

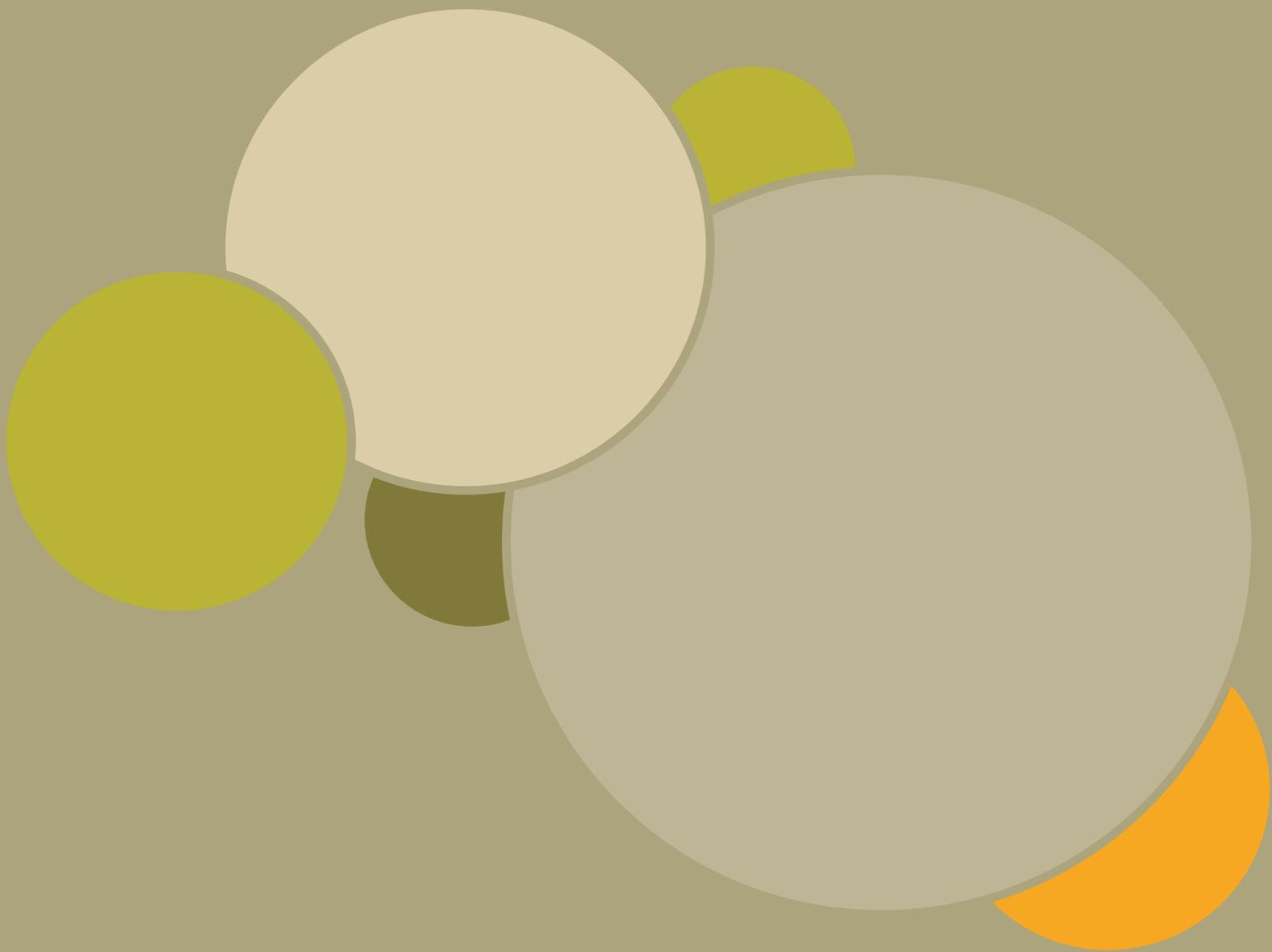
# Guidelines for the compilation of Food Balance Sheets





# **Guidelines for the compilation of Food Balance Sheets**

**October 2017**



# Contents

<b>Acronyms</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>Preface</b>	<b>ix</b>
<b>CHAPTER 1</b>	
<b>INTRODUCTION</b>	<b>1</b>
1.1. Overview	1
1.2. What is a Food Balance Sheet?	2
1.3. Potential uses of FBS	3
1.4. Caution in interpreting FBS estimates	7
1.5. Fundamental principles of FBS construction	9
1.6. Summary	10
<b>CHAPTER 2</b>	
<b>METHODOLOGICAL PRINCIPLES FOR THE CONSTRUCTION OF COUNTRY-LEVEL FBS</b>	<b>11</b>
2.1. Overview	11
2.2. The basic identity and approach	11
2.2.1. Supply and use variables	13
2.2.2. Additional variables	16
2.3. Linking Supply Utilization Accounts (SUAs) to FBS through standardization using commodity trees	19
2.3.1. Commodity trees	21
2.3.2. Processing shares	23
2.4. The recommended balancing mechanism	25
2.4.1. Approaches to distributing the imbalance at FBS level	28
2.4.2. Constraints on the balancing process	33
2.5. Summary	34
<b>CHAPTER 3</b>	
<b>DATA FOR FBS COMPILATION: CONSIDERATIONS, SOURCES AND IMPUTATION</b>	<b>35</b>
3.1. Overview	35
3.2. Assembling a TWG and assigning roles	36
3.3. Determining product scope	36
3.4. Data assessment and other preliminary considerations	37
3.4.1. Data comparability	37
3.4.2. Data quality, flags and tolerance intervals	41
3.4.3. Data search and assessment	46

3.5.	Suggested data sources and imputation	48
3.5.1.	Production	48
3.5.2.	Trade (imports and exports)	54
3.5.3.	Stocks and stock changes	57
3.5.4.	Food availability	60
3.5.5.	Food processing	64
3.5.6.	Feed	66
3.5.7.	Seed	69
3.5.8.	Tourist food	73
3.5.9.	Industrial use	75
3.5.10.	Loss	76
3.5.11.	Residual and other uses	80
3.5.12.	Additional parameters	80
3.6.	Summary	82
<b>CHAPTER 4</b>		
<b>STEP-BY-STEP GUIDE TO FBS CONSTRUCTION</b>		<b>83</b>
4.1.	Introduction	83
4.2.	Filling in an SUA table	84
4.3.	Balancing SUA accounts for derived commodities	91
4.4.	Standardization and Aggregation	92
4.4.1.	Standardization rules by variable	92
4.4.2.	Standardized and aggregated table	96
4.5.	Balancing	96
4.6.	Nutrient supplies and calorie estimates	101
4.7.	Deriving per capita estimates	103
4.8.	Validation and troubleshooting unfeasible solutions	104
4.9.	Validation through working groups	106
4.10.	Summary	106
<b>CHAPTER 5</b>		
<b>DATA QUALITY CONSIDERATIONS, DISSEMINATION AND FBS INTERPRETATION</b>		<b>107</b>
5.1.	Introduction	107
5.2.	Quality considerations	108
5.3.	Dissemination	111
5.3.1.	The case for wider dissemination of FBS data	111
5.3.2.	Suggested presentation	112
5.3.3.	Metadata	115
5.4.	Interpretation	116
5.5.	Summary	117
<b>REFERENCES</b>		<b>118</b>
<b>ANNEX 1: TOURIST FOOD</b>		<b>123</b>

## TABLES

<b>TABLE 2-1</b>	Sample blank SUA table for paddy rice.	19
<b>TABLE 2-2</b>	Unbalanced sorghum supply and utilization in Country Z.	28
<b>TABLE 2-3</b>	Step 1 – Unbalanced sorghum table with quantified error.	29
<b>TABLE 2-4</b>	Step 2 – Sum individual variable errors to calculate aggregated error.	29
<b>TABLE 2-5</b>	Step 3 – Calculate proportion of aggregated error for each individual variable.	30
<b>TABLE 2-6</b>	Step 4 – Distribute the imbalance proportionally.	31
<b>TABLE 3-1</b>	Sample flags denoting source.	42
<b>TABLE 3-2</b>	Examples of tolerance intervals by variable.	46
<b>TABLE 3-3</b>	Sample data assessment grid.	47
<b>TABLE 3-4</b>	Hypothetical sown area, harvested area and ratiosh for sunflowers in country A.	71
<b>TABLE 4-1</b>	Blank oats SUA table	84
<b>TABLE 4-2</b>	Oats SUA table with official production data added.	85
<b>TABLE 4-3</b>	Oats SUA table with official trade data added.	85
<b>TABLE 4-4</b>	Oats SUA table with imputed stock change data added.	86
<b>TABLE 4-5</b>	Oats SUA table with calculated food processing data added.	87
<b>TABLE 4-6</b>	Oats SUA table with calculated bran of oats production data added.	88
<b>TABLE 4-7</b>	Oats SUA table with imputed food data added.	88
<b>TABLE 4-8</b>	Oats SUA table with official feed data on oats added.	89
<b>TABLE 4-9</b>	Oats SUA table with imputed seed use data added.	89
<b>TABLE 4-10</b>	Oats SUA table with imputed tourist food added.	90
<b>TABLE 4-11</b>	Oats SUA with estimated industrial use added.	90
<b>TABLE 4-12</b>	Oats SUA with imputed loss added.	91
<b>TABLE 4-13</b>	Balance check on SUA accounts for commodities derived from oats.	91
<b>TABLE 4-14</b>	Balanced SUAs for products derived from oats.	92
<b>TABLE 4-15</b>	Standardization of commodities to primary equivalent, and subsequent aggregation.	94
<b>TABLE 4-16</b>	Unbalanced FBS table for oats commodity equivalent.	96
<b>TABLE 4-17</b>	Calculating the imbalance in the unbalanced FBS table.	97
<b>TABLE 4-18</b>	Tolerance interval percentages assigned.	98
<b>TABLE 4-19</b>	Errors quantified and aggregated error calculated.	98
<b>TABLE 4-20</b>	Proportion of aggregated error for each individual variable calculated.	99
<b>TABLE 4-21</b>	Distributing the imbalance proportionally.	100
<b>TABLE 4-22</b>	Food quantities scaled up based on balanced FBS.	101
<b>TABLE 4-23</b>	Converting food quantities into nutrient values.	102
<b>TABLE 4-24</b>	Nutrient values are summed to one another.	102
<b>TABLE 4-25</b>	Per capita nutrients calculated.	103
<b>TABLE 4-26</b>	Balanced FBS table for oats.	104
<b>TABLE 5-1</b>	Sample presentation 1 of data table including flags.	109
<b>TABLE 5-2</b>	Sample presentation 2 of data table detailing sources and tolerance intervals in table footnotes.	109
<b>TABLE 5-3</b>	Sample table layout of FBS data organized by year.	112
<b>TABLE 5-4</b>	Sample table layout of FBS data organized by commodity.	113

## FIGURES

<b>FIGURE 2-1</b>	Mushroom commodity tree.	22
<b>FIGURE 2-2</b>	Olive commodity tree.	22
<b>FIGURE 3-1</b>	Avocado yields in the United States of America, 1961–2013.	51
<b>FIGURE 5-1</b>	Sample visualization of FBS data, illustrating dietary composition by commodity group in a given year.	114
<b>FIGURE 5-2</b>	Sample visualization of FBS data, illustrating changing DES from 1994 to 2013.	115
<b>FIGURE 6-1</b>	Simplified representation of net visitors to country J.	124
<b>FIGURE 6-2</b>	Representation of inbound visitor flows to country J.	126
<b>FIGURE 6-3</b>	Comprehensive representation of net tourism flows for country J.	127

## BOXES

<b>BOX 2-1</b>	Sample exercise on the application of processing shares.	24
<b>BOX 3-1</b>	Example of imputation of loss using a hierarchical approach.	79

# Acronyms

<b>AFZ</b>	Association for Animal Production (Association Française de Zootechnie, France )
<b>AGA</b>	Customs General Administration (Administración General de Aduanas de Mexico, Mexico)
<b>AGMEMOD</b>	Agriculture Member States Modelling (European Union)
<b>AMIS</b>	Agricultural Market Information System
<b>CILSS</b>	Permanent Interstates Committee for Drought Control in the Sahel ( <i>Comité Permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel</i> )
<b>CIRAD</b>	Agricultural Research Centre for International Development ( <i>Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France</i> )
<b>CPC</b>	Central Product Classification
<b>DDGS</b>	Distiller's Dried Grains with Solubles
<b>DES</b>	Dietary Energy Supply
<b>EU</b>	European Union
<b>FAO</b>	Food and Agriculture Organization of the United Nations
<b>FBS</b>	Food Balance Sheet
<b>FCL</b>	FAOSTAT Commodity List
<b>FEWS NET</b>	Famine Early Warning System Network
<b>FUSIONS</b>	Food Use for Social Innovation by Optimising Waste Prevention Strategies
<b>g</b>	gram
<b>GDP</b>	Gross Domestic Product
<b>GIEWS</b>	Global Information and Early Warning System
<b>GS</b>	Global Strategy to Improve Agricultural and Rural Statistics
<b>ha</b>	hectare
<b>HS</b>	Harmonized System
<b>ICBT</b>	Informal Cross Border Trade
<b>ICEC</b>	Interagency Commodity Estimates Committee
<b>IDR</b>	Import Dependency Ratio
<b>IFPRI</b>	International Food Policy Research Institute
<b>INFOODS</b>	International Network of Food Data Systems
<b>INRA</b>	National Institute for Agricultural Research ( <i>Institut National de la Recherche Agronomique, France</i> )
<b>INTERFAIS</b>	International Food Aid Information System
<b>kcal</b>	Kilocalorie
<b>MAFW</b>	Ministry of Agriculture & Farmers Welfare (India)
<b>METIS</b>	Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata
<b>nes</b>	not elsewhere specified
<b>NCT</b>	Nutrient Conversion Table
<b>NSO</b>	National Statistics Office
<b>OCE</b>	Office of the Chief Economist
<b>OECD</b>	Organisation for Economic Cooperation and Development
<b>PoU</b>	Prevalence of Undernourishment
<b>PS&amp;D</b>	Production, Supply and Distribution
<b>SAGARPA</b>	Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food ( <i>Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, Mexico</i> )
<b>SDG</b>	Sustainable Development Goal
<b>SDMX</b>	Statistical Data and Metadata eXchange

<b>SE</b>	Secretariat of the Economy ( <i>Secretaría de Economía</i> , Mexico)
<b>SHCP</b>	Secretariat of Finance and Public Credit ( <i>Secretaría de Hacienda y Crédito Público</i> , Mexico)
<b>SNA</b>	System of National Accounts
<b>SSR</b>	Self-Sufficiency Ratio
<b>SUA</b>	Supply Utilization Account
<b>TWG</b>	Technical Working Group
<b>UN</b>	United Nations
<b>UNECE</b>	United Nations Economic Commission for Europe
<b>UNPD</b>	United Nations Population Division
<b>UNSD</b>	United Nations Statistics Division
<b>UNWTO</b>	United Nations World Tourism Organization
<b>USDA</b>	United States Department of Agriculture
<b>WASDE</b>	World Agricultural Supply and Demand Estimates
<b>WCO</b>	World Customs Organization
<b>WTO</b>	World Trade Organization
<b>WFP</b>	World Food Programme

# Acknowledgments

These Guidelines for the compilation of Food Balance Sheets have been developed under the Global Strategy to improve Agricultural and Rural Statistics (Global Strategy), with the intention to provide countries with a user-friendly handbook that can aid in the construction of country-level Food Balance Sheets (FBS) as a tool for policy analysis. Although this work was prepared by Katherine Baldwin (Statistician, Production, Trade, and Food Balance Sheets Team, FAO Statistics Division (ESS), the approaches detailed here were developed over a period of years by a multitude of other individuals within ESS. The underlying theoretical framework behind this “improved FBS methodology” is an innovation that should be credited to Josef Schmidhuber and Adam Prakash (both formerly of ESS), with additional contributions from statistical consultants, including Joshua Browning and Michael Kao. The approaches pioneered by these individuals have been further refined and adapted by the ESS Methodological Innovation Team, under the direction of Carola Fabi.

It is important to note that these guidelines suggest imputation approaches and data sources for country-level compilers, such that the strategies outlined here are not identical to those implemented indiscriminately for all countries in the FAO-compiled FBS.

