frying, used, containing, for example, rape oil, soybean oil and a small quantity of animal fat, for use in the preparation of animal feeds;

- oils, drying (excluding liquid driers);

- oils, epoxidized;

- oils, maleic;

- oils, teka.

- preparations of animal or vegetable fats or oils or fractions of different fats or oils not elsewhere specified or included, inedible;

- stand oils.

- hydrogenated oils and fats, animal and vegetable fats and oils that have been hydrogenated to raise their melting point and increase their consistency by transforming unsaturated glycerides into saturated glycerides.

Other oilcrops not elsewhere classified are also included here.

7.9 Vegetables and derived products

VEGETABLES, as classified in this group, are mainly annual plants cultivated as field and garden crops in the open and under glass, and used almost exclusively for food. Vegetables grown principally for animal feed or seed should be excluded. Certain plants, normally classified as cereals or pulses, belong to this group when harvested green, such as green maize and green peas. This grouping differs from international trade classifications for vegetables in that it includes melons and watermelons, which are normally considered to be fruit crops. However, while fruit crops are virtually all permanent crops, melons and watermelons are similar to vegetables in that they are temporary crops. Chillies and green peppers are included in this grouping when they are harvested for consumption as vegetables and not processed into spices. FAO production data for green peas and green beans refer to the total weight including pods, although some countries report on a shelled weight basis. The weight of the pods ranges from 40–50 percent for peas to up to 70 percent for broad beans. Area data on small vegetable gardens are often omitted from agricultural surveys, although production estimates may be reported. Trade data for fresh vegetables also include chilled vegetables, meaning that the temperature of the products has been reduced to about 0 °C without the products being frozen. Vegetables contain principally water, accounting for 70–95 percent of their weight. They are low in nutrients, but contain minerals and vitamins.

PRODUCTS DERIVED FROM VEGETABLES refer to processed products. Apart from a few main products, international trade classifications do not permit sufficiently detailed classification of processed products according to the primary commodity used in their preparation. A similar situation prevails for frozen vegetables.

FBS 2601 – Tomatoes

Includes tomatoes (01234) and the following processed products expressed in terms of primary equivalent: juice of tomatoes (21321); paste of tomatoes (21399.01); tomato peeled (21399.02).

Tomatoes, species of *Lycopersicon esculentum*, fresh or chilled of all kinds.

FBS 2602 – Onions

Includes onions (01253).

Onions, species of *Allium cepa*, include onion sets, Welsh and spring onions, scallions and shallots at a mature stage, but not dehydrated.

FBS 2605 – Vegetables, other

Includes cabbages (01212); artichokes (01216); asparagus (01211); lettuce and chicory (01214); spinach (01215); cassava leaves (01219.01); pumpkins, squash and gourds (01235); cucumbers and gherkins (01232); eggplants (01233); chillies and peppers, green (*Capsicum* spp. and *Pimenta* spp.) (01231); garlic (01252); leeks and other alliaceous vegetables (01254); beans, green (01241); peas, green (01242); broad beans, green (01243); string beans (01241.01); carrots and turnips (01251); okra (01239.01); green corn (maize) (01290.01); mushrooms and truffles (01270); chicory roots (01691); carobs (01356); vegetables, fresh NEC (01219.90); and the following processed products expressed in terms of primary equivalent: sweet corn, frozen (21319.01); sweet corn, prepared or preserved (21399.03); dried mushrooms (21393.01); canned mushrooms (21397.01); juice of vegetables NEC (21329); vegetables, dehydrated (21393.90); vegetables in vinegar (21340); vegetables preserved

NEC (21394, 21395, 21396, 21399.92); vegetables, frozen (21311, 21312, 21319.90); vegetables, provisionally preserved (21330); vegetables, prepared or preserved, frozen (21394, 21399.93); homogenized vegetable preparations (23991.02); coffee substitutes (23912.01).

- Cabbages, species of *Brassica chinensis* (Chinese, mustard cabbage, pak-choi), *B. oleracea* all varieties except *B. botrytis* (white, red, savoy cabbage, Brussels sprouts, collards, kale and kohlrabi). Include edible brassicas; Brussels sprouts; cabbage (Chinese, red, savoy, spring, turnip-rooted, white); collards; kale; kohlrabi.

- Artichokes, species of *Cynara scolymus*. Does not include Jerusalem artichokes (01599); artichokes, Chinese; and artichokes, globe.

- Asparagus, species of *Asparagus officinalis*.

- Lettuce and chicory, species of *Lactuca sativa, Cichorium intybus var. foliosum* (witloof chicory), *C. endivia var. crispa* (endive) and *C. endivia var. latifolia* (escarole chicory). Includes chicory (blanched, curly, escarole and witloof); endive; escarole; lettuce (cabbage, head, romaine). Excludes chicory plants and chicory roots.

- Spinach, species of *Spinacia oleracea*. Trade figures may include New Zealand spinach (*Tetragonia espansa*) and orache (garden) spinach (*Atriplex hortensis*). Includes spinach (garden, New Zealand, orache).

- Cassava leaves, species of *Manihot esculenta* and *M. utilissima.* Young cassava leaves are eaten as a vegetable in some areas of Africa.

- Cauliflowers and broccoli, species of *Brassica oleracea var. botrytis, subvariety cauliflora* and *cymosa,* including headed broccoli.

- Pumpkins, squash and gourds, species of *Cucurbita* (squash, pumpkins, zucchini, etc.), and *Lagenaria* (gourds) genus of the Cucurbitaceae family, including marrows. Includes marrows; pumpkins; squash.

- Cucumbers and gherkins, species of *Cucumis sativu*.

- Eggplants, species of *Solanum melongena*, also called aubergines.

- Chillies and peppers, green, species of *Capsicum annuum*, and *Pimenta officinalis*. Production data exclude crops cultivated explicitly as spices. In contrast, trade data include these crops, provided that they are fresh, uncrushed and unground. Includes paprika; peppers (bell, cayenne, chilli, clove, English, Indian, Jamaica, pimento, Spanish, sweet, Turkish); pimentos (*C. frutescens*).

- Onions, shallots (green), species of *Allium ascalonicum* (shallots), *A. cepa* (onions) and *A. fistulosum* (welsh onions). Young onions pulled before the bulb has enlarged. Used especially in salads. Includes onion sets.

- Garlic, species of *Allium sativum*.

- Leeks and other alliaceous vegetables, species of *Allium porrum* (leeks) and *A. schoenoprasum* (chives). Includes vegetables, alliaceous; onions and shallots, green

(01253.01); onions and shallots, dry excluding dehydrated (01253.02). Excludes green garlic (01252).

- Beans, green, species of *Phaseolus* and *Vigna*, for shelling. Includes lima or butter beans, mung beans and beans in edible pods.

- Peas, green, species of *Pisum sativum*, mostly for shelling, but including edible – podded peas or sugar peas. Includes fodder peas.

- Broad beans, green, species of *Vicia faba*, for shelling.

- String beans, species of *Phaseolus vulgaris*, not for shelling.

- Carrots, species of *Daucus carota*, trade data may include edible turnips (*Brassica rapa var. rapifera*). Excludes forage carrots.