Aside from the innovators of the approaches themselves, these guidelines would not have been possible without the valuable feedback of multiple reviewers who helped to focus and streamline the content of this text, including: Valérie Bizier – Technical Assistance and Training Coordinator, Global Strategy

Rachele Brivio – Statistical Analyst, Production, Trade and Food Balance Sheets Team, FAO

Franck Cachia – Statistician, Global Strategy

Carola Fabi – Team Leader, Methodological Innovation Team, FAO

Tomasz Filipczuk – Statistician, Production, Trade and Food Balance Sheets Team, FAO

Natalia Golini – Statistical Methodologist, FAO Production, Trade and Food Balance Sheets Team

Irina Kovrova – Statistician, Production, Trade and Food Balance Sheets Team, FAO

Aliou Mballo – Statistician, Global Strategy

Cristina Muschitiello – Statistician, Methodological Innovation Team, FAO

Vincent Ngendakumana – Coordinator, Statistical Capacity Building Division, African Development Bank

Francesca Rosa – Statistician, Methodological Innovation Team, FAO

The contribution of country-level FBS peer reviewers was also invaluable to ensuring the practicality and comprehensiveness of these guidelines. In this respect, we would like to extend our gratitude to:

P.M.P. Anura Kumara – Director, Agriculture and Environment Division, Department of Census & Statistics, Sri Lanka

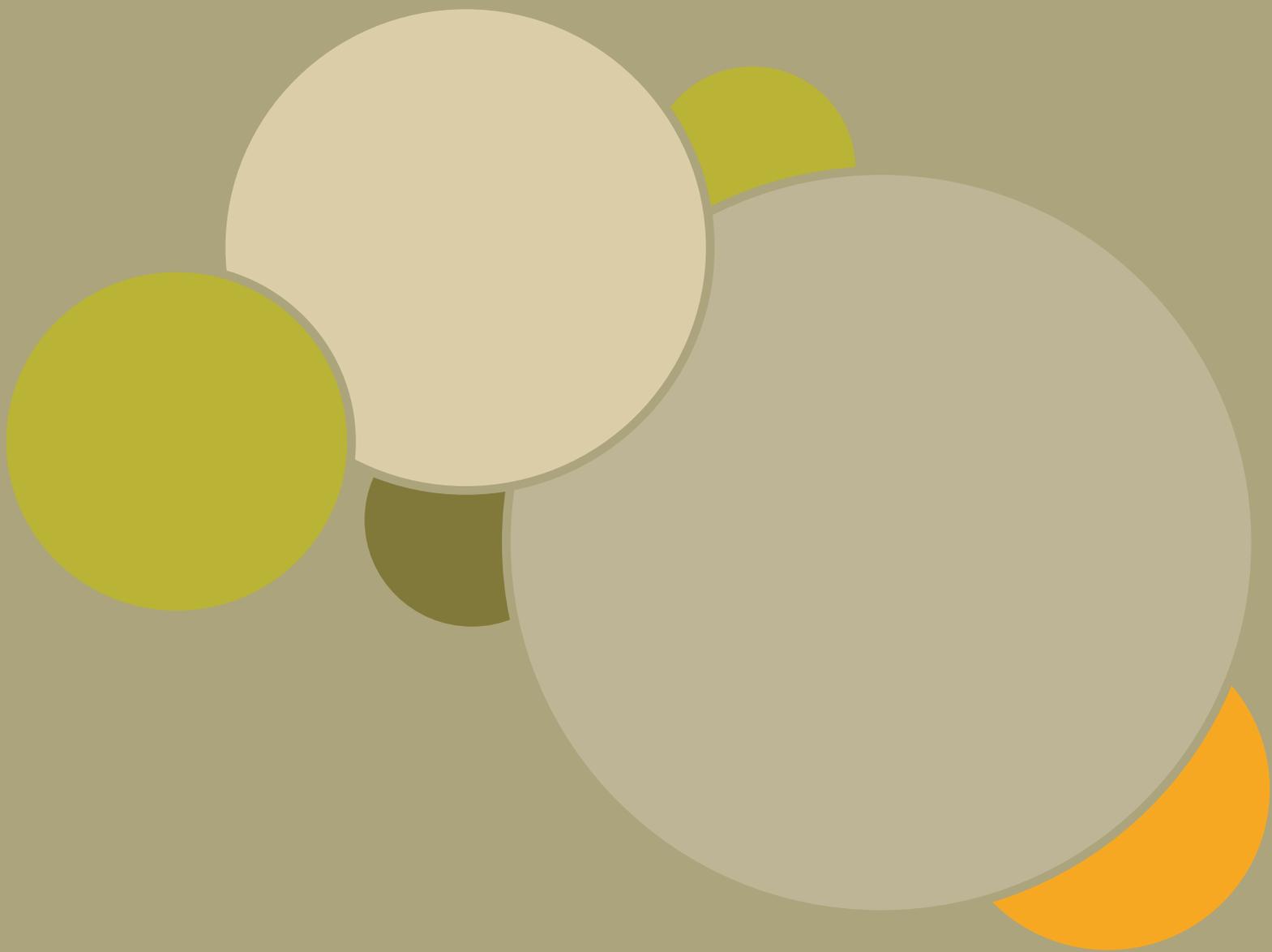
Manuela S. Nalugon – Senior Statistical Specialist, Philippine Statistics Authority

Hem Raj Regmi – Government Statistician, Nepal, and International Consultant for Statistical Capacity Building, FAO Regional Office for Asia and the Pacific

Manuel Enrique Ron – International Consultant on Food Balance Sheet Preparation, Bolivarian Republic of Venezuela

Norah de Falco coordinated the design and communication aspects.

The document was edited by Sarah Pasetto.



# Preface

These guidelines were developed under the Global Strategy to improve Agricultural and Rural Statistics (hereafter, Global Strategy), using the expertise of various persons within the Statistics Division (ESS) of the Food and Agriculture Organization of the United Nations (FAO). These guidelines on Food Balance Sheet (FBS) compilation transverse the Global Strategy's three foundational pillars – (1) produce a minimum set of core data; (2) better integrate agriculture into the national statistical systems; and (3) improve governance and statistical capacity building – as they seek to advise countries on identifying potential core data sources (alternatively, on imputing missing data), ensure that data from disparate sources are coherent by analysing them in a unified framework, and suggest mechanisms through which data quality processes and transparency can be improved.

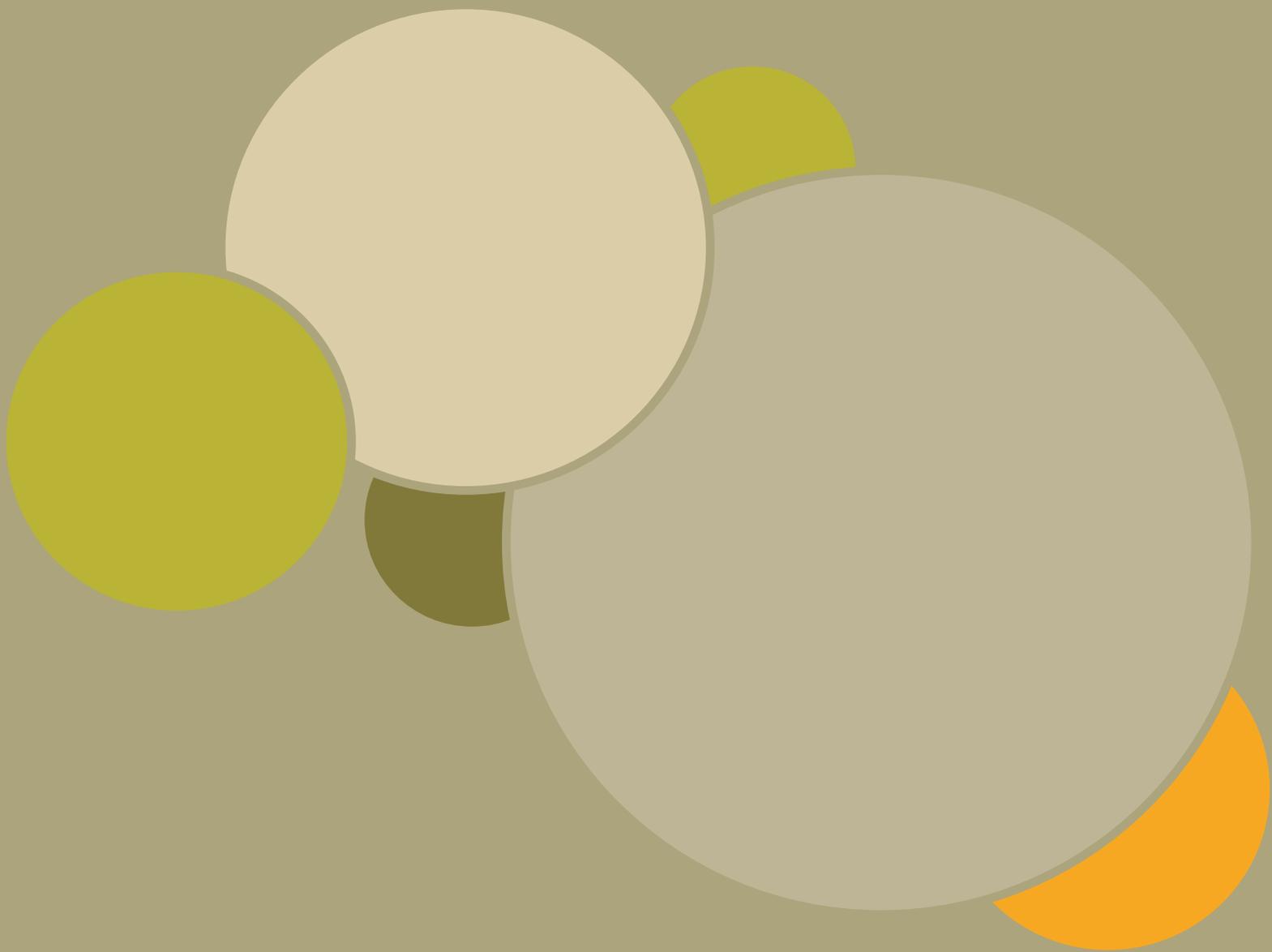
The methodologies and approaches described in these guidelines represent the latest innovations in both the imputation of missing data and the balancing of food commodity accounts. At the same time, these methods continue to be improved and refined, constantly striving towards the goal of producing consistent and replicable FBS for all countries, even in the face of limited source data. As further innovations are introduced, it is hoped that these guidelines will be updated to reflect the most up-to-date advice for country-level FBS compilers.

Even as these innovations in imputations and balancing are introduced in these guidelines, it is recognized that the new approaches may not be appropriate for or easily adopted by all countries. For this reason, these guidelines have also sought to propose alternative approaches that, although deviating from the proposed “gold standard”, may be more realistic, applicable and sustainable in the context of developing countries.

Because FBS cover so many products, industries and processes (not to mention the potential idiosyncrasies between countries), they can become extremely detailed and complicated. These guidelines strive to outline the process of FBS compilation by balancing comprehensiveness with manageability. As such, several topics relevant to FBS compilation are not covered here, or are only covered in basic detail. Most prominently, these guidelines cover only data sources and imputation procedures for crop and livestock products – no guidance on fishery or forestry data sources or imputation procedures is provided. However, recognizing that fishery and forestry products are important contributors to healthy diets, country-level FBS compilers should strive to include these commodities in their comprehensive FBS, if possible. These sectors may also be covered in future versions of these guidelines.

In addition, these guidelines provide suggestions only for the compilation of national-level FBS. Although there is a growing desire for the compilation of regional or subnational FBS accounts (with the goal of better targeting food security policies within countries that have wide regional disparities), these kinds of balance sheets require additional approaches, tools and data sources that are not covered in these guidelines. However, many of the other data source and imputation approaches suggested here could easily be adapted and applied to a regional context, should the user country desire to utilize this tool as a starting point.

Finally, it should be emphasized that while these guidelines cover the concepts and processes necessary to derive national-level FBS estimations, compilers often also seek the best tools available to help them implement these processes. At the time of writing, no such comprehensive tool is available to offer as accompaniment to these guidelines. This is partly due to a belief that compilers who have a solid understanding of the underlying processes and calculations are better able to think critically about all of the moving parts of the FBS, such that they can better identify potential sources of problems and, in turn, propose solutions. At the same time, it is envisioned that additional tools (including software, e-learning courses, and guidance on analysing FBS outputs) will be developed and made available in a comprehensive FBS toolkit in the future.



# Introduction

## 1.1 OVERVIEW

There is a global recognition that statistically sound, reliable data on food and agriculture is necessary to understand the current situation of agriculture and food supplies within any given country, track progress against established development goals, and inform future evidence-based policy decisions. Through the Global Strategy to improve Agricultural and Rural Statistics (also known as Global Strategy), efforts have been made to develop and promote the most cost-effective means by which countries can collect data on various aspects of food supply and utilization, with the end goal of providing a better foundation for evidence-based policy- and decision-making. Particularly in the context of the Sustainable Development Goals (SDGs), the value of reliable and globally comparable agricultural data to measure progress against certain targets is being realized.

While data on agricultural production, imports, exports or stocks can be valuable on their own, in isolation, none of these elements can accurately describe the complete food circumstances of a country. For example, if production of a given commodity falls, in the absence of other information, one could deduce that availability of that good would also fall. However, if imports are rising more than production is falling, an increase in availability could be expected. This simple example serves to underscore the notion that meaningful conclusions can be drawn regarding a country's food situation only when all of these elements are combined into a holistic framework. Food Balance Sheets (FBS) provide such a framework, combining all aspects of a nation's food supply and demand in a manner that allows for the validation of the underlying estimates while also providing the context necessary to enable a complete analysis of the individual elements.

The goal of these guidelines is to provide countries with the methodological framework and tools to compile high-quality FBS for crop and livestock products<sup>1</sup>. To this end, these guidelines are organized in the following manner:

- **Chapter 1** serves as an introduction to the topic, defining FBS and providing information on how they can be useful to countries.
- **Chapter 2** presents an overview of the methodological principles that underlie the construction of country-level FBS, including a discussion of the basic supply = utilization identity, an introduction to Supply Utilization Accounts (SUAs) and commodity trees, and a brief description of the recommended balancing mechanism.
- **Chapter 3** builds on the concepts introduced in chapter 2 by focusing on the data needed to compile FBS. The chapter includes a discussion of the necessity of organizing work within a Technical Working Group (TWG) that will define the scope of the task, before proceeding with information on how data should be assessed and considered before FBS compilation. This is followed by detailed discussions of suggested data sources and imputation approaches for each of the variables included in the FBS.
- **Chapter 4** brings together all of the components discussed in chapter 3 into the complete FBS framework, by walking readers through a step-by-step guide to FBS construction.
- **Chapter 5** concludes by offering some final considerations on data quality, advice on data dissemination, and guidance on FBS interpretation.

## 1.2 WHAT IS A FOOD BALANCE SHEET?

A Food Balance Sheet (FBS) can be defined as an aggregated and analytical data set that “presents a comprehensive picture of the pattern of a country’s food supply during a specified reference period.”<sup>2</sup> This is achieved within an accounting framework, wherein all potential sources of both supply and utilization of a given food product are specified. The quantities allocated to all sources of total supply – the amount of the food item produced, the amount of the food item that is imported, and the amount of the item that is either added to or taken from stocks – must be equal to the quantities allocated to all sources of utilization, which can include exports, losses along the supply chain<sup>3</sup>, livestock feed, seed use, tourist food, food processing, industrial uses, other uses, and food available for consumption by a country’s residents. This balance is compiled for every food item (estimated on a primary commodity equivalent basis) consumed within a country, and all of the primary commodity equivalent balances are then combined into a single overall FBS. An estimate of per capita supply for each food item – both in terms of quantity and, through the application of food conversion factors, in terms of caloric value, protein, and fat content – can then be derived by dividing by the country’s population. These per capita estimates of caloric value for individual food products are then summed to obtain the total daily per capita Dietary Energy Supply (DES) of a country.

Viewing the domestic food supply and demand situation within this framework allows countries to examine conditions in a holistic way, both aiding food supply analysis and facilitating food policy formulation. However, FBS have a wide range of applications, with some of the most common uses detailed in the following section.

---

1 Compilers may note that consumers also derive calories from fishery products. Although fishery products are not covered in these guidelines, the same general guidance would apply in constructing fishery product balance sheets – all elements of supply and utilization must be accounted for, and calories are derived from quantities of food. However, specific issues related to imputation and estimation of fish catch, loss and trade are not covered here.

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

3 Throughout these guidelines, the term “supply chain” is defined as in van der Vorst *et al.* (2007): “a sequence of processes and flows that aim to meet final customer requirements, and that take place within and between different stages along a continuum, from production to final consumption.” This term is often conceptually used interchangeably with “value chain”; however, in these guidelines, the reference is only to the supply chain

## 1.3 POTENTIAL USES OF FBS

An increasing number of countries compile comprehensive FBS, although the exercise can be time-consuming and require additional resources for agricultural statistics and analysis. The growing interest in compiling FBS is likely a result of both the value of the data generated through this process, and the myriad potential applications of that data. For countries currently considering whether FBS can be a positive addition to their statistical systems on food and agriculture, some possible applications are outlined here.

### **Measuring and analysing overall food supply**

FBS are perhaps most commonly applied as a mechanism to estimate a country's overall DES and macronutrient availability (fats and proteins, in addition to kilocalories, or kcal). Because FBS track overall food *availability* and not actual consumption, the DES cannot be used as an estimate of how much nourishment the average resident is consuming, but rather as an indicator of whether sufficient food is available nationally, particularly for developing countries, where undernourishment is more likely to be a problem. At the same time, in some developed countries, food availability has been used as a proxy for actual food consumption<sup>4</sup>.

In addition to providing a measure of food availability, FBS are also useful for analysing the overall content of a country's diet, including determining the availability of a certain variety of food. As an example, estimates derived from FBS were recently used to analyse declining fruit and vegetable consumption in the United States of America (Lin and Mentzer Morrison, 2016)<sup>5</sup>. In the context of developing countries, analysing shortfalls in the availability of certain foods in the FBS could be one approach to better understand the nature of malnutrition in a given country.

The framework is also flexible enough to be used in analysing the availability of additional vitamins and minerals, contingent only upon the utilization of sufficiently detailed nutrient conversion tables.

However, it should be noted that nutrient content (of both macro- and micronutrients) may vary widely within commodity categories, depending upon which variety or type of commodity dominates in domestic food supplies. For this reason, to ensure the most accurate estimates possible, compilers are encouraged to utilize their most recent country-specific nutrient conversion tables (additional information on nutrient conversion tables, including links and references to some commonly used international tables, is provided in section 3.5.12).

---

4 Documentation for the *Food Availability Data System* of the United States Department of Agriculture (USDA) specifically notes that “[t]he food availability data series is a popular proxy for actual food consumption and is especially useful for those interested in time series data.” For more information, see USDA, Economic Research Service, 2017, *Food Availability Documentation*, available at <https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/food-availability-documentation/>. Accessed on 19 January 2017.

5 This study used loss-adjusted supplies of fruits and vegetables, derived from food availability data estimated through a supply and utilization framework.

### **Food supply assessment through the calculation of indicators**

One of the primary applications of FBS is to calculate derived indicators using FBS data. These indicators can be used to analyse a wide range of concepts, including hunger, malnutrition, import dependence and food self-sufficiency. One of the most widely known of these indicators is the Prevalence of Undernourishment (PoU), which measures the “probability that a randomly selected individual from a population is found to be consuming less than her/his requirement for an active and healthy life<sup>6</sup>”. The PoU is calculated by estimating a distribution of food consumption within a country during a given reference period, where food availability (as measured by the DES) is a proxy for the mean level of consumption in that distribution.

Other well-known indicators derived from FBS data include the Self-Sufficiency Ratio (SSR), which compares the magnitude of a country’s agricultural production to its domestic utilization, and the Import Dependency Ratio (IDR), which compares the magnitude of a country’s imports to its domestic utilization<sup>7</sup>.

Although these are only a few examples, because of the wealth of data contained within the FBS framework, it can be used to derive countless indices and indicators depending upon the variables of interest to the end user. As one example, SDG 12.3 tasks countries with reducing food loss. Because loss appears as a variable in the FBS, those quantities could be useful in calculating an indicator on food loss.

### **Benchmarking and market analysis**

To the extent that the methodology for compiling FBS and deriving DES estimates is similar across countries, these estimates can be used to compare food availability from one country to another. This comparison is possible both on an aggregate level and on a product-specific level. For example, users can choose to compare the per capita food availability of almonds between countries. These types of comparisons can have various applications. First, analysing comparative diets can be a useful exercise in the area of nutrition policy. For example, assume that Country A determines that domestic per capita availability of fruit and vegetables is too low. Policy-makers in Country A can then consult contemporaneous FBS from other countries to identify other countries with a relatively higher food availability of fruits and vegetables, and then conduct research on the policy frameworks established within those other countries to determine whether there are any relevant elements that could be adopted in Country A to lead to higher fruit and vegetable availability.