- Okra, species of *Abelmoschus esculentus* and *Hibiscus esculentus*, also called gombo.

- Green corn (Maize), species of *Zea mays*, particularly *Z. mays var. saccharata*, harvested green for food. Includes *Z. mays var. saccharata*, commonly known as sweet corn, whether or not on the cob.

- Mushrooms, including *Boletus edulis*, *Agaricus campestris*, *Morchella* spp. and *Tuber magnatum*, cultivated or spontaneous. Includes truffles.

- Chicory roots, species of *Cichorium intybus* and *C. sativum*, unroasted chicory roots of a kind used primarily for human consumption, whether fresh or dried, whole or chopped.

- Carobs, species of *Ceratonia silique* (carob tree, locust bean), includes seeds. Mainly used as an animal feed and for industrial purposes. Rich in pectin. Includes locust beans (or carob), with or without seeds, fresh or dried, whether or not kibbled or ground, but not further prepared; locust endosperm (or carob) bean (excluding endosperm flour); flour of locust (or carob) bean germ or pericarp; germ, locust (or carob) bean, whether or not powdered; seeds, locust (or carob) bean, fresh or dried, not roasted, whether or not kibbled or ground, but not further prepared (excluding endosperm flour).

- Vegetables fresh NEC, includes *Bambusa* spp. (bamboo shoots); *Beta vulgaris* (beets, chards); *Capparis spinosa* (capers); *Cynara cardunculus* (cardoons); *Apium graveolens* (celery); *Anthriscus cerefolium* (chervil); *Lepidium sativum* (cress); *Foeniculum vulgare* (fennel); *Cochlearia armoracia* (horseradish); *Majorana hortensis* (marjoram, sweet); *Tragopogon porrifolius* (oyster plant); *Petroselinum crispum* (parsley); *Pastinaca sativa* (parsnips); *Raphanus sativu* (radish); *Rheum* spp. (rhubarb); *Brassica napus* (rutabagas, swedes); *Satureja hortensis* (savory); *Scorzonera hispanica* (scorzonera, black salsify); *Rumex acetosa* (sorrel); *Artemisia dracunculus* (soybean sprouts, tarragon); *Nasturtium officinale* (watercress). These are not identified separately because of their minor relevance at the international level. Because of their limited local importance, some countries report vegetables under this heading that are classified individually by FAO.

- Watermelons, species of *Citrullus vulgaris*.

- Melons, cantaloupes, species of *Cucumis melo.* Include cantaloupes; melons (casaba, citron, cranshaw, honeydew, Persian, musk).

Other vegetables not elsewhere classified are also included here.

7.10 Fruits and derived products

FRUIT CROPS consist of fruits and berries that, with few exceptions, are characterized by their sweet taste. Nearly all are permanent crops, mainly from trees, bushes and shrubs, and also vines and palms. Fruits and berries grow on branches, stalks or the trunks of plants, usually singly, but sometimes grouped in bunches or clusters (e.g. bananas and grapes). Commercial crops are cultivated in plantations, but significant quantities of fruits are also collected from scattered plants that may or may not be cultivated. Although melons and watermelons are generally considered to be fruits, FAO groups them with vegetables because they are temporary crops. Fruit crops are highly perishable. Their shelf-life may be extended through the application of chemical substances that inhibit the growth of microorganisms, and through careful control of the surrounding temperature, pressure and humidity once the fruit has been picked. Fruits and berries have a very high water content accounting for 70–90 percent of their weight. They contain, in various degrees, minerals, vitamins and organic acids, some of which reside in the peel or skin. Some fruits have a high fibre content and other inedible components, so that wastage is high – for example, 60 percent for passion fruit and 35–45 percent for pineapples. The waste in temperate zone fruit is lower, generally about 10–15 percent, while berries contain very little waste. The carbohydrate content of fruits varies widely. Protein content is very low, averaging less than 1 percent, and less than that of vegetables. Fat content in fruit is negligible, with the notable exception of avocados. Fruit crops are consumed directly as food and are processed into dried fruit, fruit juice, canned fruit, frozen fruit, jam, alcoholic beverages, etc. Fruit crops are not normally grown for animal feed, although significant quantities of diseased and substandard fruits, and certain by-products of the fruit processing industry are fed to animals. Production data for fruit crops should relate to fruits actually harvested. Data on bananas and plantains should relate to the weight of single bananas or banana hands, excluding the weight of the central stalk.

FRUIT CROP PRODUCTS. Apart from a few main products, international trade classifications do not permit sufficiently detailed classification of processed products according to the primary commodity used in their preparation. Fruit crops are processed for preservation and conservation, or for transformation from one substance into another, such as sugar into alcohol. Drying and wine making are two of the oldest methods of preservation. The manufacture of fruit syrups and juices, jams, jellies, marmalades, chutneys and sauces are other traditional methods of preservation. Modern processes include canning, freezing, quick-freezing and dehydration. Other fruit products include fruit squashes (juice with some fruit tissues included), fruit nectars containing at least 30 percent fruit solids, and soft drinks that contain a very small amount of fruit juice. Essential oils are extracted from some fruits and fruit peels, while the peel of some fruit is also used in confectionery.

FBS 2611 – Oranges, mandarins

Includes oranges (01323) and the following processed products expressed in terms of primary equivalent: orange juice, single strength (21431.01); orange juice, concentrated (21431.02); tangerine juice (21439.01).

Oranges and mandarins, species of *Citrus sinensis* (common, sweet orange), *C. aurantium* (bitter orange), *C. reticulate* (mandarin, tangerine) and *C. unshiu* (clementine, satsuma), fresh or chilled. Bitter oranges are used primarily in the preparation of marmalade. The category includes oranges, green for preserving; oranges, Seville and Wilkings.

FBS 2612 – Lemons, limes

Includes lemons and limes (01322) and the following processed products expressed in terms of primary equivalent: lemon juice, single strength (21439.02); lemon juice, concentrated (21439.03).

Lemons and limes, species of *Citrus limon* (lemon), *C. latifolia* (lime), *C. aurantifolia* (sour lime) and *C. limetta* (sweet lime), fresh or chilled.

FBS 2613 – Grapefruit

Includes pomelos and grapefruits (01321) and the following processed products expressed in terms of primary equivalent: juice of grapefruit (21432); grapefruit juice, concentrated (21432.01).

Grapefruit, species of *Citrus maxima* and *C. grandis* (pomelos, shaddocks) and *C. paradise* (grapefruits, fruit of the grapefruit tree).

FBS 2614 – Citrus, other

Includes citrus fruit, NEC (01329) and the following processed products expressed in terms of primary equivalent: citrus juice, single strength (21439.04); citrus juice, concentrated (21439.05).

Citrus, other, species of *Citrus bergamia* (bergamot), *C. medica var. cedrata* (citron), *C. myrtifolia* (chinotto, fruit of the myrtle-leaved orange) and *Fortunella japonica* (kumquat), fresh or chilled. Some minor varieties of citrus are used primarily in the preparation of perfumes and soft drinks.

Other citrus fruits not elsewhere classified are also included here.