Another application of FBS in the context of comparative food availability is the potential for market research. For instance, assume that Country B is a small exporter of pistachios. Firms or export promotion agencies within Country B could then use FBS to identify countries with either a comparatively high food availability of pistachios as potential export markets with existing consumer demand, or identify countries with a comparatively low food availability of pistachios as possibilities for future market development.

### **Comparing food availability across time**

Just as an FBS compiled for one year will provide a snapshot of a country’s food supply, compiling FBS over a period of years will enable users to track changes in the food supply over time, including estimated total caloric availability, growth of consumption in new products, and general changes in dietary composition. Such time series data sets are extremely useful to both academics and policy-makers. For example, rising obesity rates within a country could be tracked using rising overall per capita food availability estimates drawn from the FBS. Governments can then use this information to support certain policy measures to stem the rise in obesity.

---

6 For more information on the methodology underlying PoU calculation, see FAO, 2017, *Food security methodology*, available at <http://www.fao.org/economic/ess/ess-fs/fs-methods/fs-methods1/en/>. Accessed on 19 January 2017.

7 For more details on the methodology underlying the calculation of the SSR and the IDR, see, for example, FAO, 2012, *FAO Statistical Yearbook 2012*, available at <http://www.fao.org/docrep/015/i2490e/i2490e00.htm>. Accessed on 27 April 2017.

FBS time series data on per capita food availability can also be valuable to forecast future consumption trends. This application could be useful for both policy-makers and firms seeking new market opportunities.

### **Improving national statistical integration**

By their very nature, FBS provide a framework for reconciling data, as total supply must equal total utilization. This can be a challenging exercise, because for most countries and most products, the necessary input data will come from a variety of different sources and agencies within the government, and potentially even from semi-official actors providing information on a single commodity. While reconciling these data may be time-consuming, the process provides a unique opportunity to both harmonize data collection efforts across agencies (ensuring that bodies are measuring or estimating quantities in apples-to-apples comparisons), and validate estimates by placing them within the context of the commodity's overall supply and demand picture. By bringing together all of the relevant stakeholders at this stage, problems or inconsistencies in data collection and estimation can be identified early, and a country's overall agricultural statistical program can be improved.

Two examples of this improved national statistical integration can be highlighted. The first is the use of so-called Interagency Commodity Estimates Committees (ICECs) at the United States Department of Agriculture (USDA) in their formulation of supply and demand balances for the monthly World Agricultural Supply and Demand Estimates (WASDE) report. The ICECs are composed of representatives of various agencies within the USDA, and the data used by the ICECs to formulate the commodity balances come from a variety of additional sources (including the National Agricultural Statistics Service, or NASS, and the Census Bureau)<sup>8</sup>.

Recently, a similar system has been established in Mexico: the Dashboard for Strategic Agrifood Products (*Tablero de Control de Productos Estratégicos Agroalimentarios*). Under this system, a working group meets periodically to compile supply/demand balances for strategic agricultural products. This working group composes various commodity balances using information from the Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food (SAGARPA), the Secretariat of the Economy (SE), and the Customs General Administration (AGA) housed within the Secretariat of Finance and Public Credit (SHCP)<sup>9</sup>.

While both of these examples refer to statistical integration in the context of deriving commodity balances, at least in the USDA example, the data on "food" derived through the commodity balance estimation process is sometimes used at a subsequent stage to estimate per capita availability of the relevant food commodities<sup>10</sup>.

It should be noted that the Global Strategy already recommends that countries set up a permanent steering committee on agricultural statistics and a technical working committee on agricultural statistics in the context of their individualized Strategic Planning for Agriculture and Rural Statistics (SPARS) plans, specifically for the purpose of bringing together all relevant actors in the agricultural statistics space (Global Strategy, 2014). In particular, the technical working committee could be utilized as a vehicle for improved national statistical integration in this context, as its role is to bring together both strategic stakeholders and representatives of subsectors who could then participate in a collective FBS validation exercise.

<sup>8</sup> For more information on the methodology of compiling the commodity balances used in the WASDE using various sources, see USDA, Office of the Chief Economist (OCE), *How the WASDE is Prepared*, available at: <http://www.usda.gov/oce/commodity/wasde/prepared.htm>. Accessed on 19 January 2017.

<sup>9</sup> While all the mentioned agencies contribute data, the Agrifood and Fisheries Information Service (SIAP) of Mexico states that SAGARPA and the SE are the primary governmental bodies compiling the balances (see SIAP, *Cosechando Números del Campo*, available at: <http://www.numerosdelcampo.sagarpa.gob.mx/publicnew/>; accessed on 19 January 2017). However, the individual commodity balances also cite AGA as an additional source of information (for example, see the sorghum balance, available at: <http://www.numerosdelcampo.sagarpa.gob.mx/publicnew/productosAgricolas/cargarPagina/5>, accessed on 19 January 2017).

<sup>10</sup> In this respect, the best examples are the balances for food grains. For more information on this point, see USDA, 2017, Economic Research Service, *Food Availability Documentation* (entry on grains), available at: <https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/food-availability-documentation/#grains>. Accessed 19 January 2017.

### **Inputs in national accounts**

Fundamentally, FBS are an accounting framework specific to food and agricultural products. In fact, the FBS framework is similar to the System of National Accounts (SNA) supply-use framework<sup>11</sup>. As such, FBS are naturally complementary to the estimation of national accounts. National accounts typically include estimations of household and collective consumption, trade in goods and services, and output and value-added by industry. Improved measurement or estimation of agricultural production and utilization – which are validated in conjunction with import and export data, as well as other components of supply and demand – also lead to improved national account estimates. In turn, because national accounts data are used for all sorts of economic and policy analysis, improved data in this area could lead to more targeted and more effective policies.

It is worth noting that just as FBS can be an input into national accounts, national accounts can also provide helpful data for the compilation of FBS.

### **Providing input for economic models**

FBS are accounting frameworks. Therefore, FBS data can be used by several models that are structured using the supply-use format of national accounts data described above. Most partial equilibrium models for agriculture utilize commodity balances in their data structures. The AGLINK-COSIMO model of the Organisation for Economic Co-operation and Development (OECD) and the IMPACT model of the International Food Policy Research Institute (IFPRI) are two of the main models used at the international level; however, country-specific models also exist. One example is the partial equilibrium Agricultural Member States Modelling (AGMEMOD) project of the European Union (EU), which has been expanded to include Ukraine and Turkey to inform analyses of how developments within either of these countries may affect global markets<sup>12</sup>. These inputs from FBS can then serve as base-year data sources, either enabling simulation analyses or facilitating projections. While FBS data may be the primary data source for many of these models, the final input data set is unlikely to be identical to the data in the FBS, as many of these types of models impose a zero net trade rule as their solving algorithms. Nonetheless, because these models are widely used for agricultural projections and policy analysis (by academics and governments alike), improved FBS estimates can improve the overall quality of a country's agricultural sector modelling capabilities.

---

11 For more information on national accounts, see United Nations Statistics Division (UNSD), 2017, *The System of National Accounts (SNA)*, available at: <http://unstats.un.org/unsd/nationalaccount/sna.asp>. Accessed 19 January 2017.

12 For more information, see van Leeuwen *et al.* (2012) and van Leeuwen *et al.* (2011).

## 1.4 CAUTION IN INTERPRETING FBS ESTIMATES

As noted above, FBS provide a framework for understanding general trends in changing food consumption patterns by measuring average food availability. However, while FBS can be useful tools for assessing the food situation of a country, users must also be aware of their limitations if they are to accurately interpret FBS estimates.

### **Food availability, not food consumption**

The most important caveat to bear in mind when analysing FBS is probably that concerning the meaning of the estimates of per capita dietary energy supply. It should be emphasized that these are estimates of the *available* dietary energy supplies, and can thus be described as the *apparent food consumption*<sup>13</sup>. In other words, FBS estimates represent food that is intended for human consumption available for purchase by consumers at the point of sale. This concept is distinct from *effective food consumption*, which is the actual quantity of food consumed. In practice, this means that FBS food use estimates are not net of food waste at the retail or household levels. Particularly in developed countries, food waste at the household level can be substantial<sup>14</sup>, such that the DES is likely to overestimate the amount of food actually consumed. This is less of a problem in developing countries, where food waste at the household level is limited compared to food loss at other stages of the supply chain.

### **Lack of accounting for distributional differences**

Estimates derived through an FBS framework represent only an overall average amount of either food or nutrient availability within a country. While useful for certain types of analysis, these average estimates may mask other underlying consumption trends among disparate groups of people within the same country. This is particularly the case with the most food-insecure segment of the population. In many developing countries, for example, consumption patterns may more accurately be described with bimodal distributions: one portion of the population will consume a certain food at an average value of a certain level, and another portion of the population will consume that same food at a higher average level. The resulting overall countrywide average would not reflect the consumption levels of either group. Particularly when analysing FBS for countries with very unequal income distributions, it is important to keep these distributional limitations of the FBS in mind.

### **Compatibility with household surveys**

While household surveys are another useful tool in analysing the consumption habits of individuals and households, users are advised that food availability estimates derived from FBS can be expected to differ from those indicated in household surveys. This discrepancy has several sources. The first results from the fact that FBS estimate food availability, while household surveys are more likely to measure actual food consumption levels. A second source of discrepancy is that FBS take into account food availability at all locations within a country (including schools, hospitals, military, prisons, restaurants and food service), and not only household food availability<sup>15</sup>. A third source of discrepancy may be survey coverage and representativeness: household surveys may only reflect the consumption patterns for a given region or income group, and thus may not be statistically representative of the society's consumption patterns as a whole. A fourth source of potential discrepancy may be the number of commodities and the aggregation level considered in the household survey. Household surveys may record the consumption of specific derived goods, whereas FBS may have little information on those derived goods and, instead, report availability for a less processed version (for example, a household survey may report the consumption

---

13 The term “apparent food consumption” is used here to avoid confusion with the term “apparent consumption”, which is commonly used in the literature as a measure to indicate the supplies available for all uses, and calculated as “production + imports - exports (and sometimes adjusted for stock changes)”. See, for example, UNSD (2016).

14 Buzby *et al.* (2014) and FUSIONS (2016) detail the scale of developed country food waste. Buzby *et al.* estimate that consumer-level losses were equivalent to 21 percent of total food supplies in the United States of America in 2010, while FUSIONS estimate that households wasted approximately 11 percent of the food produced in the EU in 2011 (an estimated 92 kg per capita were wasted at the household level in the EU, out of 865 kg of food produced per person).

15 It bears noting that the preceding two factors increase the likelihood that food supplies estimated from an FBS framework will be higher than those obtained from household surveys.

of tortillas, while the FBS may consider the flour used to make those tortillas). The last source of discrepancy may be timing. FBS account for food availability over a full year, while household surveys are much more likely to capture a shorter period of time. Especially in countries experiencing a lean season, household survey results may vary widely, depending upon when the survey takes place.

Nevertheless, the general trends and overall dietary patterns represented in the two instruments should typically be similar. In addition, household surveys can theoretically serve as either a key source of data for food estimates within FBS or as a validation tool for FBS accuracy, provided that the underlying differences in the two sources of data are well understood. At least one scholarly attempt has been made to understand and reconcile these two data sources, and countries may find that exercise useful when comparing their own data<sup>16</sup>.

Country-level FBS compilers should note that FBS and household surveys should be treated as complementary tools and not as substitutes of one another; indeed, household surveys are costly and lengthy endeavours, such that most countries can only afford to run them once every four or five years. This leaves researchers and policy-makers without any data in the interim period to inform analysis. In comparison, FBS can be compiled every year at a much lower cost, generating an annual data set that can be used by various actors for research and analysis.

### **Commodity balances and FBS**

As described above, FBS can be considered as one type of accounting framework from which the overall supply and demand situation of a particular commodity can be analysed. For this reason, FBS can be said to belong to the general family of commodity balances (indeed, the process of deriving FBS begins with the compilation of a commodity balance). However, there are some differences between the two instruments that should be noted. The most obvious difference is that FBS are composed for food-related commodities only. For example, although a country may wish to publish a commodity balance for natural rubber, this balance would not be considered an FBS because there is no food element to rubber. The second difference regards the assignment of nutrient value to the “food” element in a commodity balance. Although any country may publish a commodity balance for a product that is consumed as food, this balance is not considered an FBS until the quantity estimates of food have been reported in their nutritive equivalent. The final difference is more technical and concerns the mechanics of composing FBS. FBS should contain aggregated estimates of both a primary commodity and all of its derived products, expressed at the primary commodity equivalent level. While many countries produce commodity balances for primary products, they often do not account for goods derived from those primary products, and may thus underestimate total availability and consumption for groups of certain commodities. Countries may wish to publish separate accounts for primary and derived commodities rather than aggregating them; however, aggregated primary equivalent accounts are preferred for purposes of comparability across countries.

---

<sup>16</sup> For further information, see Grünberger (2014).

## 1.5 FUNDAMENTAL PRINCIPLES OF FBS CONSTRUCTION

While the general overall approach of FBS construction is outlined in these guidelines, the exact process to be followed will differ from country to country, depending upon data availability, food supply chain structure, expertise, and other potential resource constraints. Nevertheless, compilers should adhere to certain basic principles to ensure that their country-level FBS are reproducible, coherent and transparent. These principles, discussed in further detail below, are:

- Measurement first
- Document data and processes
- Feedback and collaboration

Country-level compilers are encouraged to consider these fundamentals when drafting their country's overall FBS construction plan, and to always recall them when updating or improving their processes.

### **Measurement first**

Without question, the key to producing FBS that provide consistent and reliable estimates that can effectively aid analysis and policy-making is to ensure the accurate measurement and availability of key input data on all supply and use variables. While countries can theoretically produce FBS with a limited number of data inputs, estimates of food availability derived from sparsely populated input data frameworks may be accompanied with large margins of error, which limits their usefulness as an analytical tool. At the same time, most imputation modules for missing data rely on some measurement of data having been made in the past, such that it may not be possible to impute missing data without the use of some underlying source data. For these reasons, it is recommended that countries invest resources in improving the measurement of input data before attempting to compile country-level FBS. There is no substitute for input data measurement, and reliable FBS depend upon reliable input data.

### **Document data and processes**

These guidelines outline the general process of FBS compilation. However, data sources and approaches will vary from country to country, and may even change within a country over time as new data sources become available and new imputation approaches are developed. Staff turnover is also a common occurrence in the agencies charged with developing FBS. For these reasons, it is essential that compilers document the data sources, applied methodologies, and solutions to identified data inconsistencies on a real-time basis. This should include the drafting of accompanying metadata (data that describes or gives other background information about the data set). Such documentation is essential both for the sake of continuity (to ensure that new compilers are able to produce estimates in a manner consistent with the previous data series), and to ensure that future users can understand the rationale behind the development of certain estimates or the choice of one methodology over another.

### **Feedback and collaboration**

FBS should not be compiled in a vacuum. Instead, the process should be collaborative, involving all relevant actors within the supply chain, and, to the extent possible, users of the resulting data sets. Ideally, all relevant actors should collaborate within a permanent TWG; at the very least, a TWG should validate the final output. Validation by multiple actors is important because experts in certain supply chains are likely to possess additional knowledge on product specifics that will contribute to the accuracy of the balance sheets. This practice can also lead to the improvement of input data: if the relevant actors are aware that the accuracy of balance sheet estimates is limited by the quality of the input data, they may be more likely to participate in efforts to improve such data.

This feedback process also serves to build confidence in the estimates produced with the FBS framework: if the sources of data are explained and understood, users are more comfortable with using those estimates for their own data needs.

## 1.6 SUMMARY

This chapter introduces the overall concept of FBS to the lay user or novice compiler. The overall framework is introduced, some potential uses of FBS are noted, and some notes of caution on FBS interpretation are described. The chapter closes by highlighting the fundamental principles of FBS construction: measurement first, document data and processes, and feedback and collaboration. Chapter 2 focuses on the preparatory steps to the construction of country-level FBS.

# 2

## Methodological principles for the construction of country-level FBS

### 2.1 OVERVIEW

The process of constructing FBS at the country level should not be undertaken without a thorough understanding, by all compilers, of the mechanics of the FBS. This chapter introduces the relevant concepts involved in compiling an FBS: the basic underlying identity, the supply and use variables, the additional necessary variables, the difference between SUAs and FBS, how commodity trees link SUAs back to the primary commodity equivalent-level FBS, and the balancing mechanism.

### 2.2 THE BASIC IDENTITY AND APPROACH

FBS are built on the basic premise that within a given country in a given year, the sum of all aspects of the supply of a given food product must be equal to the sum of utilizations of that product. This concept is commonly expressed with one of two different basic identities: total domestic supply is equal to total domestic utilization, or total supply is equal to total utilization (equations (2.1) and (2.2) respectively).

**Domestic supply = domestic utilization:**

$\begin{aligned} & \textit{Opening Stocks} + \textit{Production} + \textit{Imports} - \textit{Exports} \\ & = \textit{Food} + \textit{Feed} + \textit{Seed} + \textit{Tourist Food} + \textit{Industrial Use} \\ & + \textit{Loss} + \textit{Residual Use} + \textit{Closing Stocks} \end{aligned}$	(2-1)
---	-------

**Total supply = total utilization:**

$ \begin{aligned} & \textit{Opening Stocks} + \textit{Production} + \textit{Imports} \\ & = \textit{Exports} + \textit{Food} + \textit{Feed} + \textit{Seed} + \textit{Tourist Food} \\ & + \textit{Industrial Use} + \textit{Loss} + \textit{Residual Use} + \textit{Closing Stocks} \end{aligned} $	(2-2)
---	-------

More precise definitions of these variables follow below (section 2.2.1).

The only difference between the two equations is the placement of exports. In equation (2-1), supply is defined in terms of net trade (imports minus exports). In equation (2-2), imports are registered as a supply variable, while exports are recorded as a utilization variable<sup>1</sup>. Countries are free to choose the conceptualization that is most suitable to their own FBS.

At the same time (as will be described in subsequent sections), few countries collect data on stock levels for the majority of products. For this reason, the supply = utilization identity is often expressed using an estimate of the change in stock levels during the reference period (that is, either a stock buildup or a stock drawdown), rather than by including estimates of absolute opening and closing stock levels. With regard to equations (2-1) and (2-2), this can be expressed as:

**Domestic supply = domestic utilization:**

$ \begin{aligned} & \textit{Production} + \textit{Imports} - \textit{Exports} - \Delta\textit{Stocks} \\ & = \textit{Food} + \textit{Feed} + \textit{Seed} + \textit{Tourist Food} + \textit{Industrial Use} \\ & + \textit{Loss} + \textit{Residual Use} \end{aligned} $	(2-3)
---	-------

**Total supply = total utilization:**

$ \begin{aligned} & \textit{Production} + \textit{Imports} - \Delta\textit{Stocks} \\ & = \textit{Exports} + \textit{Food} + \textit{Feed} + \textit{Seed} + \textit{Tourist Food} \\ & + \textit{Industrial Use} + \textit{Loss} + \textit{Residual Use} \end{aligned} $	(2-4)
---	-------

where  $\Delta\textit{Stocks}$  is defined as  $\Delta\textit{Stocks} = \textit{Closing Stocks} - \textit{Opening Stocks}$ .

The basic identity can also be specified with an additional utilization variable, Food Processing, as shown below in equation (2-5):

Alternative specification of total supply = total utilization, with food processing included:

$ \begin{aligned} & \textit{Production} + \textit{Imports} - \Delta\textit{Stocks} \\ & = \textit{Exports} + \textit{Food} + \textit{Food Processing} + \textit{Feed} + \textit{Seed} \\ & + \textit{Tourist Food} + \textit{Industrial Use} + \textit{Loss} + \textit{Residual Use} \end{aligned} $	(2-5)
--	-------

<sup>1</sup> It should be noted that countries in which stock levels are measured or estimated may include “beginning stocks” as a supply variable and “ending stocks” as a utilization variable, thus effectively eliminating  $\Delta\textit{Stocks}$  from the equation. However, for the purposes of these guidelines, the specifications provided herein are preferred because the imputation of  $\Delta\textit{Stocks}$  is a much simpler process with stronger theoretical underpinnings. Further details on the imputation of stock level changes are available in section 3.5.3.33.5.3.3.

The reason why food processing is not always included in expressions of the basic underlying identity is that this variable is typically dropped in the final stages of FBS compilation (standardization and aggregation) to avoid double-counting (see chapter 4 of these guidelines for a more in-depth explanation). However, food processing should be included as a utilization variable in the specification of the preliminary individual commodity balances (covered below, section 2.3); therefore, it may be helpful to conceptualize the basic identity in this manner.

Ideally, countries should measure each of these variables for each product, with the measurement including both an expected value of the estimate and a quantification of the confidence of said estimate, deemed here as a “tolerance interval” – defined as an interval within which the actual point estimate is likely to lie (for measured values, this tolerance interval is equal to the measurement error). The rationale behind the provision of both the expected value and the tolerance interval will be explored further in section 2.4. For current purposes, it suffices to state that both components are key to balancing the overall identity.

### 2.2.1 Supply and use variables

The variables elaborated in the basic supply = utilization identity should be mostly intuitive concepts for FBS compilers. However, a more exact definition of these concepts is warranted to ensure that the scope of what falls within the calculations is clear. Countries should try to adhere to these definitions where possible, to ensure that calculations of dietary energy supply derived from country-level FBS are a faithful approximation of the real food supply situation. For example, reporting only commercial production would lead to an underestimation of the supply of some products in countries where the product is commonly grown in household gardens or, more generally, for own consumption, which in turn would cause an underestimation of total per capita availability. All definitions used here are adapted from the FAOSTAT Food Balance Sheets Glossary<sup>2</sup> and Schmidhuber (2016).

#### **Production**

Data in the FBS for production should include all production quantities of a given commodity within the country in question, including both commercial and non-commercial production (such as that from home gardens or subsistence agriculture). The production of primary products should be reported at the farmgate level, so that it does not include harvest loss. In theory, it should include any post-harvest on-farm loss occurring during the different farm operations, such as threshing, cleaning/winnowing or storage. The data for meat production should include both commercial and farm slaughter, and production should be expressed in terms of carcass weight. Production of any derived or processed commodities refers to the total output of the product after transformation. This transformation may occur either within the household or at a commercial establishment. It should be noted that the standard unit for the reporting of agricultural production at the international level is the metric tonne; however, many countries also use local units. Compilers should also consider that for production straddling calendar years, production should be assigned to the year in which most of the crop will be consumed (more details on this concept are given in section 3.4.1.3).

#### **Imports and exports**

Imports and exports are the two primary types of foreign trade, which can be defined as the exchange of goods (and services) across international borders. More precisely, imports are transboundary flows of goods destined for a given final destination country that add to the total supply of goods available in that country, while exports are transboundary flows of goods from a given country of origin that take away from the total availability of goods in that country. Goods that enter and exit a given country without having undergone any type of transformation are categorized separately as re-exports. In the context of the FBS, re-exports should be added to exports to fully

<sup>2</sup> The glossary may be accessed from the metadata for the FAO FBS published on FAOSTAT: FAO, 2017, *FAOSTAT: Food Balance Sheets*, available at <http://www.fao.org/faostat/en/#data/FBS>. Accessed on 19 January 2017.

account for all outgoing trade flows. This is particularly important for those countries that act as transportation hubs (countries that import goods from one destination and then pack them onto a different vessel to ship elsewhere). If re-exports are not counted in these instances, the FBS would describe a distorted picture of the food situation: one in which all goods imported remain part of domestic supplies and are subsequently consumed.

It should be noted that, to the extent possible, imports and exports estimates should endeavour to cover both official and unofficial trade flows, including food aid shipments. For certain countries and commodities, unrecorded trade flows may be substantial, thus potentially having a large knock-on effect on food availability estimates.

### **Stocks**

*Stocks* are defined as the aggregate total of product allocated to storage for use at some future point in time (regardless of their intended future utilization). Stocks can be held by a variety of actors (governments, manufacturers, importers, exporters, wholesale merchants and farmers) at any level of the supply chain, from production up to – but excluding – retail<sup>3</sup>. As noted above, stocks can be accounted for in one of two ways in the balance sheet setting. First, stock levels at both the beginning and the end of the period can be noted (on the left- and right-hand side of the equation, respectively). Alternatively, the FBS framework can be elaborated by estimating the change in stocks from one time period to the next as a component of supply. If closing stocks are lower than opening stocks, it is implied that stock withdrawals were made during the period, thus increasing supplies.

### **Food availability**

*Food availability* is defined as the quantity of any substance, whether raw, processed or semi-processed (including drinks), that is available for human consumption during a given reference period. As mentioned in the foregoing discussion on what quantities reported in the FBS represent, the quantities defined by this variable represent the amount of food that is available for consumption at the retail level. For this reason, any waste (and/or loss) occurring at the retail or consumer levels is included in this quantity, because that food was technically available for human consumption. It is important to note that because the quantities reported under this variable represent those available as food, they will typically be higher than the quantities reported through household-level consumption surveys. In addition, the quantities reported here represent food available for consumption not only in households, but also in restaurants and institutions (hospitals, schools, military bases, prisons, etc.). Finally, as these quantities should represent how the product is sold, they are expressed in gross weight, and as such may include inedible parts. Therefore, when these quantities are converted into nutrients, either some conversion factor should be applied that calculate edible quantities before conversion into nutrients occurs; otherwise, the nutrient conversion factors should take into account the fact that the quantities are in gross weight and not net of inedible parts.

### **Food processing**

*Food processing* refers to the quantities of a food product that are directed towards a manufacturing process and are then transformed into a different edible commodity with a separate entry in the FBS. These separate commodities may be structured within the same commodity tree or food group (for example, tomatoes may be processed into tomato paste), or completely separately from it (for instance, barley is processed into beer, which is typically aggregated into an alcoholic beverages category and not into the barley balance). For derived commodities in the same food group, the food processing variable should disappear in the final stages of FBS compilation to avoid double-counting. For quantities that were used in the production of derived commodities in different groups, the food processing variable should remain in the final account.

It should be noted that quantities devoted to the manufacture of inedible products (such as soap or biofuels) should be accounted for under industrial use and not *food processing*.

---

<sup>3</sup> Food availability is defined at the retail level; therefore, there is no need to account for retail-level stocks.

**Feed**

*Feed* is defined as all quantities of commodities, both domestically produced and imported, that are available for feeding to livestock and poultry. Many commodities that are used as feed are byproducts of industrial processes, such as oilcake, dregs or distiller's dried grains with solubles (DDGS). While they are included in initial calculations, most will not be aggregated up to the primary commodity level to avoid double counting.

**Seed**

*Seed* is defined as any quantity of a commodity set aside for reproductive purposes in the following year. This can include seed for sowing, plants for transplanting, eggs for hatching, and fish used as bait. This quantity should also take into account double or successive sowing.

**Tourist food**

*Tourist food* refers to food that is available for consumption by non-resident visitors to a given country during the course of their stay. While the term "tourist" is used here, it is emphasized that this variable covers food availability for all non-residents, including tourists, business travelers and non-resident migrants, where the latter are not counted as part of the country's population. This variable is expressed in the FBS in net terms (as food available for consumption by incoming visitors minus the food that would have been consumed by residents who have travelled to other countries).

Countries with negligible numbers of visitors may choose to not estimate tourist food as a separate FBS component, preferring to capture it instead in other residual uses. However, for certain countries (particularly Small Island Developing States), accounting for tourist food is essential to accurately estimate local consumption patterns.

**Industrial use**

*Industrial use* is defined as any quantity of a given product used in a non-food transformation or manufacturing process, including for biofuels, cosmetics, detergents, or paints.

## Loss

In the context of FBS, *loss* refers to the quantities of a product that leave the supply chain and are not diverted to other uses. Loss results from an involuntary activity and can occur at any node of the supply chain after the harvest and up to (but excluding) the retail/consumption stage. This variable may also be referred to as *post-harvest or post-slaughter* loss. The category does not include any quantities of food wasted or lost at the consumer or retail level. Additionally, for consistency with other variables in the FBS, the quantities reported here should include both edible and inedible parts. It is also important to note that any volume of product lost in the transformation of primary products into processed products is accounted for through extraction rates and conversion factors, and for this reason is excluded from loss. However, loss that occurs in all other utilizations (particularly during storage and transportation) is included.

## Residual and other uses

*Residual and other uses* can, in most cases, be defined as the combined imbalance and accumulated error in the supply equals utilization equation. As such, this category is computed *ex post* as a balancing item and is not independently estimated. If all of the other utilizations within the equation are accounted for and there is no measurement error, then the residual would be calculated as zero. The decision on whether to include a Residual and other uses category is up to the compiling country; however, it is noted that the use of such a category is a simple way to acknowledge and account for small measurement errors.

Countries may also choose to utilize this category to account for quantities that are designated for any uses other than those described above. As such, this variable is more appropriately defined by what it is not, rather than what it is. To the extent possible, countries should account for all possible uses of a given product when elaborating an FBS; however, countries may wish to utilize this category to capture small amounts of product use that are otherwise unaccounted for, such as tourist food for countries with negligible numbers of tourists. Countries with large refugee populations may also wish to utilize this category to capture food availability for refugees (such countries may also wish to add a designated category for “food available for refugees”).

## 2.2.2 Additional variables

While the basic supply and utilization variables outlined above cover all of the aspects of the basic identity, composing the complete FBS (including estimates of per capita nutrient availability) requires several additional variables. A practical example more clearly demonstrates the precise manner in which these variables fit into the balance sheet (see chapter 4). However, they are defined here for convenience.

## Population

*Population* is defined following the definition given by the UN Population Division (UNPD): the “de facto population in a country, area or region as of 1 July of the year indicated”<sup>4</sup>. The adjectival *de facto* is important, because it indicates that citizens *as well as all* residents should be counted in the population, thus potentially including refugees or resident migrant workers. In addition, the persons who are not counted as part of the “population” should be conceptualized as “visitors”, so that their food availability can be appropriately captured under tourist food. Estimates of the population are necessary to convert aggregate national nutrient supplies into per capita nutrient supplies.

## Nutrient estimates

*Nutrients* are substances that the body needs to function properly. One of the primary motivations for compiling an FBS is to derive estimates of the amount of calories, fat, and protein available for consumption by a country’s

4 UNPD, *Glossary of Demographic Terms*, available at <https://esa.un.org/unpd/wpp/General/GlossaryDemographicTerms.aspx>. Accessed 19 January 2017.

population. These estimates are derived from the final food estimates in the balance sheet for each product by applying certain conversion factors to those quantities. Currently, the following nutrient-related variables are commonly derived from food estimates using nutrient conversion tables<sup>5</sup>:

- Food: total calorie equivalent
- Calories per capita per day
- Food: total protein equivalent
- Proteins per capita per day
- Food: total fat equivalent
- Fats per capita per day

However, additional nutrient estimates can be calculated if country-level FBS compilers wish to analyse the availability thereof, provided that sufficiently detailed nutrient conversion tables are available. It should also be noted that the food estimates in the balance sheet are expressed in gross weight, such that they include edible parts. If country-specific nutrient conversion tables are specified for commodities net of inedible parts, then country-level compilers must first apply a refuse conversion factor to calculate the weight of edible product before applying the nutrient conversions.

### **Activity and productivity variables**

Compilers should also collect data on other relevant variables that may be necessary for the imputation of missing values. The most widely recognized of these are activity and productivity variables. With respect to primary crops, the relevant activity variables are area sown and area harvested (most commonly measured in hectares, or ha). For livestock, the activity variable refers to the number of animals – both the total number of animals and the numbers of animals kept for specific purposes, including for dairy production, breeding, or draft purposes. With regard to productivity, the most common productivity indicator for crops is yield (which is often expressed in tonnes per ha); for livestock, relevant productivity indicators include carcass weight (sometimes also referred to as slaughter weight) and take-off (sometimes referred to as off-take) rate.

In addition to being necessary for the imputation of missing values, these activity and productivity variables can also be useful in validating the main production variables. For example, the production of crops is typically the result of a harvested area and a crop yield. To verify the feasibility of the production estimate, compilers may analyse the area and the yield required to achieve the stated production value, potentially also comparing yields to historic trends or agronomic potential, and analysing the harvested area in the context of current market conditions or the land suitable for the production of said commodity. In the same vein, carcass weight may also serve as a simple validation check on the quantity of meat produced from a given number of animals in the livestock data domain.

---

5 The nutrient conversion tables used by FAO ESS in composing FBS are available at: [http://www.fao.org/fileadmin/templates/ess/ess\\_test\\_folder/Food\\_security/Excel\\_sheets/Nutritive\\_Factors.xls](http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Food_security/Excel_sheets/Nutritive_Factors.xls). Because food quantities are reported in gross weight, these nutritive factors take into account the fact that the weights contain inedible parts. It should be noted that the calorie contents for certain products may vary slightly from those reported in the spreadsheet. In fact, various other countries publish their own versions of nutrient conversion tables. Countries may use the FAO factors as a default; however, they are encouraged to investigate the calorie contents of their largest food staples to ensure an accurate estimation of food availability.

## Extraction rates

*Extraction rates* are parameters that reflect the loss in weight during the conversion (or processing) of one product into another. They are one example of a “technical conversion factor”, a numerical factor that can be applied to one quantity to convert it to a different unit of measure. Extraction rates are typically expressed as a percentage, and are calculated as the amount (by weight) of derived product that is produced using a given amount of input product, as below in equation (2-6).

$$\text{Extraction rate} = \frac{\text{Quantity of output}}{\text{Quantity of input}} \quad (2-6)$$

For example, to produce 80 tonnes of maize flour, 100 tonnes of maize are required: the extraction rate for this transformation would calculate to 80 percent, expressed as follows:

$$\begin{aligned} \text{Extraction rate} &= \frac{80 \text{ MT maize flour}}{100 \text{ MT maize}} \\ \text{Extraction rate} &= 0.80 \end{aligned} \quad (2-6)$$

Note: MT = metric tonnes

Extraction rates are key components of the balance sheet, both when calculating the production of processed products from primary ones, and when converting derived product quantities back up to the primary product equivalent.

It should be noted that several output products may be produced from a single transformation process of one input good. In these cases, it is important to check that the cumulative extraction rate is lower than 100 percent, because it is not possible to produce, by product weight, more output than was originally input into the process<sup>6</sup>. Carrying forward the example of maize flour from above, the same transformation process that produces flour also produces both maize bran and maize germ.

## Processing shares

In the context of the FBS, *processing shares* are percentages of the amount of a given commodity sent to processing that are thought to be dedicated to a specific transformation process. They are often necessary for the composition of FBS because goods can be processed into an array of derived products, and the input used for the production of these derived goods is seldom known with certainty. As such, shares can be applied to the amount of a good sent for processing to calculate the amount of input into a given transformation process, and an extraction rate can then be applied to those input quantities to derive a production estimate. Thus, by using processing shares and extraction rates in concert, FBS compilers can obtain an estimate of the production of derived goods when very little information exists. An example of the application of shares will be elaborated in the following section after the concept of linking primary and derived goods through commodity trees has been laid out in further detail.

<sup>6</sup> The only exception is in cases where water, vinegar or other products have been added during the transformation process.

## 2.3 LINKING SUPPLY UTILIZATION ACCOUNTS (SUAS) TO FBS THROUGH STANDARDIZATION USING COMMODITY TREES

While FBS are typically only published at the primary commodity equivalent level to facilitate interpretation and policy formation, accounting for supply and use only for the primary commodity would not provide a holistic picture of how the commodity is being consumed, traded or otherwise used after being processed into various derived products. For example, in most cases, a balance for wheat alone would include little or no food use, because wheat is commonly processed into flour before it is consumed by humans, and flour is then used to produce various other derived products such as bread, pastries and pasta. Because there is both supply and demand for each of these products (both primary and derived), individual accounts should be kept for both the primary product and all of its derived products. These accounting balances for individual products are called Supply Utilization Accounts (SUAs). For the purposes of deriving FBS, SUAs are typically organized into tables where the SUA for the primary commodity is at the top, and the SUAs for all of the products derived from that commodity follow (Table 2-1).

**TABLE 2-1. SAMPLE BLANK SUA TABLE FOR PADDY RICE.<sup>A</sup>**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss	ROU
Paddy rice	-	-	-	-	-	-	-	-	-	-	-	-
Husked rice	-	-	-	-	-	-	-	-	-	-	-	-
Milled paddy rice	-	-	-	-	-	-	-	-	-	-	-	-
Rice bran	-	-	-	-	-	-	-	-	-	-	-	-
Broken rice	-	-	-	-	-	-	-	-	-	-	-	-
Rice flour	-	-	-	-	-	-	-	-	-	-	-	-

<sup>A</sup>Compilers should note that this is only a sample SUA table, in that (1) not all of these products will be produced from paddy rice in every country, and (2) additional products may also be produced from paddy rice, including rice bran oil, starch, beer, breakfast cereals, rice fermented beverages and cereal preparations not elsewhere specified (nes). For sample commodity trees, see FAO (1999).