FBS 2615 – Bananas

Includes: bananas (01312).

Bananas, species *Musa sapientum*, *M. cavendishii* and *M. nana* (sweet/dessert bananas), normally eaten without further preparation. Trade figures may include dried bananas. Data should be reported excluding the weight of the central stalk.

Plantains (*M. paradisiaca*), cooking bananas, are not included (01313).

FBS 2616 – Plantains

Includes: plantains and others (01313).

Plantains, species of *Musa paradisiaca*, starchy bananas that are less sweet than other bananas generally known as a cooking banana, primarily consumed after being fried, roasted, steamed, boiled or otherwise cooked. Data should be reported excluding the weight of the central stalk.

Bananas (*M. sapientum, M. cavendishii*, *M. nana*), cooking bananas, are excluded (01312).

FBS 2617 – Apples

Includes apples (01341) and the following processed products expressed in terms of primary equivalent: apple juice, single strength (21435.01); apple juice, concentrated (21435.02).

Apples, species of *Malus pumila*, *M. sylvestris*, *M. communis* and *Pyrus malus*, suitable for dessert, making beverages or industrial purposes.

FBS 2618 – Pineapples

Includes pineapples (01318) and the following processed products expressed in terms of primary equivalent: pineapples, otherwise prepared or preserved (21491); pineapple juice (21433); juice of pineapples, concentrated (21433.01).

Pineapples, species of *Ananas comosus* and *A. sativus*, fresh, dried or chilled. Trade figures may include dried pineapples.

FBS 2619 – Dates

Includes dates (01314).

Dates, species of *Phoenix dactylifera*, fresh, dried or chilled.

FBS 2620 – Grapes

Includes grapes (01330) and the following processed products expressed in terms of primary equivalent: raisins (21411); grape juice (21434); must of grapes (24212.01).

Grapes, species of *Vitis vinifera*, fresh or chilled, whether or not rough-packed in barrels, for dessert purposes or for wine‑production, whether grown outdoors or under glass.

Includes both table and wine grapes.

FBS 2625 – Fruits, other

Include pears (01342.01); quinces (01342.02); apricots (01343); sour cherries (01344.01); peaches and nectarines (01345); plums and sloes (01346); pome fruit NEC (01349.10); strawberries (01354); raspberries (01353.01); gooseberries (01351.02); currants (01351.01); blueberries (01355.01); cranberries (01355.02); other berries and fruits of the genus *Vaccinium*, NEC (01355.90); watermelons (01221); cantaloupes and other melons (01229); figs (01315); mangoes (01316.01); guavas (01316.02); mangosteens (01316.03); avocados (01311); persimmons (01359.01); cashew apples (01359.02); kiwi fruit (01352); papayas (01317); other tropical and subtropical fruits, NEC (01319); other fruit; NEC (01359); and the following processed products expressed in terms of primary equivalent: dry apricots (21419.01); plums, dried (prunes) (21412); plum juice, single strength (21439.06); plum juice, concentrated (21439.07); mango juice (21439.08); fruit, tropical, dried NEC (including mangoes and pineapples) (21419.90); fruit, dried NEC (21419.05); fruit juice, NEC (21439.90); peaches, otherwise prepared or preserved (21492); fruits, uncooked or cooked, frozen (21493.01); jams, fruit jellies, marmalades, fruit or nut puree and fruit or nut pastes (21494); jams, fruit jellies, marmalades, fruit puree and fruit pastes (21494.01); fruit, provisionally preserved, not for immediate consumption (21496.01); other prepared and preserved fruit, NEC (21499.02); flour of fruits (23170.04); fruit, nuts, peel, sugar-preserved (23670.02); homogenized cooked fruit, prepared (23991.03).

- Pears, species of *Pyrus communis*, suitable for dessert, making beverages or industrial purposes.

- Quinces, species of *Cydonia oblonga*, *C. vulgaris* and *C. japonica*, suitable for dessert, making beverages or industrial purposes. Used mainly for making jam or jelly.

- Apricots, species of *Prunus armeniaca*.

- Sour cherries, species of *Prunus cerasus* and *Cerasus acida*.

- Cherries, species of *Prunus avium*, *Cerasus avium* (mazzard, sweet cherry) *var. duracina* (hard-fleshed cherry) and *var. juliana* (heart cherry). Includes whiteheart cherries and morello cherries.

- Peaches and nectarines, species of *Prunus persica*, *Amygdalus persica* and *Persica laevis*, including nectarines.

- Plums of all kinds (greengages, mirabelles, damsons, etc.) and sloes, species of *Prunus domestica* (greengages, mirabelles, damsons) and *P. spinosa* (sloe)

- Stone fruit fresh NEC. Other stone fruit not separately identified. In some countries, apricots, cherries, peaches, nectarines and plums are reported under this general category.

- Pome fruit, NEC. Other pome fruit not separately identified. In some countries apples, pears and quinces are reported under this general category.

- Strawberries, species of *Fragaria*.

- Raspberries, species of *Rubus idaeus.* Trade data may include blackberries, mulberries and loganberries (a cross between the raspberry and the blackberry).

- Gooseberries, species of *Ribes grossularia*. Trade data may sometimes include black, white or red currants.

- Currants, species of *Ribes nigrum* (black) and *R. rubrum* (red and white). Trade data may sometimes include gooseberries.

- Blueberries, species of *Vaccinium myrtillus* (European blueberry, wild bilberry, whortleberry) and *V. corymbosum* (American blueberry). Trade data may include cranberries, myrtle berries and other fruits of the genus *Vaccinium*.

- Cranberries, species of *Vaccinium macrocarpon* (American cranberry) and *V. oxycoccus* (European cranberry). Trade data may include blueberries, myrtle berries and other fruits of the genus *Vaccinium*.

- Berries, NEC, including species of *Morus nigra* (blackberry), *M. alba*, *M. rubra* (loganberry, white and red mulberry), *Myrtus communis* (myrtle berry) and *Gaylussacia* spp. (huckleberry, dangleberry). Other berries not separately identified. In some countries, some or all of these berries are reported under this general category.

- Watermelons, species of *Citrullus vulgaris,* and melons, cantaloupes, species of *Cucumis melo.* Includes casaba, citron, cranshaw, honeydew, Persian and musk melons.

- Figs, species of *Ficus carica*, whether or not for use in distillation.

- Mangoes, species of *Mangifera indica*. Trade figures may include dried mangoes, guavas and mangosteens, fresh and dried.

- Avocados, species of *Persea Americana*.

- Persimmons (kakis), species of *Diospyros kaki* and *D. virginiana*.

- Cashew apples, species of *Anacardium occidentale*, the thickened, fleshy stem below the cashew nut. When soft, used for jam.

- Kiwi fruit, species of *Actinidia chinensis* or *A. deliciosa*.

- Papayas, species of *Carica papaya*.