For each primary commodity family, compilers should elaborate SUAs for both the primary commodity in question and all of its derived subproducts, which can include several different levels of processing. To continue the paddy rice example above in Table 2-1, primary-commodity paddy rice that is processed can be both processed to produce husked rice, or milled into milled paddy rice and rice bran. However, once the paddy rice has been milled to produce milled paddy rice, broken rice can then be separated from that milled paddy rice, and then used to produce rice flour. Each of these subsequent processing levels is linked back to the previous level through an extraction rate.

Once individual SUAs have been elaborated for the primary commodity and all of its derived products, the accounts cannot simply be added together to obtain one primary equivalent balance. This is because the balances are elaborated by weight in tonnes of primary-commodity equivalents, and 1 tonne of a derived product is not equivalent to 1 tonne of the primary commodity.

Indeed, consider a hypothetical customer who is interested in buying a large quantity of orange juice. Company X sells both fresh oranges and orange juice, and they offer to sell the customer either 100 tonnes of fresh oranges

or 100 tonnes of orange juice for the same price. Furthermore, Company X offers to process the fresh oranges into juice free of charge at an extraction rate of 55 percent (that is, 0.55 tonne of juice output per 1 tonne of fresh orange input). FBS compilers will recognize that the customer should definitely choose the juice instead of the fresh oranges, because 100 tonnes of fresh oranges will only yield approximately 55 tonnes of juice after processing. This calculation can be made by rearranging equation (2-6) as:

$\text{Quantity of output} = \text{Quantity of input} * \text{Extraction rate}$	(2-7)
---	-------

$\text{Quantity of output} = 100 \text{ MT fresh orange} * \frac{0.55 \text{ MT orange juice}}{1 \text{ MT fresh orange}}$	
$\text{Quantity of output} = 55 \text{ MT orange juice}$	

Note: MT = metric tonnes

This non-equivalence between 100 tonnes of orange juice and 100 tonnes of oranges is also evident by working backwards from the amount of juice. To produce 100 tonnes of orange juice, approximately 182 tonnes of fresh oranges would be necessary as input. Again, this calculation can be made by rearranging equation (2-6) as:

$\text{Quantity of input} = \frac{\text{Quantity of output}}{\text{Extraction rate}}$	(2-8)
---	-------

$\text{Quantity of input} = \frac{100 \text{ MT orange juice}}{\frac{0.55 \text{ MT orange juice}}{1 \text{ MT fresh orange}}}$	
$\text{Quantity of input} = 182 \text{ MT fresh orange}$	

Note: MT = metric tonnes

This illustration underscores the concept that it is incorrect to simply add the quantities of primary and derived products together. Instead, derived products must first be converted back to their “primary commodity equivalent” (the amount of primary commodity input that would be required to produce a given amount of derived product output), and then all of the primary commodity equivalents can be added together to obtain a single overall balance. As seen in the example, derived products can be converted back to their primary commodity equivalents simply by dividing by the extraction rate.

This process of converting derived products to a primary equivalent so that they can be added up is called “vertical standardization”. The FBS for primary equivalent products are created by standardizing and adding up individual SUAs for derived products. Graphical representations of this process follow in section 2.3.1 on commodity trees, and the calculations behind standardization will be covered in greater detail in chapter 4. For current purposes, the concept has been simply introduced and its links to extraction rates explored. Following up on this discussion, it is possible to also introduce an equation to be used in the standardization process based on the preceding discussion (equation (2-9)).

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

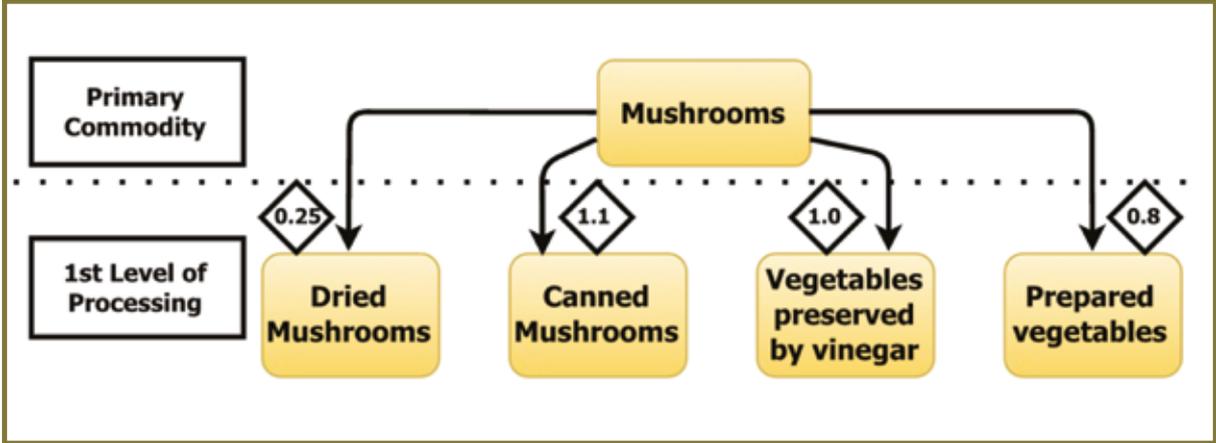
This linking of primary to derived commodities using extraction rates is fundamental to the FBS compilation process. This process is simple to understand if there is only one derived good. However, this is typically not the case: most food manufacturing commodities produce multiple outputs, and it is even possible for those outputs to undergo further transformation into second-tier derived goods. To better conceptualize these complicated primary/derived product relationships and better organize the standardization work, commodities and their derived products are organized into “commodity trees.” These constructs are so useful and crucial to the FBS process that a separate section describing them is warranted.

**2.3.1 Commodity trees**

Commodity trees are thus called because they “stem” from one primary product and then branch out into one or successive levels of processed products, with each level linked by extraction rates. Commodity trees are designed to be exhaustive, such that all processing uses of a particular commodity are covered. This means that they can be more or less complicated depending upon the number of derived products, the number of processing levels, and the creation of co-products during processing.

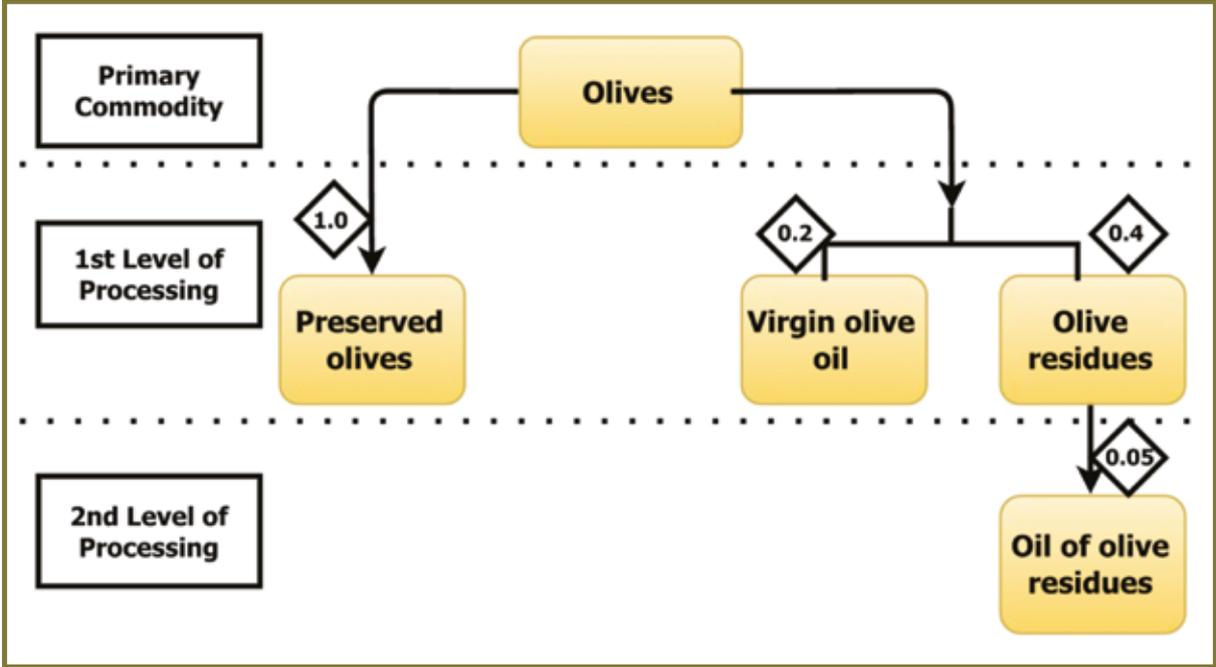
Two sample commodity trees can better illustrate some of these concepts. In Figure 2 1, it may be seen that the primary commodity of “mushrooms” can be processed into four different derived products: dried mushrooms, canned mushrooms, vegetables preserved by vinegar, and prepared vegetables. Note that each arrow for all of these products is separate: the quantities that enter the production process to become one of the derived goods cannot enter the production process to become any other of these goods. The extraction rate for each of these conversion processes is noted in the diamond above each derived product. In the case of dried mushrooms, the extraction rate “0.25” indicates that for every 100 tonnes of mushrooms that enter the process to become dried mushrooms, 25 tonnes of dried mushrooms will be produced. Note the extraction rate of “1.1” for the production of canned mushrooms from fresh mushrooms, indicating that for every 100 tonnes of mushrooms entering the canning process, 110 tonnes of canned mushrooms result. While this rate may at first appear to be nonsensical, this is because brine is added during the canning process, which increases the total product weight. Compilers should note that while cases such as these exist for a few products, extraction rates for most processes are lower than 1.

FIGURE 2-1. MUSHROOM COMMODITY TREE.



While transformation processes for mushrooms produce only one output from each process, compilers should note that the transformation processes for many derived products commonly have two (or even three) outputs. Multiple products that are produced from a single transformation process are called co-products. Co-products are commonly produced in the milling of cereals (flour milling also produces bran) and in oilseed crushing (where both oil and cake are produced), among others. A simple example of co-products produced from a single transformation process may be found in the olive commodity tree. Olives can be either diverted to the production of processed olives, or milled to produce olive oil. As may be seen in the olive oil processing branch of Figure 2-2, the extraction rate of “0.2” indicates that for every 100 tonnes of olives milled, 20 tonnes of olive oil are produced. However, the presence of the “olive residues” co-product indicates that this same process also creates 40 tonnes of olive residues. It is important to note that only one derived commodity from each transformation will be standardized back up to the primary commodity balance. This concept will become more evident in the product-specific balance sheet illustrations at the end of these guidelines.

FIGURE 2-2. OLIVE COMMODITY TREE.



Compilers at the country level are encouraged to familiarize themselves with the concept of commodity trees before starting to compile FBS. The tree structures for almost all commodities that undergo processing are available on FAO’s website<sup>7</sup>. Countries are encouraged to review those trees and update them as necessary for their purposes. At a minimum, country-specific extraction rates should be added. In the absence of extraction rate estimates from the country for which the FBS is being compiled, the extraction rates of neighbouring countries can certainly be adopted as a next-best option, particularly if the neighbouring country utilizes similar technologies in their agroprocessing industries.

**2.3.2 Processing shares**

These sample commodity trees also provide an excellent starting point for better comprehending the concept of processing shares, mentioned in section 2.2.2. It should be recalled that a processing share is the percentage of the amount of a given commodity sent to processing that is thought to be dedicated to a specific transformation process. These shares are then used to calculate the amount of input used for a given transformation process, as below in equation (2-10). Specifically, the quantity of input required for any processed Good B is equivalent to the quantity of its source Good A that is sent to processing, multiplied by the a priori processing share.

<i>Q Input for B = Q of A sent to processing * B's Processing share</i>	(2-10)
---	--------

Some key points must be kept in mind when working with shares. First, the processing shares of co-products (two or more products that are outputs of the same transformation process of a single input good) will be identical. In addition, the processing shares must add up to 100, given that the entire amount of the higher-level good sent to processing is, by definition, transformed into another good. A brief example on the use of processing shares to calculate processing input is provided below in box 2-1.

<sup>7</sup> These trees may be found in FAO, *Technical Conversion Factors for Agricultural Commodities*, available at: <http://www.fao.org/fileadmin/templates/ess/documents/methodology/tcf.pdf>. Accessed on 19 January 2017.

### BOX 2-1. SAMPLE EXERCISE ON THE APPLICATION OF PROCESSING SHARES.

Using only the first level of processing illustrated in the olive commodity tree described above, consider the following example. FBS compilers in Country A know that olives are processed domestically into both preserved olives and virgin olive oil. Although the FBS compilers know that the amount of olives sent to processing is 150 000 tonnes (because olives are not consumed fresh, all olives produced, net of trade, are sent to processing), they do not know the exact quantities directed to each of the different transformation processes. By using supply chain studies and consulting with market experts, the compilers learn that only a small fraction of olives are processed into preserved olives: approximately 10 percent, thus indicating that 90 percent of olives are milled for olive oil. Two points are noteworthy here. First, co-products from the same transformation process will have identical processing shares, because they are two goods derived from a single input. In this example, this means that the processing shares for virgin olive oil and olive residues will both be 90 percent, since the two are outputs of a single transformation process. Second, the processing shares for the different transformation processes must sum to 100; in other words, FBS compilers must ensure that all transformation processes are accounted for. For this example, although there are three output goods, there are only two transformation processes; therefore, it is only necessary to add the 10 percent processing share for preserved olives to the 90 percent processing share for olive oil as a check.

Using this information on the amount of the primary good sent to processing and the processing shares for the different transformation processes, FBS compilers can calculate the amount of input for each good, as below in BoxTable 1. Applying equation (2-10), the amount of olives processed in line A is multiplied by the processing shares for the various derived goods in line B to calculate the amount of input for each transformation process in line C. Note that the inputs for co-products virgin olive oil and olive residues are identical.

**BOX TABLE 1. SAMPLE CALCULATION OF PROCESSING INPUT FOR DERIVED OLIVE GOODS USING PROCESSING SHARES.**

		Olives	Preserved olives	Virgin olive oil	Olive residues
A	Amount processed	150 000			
B	Processing share		10%	90%	90%
C	Amount of input		15 000	135 000	135 000

Using these input quantities, it is possible to go one step further and add the product-specific extraction rates noted in the olive commodity tree above to calculate the production of the derived good, by multiplying the amount of input on line C by the product-specific extraction rates on line D (BoxTable 2). This concept of calculating production from input quantities and extraction rates is further elaborated upon in section 3.5.1.3 on production imputation.



**BOX TABLE 2. SAMPLE CALCULATION OF PRODUCTION OF DERIVED GOODS USING INPUTS AND EXTRACTION RATES.**

		Olives	Preserved olives	Virgin olive oil	Olive residues
A	Amount processed	150 000			
B	Processing share		10%	90%	90%
C	Amount of input		15 000	135 000	135 000
D	Extraction rate		100%	20%	40%
E	Production		15 000	27 000	54 000

## 2.4 THE RECOMMENDED BALANCING MECHANISM

In its simplest form, the exercise of compiling FBS merely uses measured and imputed data to balance the supply = utilization identity. While this is a simple and intuitive construct, balancing the equation in practice is not so straightforward, for many reasons. First, countries measure all supply and demand variables only in a very limited number of cases. Instead, what tends to occur is that supply-side variables are measured, while most demand-side variables are imputed using statistical models or estimated by subject matter experts. If all demand-side variables were to be estimated to balance out the supply side, all of the measurement error of the supply-side variables would accumulate in the demand-side variables, affecting the accuracy and potentially increasing the uncertainty or volatility of demand-side estimates. Second, in the rare cases when all supply and demand variables are measured independently, the point estimates alone are unlikely to lead to a precisely balanced supply and demand equation, because of discrepancies in data sources, data collection and compilation methods, reference periods and the measurement errors occurring at any of these stages. For these reasons, an overall strategy for balancing the equation must be found.

Historically, the approach has been to assign one element of the equation as the balancing item: a certain variable is selected to represent the combination of its estimate and the residual imbalance of the equation. For example, in their balance sheets for coarse grains, USDA estimates “feed and residual” use as the balancer to the equation, such that after production, trade, stocks, food, seed, and industrial use are measured or estimated, “feed and residual” is estimated as the remainder: that is, the supply minus all measured uses except the unmeasured feed and residual component. The variable used as the balancing item may vary depending upon the nature of the product and the country’s agricultural statistical systems; however, feed and food are commonly used as balancers. It should be noted that this approach is most appropriate for statistical systems in which all variables are measured except for the balancing item.

This “one balancer” approach has been popular for decades, with its main advantage being convenience (with this approach, one variable of the equation does not have to be measured). However, the approach also has various drawbacks. First, in most countries, few utilization variables are measured, such that the supply = utilization equation will actually present more than one unknown, which can complicate the assigning of a single balancing item. Second, assigning one variable to be the balancer lumps the entire measurement error occurring in each one of the other variables onto that single balancer variable, such that estimates for the balancing item may fluctuate

wildly from year to year; this could be inaccurately interpreted by laypeople as fluctuation in the variable itself. Additionally, over time, if the errors are biased, those annual errors accumulate, such that the underlying variable that is purported to be estimated may become difficult to distinguish from the error itself. Finally, the choice of variable to use as the balancer can be problematic. For example, using the “food” variable as a balancer would then tend to show fluctuation in food availability, which should theoretically change little from year to year. In some cases, countries have begun to measure a variable that had previously been assigned as the balancing item, such that a new balancer must be assigned which could potentially be even less suitable for that function. For example, consider a country that has designated feed as the balancing item in their maize balance. If the country begins to survey farmers and feeding operations to derive an official estimate of the feed use of maize, a new balancer will have to be identified, since feed will now be measured.

Keeping these issues of accumulated error in mind, the preferred approach to balancing the supply = utilization identity is one that not only acknowledges measurement error, but also seeks to use these errors of individual variables to help balance the overall identity. This is accomplished by specifying each of the variables as a range of possible values according to their measurement errors. In this way, the basic identity should be expanded, as:

$  \begin{aligned}  & Production^* + Imports^* - \Delta Stocks^* \\  & = Export^* + Food^* + Food Processing^* + Feed^* + Seed^* \\  & + Tourist Food^* + Industrial Use^* + Loss^* + Residual Use^*  \end{aligned}  $	(2-11)
--	--------

where:

$  \begin{aligned}  & Production^* = Production \pm e(Production) \\  & Imports^* = Imports \pm e(Imports) \\  & etc. \text{ for the remaining variables.}  \end{aligned}  $	
--	--

In contrast to the point estimates of the variables described previously in equation (2-4), the terms in equation (2-11) represent the possible range – or tolerance interval – of each of the variables, which is specified as the sum of the original expected value for that variable and the measurement error of that variable, *e*.<sup>8,9</sup> Further guidance on the estimation of these tolerance intervals is detailed in section 3.4.2.3. However, for current purposes, the interpretation of this equation is facilitated by emphasizing that variables that are measured (by surveys, for instance) are likely to have narrower tolerance intervals than the variables that are imputed or estimated. The underlying principle of a statistically sound balancing method is that the elements with the largest tolerance intervals should be those in which the bulk of the imbalance in the identity is distributed.

Given this interpretation of the supply = utilization identity, the balancing process becomes merely a matter of distributing the equation’s imbalance. Several approaches to this imbalance distribution are offered below in section 2.4.1, and all follow three basic steps:

8 In this context, “measurement error” includes both non-random error (systematic errors that always bias the estimates in a particular direction) and random error (including sampling errors, which are not biased in any one direction).  
 9 Food processing is not included in this identity. This is because when balancing occurs (after standardization and aggregation), in most cases, food processing will already have been dropped from the balance to avoid double counting. More details on this process can be found in chapter 4.

**Step 1. Calculate the imbalance from the supply = utilization identity, where the imbalance for a given commodity in the country in question, *Imbalance*, is defined as:**

$  \begin{aligned}  \textit{Imbalance} = & \textit{Production} + \textit{Imports} - \Delta\textit{Stocks} - \textit{Exports} - \textit{Food} \\  & - \textit{Food Processing} - \textit{Feed} - \textit{Seed} - \textit{Tourist Food} \\  & - \textit{Industrial Use} - \textit{Loss} - \textit{Residual Use}  \end{aligned}  $	(2-12)
---	--------

It is important to note that in this step, the imbalance is calculated from the variable point estimates. No accounting has yet been made for the measurement error; this is done in step 2. As noted above, regardless of the method chosen to distribute the imbalance throughout the identity, step 1 must be performed and does not change in any of the proposed methods.

**Step 2. Distribute the imbalance throughout the supply = utilization identity:**

This step can be more or less complicated or computationally demanding, and it is at this stage that the methodological approaches of individual countries may differ. The optimal approach will consider all of the information contained within the underlying variable estimates – that is, balancing will be achieved using the information contained in the tolerance intervals to distribute the equation’s imbalance throughout the variables.

Additionally, it is important to note that the direction of adjustments in the point estimates depends upon the sign on the imbalance calculated in step 1. In other words, if the calculated imbalance is positive (indicating that supply is greater than utilization), any adjustments in the supply variables “production” and “imports” should be downward, while adjustments in the remainder of variables (the utilization variables, as well as stock changes given its opposite sign) must be positive. In contrast, if the calculated imbalance is negative (that is, utilization is greater than supply), then any adjustments in production and imports should be positive, while adjustments in the remaining variables should be negative.

**Step 3. Check that all balanced quantities are within any set bounded values, and rebalance if necessary:**

In some instances, the balancing process will produce results where certain balanced quantities are estimated outside of bounded (or likely) values. In these cases, this problem is resolved by setting the value in question at the boundary level and assigning that value a tolerance interval of zero (thus, a fixed, “balanced” value), and then repeating steps 1 and 2 to redistribute the imbalance. For example, if there is an upper bound for industrial use of 50 tonnes because of the available processing capacity, but the balancing calculation estimates industrial use at 100 tonnes, then the value for industrial use is set at the bounded level of 50 tonnes with a tolerance interval of zero, and the balancing calculation is performed again.

## 2.4.1 Approaches to distributing the imbalance at FBS level

This section presents the most common methods that can be used to distribute the imbalances, along with what can be considered as the “gold standard” approach. The countries’ choice of which approach to utilize may be influenced by their statistical capacity, time constraints, desire for replicability, or structural constraints. The advantages and disadvantages of each of these approaches is included below, with a general description of the approach.

### 2.4.1.1 Recommended approach – distribute imbalance proportionally based on aggregated error

The recommended imbalance distribution approach for country-level FBS compilers seeks to use the information on the uncertainty of point estimates to proportionally distribute the imbalance; that is, the variables with the highest tolerance intervals (considered the least reliable) are adjusted proportionally more than variables with a lower tolerance interval. The necessary steps in this method include:

**Step 1: Use tolerance interval percentages and point estimates to quantify the error of each variable. If the quantity for a given variable should remain fixed because it is an official estimate, a tolerance interval of zero can be assigned.**

**Step 2: Sum up the individual errors of each variable to calculate an aggregated error for the equation.**

**Step 3: Calculate the proportion of aggregated error for each of the elements.**

**Step 4: Distribute the imbalance proportionally, recalling that a negative imbalance indicates that production and import variables must be increased and the remaining variables must be reduced from their pre-balanced values, while the opposite is true for a positive imbalance.**

**Step 5: Ensure that any constraints are met, and recalculate if necessary.**

A simple example case balancing sorghum supply and utilization in Country Z can be used to illustrate this method (compilers may note that several variables have been eliminated from this brief example, for the sake of simplicity). FBS compilers in Country Z have produced the following unbalanced supply and utilization table for sorghum in their country (for the purposes of this illustration, sorghum is not consumed as food or processed in Country Z but is mostly destined for feed, and the country does not utilize a Residual and other uses category). Note the point estimates on line A and the a priori tolerance intervals on line C below in table 2-2. Given the data on line A, the imbalance in the supply and demand account is calculated in line B (as  $Imbalance = Production + Imports - Exports - Feed - Seed - Loss$ , or  $892 + 307 - 48 - 1061 - 3 - 44 = 43$ ).

**TABLE 2-2. UNBALANCED SORGHUM SUPPLY AND UTILIZATION IN COUNTRY Z.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Feed (4)	Seed (5)	Loss (6)	
A	Sorghum	892	307	48	1 061	3	44	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6]							43
C	Tolerance Interval (in %)	±15.0%	±0.0%	±0.0%	±40.0%	±15.0%	±15.0%	

Beginning with this table, the first step is to quantify the error into units instead of percentages. This can be done by multiplying the values in line A by the percentages in line C to obtain the values in line D (table 2-3).

**TABLE 2-3. STEP 1 – UNBALANCED SORGHUM TABLE WITH QUANTIFIED ERROR.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Feed (4)	Seed (5)	Loss (6)	
A	Sorghum	892	307	48	1 061	3	44	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6]							43
C	Tolerance interval (in %)	±15.0%	±0.0%	±0.0%	±40.0%	±15.0%	±15.0%	
D	Error [D=A*C]	133.8	0	0	424.4	0.5	6.6	

Next, the individual error estimates for each of the variables are added to obtain an estimate of the equation's aggregated error. In this case, all of the values highlighted in green on line D are added together to obtain the aggregated error of 565.3 on line E (table 2 4).

**TABLE 2-4. STEP 2 – SUM INDIVIDUAL VARIABLE ERRORS TO CALCULATE AGGREGATED ERROR.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Feed (4)	Seed (5)	Loss (6)	
A	Sorghum	892	307	48	1 061	3	44	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6]							43
C	Tolerance interval (in %)	±15.0%	±0.0%	±0.0%	±40.0%	±15.0%	±15.0%	
D	Error [D=A*C]	133.8	0	0	424.4	0.5	6.6	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6]							565.3

From here, the proportion of the aggregated error pertaining to each of the individual variables is calculated in line F (table 2-5). This is achieved by dividing the error of each individual variable by the aggregated error estimate. For example, the proportion of aggregated error that is attributed to production in this scenario is 23.7 percent ( $\frac{133.8}{565.3} = 0.237$ ).

**TABLE 2-5. STEP 3 – CALCULATE PROPORTION OF AGGREGATED ERROR FOR EACH INDIVIDUAL VARIABLE.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Feed (4)	Seed (5)	Loss (6)	
A	Sorghum	892	307	48	1 061	3	44	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6]							43
C	Tolerance interval (in %)	±15.0%	±0.0%	±0.0%	±40.0%	±15.0%	±15.0%	
D	Error [D=A*C]	133.8	0	0	424.4	0.5	6.6	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6]							565.3
F	Proportion of aggregated error [F=D/E]	23.7%	0.0%	0.0%	75.1%	0.1%	1.2%	

Next, the imbalance in the equation is assigned proportionally based upon these percentages. In other words, the total imbalance in the equation (in this case, 43) is proportioned across the different variables according to their share of the aggregated error: 23.7 percent of the aggregated error is accounted for in production, 0 percent is accounted for in imports, 0 percent is accounted for in exports, 75.1 percent is accounted for in feed, 0.1 percent is accounted for in seed, and 1.2 percent is accounted for in loss (table 2-6). To calculate the adjustment to be applied to the unbalanced quantities, the imbalance in line B is simply multiplied by the percentages in line F to obtain the adjustments in line G. Then, the unbalanced values in line A are adjusted by the values in line G to achieve the balanced equation in line H. Recall that because there is a positive imbalance, the supply variables production and imports must be adjusted downward, and the utilization elements should all be adjusted upward. If the imbalance were to be negative, production and imports would instead be adjusted upward and the remaining elements adjusted downward.

**TABLE 2.6. STEP 4 – DISTRIBUTE THE IMBALANCE PROPORTIONALLY.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Feed (4)	Seed (5)	Loss (6)	
A	Sorghum	892	307	48	1 061	3	44	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6]							43
C	Tolerance interval (in %)	±15.0%	±0.0%	±0.0%	±40.0%	±15.0%	±15.0%	
D	Error [D=A*C]	133.8	0	0	424.4	0.5	6.6	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6]							565.3
F	Proportion of aggregated error [F=D/E]	23.7%	0.0%	0.0%	75.1%	0.1%	1.2%	
G	Adjustment [G=B*F]	10.2	0.0	0.0	32.3	0.0	0.5	
H	Sorghum adjusted values [for (1) and (2), H=A-G, for remaining, H=A+G]	881.8	307.0	48.0	1 093.3	3.0	44.5	
I	Imbalance for H [I=H1+H2-H3-H4-H5-H6]							0

The final step in this method is to ensure that any constraints are met, and recalculate if necessary. In this particular simplification, none of the values violate any constraints; therefore, it is not necessary to perform any rebalancing and the account can be considered to be balanced.

Although this approach requires several steps, it is not computationally demanding and can be easily replicated. This approach also takes into account the imprecision of the point estimates by calculating the adjustments based on confidence in the individual estimates (as expressed by the tolerance interval), using all information available to achieve a balanced account. At the same time, the balanced equation produced by this approach will vary slightly depending upon the a priori assigned tolerance intervals. The final limitation to this method is that it may not be feasible in countries that wish to publish detailed accounts for derived products: here, the balancing occurs at the primary equivalent; therefore, it is not straightforward to go back down the tree and recreate the SUAs.

#### **2.4.1.2 Assigning small positive imbalances to a residual use category**

An additional alternative approach (ideally for situations when the imbalance is small and positive) is the use of the Residual and other uses category, instead of distributing the imbalance across variables. This approach could be utilized on its own where a positive imbalance is lower than an a priori threshold (for example, less than 5 percent of total supply or total demand). In this way, the error does not accumulate in any of the other variables, and is addressed in a transparent way. At the same time, this approach should not be used for imbalances greater than a low threshold level, as imbalances of such dimensions would indicate either that a certain form of utilization is being missed, or that the estimation of at least one utilization is highly imprecise. In such cases, countries may wish to continue using the Residual and other uses category to account for the imbalance up to the established threshold, and then redistribute the remaining imbalance with proportional balancing. However, it should be emphasized that if the TWG determines that the imbalance is very large, leaving a sizeable quantity of utilization unexplained, then the data of all the other utilizations should be checked (and, potentially, other data sources should be consulted).

#### **2.4.1.3 Single balancer approach**

The single balancer approach – in which one utilization variable is calculated as the remainder after all other utilizations are accounted for – remains a viable option for countries elaborating their country-level FBS. However, particularly for countries where multiple variables are not measured, country-level FBS compilers are encouraged to attempt the proportional balancing approach first, as that approach prevents the systematic accumulation of error in a single variable over time, allows for the quality of the data for each of the variables to be recognized in the final overall commodity balance, and enables the imposition of certain feasibility constraints. If, for whatever reason, country-level FBS compilers decide that the single balancer approach is the most feasible for their particular circumstances, they should be aware of and publicize the caveats and shortcomings of the estimates produced using this methodology.

In addition, it should again be emphasized that not all variables are appropriate as balancers in the single balancer approach, and the degree of appropriateness may even differ from product to product. For example, feed may be considered as the preferred balancing item for the maize balance sheet; however, for a commodity such as apple juice, which is unlikely to ever be used as feed, it is not possible to use feed as the balancer, and food use is instead the preferred balancer.

## 2.4.2 Constraints on the balancing process

Step 3 of the recommend balancing approach alludes to the idea that the balancing process should take into account certain constraints on the values – a topic explored in further detail in this section. While some constraints may be universal across countries and products, country compilers may find it necessary to impose additional constraints based on the prevailing conditions within their individual countries (as in the above example on processing capacity). Because it is not possible to foresee all possible constraining situations, only the universal constraints are explored below.

### Row constraint

The most obvious constraint on the process of allocation of the imbalance is that the supply for each commodity must be equal to the utilization for that commodity, referred to as a “row constraint.” As an extension of this row constraint, a country’s exports of a given commodity cannot exceed their supply of that commodity. Mathematically, this can be expressed as:

$Production + Imports - \Delta Stock > Exports$	(2-13)
---	--------

This particular row constraint can be a useful way of either identifying errors in trade data, or otherwise alerting country-level FBS analysts to the fact that a new commodity is being produced in their country. After all, a country cannot export a product if they are not producing it or importing it from elsewhere before exporting it (the latter case would technically be defined as a re-export; however, in trade data reporting, exports may not always be clearly differentiated from re-exports).

### Column constraints

Compilers should also be aware that constraints may have to be imposed on changes in quantities over time. Two examples of this type of constraint should be highlighted: single-year constraints and multiple-year constraints. With respect to single-year constraints, compilers should note whether year-to-year changes are feasible. One example is changes in food availability and derived DES estimates. Barring catastrophe (war, natural disasters, etc.), DES estimates are unlikely to vary greatly on an annual basis, with aggregate changes of 100 calories per capita considered the absolute upper bound. Stocks represent another obvious example of the need for a single-year column constraint, as subtraction from stocks in a given year cannot be greater than the overall level of stocks.

Multiple-year column constraints should also be considered. In this case, stocks are again the most prominent example, as it is considered highly unlikely that a country would either add to stocks or take away from stocks for many years in a row. If compilers find this to be the case, then they should consider imposing a bound on the stocks changes in the balancing process, to ensure that the pattern is discontinued.

### **“Vertical standardization” constraint**

Especially when information on production, trade and other utilizations of derived products are drawn from official data, it may be useful for countries to also specifically apply a “vertical standardization” constraint. That is, they must ensure that there is a sufficient quantity of primary product sent to processing, to ensure that none of the derived product accounts has any negative discrepancies (the “row constraint” mentioned above).

If countries find that their input for the production of derived commodities is insufficient, this is commonly an indication that the estimated extraction rate may be too low (therefore, a lower amount of input would be necessary to generate the same amount of output).

### **Imbalance exceeds aggregate measurement error**

Country-level FBS compilers may identify cases in which the imbalance in the equation exceeds the aggregate measurement error. These instances can result from a much larger error in one of the point estimates than is indicated by the assigned tolerance intervals. In itself, this situation is not problematic for any of the approaches outlined above. However, it does indicate that the uncertainty levels are set too conservatively (guidance on setting tolerance intervals is provided in section 3.4.2.3). As such, these cases are an opportunity for countries to re-examine official estimates for their accuracy, and in some cases, to consider assigning a higher tolerance interval to those estimates if their precision is deemed to be questionable.

## **2.5 SUMMARY**

This chapter has sought to outline the basic approach to constructing an FBS. FBS are based on an overall supply = utilization identity, in which accounts of primary and derived products are organized into commodity trees and linked by extraction rates. Individual supply utilization accounts of derived products are filled and balanced, and then aggregated up to the primary commodity equivalent level. The accounts at the primary commodity equivalent level are then balanced. The recommended approach to balancing the account at the primary commodity equivalent level involves taking into account the precision of the point estimates by using tolerance intervals to distribute the imbalance in the equation. Compilers should use this knowledge – that both expected value and tolerance interval are necessary as inputs – in their search for reliable input data. These concepts are covered in more detail in chapter 3. A practical example utilizing the balancing mechanism will follow in chapter 4

# 3

## Data for FBS compilation: considerations, sources and imputation

### 3.1 OVERVIEW

The bulk of the work of compiling an FBS involves the compilation and reconciliation of data. This process begins with an exercise in determining which products are most important to include in the assessment of data availability, including identifying all possible data sources, assessing data quality, and formulating a plan for filling in missing or unknown values. This assessment is fundamental to the construction of country-level FBS, as it allows the user to assemble the expected values of all variables in the balance and assign an a priori tolerance interval based on the quality of the data. Additionally, this process improves the user's understanding of the dynamics of each supply chain considered, to ensure that all uses are accounted for. Furthermore, the process facilitates the creation of a network of experts who can be consulted when doubts arise and may assist in the validation of the final product.

In the process of compiling data, country-level FBS compilers will consult a variety of sources, based on the variable and product in question. Guidance on the potential sources available to country-level FBS compilers is provided in section 3-5, which also contains suggestions on modelling or estimation approaches that country-level FBS compilers may find useful when attempting to impute or estimate missing values.

## 3.2 ASSEMBLING A TWG AND ASSIGNING ROLES

As emphasized in chapter 1, the ideal setup for FBS compilation is a dedicated TWG involving relevant actors from government institutions (potentially including the National Statistics Office (NSO), the ministry of agriculture, the customs office/ministry of trade, the ministry of industry/commerce and the ministry of health), research organizations, industry groups and producer groups. Thus, the first step toward FBS compilation is the assembling of such a TWG, which should include all organizations that produce cross-commodity data that will be included in the balance sheet. The effort should most likely be led by a group within the ministry of agriculture or the NSO; however, because institutional structures differ substantially between countries, it is not possible to advise with certainty which organization should lead this effort. Countries may also find that it is useful to put together commodity-specific committees to analyse the balances for certain products; in any case, it is recommended that an overarching FBS TWG be assembled to bring all of the individual primary-equivalent balances together at the end.

Once the TWG has been established, the roles of each organization should be defined. For example, some TWGs may prefer that the whole start-to-finish process take place within the entire TWG. Other TWGs may wish to assign the responsibility of producing data for certain components or variables in the balance sheet to a certain institution, with the TWG only meeting to validate the final output. Again, these arrangements depend upon the resources, technical capacities and desires of the compiling country, although it is highly recommended that final approval of the national FBS be endorsed by a TWG representing a collaboration of stakeholders.