Fruit, tropical, fresh NEC. Includes:

- *Artocarpus incisa* (breadfruit);
- *Averrhoa carambola* (carambola);
- *Annona* spp. (cherimoya, custard apple);
- *Durio zibethinus* (durian);
- *Feijoa sellowiana* (feijoa);
- *Psidium guajava* (guava);
- *Spondias* spp. (hog plum, mombin);
- *Artocarpus integrifolia* (jackfruit);
- *Nephelium longan* (longan);
- *Mammea americana* (mammee);
- *Garcinia mangostana* (mangosteen);
- *Solanum quitoense* (naranjillo);
- *Passiflora edulis* (passion fruit);
- *Nephelium lappaceum* (rambutan);
- *Calocarpum mammosum* (sapote, mamey colorado);
- *Achras sapota* (sapodilla);
- *Chrysophyllum* spp. (star apple, cainito).

  Other tropical fresh fruit are not identified separately because of their minor relevance at the international level. In some countries, mangoes, avocados, pineapples, dates and papayas are reported under this general category.

Other fruit includes *Crataegus azarolus* (azarole); *Carica pentagona* (babaco); *Sambucus nigra* (elderberry); *Zizyphus jujuba* (jujube); *Nephelium litchi* (litchi); *Eriobotrya japonica* (loquat); *Mespilus germanica* (medlar); *Asimina triloba* (pawpaw); *Punica granatum* (pomegranate); *Opuntia ficus-indica* (prickly pear); *Rosa* spp. (rose hips); *Sorbus aucuparia* (rowanberry); *Sorbus domestica* (service apple); *Tamarindus indica* (tamarind); *Arbutus unedo* (tree-strawberry).

Other fruits not elsewhere classified are also included here.

7.11 Stimulant crops and derived products

COFFEE is a tropical shrub that yields fruits or cherries that are processed to free the seeds or "beans" from the fruit pulp and then from the mucilage and silver skin covering the beans. Coffee with the mucilage and skin retained is called parchment coffee. By weight, the fresh cherries consist of 45–55 percent pulp, mucilage and skin, and 45–55 percent beans. The clean beans are called "green coffee" or "clean coffee" and this is considered to be a primary crop. Coffee contains caffeine, an alkaloid. Coffee is a stimulant, not a food crop.

COCOA is a rain forest tree that is cultivated for its beans. The beans are contained in ovoid pods that grow directly on the trunk and on major branches. The beans and the white mucilage or pulp that surrounds them represent about one-third of the total weight of the pods. The fermented and dried beans are considered to be a primary crop from which various processed products are derived, including roasted beans (still in the shell) and nibs, or fragments of roasted, shelled and crushed beans. The nibs are ground to give cocoa mass, from which cocoa fat or butter is extracted by pressing. Pods, shells, pulp and cake have only limited use as an animal feed owing to their high alkaloid content. Cocoa beans contain carbohydrates, protein and – particularly – fat, making them a food crop as well as a stimulant.

TEA is a shrub of the Camellia family that is cultivated for its tender leaves. The two main varieties are *Camellia assamica* and *Thea sinensis*. The primary crop consists of the tender leaves, which may be withered, rolled, fermented and dried (black tea). Green tea is not fermented. Tea is a stimulant, not a food crop.

FBS 2630 – Coffee

Includes coffee, green (01610) and the following processed products expressed in terms of primary equivalent: coffee, decaffeinated or roasted (23911); coffee extracts (23912.02).

Coffee, species of *Coffea Arabica*, *C. robusta* and *C. liberica*. Raw coffee in all forms.

Includes coffee beans, not roasted, with or without their skins, not decaffeinated; coffee berries, as gathered from the shrub; green coffee; coffee, not roasted, not decaffeinated; coffee raw, in all forms, not decaffeinated; coffee seeds, with or without skins

FBS 2633 – Cocoa beans

Includes cocoa beans (01640) and the following processed products expressed in terms of primary equivalent: cocoa paste, not defatted (23610.01); cocoa paste, defatted (23610.02); cocoa powder, sweetened (23640); chocolate and other food preparations containing cocoa (except sweetened cocoa powder), in bulk forms (23650); chocolate and other food preparations containing cocoa (except sweetened cocoa powder), other than in bulk forms (23660).

Cocoa beans, species of *Theobroma cacao*, the seeds contained in the fruit of the cacao tree, whether or not separated from their shells, husks, skins or germs, including whole or broken, raw or roasted.

FBS 2635 – Tea

Includes tea (01620) and maté leaves (01630); and the following processed products expressed in terms of primary equivalent: extracts, essences and concentrates of tea or maté, and preparations with a basis thereof or with a basis of tea or maté (23914).

Tea, species of *Camellia sinensis*, *Thea sinensis* and *T. assaamica*. Includes green tea, unfermented; black tea, fermented; and partially fermented tea. Excludes green tea eaten as a vegetable.

Maté, species of *Ilex paraguayensis*, the dried leaves of certain shrubs of the holly family, which grow in South America. Sometimes known as “Paraguay tea” or “Jesuits’ tea”, it is prepared by infusion, in a way similar to tea, and used for drinks containing a little caffeine.

7.12 Spices

SPICES are vegetable products such as leaves, flowers, seeds and roots that are rich in essential oils and aromatic principles. They are used mainly as condiments. The FAO definitions include ten spices. For practical reasons, spices are considered to be primary crops.

Production data on spices should be reported in terms of ripe, dried or powdered products. Essential oils extracted from spices are included under tobacco and rubber, and among other crops, along with other essential oils.

FBS 2640 – Pepper

Includes pepper (*Piper* spp.), raw (01651), and the following processed products expressed in terms of primary equivalent: pepper (*Piper* spp.), processed (23921).

Pepper, species of *Piper nigrum* (black, white pepper) and *P. longum* (long pepper), is a perennial climbing vine. Includes whole, crushed or ground berries, and pepper dust and sweepings. Black pepper is produced from partially ripe berries, while white pepper is from fully ripe berries that have had the outer hull removed.

Cubeb pepper (*Piper cubeba*) is not included.

FBS 2641 – Pimento

Includes chillies and peppers, dry (*Capsicum* spp. and *Pimenta* spp.), raw (01652), and the following processed products expressed in terms of primary equivalent: chillies and peppers, dry, processed (23922).

Pimento, species of the genus *Capsicum* (capsicum sweet pepper and chilli pepper), *C. frutescens, C. annuum* (red and cayenne pepper, paprika, chillies, jalapeno pepper, anaheim pepper and pimento) and *Pimenta officinalis* (allspice, Jamaica pepper). May include sweepings, of pepper of the genus *Capsicum* or the genus *Pimenta*.

Includes fresh or dried, whether or not crushed or ground Hungarian paprika; pepper (clove, English, Indian, Sierra Leone, Spanish, sweet, Turkish, Zanzibar pepper).

Uncrushed or unground fresh pimentos are considered to be vegetables.

FBS 2642 – Cloves

Includes cloves (whole stems), raw (01656), and the following processed products expressed in terms of primary equivalent: cloves (whole stems), processed (23926).

Cloves, fruits of evergreen trees *Syzygium aromaticum*, *Eugenia caryophyllata* and *Caryophyllus aromaticus*, the whole fruit of the clove tree, including the flowers picked before maturity and dried in the sun, and the stems of the clove flowers, whether or not crushed or ground.

FBS 2645 – Spices, other

Includes vanilla, raw (01658); cinnamon and cinnamon tree flowers, raw (01655); nutmeg, mace and cardamoms, processed (23923); anise, badian, coriander, cumin, caraway, fennel and juniper berries, raw (01654); ginger, raw (01657); other stimulant, spice and aromatic crops, NEC (01699); and the following processed products expressed in terms of primary equivalent: vanilla, processed (23928); cinnamon (canella), processed (23925); nutmeg, mace and cardamoms, processed (23923); anise, badian, coriander, cumin, caraway, fennel and juniper berries, processed (23924); ginger, processed (23927); other spices and aromatics, processed (23929).

Includes the following, raw or processed:

- Vanilla, species of *Vanilla planifolia* and *V. pompona.* The fruit (or bean) of a climbing plant of the orchid family, whole, crushed or ground. Includes, whether or not fresh, crushed or ground, vanilla (pompon, long, short); vanillin.

- Cinnamon (canella), species of *Cinnamomum zeylanicum* (Ceylon cinnamon) and *C. cassia* (Chinese, common cinnamon, cassia). The inner bark of young branches of certain trees of the Laurus family. Includes cinnamon tree flowers, cinnamon fruit and cinnamon waste (chips), whether whole, crushed or ground. Includes, whether or not fresh, crushed or ground, cinnamon (bark, Ceylon dried, Chinese dried, common or fine); cinnamon tree flowers; cinnamon fruit.

- Nutmeg, whether shelled or unshelled, mace and cardamoms, species of *Myristica fragrans* (nutmeg, mace), *Elettaria cardamomum* (cluster cardamom), *Aframomum angustifolium*, *A. hambury*, *Amomun aromaticum*, *A. cardamomum* (other cardamoms) and *Aframomum melegueta* (Malaguetta pepper, grains of paradise). Nutmeg is the inner brown kernel of the fruit of the nutmeg tree. Mace is the net-like membrane between the outer shell and the kernel. Cardamom seeds are enclosed in the capsule produced by perennial herbs of the Zingiberaceae family.

- Anise, badian, fennel, corian, whether or not raw, crushed or ground, fresh or chilled.
  Includes:
  \- *Pimpinella anisum* (anise);
  \- *Illicium verum* (badian or star anise);
  \- *Carum carvi* (caraway);
  \- *Coriandrum sativum* (coriander);
  \- *Cuminum cyminum* (cumin);
  \- *Foeniculum vulgare* (fennel);
  \- *Juniperus communis* (juniper berries).

  Seeds and berries from the various plants listed are normally used as spices, but also have industrial (e.g. in distilleries) and medicinal applications. Fennel seeds, raw, are used as spice.

- Ginger, whether or not crushed or ground, species of *Zingiber officinale*. Rhizome of a perennial herb. It is also used for making beverages. Includes fresh, provisionally preserved or dried ginger. Excludes ginger preserved in sugar or syrup.

  Other spices not elsewhere classified are also included here.

7.13 Alcoholic beverages

BEVERAGES include five main groups of commodities that differ by source, use, nutritive value and commercial importance. The first group includes products usually found in nature and used mainly for drinking purposes, such as water, ice and snow. Mineral water and aerated water, even when artificially produced, are also included here. The second group includes water to which sweeteners and flavourings have been added. These beverages have been gaining large markets in recent years and represent an important contribution to food consumption in some areas because of their sweetener content (up to 20 percent by weight). The third group includes the most traditional alcoholic beverages consumed by humans. Typically, the alcohol content of these beverages, which is obtained through fermentation of many vegetable crops, varies from 3 to 25 percent. The fourth group refers to un-denatured ethyl alcohol with alcoholic strength by volume of less than 80 percent, and usually between 40 and 50 percent. This category includes all distilled alcoholic beverages, whether or not sweeteners and/or flavourings have been added. The fifth group includes products that are not for human consumption, but are included here because they are closely related to alcoholic beverages. In these products, the strength of alcohol by volume is at least 80 percent. This group includes both un-denatured and denatured alcohol.

FBS 2655 – Wine

Includes wine (24212.02); sparkling wine of fresh grapes (24211); vermouth and other wine of fresh grapes flavoured with plants or aromatic substances (24220).

Wine of fresh grapes. The final product of the alcoholic fermentation of the must of fresh grapes of all qualities, including sparkling (charged with carbon dioxide), fortified and dessert wines (generally obtained from must with a high sugar content, only part of which is converted to alcohol by fermentation).

Dessert (or liqueur) wines. Includes Canary, Cyprus, Lacryma Christi, Madeira, Malaga, Malmsey, Marsala, Port, Samos and Sherry. Includes champagne.

Vermouths and other wine of fresh grapes flavoured with plants or aromatic substances. Includes beverages made with wine of fresh grapes and flavoured with aromatic substances.

FBS 2656 – Beer of barley

Includes beer of barley, malted (24310.01); other non-alcoholic caloric beverages (24490).

Beer of barley is a beverage that may be alcoholic or non-alcoholic, and is made from fermented malted cereals (mainly barley), water and hops. Non-malted cereals may also be used. The FAO definition differs from the main international classifications in that it includes non-alcoholic beer.

FBS 2657 – Beverages, fermented

Includes wheat-fermented beverages (24230.01); rice-fermented beverages (24230.02); beer of maize, malted (24310.02); beer of millet, malted (24310.03); beer of sorghum, malted (24310.04); cider and other fermented beverages (24230.03); other non-alcoholic caloric beverages (24490).

- Wheat-fermented beverages: Low-alcohol beverages from fermented flour (e.g. Korean *jakju* and *takju*), either naturally sparkling or artificially charged with carbon dioxide. May also contain added vitamins or iron compounds. Fruit juices are excluded.

- Rice-fermented beverages: Low-alcohol beverages, such as rice wine and sake.

- Beer of maize. Prepared from either malted or unmalted cereal. Beer of millet and sorghum is a traditional beer prepared in African countries where millets are cultivated. It is normally consumed while still fermenting.

- Cider and other fermented beverages NEC (e.g. cider, perry, mead). Includes alcoholic beverages (that are not distilled) made from cereals, roots and fruits, which are not included under the other headings, such as beer from plantains and ginger.

FBS 2658 – Beverages, alcoholic

Includes spirits, liqueurs and other spirituous beverages of an alcoholic strength by volume of about 40 percent volume (24131); other spirituous beverages and un-denatured ethyl alcohol of an alcoholic strength by volume of less than 80 percent volume (24139).