## 3.3 DETERMINING PRODUCT SCOPE

Although FBS should strive to be comprehensive and include all products consumed within a country, the reality is that, for many countries, it may not be possible or practical to elaborate balances for every single product. For this reason, it is recommended that FBS compilers first determine the product scope of their national-level FBS before endeavouring to produce an FBS, with the ambition of reaching an approximation of national food availability. As a general rule, countries should strive to cover products that represent at least 90 percent of total caloric consumption, as identified in household consumption surveys. At the same time, compilers should seek to ensure that at least the most-consumed commodities in each commodity group are represented. In this exercise, data availability may be a constraining factor.

Although compilers may begin with an abbreviated list of commodities, they are encouraged to add products over time, as access to different source data increases and capacity for FBS compilation within the country improves. Country-level compilers are both welcomed and encouraged to use FAO's list of commodity items and FBS aggregates as a starting point in determining which products are most relevant for domestic FBS compilation<sup>1</sup>.

---

<sup>1</sup> The list of aggregates and their component commodities is available at: [http://www.fao.org/fileadmin/templates/ess/ess\\_test\\_folder/Food\\_security/Excel\\_sheets/Commodities\\_which\\_are\\_aggregated\\_or\\_standardized.xls](http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Food_security/Excel_sheets/Commodities_which_are_aggregated_or_standardized.xls).

## 3.4 DATA ASSESSMENT AND OTHER PRELIMINARY CONSIDERATIONS

After the scope of the products to be included in the FBS has been determined, the next step of data assessment is crucial to the FBS compilation process, as it will document all of the data sources used and help country-level FBS compilers to ensure data comparability. This assessment begins with the inventory of potential data sources for all relevant variables for each commodity, accompanied by an inventory of the various attributes of each data series. This assessment should document the basic accessibility of the data (including the source of the data, its typical release date and/or the frequency of its publication), attributes that establish the comparability of the various data series (including product classification, unit and reference period), and an assessment of the quality of the data, largely for the purposes of assigning a tolerance interval for each variable. Before providing guidance on the assessment itself, the concepts of data comparability and data quality will first be explored in further detail.

### 3.4.1 Data comparability

Throughout the process of gathering data, users must recall that for an “apples-to-apples” comparison to take place in the context of the balanced supply and utilization equation, the data must be fully comparable. This comparability includes various levels, including comparability of the item itself, comparability of the chosen reference period, and comparability of the units in which the item is measured. Each of these aspects is explored in further detail below.

#### 3.4.1.1 Ensuring product comparability through the use of statistical classifications

Perhaps the most obvious starting point for data comparability is to ensure that the products being compared are actually the same. As an example, production quantities for rice can be reported on either a paddy (unmilled) or milled basis. If production were to be recorded on a paddy basis in the supply and utilization balance, but another variable such as tourist food were to be recorded on a milled basis, an unintentional error would be introduced into the balancing process that could easily have been avoided if an analyst had checked that the products were strictly comparable. Thus, it should be evident why users must be aware of these types of distinctions during the process of compiling data.

To avoid these types of situations, countries are advised to express quantities of products using an international statistical classification structure. Such a structure will not only ensure the comparability of products within a balance sheet framework, but also facilitate the comparability of data between countries that utilize the same structures. Countries are of course free to develop their own statistical classification schemes, but numerous international statistical classifications already exist for agricultural statistics that can, in most cases, be adapted to fit the needs of any given country<sup>2</sup>.

For the purposes of these guidelines, two classification structures should be highlighted: the UN Central Product Classification (CPC) and the Harmonized Commodity Description and Coding System (HS).

---

<sup>2</sup> For more information on international classifications for agricultural statistics, see Global Strategy (2015b).

## UN Central Product Classification

The CPC is an international statistical classification for products that is maintained by the UN Statistics Division (UNSD). The most recent version of the CPC (Version 2.1) includes an annex on agricultural statistics that was developed by FAO to facilitate agricultural data collection and harmonization efforts. FAO uses this Version 2.1 Expanded of the CPC for its production surveys, and production data is available in both CPC and the designated FAO production classification (FAOSTAT Commodity List, also known as FCL)<sup>3</sup> on FAOSTAT<sup>4</sup>. It should be noted that the two lists have a 1-to-1 link for almost all commodities. For this reason, adopting the CPC at the country level would benefit the comparability of both in-country FBS with FAO-produced FBS, as well as in-country FBS with those of other countries.

The CPC organizes products into a five-level hierarchical structure, with the CPC expanded adding two more digits at the lower level to better account for some minor agricultural products. This structure is well-suited to the work of FBS compilation, as products are aggregated at the primary equivalent level in the commodity trees covered in section 2.3.1. An additional benefit of the CPC is that it is mapped to the Harmonized Commodity Description and Coding System (commonly referred to as the *Harmonized System*, or HS) classification for international trade (covered below), which greatly facilitates the comparison of production and trade data within the FBS context.

## The Harmonized Commodity Description and Coding System

The HS is a classification developed by the World Customs Organization (WCO). The HS is the most widely utilized classification in the context of international trade, as it is used by over 200 countries and covers 98 percent of international merchandise trade (Global Strategy, 2015b). The classification is updated every five years, with the 2017 version (HS 2017) having entered into force on 1 January 2017<sup>5</sup>.

The HS is also a hierarchical structure, comprised of 5 000 6-digit product groups contained within 97 chapters<sup>6</sup>. As with the CPC, expansions on this system are also possible. Indeed, many countries add further classifications at the 8-digit, 10-digit, or even 12-digit levels.

The use of the HS as a format for trade data within the FBS context is recommended, primarily for data comparability purposes (given that more than 200 countries already utilize this classification) and because of its ease of concordance with the CPC, as noted above<sup>7</sup>. FAO has also produced direct HS6 to FCL concordance tables (publicly available on the FAO website), which can facilitate classifications of imports and exports during FBS compilation<sup>8</sup>. As with the CPC to FCL classifications, most concord on a one-to-one basis.

3 The complete FCL, and a search for FCL code by keyword, are provided by FAO, 2017, FAOSTAT *commodity list*, available at: <http://www.fao.org/economic/ess/ess-standards/commodity/en/>. Accessed on 25 January 2017.

4 The most up-to-date correspondence table between the FCL and the CPC is available at: [http://www.fao.org/fileadmin/templates/ess/classifications/Corr\\_11Jan2017.xlsx](http://www.fao.org/fileadmin/templates/ess/classifications/Corr_11Jan2017.xlsx). Accessed on 22 February 2017.

5 The full HS 2017 is available at the WCO website. See WCO, 2017, *HS Nomenclature 2017 edition*, available at: <http://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs-nomenclature-2017-edition/hs-nomenclature-2017-edition.aspx>. Accessed on 17 February 2017.

6 Because data in the HS is harmonized at the 6-digit level, most FBS are compiled using trade at the 6-digit level (although an FBS commodity may include trade flows of multiple HS6 codes). For certain countries and commodities, it may be possible to compile FBS at a more detailed level, provided that data is collected at a more detailed level.

7 The correspondence table between the HS 2012 and CPC Ver.2.1 can be found at UNSD, 2017, *Correspondence between HS 2012 and CPC Ver.2.1.*, available at <http://unstats.un.org/unsd/cr/registry/regso.asp?Ci=81&Lg=1>. Accessed 25 January 2017.

8 The correspondence table for FCL to HS6 can be found at FAO, 2017, FAOSTAT *commodity definitions and correspondences: FCL – HS 2007*, available at: <http://www.fao.org/economic/ess/ess-standards/commodity/item-hs/en/> (accessed 22 February 2017), while the complementary table of correspondences for HS6 back to FCL is available at FAO, 2017, FAOSTAT *commodity definitions and correspondences: HS 2007 – FCL*, <http://www.fao.org/economic/ess/ess-standards/commodity/hs-item/en/> (accessed 22 February 2017).

In some cases, either multiple FCL or CPC products are covered under a single HS6 category, or multiple HS6 products feed into a single FCL category. This typically occurs for “baskets” of similar commodities that are not elsewhere specified (nes), such as other fresh fruits, fat preparations nes, or food preparations nes. In these cases, countries may have more specific 8- or 10-digit HS codes that can facilitate the linking of a one-to-one comparison. In other cases, it may be necessary for country-level compilers to conduct more careful research on the types of products that predominate in the basket categories identified. This may be accomplished by speaking with customs brokers, analysing customs declarations, or even by consulting mirror trade data from other countries that may offer a more detailed product description at the HS8 or HS10 level. If compilers specifically wish to categorize import data for which one HS is mapped to many FCLs or CPCs, it may be possible to use the country of origin as an additional strategy to identify which products are entering under a basket category. For example, the HS6 code 0810.90 covers “Other, other fresh fruit,” which includes stone fruit nes, pome fruit nes, tropical fruit nes and other fruit nes. An analysis of the source countries for trade flows under 0810.90 could then be cross-referenced with the reported production from each source country as a way to indicate where a given trade flow should be classified.

#### **3.4.1.2 Common units**

Aside from ensuring that the data compiled for the different FBS variables refer to the same products, by means of using an international statistical classification structure, it is also important to recall that product values must be reported in common units so that the equation can be balanced. Many countries report production for the majority of agricultural products in terms of tonnes; however, users should also consider that some quantities may be reported in 1 000 tonnes. Adjustments should be made accordingly when compiling the balance sheet, to ensure usage of a common unit.

In some countries, however, agricultural production is measured and reported in other units specific to those countries. As two examples, El Salvador reports production of most crops in terms of quintals, and the United States of America reports production for most grain crops in terms of bushels. At the same time, most trade data is reported in tonnes, and most calorie conversion tables are elaborated in terms of calories per kilogram. For this reason, it is recommended that countries elaborate balance sheets in tonnes. Aside from the ease of converting quantities into calories equivalents, compiling balance sheets in tonnes has other advantages – primarily, country balance sheets compiled in tonnes are directly comparable to those compiled by FAO, and are also more likely to be comparable to balance sheets compiled by other countries. At the same time, balance sheets compiled in other units may be more easily understood by stakeholders within the country. Whichever unit is chosen, compilers should be careful to convert all data expressed in different units of measure using internationally accepted conversion rates.

Last, for some liquid products, certain variables in the balance sheets may be reported in litres, while other variables may be reported in tonnes. For these conversions, it should be noted that compilers must use conversion factors specific to the product in question (due to differences in densities) to avoid introducing additional error to the estimates. FBS compilers should consult their relevant national authorities to determine applicable conversion factors for liquid products. Absent data from national authorities, approximate densities for similar products are available at the databases created and managed by FAO and the International Network of Food Data Systems (INFOODS)<sup>9</sup>.

---

<sup>9</sup> The FAO/INFOODS Databases can be accessed here: FAO, 2017, FAO/INFOODS Food Composition Databases, available at: <http://www.fao.org/infoods/infoods/tables-and-databases/faoinfoods-databases/en/>. Accessed 19 January 2017.

### **3.4.1.3 Reference period**

After the comparability of the item and the unit of measurement have been ensured, compilers should decide upon a consistent reference period for the balance sheet. Common reference periods include the marketing year (sometimes also referred to as the agricultural year, harvest year or crop year), the calendar year and the fiscal year (which may also be referred to as the financial year or budget year): the marketing year begins in the month when the bulk of the crop in question is harvested, the calendar year begins in the first month of the calendar (January–December, for countries observing the Gregorian calendar), and the fiscal year begins whenever defined by the country in question. For the sake of data comparability, it is recommended that countries compile their balance sheets on a calendar-year basis. However, depending upon agricultural data collection programs, compiling sheets on a marketing-year or fiscal-year basis may be more feasible. Each of these options is explored in further detail below.

#### **Marketing year**

Marketing years have the advantage that they closely follow the cycle of each season, such that the utilization of each year's production follows a logical conceptual temporal flow, ending with the harvest of the subsequent season. However, this type of temporal aggregation may be problematic if different crops are harvested at different points in the year – for example, if the balance sheet for corn is compiled on an April–March basis, and the balance for oranges is compiled on a July–June basis, then the calories from the two products cannot be added together for the purposes of obtaining a total DES estimate. For this reason, marketing years tend to be more appropriate for estimating single, stand-alone commodity balances rather than an overall FBS.

An additional complication posed by the marketing year involves countries that experience multiple harvests. For example, in several tropical countries, rice can be harvested at three separate times of the year, which complicates the process of defining a “marketing year” for its production. The final consideration with regard to the use of marketing-year data is that trade data is often by default aggregated into calendar years. For most countries, trade data is also simultaneously published on a monthly basis, such that compilers may aggregate the monthly data to fit the marketing year, if they see fit.

#### **Calendar year**

Each of the limitations delineated for the crop year may be considered as points in favour of compiling FBS on the basis of calendar years; indeed, calendar years provide a “neutral” reference period that is consistent across commodities, and are the default reporting periods for trade data. Additionally, calendar years are consistent across countries, which facilitates data comparability. Indeed, FAO's global coverage of FBS is calculated on a calendar-year basis for this very reason. The final advantage of using calendar years for FBS compilation is to facilitate the comparability of FBS data with national accounts data, which are typically compiled on a calendar-year basis (although they may also be compiled on a fiscal-year basis).

The only limitation of using the calendar year as a reference period is that it may be difficult to understand conceptually. Countries should follow this general rule when converting marketing years to calendar years: production should be assigned to the calendar year in which most of the crop will be consumed. This should be a relatively intuitive construct for crops whose marketing years start either late or early in the year. For example, if the marketing year begins in February 2015, then production should be assigned to calendar year 2015. Likewise, if the marketing year begins in December 2015, then production should be assigned to calendar year 2016. The complications occur for crops whose harvests occur toward the middle of the calendar year, particularly from May to August. In these cases, countries should consider both the reliability of their estimates on stocks, and the potential implications for trade of assigning production to one calendar year or another. For example, if a crop is harvested in June 2015, but the supply chain (potentially including drying, processing, storage and aggregating) is such that it is not exported until February 2016, then country-level FBS compilers have two options: they can either assign the production to 2015, when the crop was harvested and carry over the production into the following calendar year using stocks; or, if estimating stocks is problematic, they can instead assign production to calendar year 2016 to

ensure sufficient exportable supplies. In making these determinations, it would also be beneficial for country-level FBS compilers to consult industry experts for their opinions.

### **Fiscal year**

The primary motivation for compiling FBS based on the fiscal year is if the country also collects agricultural production and other data on the basis of a fiscal year (as is the case in Nepal, for example). Because fiscal years are periods of time defined by governments for accounting purposes, they present some of the same advantages of the calendar year: they provide a neutral reference period (which is particularly helpful in countries where multiple crops are harvested), and the data are compatible with national accounts compiled on a fiscal year basis. At the same time, fiscal-year FBS may be difficult to understand conceptually (depending upon when the bulk of the crop is harvested and traded, as above). An additional drawback to fiscal-year FBS compilation is that fiscal years vary by country, such that it is difficult to compare fiscal-year FBS across countries.

## **3.4.2 Data quality, flags and tolerance intervals**

By definition, FBS are analytical data sets, comprised of data extracted from a variety of different sources. This is of course the nature of the work; however, compilers should also recognize that data taken from different sources are likely to vary in terms of quality. For example, official sources (such as government agencies) are likely to have a publicly available methodology detailing how certain data was collected. Other sources may be less transparent about how their data was derived, and their processes may not be subject to the same rigorous standards. For this reason, there is a preferred hierarchy of data sources: official data, followed by semi-official data, data imputation, and data from expert estimation. Each of these categories is examined in further detail below.

### **3.4.2.1 Hierarchy of data sources**

#### **Official data**

For each variable in the framework, official data sources are always preferred for expected values, as it is assumed that such data are most likely to have been compiled according to sound statistical methodologies. However, for many countries, there is no single “official” estimate, as multiple agencies may publish data relating to agricultural output. For example, if an NSO publishes an estimate of a country’s total maize production, and the ministry of agriculture releases a different estimate of maize production, confusion will arise on the part of both the FBS compiler and external users as to which estimate is the *real* “official” production estimate. To avoid this confusion, it is suggested that the process of compiling FBS be preceded by the reconciliation of estimates between different official sources. Optimally, only one agency would be tasked with producing an “official” estimate. Where this may not be possible for whatever reason, it is recommended that different sources explore the methodological reasons for the divergence in the estimates and agree on which number is most appropriate for the purposes of compiling FBS. These guidelines do not venture to recommend one source over another, as the situation inside individual countries may make the estimates of one agency more appropriate for FBS compilation than another. However, it is recommended that the NSO coordinate data reconciliation activities as part of an integrated national statistical framework.

#### **Semi-official data**

Where official data are not available (particularly for utilization variables), it may become necessary to consult and consider alternative sources. These sources – deemed “semi-official” – may include industry groups, trade publications and investigations conducted by product supply chain experts. In ideal circumstances, the expertise uncovered in a search for alternative sources may help to inform or reform processes for collecting official data for the variable in question, as semi-official sources may not collect data with the same regularity as an official source.

### Imputation

When no official or semi-official sources are available for the data in question, the next alternative is the model-based imputation of missing data. It should be noted that in most cases, data imputation relies on a historical data series, such that the quality of imputed data will be highly dependent upon the quality of the source data – often referred to as the “garbage-in, garbage-out” phenomenon. Because the respective variables are influenced by different underlying factors, separate imputation approaches are suggested for all variables in the balance sheet. As such, these guidelines have sought to detail how country-level compilers may approach imputation for each variable; in many cases, this occurs by offering more than one imputation approach per variable. At the same time, users should keep in mind that other methodologies may be more appropriate for certain products in their home countries.

### Expert estimation

The lowest quality level of source data is that derived through expert estimation. Expert estimation is different from imputation in that it relies not on a model, but rather on expert judgment. Because this approach is not based on statistical methodology and is not replicable, the error on values estimated by experts is assumed to be high.

#### 3.4.2.2 Application of flags to denote data source

In the compilation of FBS, using data from a variety of sources is very common. However, as may be seen in the discussion above on the quality hierarchy of data sources, not all data in the FBS should be assumed to be of equal quality. For this reason, it is recommended that countries not only keep track of the data sources used, but also publish a flag denoting the data source alongside the final estimates. This allows users to immediately know the source of the data, and to understand that some data points may be more reliable than others. Additionally, the flags can be used to assist compilers in the process of assigning a priori tolerance intervals for the purposes of the balancing process (detailed below in section 3.4.2.3). Table 3 1 lists some sample flags that can be utilized to denote the data source.