Distilled alcoholic beverages. Includes un-denatured ethyl alcohol (strength by volume less than 80 percent), spirits, liqueurs and other spirituous beverages and preparations. Includes:

- alcohol, ethyl, un-denatured, of an alcoholic strength by volume of less than 80 percent;
- anisette;
- aperitifs (excluding those with a basis of wine of fresh grapes);
- aquavit;
- Armagnac;
- Arrack;
- beverages, spirituous, obtained by distilling alcohol with fruits or other plant parts;
- bitters;
- brandy (excluding from wine or grape marc);
- brandy obtained by distilling wine or grape marc;
- calvados;
- cocktails, alcoholic, ready-mixed;
- Cognac;
- cordials, alcoholic;
- crèmes (liqueurs);
- Curaçao;
- eggnog, alcoholic;
- Geneva;
- gin;
- grappa;
- juice, fruit (excluding fermented grape juice and grape must), with added alcohol; juice, grape, unfermented, with added alcohol; juice, vegetable, with added alcohol;
- kirsch;
- kummel;
- lemonade, alcoholic, unmedicated;

- liqueurs;
- mirabelle spirits;
- quetsch;
- rum;
- spirits (excluding whisky) obtained by distilling fermented mash of cereal grains; consisting of emulsions of spirit with egg yolk or cream; flavoured with caraway or cumin seeds; from bitter orange peel; from cherries; from cider; from green anise or badian; from juniper berries; from palm wine; from plums; from rice wine; obtained by distilling fermented locust bean juice, mash of potatoes, sugar cane molasses or sugar cane juice, fruits (excluding grapes); wine or grape marc; neutral, un-denatured, of an alcoholic strength by volume of less than 80 percent;
- tafia;
- vodka;
- whisky, bourbon;
- whisky, rye;
- whisky, Scotch;
- wine, distilled (excluding wine of fresh grapes).

7.14 Products from slaughtered animals

MEAT AND EDIBLE OFFALS. FAO defines MEAT as the flesh of animals used for food. In production data, meat is normally reported inclusive of bone and exclusive of meat that is unfit for human consumption. As reported by individual countries, meat production data may refer to commercial production (meat entering marketing channels), to inspected production (from animals slaughtered under sanitary inspection) or to total production (the total of these two categories plus slaughter for personal consumption). All FAO annual production data refer to total production.

Country statistics on meat production are in terms of one or more of the following concepts:

1. Live weight: the weight of the animal immediately before slaughter.

2. Killed weight: the live weight less the uncollected blood lost during slaughter.

3. Dressed carcass weight: the weight minus all parts – edible and inedible – that are removed in dressing the carcass. The concept varies widely from country to country and according to the various species of livestock. Edible parts generally include edible offals (head or head meat, tongue, brains, heart, liver, spleen, stomach, tripe and, in a few countries, other parts such as feet, throat and lungs). Slaughter fats (the unrendered fats that fall in the course of dressing the carcasses) are recorded as either edible or inedible according to country practice. Inedible parts generally include hides and skins (except in the case of pigs), hoofs and stomach contents.

Meat production data for minor animals (poultry, rabbits, etc.) are reported in one of the following three ways: ready-to-cook weight (giblets are sometimes included and sometimes excluded); eviscerated weight (including the feet and head); or dressed weight, i.e. the live weight less the blood, feathers and skin.

FAO data relate to dressed carcass weight for livestock and, wherever possible, ready-to-cook weight for poultry.

Among individual countries, one of the following three concepts is used to measure production:

(A) Production from all animals, of both indigenous and foreign origin, that are slaughtered within national boundaries.

(B) Production from the slaughter of indigenous animals plus exports of live indigenous animals during the reference period. Derived from meat production as follows: production from slaughtered animals plus the meat equivalent of all animals exported live, minus the meat equivalent of all animals imported live. As imports/exports of live animals are recorded by FAO in numbers, not weight, animal type and size are of significance.

(C) The biological production concept covers indigenous animals that are either slaughtered or exported live, plus net additions to the stock during the reference period. Derived from indigenous production as follows: indigenous production plus (or minus) the meat equivalent of the change in stock numbers during the reference period. Production is expressed in terms of live weight. Changes in the total live weight of all animals are not taken into account.

FAO uses the first concept of meat production in the construction of its FBS and for related indicators. The second concept, indigenous meat production, which measures the output of the national livestock sector, is useful mainly in the construction of index numbers of agricultural production. The third concept, biological production, is the most complete as it also reflects changes in the livestock herd, but it is not used by FAO because of difficulties in obtaining information from national reporting offices. The prices applied to indigenous meat production are derived from prices of live animals, which cover not only the value of meat, but also the value of offal, fats, hides and skins.

PROCESSED PRODUCTS FROM SLAUGHTERED ANIMALS. Meat (including chilled or frozen), edible offal, fats, and hides and skins are considered primary products. The main processed meat products are the following:

1. Cured meats include meats processed with salt and usually containing various additives (such as flavouring and preserving agents), and dried or smoked meat, such as bacon and ham made from pig meat. Pâté is a spread of finely mashed, seasoned and spiced meat or liver of pigs and poultry.

2. Sausages are highly seasoned products made from meat (usually beef or pork) that has been ground, chopped and encased. Sausages may be fresh, pickled, dry or semi-dry, cooked or uncooked and smoked or unsmoked. Sausages usually contain various additives, such as salt, onions and spices. The casings are made of either prepared animal intestines or synthetic material.

3. Other preserved meats include meat and meat offal that have been boiled, steamed, grilled, fried, roasted or otherwise cooked.

The codes and names of all livestock products – with primary products in uppercase letters and processed products in upper and lower case letters – are shown in the following list, along with any accompanying remarks.

FBS 2731 – Bovine meat

Includes meat of cattle, fresh or chilled (21111); meat of cattle, boneless (21111.01; 21131.02); beef and veal, dried, salted, smoked (21182); meat extracts (21185); sausages of beef and veal (21184.01); beef and veal preparations NEC (21186.01; 21189.01); homogenized meat preparations (23991.04).

Bovine meat, including meat of bovine animals (common trade names are beef and veal), fresh, chilled or frozen, with bone in or boneless; and buffalo meat fresh, chilled or frozen, with bone in or boneless.

Includes:

- beef, of buffalo, with bones or boneless, fresh, chilled or frozen;
- meat of buffalo, with bones or boneless, fresh, chilled or frozen; boneless, packed with salt as a temporary preservative during transport;
- meat of buffalo, or bovine animals, with bones or boneless, fresh or chilled;
- Meat of buffalo, or bovine animals, with bones, packed with salt as a temporary preservative during transport.

FBS 2732 – Mutton and goat meat

Includes meat of sheep, fresh or chilled (21115); meat of sheep, frozen (21135); meat of goat, fresh or chilled (21116); meat of goat, frozen (21136).

Mutton and goat meat, including meat of sheep (rams, ewes and lambs) and goat, whether domestic or wild, fresh, chilled or frozen, with bone in or boneless.

Meat of lamb comes from animals of the ovine species not more than 12 months of age.

FBS 2733 – Pig meat

Includes meat of pigs with the bone, fresh or chilled (21113.01); meat of pigs, frozen (21133); meat of pigs, boneless, fresh or chilled (21113.02); pig meat, cuts, salted, dried or smoked (bacon and ham) (21181); sausages and similar products of meat, offal or blood of pigs (21184.02); prepared dishes and meals based on meat of pigs (21186.02); prepared or preserved meat, meat offal or blood of pigs (21189.02).

Pig meat, with the bone in, of domestic or wild pigs (wild boars, swines, etc.), whether fresh, chilled or frozen. Includes pig meat, excluding butcher fat and bones. Includes ham, fresh or chilled.

FBS 2734 – Poultry meat

Includes meat of chickens, fresh or chilled (21121); meat of guinea fowl, fresh or chilled (21125); meat of chickens, frozen (21141); meat of guinea fowl, frozen (21145); fatty liver preparations (21189.05); prepared dishes and meals based on meat poultry (21186.03); prepared or preserved meat, meat offal or blood of poultry (21189.03); meat of ducks, fresh or chilled (21122); meat of ducks, frozen (21142); meat of geese, fresh or chilled (21123); meat of geese, frozen (21143); meat of turkeys, fresh or chilled (21124); meat of turkeys, frozen (21144).

Poultry meat may include all types of poultry meat if national statistics do not report separate data.

Includes capons, chickens, fowls, domestic (*Gallus domesticus* spp.); guinea fowl; ducks, geese and turkeys; whether or not cut up, fresh, chilled or frozen.

Includes chicken, goose and duck livers; fatty livers of geese and ducks, which are distinguished from other livers by being much larger and heavier, firmer and richer in fat.

FBS 2735 – Meat other

Includes meat of pigeons and other birds NEC, fresh, chilled or frozen (21170.01); horse meat, fresh or chilled (21118.01); meat of asses, fresh or chilled (21118.02); meat of mules, fresh or chilled (21118.03); meat of horses and other equines, frozen (21138); meat of camels, fresh or chilled (21117.01); meat of other domestic camelids, fresh or chilled (21117.02); meat of camels and camelids, frozen (21137); meat of rabbits and hares, fresh or chilled (21114); meat of rabbits and hares, frozen (21134); meat of other domestic rodents, fresh or chilled (21119.01); game meat, fresh, chilled or frozen (21170.02); other meat of mammals, frozen (21139); other meat and edible meat offal, salted, in brine, dried or smoked, edible flours and meals of meat or meat offal (21183); other meat of mammals, fresh or chilled (21119.90); other meat NEC (excluding mammals), fresh, chilled or frozen (21170.92); other sausages and similar products of meat, offal or blood NEC (21184.03); other prepared dishes and meals based on meat (21186.90); other prepared or preserved meat, meat offal or blood NEC (21189.90); snails, fresh, chilled, frozen, dried, salted or in brine, except sea snails (02920).

Other meat includes fresh, chilled or frozen meat of the following animals: pigeons and other birds NEC; horses; asses; mules; camels; rabbits (may include hare meat); other domestic rodents and camelids; game (meat and offal of wild animals); snails, other than sea snails.

Meat NEC includes frog legs and marine mammals, fresh, chilled or frozen. Some countries include meats that are listed above under this heading, rather than reporting them separately.

Other meats not elsewhere classified are also included here.

FBS 2736 – Offal, edible

Includes edible offal of cattle, fresh, chilled or frozen (21151); liver preparations (21189.04); edible offal of buffalo, fresh, chilled or frozen (21152); edible offal of sheep, fresh, chilled or frozen (21155); edible offal of goats, fresh, chilled or frozen (21156); edible offal of pigs, fresh, chilled or frozen (21153); edible offal and liver of chickens and guinea fowl, fresh, chilled or frozen (21160.01); edible offal and liver of geese, fresh, chilled or frozen (21160.02); edible offal and liver of ducks, fresh, chilled or frozen (21160.03); edible offal and liver of turkey, fresh, chilled or frozen (21160.04); edible offal of horses and other equines, fresh, chilled or frozen (21159.01); raw hides and skins of goats NEC (02954.90); edible offal and liver of chickens and guinea fowl, fresh, chilled or frozen (21160.01); edible offal of mammals NEC, fresh, chilled or frozen (21159.90); offal NEC (excluding mammals), fresh, chilled or frozen (21170.93).

Edible offal includes fresh, chilled or frozen offal of the following animals: cattle, bovine animals, buffaloes, sheep, goats, pigs, horses, camels; tongues, livers, heads and cuts thereof (including ears), feet, tails, hearts, udders, livers, kidneys, sweetbreads (thymus glands and pancreas), brains, lungs, throats, thick skirts, thin skirts, spleens, tongues, caul, spinal cords, edible skin, reproductive organs (e.g. uteri, ovaries and testes), thyroid glands, pituitary glands; liver of any animal (excluding fatty livers of ducks and geese when cooked, prepared or preserved, such as pâté), chickens, guinea fowls, turkey, geese and ducks (the poultry offal of greatest importance in international trade is chicken, goose or duck livers, including the fatty livers of geese or ducks).

Edible offal includes offal of primates; whales, dolphins and porpoises, fresh; manatees and dugongs; seals, sea lions and walruses; reptiles; rabbits; hares; frogs; reindeer; beaver; turtle.

7.15 Animal fats and oils

Includes animal fats that are obtained in the course of dressing the carcasses of slaughtered animals (slaughter fats), or at a later stage in the butchering process when meat is being prepared for final consumption (butcher fats). Butter and similar products obtained from milk are included in products from live animals. Processed animal fats include lard obtained by melting raw pig fat, and tallow obtained from raw fat of other animal species. Animal fats are largely used in the production of margarine, shortening and compound fat. They are also used in many processed food products. Industrial and non-food uses of animal fats include the production of soaps, fatty acids, lubricants and feedstuffs.

FBS 2740 – Butter, ghee

Includes butter of cow milk (22241.01); ghee from cow milk (22241.02); butter of buffalo milk (22242.01); ghee from buffalo milk (22242.02); butter and ghee of sheep milk 922249.01); butter of goat milk (22249.02).

Butter of cow milk is an emulsion of milk fat and water that is obtained by churning cream. Trade data cover butter from the milk of any animal.

Ghee from cow milk derives from butter from which the water has been removed and is very common in hot countries. Includes also anhydrous butterfat or butter oil.

Includes natural butter, whey butter and recombined butter (fresh, salted or rancid, including canned butter). Butter must be derived exclusively from milk, with no added emulsifiers, but may contain sodium chloride, food colours, neutralizing salts and cultures of harmless lactic acid-producing bacteria.

Butter and ghee obtained from goat or sheep milk is also reported in this group.

FBS 2737 – Fats, animals, raw

Includes cattle fat, unrendered (21512); cattle, butcher fat (21512.01); buffalo fat, unrendered (21513); sheep fat, unrendered (21514); wool grease (21519.01); lanolin (21529.01); goat fat, unrendered (21515); fat of pigs (21511.01); pig, butcher fat (21511.02); pig fat, rendered (21521); fat of poultry (21511.03); poultry fat (21522); fat of camels (21519.02); edible offal of poultry, fresh, chilled or frozen (21160); animal oils and fats NEC (21529.03); lard stearine and lard oil (21529.02); degras (21932.01); tallow (21523); fat preparations NEC (21693.01); animal fats and their fractions, partly or wholly hydrogenated, inter-esterified, re-esterified or elaidinized, whether refined or unrefined, but not further prepared (21590).

Animals fats, raw are unrendered slaughter fats from different animals, including edible and inedible fats that are removed in the course of dressing the carcass, whether refined or unrefined, but not chemically modified.

Includes other animal oils and fats NEC obtained from other animal species, and oils and fats recovered from guts, feet, sweepings, hide trimmings, etc.

Includes fat (from bone of bovine animals, sheep or goats, rendered or solvent-extracted; from waste of bovine animals, sheep or goats, rendered); premier jus (oleo stock); tallow (beef and mutton, whether or not fit for human consumption).

7.16 Products from live animals

Includes milk, eggs, honey and beeswax. Fibres of animal origin (mainly wool and silk) are included in fibres of vegetal and animal origin.

MILK AND DAIRY PRODUCTS. Estimates of milk production reported by countries refer to one or more of the following three concepts: gross production is milk production plus milk sucked by young animals; net production, which excludes milk sucked by young animals but includes milk fed to livestock; and production available for consumption, which is net production less milk fed to animals, milk retained by farmers for food and feed, direct sales to consumers and farm waste.

The FAO data relate to net milk production. Data should be reported by kind of milking animal (cow, sheep, goat, etc.) in terms of whole milk and by weight.

In most developed countries, only 5–10 percent of whole milk is used directly for human consumption. The bulk of milk production is processed before being marketed as liquid milk (e.g. standardized, pasteurized, skimmed, etc.), or is manufactured into products such as cream, butter, cheese, evaporated and condensed milk, milk powder, casein, yogurt, ice cream, etc. About 70 percent of whole milk is processed into dairy products; the by-products of these processes (e.g. skim milk, buttermilk and whey) are used either for feed or are manufactured into other dairy products, such as dry skim milk and low-fat cheese. Processed milk and dairy products are often supplemented with vitamins, minerals and various additives.

The following is FAO's list of 50 milk and dairy product items, of which five are primary products. Some food products containing milk are not listed separately by FAO, such as eggnog, sherbet, malted milk, chocolate milk drink and mellorine.

EGGS AND EGG PRODUCTS. Egg production by type of poultry should refer to the total production of eggs in the shell by all types of hens in both the traditional sector (individually owned small flocks) and the modern sector (large-scale, intensive commercial poultry farms). Total production includes eggs for hatching, but excludes waste on farms. Countries should report in terms of both numbers and weight.

FAO lists seven egg and egg product items, including four primary and three processed products.

HONEY AND BEESWAX. Honey is the nectar of flowers collected and processed by certain insects, especially the honey bee. Production data should cover the amount sold by beekeepers plus other recorded collection of honey. Bees store honey in honeycombs that consist of hexagonal wax cells. The beeswax that is obtained by melting honeycombs with boiling water is used in candles, cosmetics and other non-food products.

The FAO codes and names of milk and dairy products, eggs and egg products, and honey and beeswax are listed in the following, along with necessary remarks.

FBS 2744 – Eggs

Includes hen eggs, in shell, fresh, for hatching (02311); other hen eggs, in shell, fresh (02312); and the following processed products expressed in terms of primary equivalent: eggs, liquid (23993.02); eggs, dried (23993.03); eggs from other birds, in shell, fresh, for hatching (02321); other eggs from other birds, fresh, in shell (02322); egg albumin (23993.01).

Eggs of chickens, ducks, geese, ostriches, quail and turkeys, fresh, in shell, not for hatching, weight in shell.

Fertilized eggs for incubation and other fresh (including chilled) eggs of all birds are included. It also covers preserved or cooked eggs, in shell.

FBS 2848 – Milk, excluding butter

Includes raw milk of cattle (02211) and the following processed products expressed in terms of primary equivalent: skim milk of cows (22110.02); whole milk, condensed (22222.01); whey, condensed (22130.03); yoghurt (22230.01); yoghurt with additives (22230.02); buttermilk (22230.03); whole milk, evaporated (22221.01); skim milk, evaporated (22221.02); skim milk, condensed (22222.02); whole milk powder (22211); skim milk and whey powder (22212); buttermilk, dry (22230.04); whey, dry (22130.02); cheese from whole cow milk (22251.01); cheese from skimmed cow milk (22251.02); whey cheese (22251.03); processed cheese (22251.04); reconstituted milk (22110.03); casein (22260); raw milk of buffalo (02212); skim milk of buffalo (22110.04); cheese from milk of buffalo, fresh or processed (22252); raw milk of sheep (02291); cheese from milk of sheep, fresh or processed (22253); skim sheep milk (22110.05); raw milk of goats (02292); cheese from milk of goats, fresh or processed (22254); skim milk of goat (22110.06); raw milk of camel (02293); whey, fresh (22130.01); dairy products NEC (22290); ice cream and other edible ice (22270).

Fresh milk of cattle (cow and yak), buffalo, sheep, goat and camel.

Production data refer to raw milk containing all its constituents. Trade data normally cover milk from any animal, and refer to milk that is not concentrated, pasteurized, sterilized or otherwise preserved, homogenized or peptonized. It includes raw milk.

Butter is excluded (cf. FBS 2740: 22241.01, 22242.01, 22249.01, 22249.02).

References

Buiten, G., Kampen, J. & Vergouw, S. 2009. *Producing historical time series for STS-statistics in NACE Rev.2: Theory with an application in industrial turnover in the Netherlands (1995–2008).* Discussion Paper No. 09001. The Hague, Statistics Netherlands. HTTP://WWW.CBS.NL/NR/RDONLYRES/A8A9AB3B-37F6-480A-BA76-253979DED22D/0/200901X10PUB.PDF

FAO. 2001. *Food balance sheets: A handbook*. Rome. HTTP://WWW.FAO.ORG/DOCREP/003/X9892E/X9892E00.HTM

FAO. no date. Definition and classification of commodities. HTTP://WWW.FAO.ORG/WAICENT/FAOINFO/ECONOMIC/FAODEF/FAODEFE.HTM

FAO. no date. Metadata of food balance sheets. HTTP://FAOSTAT.FAO.ORG/SITE/655/DEFAULT.ASPX

Hancock, A. 2013. Best practice guidelines for developing international statistical classifications. Presented at the Meeting of the Expert Group on International Classifications, UNSD, New York, 13–15 May 2013. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2013/AC267-5.PDF

Ramaschiello, V. 2011. CPC Ver.2 review and harmonization with food and agriculture statistics in FAO. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 18–20 May 2011. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2011/AC234-15.PDF

Ramaschiello, V. 2013. Aligning classifications for agricultural statistics with other international standards. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 13–15 May 2013. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2013/AC267-21.PDF

Ramaschiello, V. 2015. CPC implementation and other activities on classifications in FAO. Presented at the Expert Group Meeting on International Classifications, UNSD, New York, 19–22 May 2015. HTTP://UNSTATS.UN.ORG/UNSD/CLASS/INTERCOP/EXPERTGROUP/2015/AC289-16.PDF

UNSD. 2015. Central Product Classification (CPC) Ver.2.1. HTTP://UNSTATS.UN.ORG/UNSD/CR/REGISTRY/CPC-21.ASP

WCO. 2012. Harmonized System 2012. HTTP://WWW.WCOOMD.ORG/EN/TOPICS/NOMENCLATURE/INSTRUMENT-AND-TOOLS/HS_NOMENCLATURE_2012/HS_NOMENCLATURE_TABLE_2012.ASPX