**TABLE 3-1: SAMPLE FLAGS DENOTING SOURCE.**

Source	Flag
Official	
Semi-official	T
Imputed	I
Expert estimation	E

Country-level compilers can expand on the list of flags suggested here if they consider that more detail is necessary to properly communicate necessary information to users. In fact, a list of additional suggested flags, as well as international guidelines on observation dissemination flags, are available from the Statistical Data and Metadata eXchange (SDMX)<sup>10</sup>. However, two additional recommendations should be made. First, compilers should strive to use codes contained within the SDMX list, as it is both comprehensive (that is, it should cover all possible data contingencies) and internationally recognized (such that data flags would be understood by users from different countries). Second, compilers should strive to keep their data flags listing as short as possible, for reasons of clarity and data manageability. Compilers should also keep in mind that additional information on the data sources and methodologies employed can be documented in accompanying metadata (for additional information on metadata, see section 5.3.3).

### **3.4.2.3 Confidence and tolerance intervals**

As noted above, it is crucial to the balancing process that each point estimate also be accompanied by an estimate of confidence, as signified by the tolerance interval. Estimates that come from surveys are likely to be published with their measurement errors, which country-level FBS compilers should note and use in the balancing phase as the tolerance interval. For estimates that are not published with a measurement error, it is necessary for the balancing phase to assign an a priori tolerance interval that denotes the perceived quality of the estimate: the highest-quality data can be assumed to have the highest confidence and the lowest tolerance interval, while data of lower quality can be assumed to have a lower confidence and higher tolerance interval. Because quantities for different commodities in the SUA table will be standardized and aggregated, the tolerance intervals should be assigned by variable. At the same time, the sources of the data should influence the a priori tolerance interval value assigned to each variable, with the lowest tolerance intervals assigned to those variables for which official data are most likely. Country-level FBS compilers are encouraged to thoroughly examine their own supply chains to understand the dynamics at play; however, the rationale behind the likely tolerance intervals for each variable of the balance is laid out in the following subsections.

#### **Production**

At least for the main commodities, most countries will be measuring production through agricultural surveys. For this reason, there should be high confidence in the production estimate. Whether this confidence is 100 percent or slightly lower is for the judgement of the FBS compilers, based on an assessment of the data collection processes under which the production estimate was derived. For example, if the compilers wish for the “official” estimate to appear in the final balance, they should assign a tolerance interval of 0 percent. If the estimates come directly from surveys published with their own confidence intervals, compilers should feel free to use that data in this process.

#### **Trade**

As with production, most countries should have official data on imports and exports published by the relevant agency (customs, ministry of trade, etc.). However, it may be the case that sizeable quantities of cross-border flows, for whatever reason, are not included in official trade data (more information on this possibility is provided in section 3.5.2); therefore, compilers may not have 100 percent confidence in the trade estimates, and may instead assign trade a lower degree of confidence and a higher tolerance interval.

#### **Stocks**

Stocks are kept to smooth consumption levels between harvests; therefore, by their very nature, they may fluctuate wildly from year to year. At the same time, most countries do not measure stock levels, such that most estimates

---

10 See SDMX, 2015, *Guidelines for the Creation and Management of SDMX Cross-Domain Code Lists: Version 2.0*, available at [http://sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_CDCL.doc](http://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_CDCL.doc) (accessed on 26 April 2017); and SDMX, 2016, *Code List for Observation Status*, available at [https://sdmx.org/wp-content/uploads/CL\\_OBS\\_STATUS\\_v2\\_1.docx](https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_1.docx) (accessed 26 April 2017).

on stocks are already based on a form of expert judgement. Given these factors, confidence in most stock level and stock change estimates is likely to be lower.

### **Food availability**

Particularly for staple foods, consumption levels are unlikely to fluctuate greatly from one year to the next, as even in the face of higher prices, households are more likely to cut spending on other goods rather than substantially change their food consumption patterns. For this reason, although food availability is not typically measured by countries, it is also likely to be fairly stable over time. Thus, although some flexibility in food availability may be necessary for the purposes of balancing the supply = utilization identity, the confidence in the food estimate should be relatively high – perhaps not at the levels estimated for production and trade, but nevertheless higher than the confidence in the other variables of the balance sheet.

### **Food processing**

In most cases, food processing is dropped from the balance before the balancing process to avoid double-counting. As such, in most cases it will not have to be assigned a tolerance interval for the balancing process. Where it is not dropped, its tolerance interval may vary, depending upon how production quantities of derived commodities have been estimated. Recall that food processing quantities are related to production quantities of derived commodities, through extraction rates. When there is high confidence in derived commodity production estimates, there should also be high confidence in the food processing estimate. At this extreme, when the official production data of derived commodities is available, the tolerance interval for food processing should be 0 percent. Alternatively, if there is no information on the production quantities of derived commodities, countries may wish to assign a higher tolerance interval to this variable.

### **Feed**

Depending upon how the feed estimate is derived, it may have a larger or smaller implied tolerance interval. This is for two reasons. First, most countries do not measure feed use, implying that confidence in the point estimate is low to begin with. Second, although the total feed demand for a country in terms of nutrient value should be a hard number (if an accurate livestock count is available), the substitutability of feedstuffs ensures that for each individual commodity, the amount of feed might vary greatly from year to year depending upon relative pricing. For this reason, the tolerance interval of the feed estimate may be quite high for certain countries. For other countries where good data is available from feed industries and there is high confidence in data on estimated numbers of livestock, there may be high confidence in livestock data, implying a lower tolerance interval.

### **Seed**

The quantities of seed needed for the following year are solely a function of planted area and seeding rates<sup>11</sup>. Although the planted area may fluctuate depending upon crop prices, the seeding rate should remain stable. For this reason, even if estimates of seed use are imputed, if there is a solid estimate of planted area in the following year, seed use estimates should be fairly rigid to maintain a plausible seeding rate. However, if no estimate of area is available, then seed use may have a lower confidence.

### **Tourist food**

Estimates of tourist food are somewhat flexible. This is because the proposed imputation methodology assumes a level of food consumption by tourists that is a rough approximation, not being based on any measurements. As such, the confidence in this variable should most likely be lower.

---

<sup>11</sup> It should be noted that seeding rates may vary according to production system even within a country. In these cases, an average seeding rate should be used, taking into account the prevalence of the various production systems.

### **Industrial use**

In most cases, the confidence in industrial use estimates will be fairly low. This is because only limited data is available on the industrial uses of commodities, and most of these estimates are limited to biofuel usage. In reality, there are many more industrial uses of commodities (including starches, paints and cosmetics); however, measurements or estimates of the quantities of commodities used in these processes is scarce. For this reason, the tolerance interval of these estimates or imputations is likely to be higher.

### **Loss**

Despite the push for better data on loss in the context of the SDGs, for most countries and commodities, data on loss is very limited. In addition, the quantity of loss may not be consistent from year to year, due to crop size, constraints in storage, transportation and infrastructure constraints, weather and dozens of other factors. For this reason, as with industrial use, confidence in these estimates is likely to be low.

### **Sample confidence and tolerance intervals given a priori knowledge of variables**

Based on the previous considerations, an a priori assignment of sample confidence and tolerance intervals by variable may produce a table such as table 3-2 below. Users should keep in mind that the values in table 3 2 should be based on a discussion of the quality of data within the country compiling the balance sheets: they should not feel bound to use the values suggested here<sup>12</sup>. However, it is advised that compilers fix at least one element in the supply = utilization identity (likely production and/or trade).

---

12 Indeed, the astute compiler will recognize that these sample tolerance intervals are not those utilized in the step-by-step example laid out in chapter 4.

**TABLE 3-2. EXAMPLES OF TOLERANCE INTERVALS BY VARIABLE.**

Variable	Confidence	Tolerance interval
Production	1.0	± 0%
Trade	1.0	± 0%
Stocks	0.75	± 25%
Food	0.90	± 10%
Food processing	1.0	± 0%
Feed	0.75	± 25%
Seed	0.90	± 10%
Tourist food	0.75	± 25%
Industrial use	0.75	± 25%
Loss	0.75	± 25%

### 3.4.3 Data search and assessment

Keeping in mind the considerations for data comparability and data quality laid out in the previous sections, the first step in compiling balance sheets is to search for all possible data sources of the variables described in section 2.2.1. Compilers are encouraged to start by inventorying official data sources, and then reach out to commodity or supply chain experts (in both the public and private sector) for their input on alternative available data sources (the suggested data sources will depend upon the variable in question, and are elaborated below in section 3.5).

Compilers should assess each data source identified for both data comparability and data quality. For all sources, compilers are encouraged to note the release date or frequency with which the data is produced, the product classification system used, the unit of the data, the reference period, and the data quality or flag. To the extent possible, this assessment should be documented, to ensure both transparency and institutional memory. A sample data assessment grid is provided below in table 3-3. It should be further noted that many of the data parameters documented in the data assessment grid will be helpful in ensuring that the underlying metadata of the final FBS product is complete. FBS compilers should be aware that as data sources may, in some cases, be specific to certain commodities, it may be more appropriate to either fill out a separate assessment grid for that commodity, or to note the relevant commodity for the data source within the general assessment grid.

**TABLE 3-3. SAMPLE DATA ASSESSMENT GRID.**

Variables	Sources	Release date/ frequency	Classification	Unit	Reference period	Quality/flag
Production						
Trade						
Stocks						
Food						
Food processing						
Feed						
Seed						
Tourist food						
Industrial use						
Loss						
Additional parameters						

## 3.5 SUGGESTED DATA SOURCES AND IMPUTATION

During the course of the data inventory and assessment, compilers will have to consult a variety of sources that may offer information on each of the variables in the FBS. Additionally, when a data source cannot be found, missing values must be imputed. The guidance on both potential data sources and suggested imputation methodologies, however, will differ based on the variable in question. For this reason, the following section explores FBS data input needs on a variable-by-variable basis.

### 3.5.1 Production

Data on agricultural production is one of the foundations of the FBS framework. Arguably, it is so important that countries not currently collecting agricultural production data (at least for the major food crops) should first consider investing their resources in generating reliable data on production before endeavouring to compile FBS.

At the same time, it is likely that even countries with highly developed official survey methodologies may not collect production data on every single commodity, with minor crops often overlooked in annual surveys because including them may be cost-prohibitive<sup>13</sup>. For this reason, some suggestions on alternative data sources and imputation strategies are also included.

Before proceeding to the discussion on data sources, one point touched upon in the previous section bears repeating. For crops where harvests straddle two calendar years, compilers should assign production to the calendar year during which most of the crop will be consumed (for a more extensive treatment of this topic, see section 3.4.1.3).

#### 3.5.1.1 Official data sources

The preferred source of data on agricultural production – including the production of crops, livestock and derived goods – is survey-based official data. It is highly recommended that, at the very least, countries conduct annual production surveys for major commodities, and endeavour to measure all commodities in less frequent agricultural censuses or structural surveys. Additionally, it is recommended that official sources collect not only information on production output, but also on activity (sown area, harvested area and number of animals) and productivity (crop yield per unit of harvested area, milk yield per milking animal, meat yield per animal slaughtered, etc.) variables. This information is helpful for two main reasons. First, it is useful for validating production data: higher crop production, for example, results only from some combination of increased cropping area or increased yield. Therefore, checking the feasibility of growth in area and yield may serve as a rudimentary but accessible check on the feasibility of the overall production estimate. The second reason for measuring activity and productivity variables is to assist in the imputation of missing data in future years or in years when surveys do not take place.

Beyond surveys, administrative data may be another potential source of data for certain products. The most common example of this is data from slaughterhouses, which may be required to keep records on the numbers of animals slaughtered and whether the carcasses have been inspected. Industrial output surveys may also be useful sources for data on the production of derived products, such as flour or beer.

---

13 However, ideally, these minor products are still captured in periodic agricultural censuses or structural surveys.

Specific guidance on improving the collection of official data on either crop or livestock production is beyond the scope of these guidelines. However, the Global Strategy has produced several publications that countries may find useful in their efforts to improve production data for both crops and livestock<sup>14</sup>.

**3.5.1.2 Alternative data sources**

Country-level FBS compilers can consult two additional potential data sources in their search for production data: records of private firms, and commodity organizations.

Particularly in cases where crop production is concentrated in a small area, or where production is delivered to a handful of firms for further processing, an aggregated production figure for the country as a whole can be derived from simply adding up the individual purchase records of said companies. These records may be accessible through tax authorities or through an agreement with an industry or commodity organization.

Even when records from firms are not available, direct production estimates from commodity organizations could also prove useful if their members represent almost all production. They can be a particularly helpful source for information on the production of minor or cash crops. Some of these commodity organizations are international in scope, and publish data on a variety of countries. Examples include the International Coffee Council, the International Cotton Advisory Committee, the International Sugar Organization, and Oil World. In other cases, commodity or industry groups will be specific only to the country in question, and may even only specialize in a certain production area.

**3.5.1.3 Imputation and estimation**

The suggested imputation strategy for missing production data at the country level depends, to a certain extent, on the commodity for which production is to be estimated, with different approaches to imputation suggested for crops, processed products derived from crops, and livestock-derived products. Individual strategies for these three groups are laid out below:

**Crop production imputation**

When estimating production of crops, imputation is based upon the following identity:

$Production (MT) = Yield \left( \frac{MT}{HA} \right) * Harvested Area (HA)$	(3-1)
--	-------

Note: MT = metric tonnes

<sup>14</sup> For guidelines on how to collect data on agriculture by including production modules in other household surveys, see Global Strategy (2015a). For information on estimation of crop production, see Sud et al. (2016). For further information on estimation of livestock production, see Moss et al. (2016).

As noted above, it is highly recommended that country-level crop production series publish estimates not only for production, but also for yield and harvested area. This is done partly as a data validation mechanism, but also to ease crop forecasting, given that yields for most crops follow a definite positive trend over time. It should be noted that in the agricultural survey program of many countries, data is collected on sown area, but not on harvested area. While data on harvested area is preferred for the purposes of imputing missing production data, sown area estimates can also be adapted and used for this purpose.

As should be evident from equation (3-1), calculating any of the three unknowns (production, yield or area) requires only an estimate of the other two terms. Thus, for production, the suggested imputation approach is a three-step procedure:

**Step 1: Measure, impute or approximate a yield estimate.**

**Step 2: Measure, impute or approximate an estimate of harvested area.**

**Step 3: Multiply yield by harvested area estimates to obtain a production estimate.**

Each of these steps is detailed below.

### **Step 1: Yield estimate**

Yields are sometimes measured directly by government agencies (from objective yield surveys, for example), and if such measurements have been made on a representative basis, then these can serve as an overall yield estimate. Alternatively, if production and harvested area have been measured, yields can simply be calculated by rearranging equation (3-1) as follows:

$$Yield \left( \frac{MT}{HA} \right) = \frac{Production (MT)}{Harvested Area (HA)} \quad (3-2)$$

Note: MT = metric tonnes

If governments do not directly measure yield or production, analysts should also consult any other available sources that may be collecting or estimating quantitative information on yields. For example, early warning system reports (such as the Country Briefs produced by the FAO Global Information and Early Warning System (GIEWS) group)<sup>15</sup> may either contain information on yields or production and area (such that an estimate of crop yield may be derived).

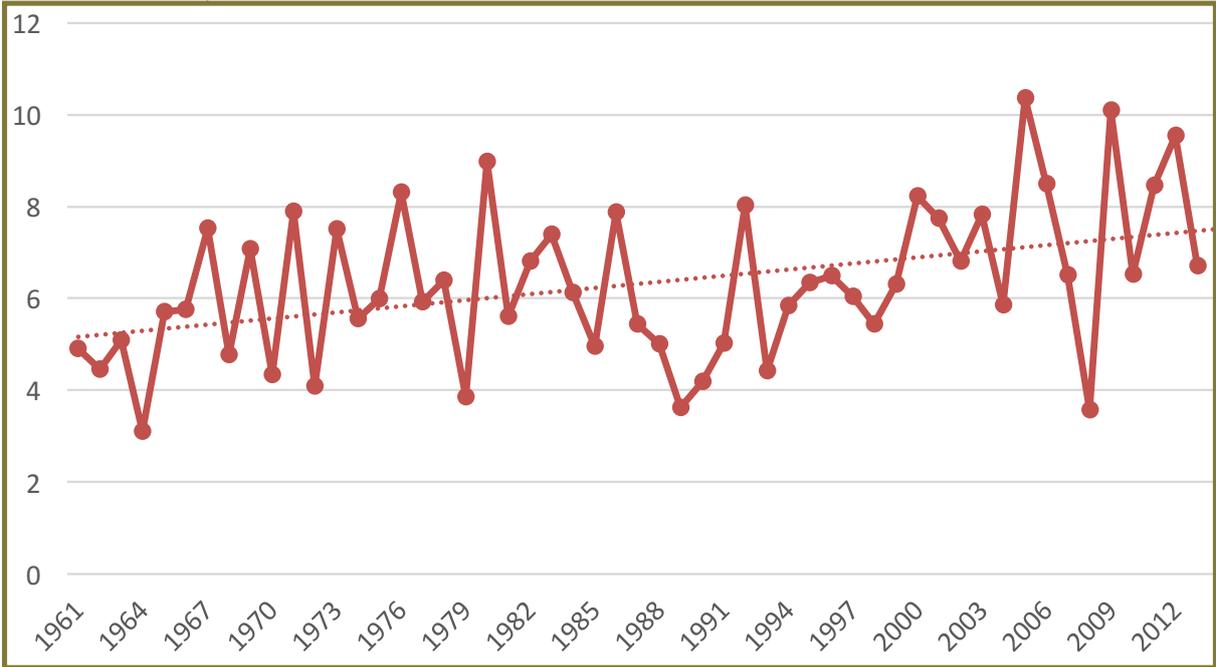
If yields are not measured, imputation using time series estimation is recommended. This approach is preferable because although crop yields in any given year may depend upon a host of factors (including temperature, rainfall, pests, diseases and production management), over time, yields tend to follow trends. Three principles should guide this estimation process:

#### **Principle 1: Understand the nature of yields for the crop being modelled**

Any yield modelling exercise should be preceded by a graphing of historical yields and general research into the typical yield characteristics of the commodity in question. For example, many fruit and nut trees follow alternate bearing cycles, such that although the general trend is towards higher yields over time, a year of high yields will often be followed by a year of low yields. Avocado yields in the United States of America behave in this manner (figure 3 1): although the linear trend of avocado yields over time (the dotted line) has been positive, the tendency of yields to fluctuate from year to year is evident. Failure to take into account these types of idiosyncrasies may lead to highly erroneous yield estimations.

<sup>15</sup> For more information, see the GIEWS webpage, available at: <http://www.fao.org/giews/en/>. Accessed on 10 June 2017.

**FIGURE 3-1. AVOCADO YIELDS IN THE UNITED STATES OF AMERICA, 1961–2013.**



Source: FAOSTAT.

**Principle 2: Use the appropriate functional form**

The graphing of historical yield data should be followed by an analysis to determine which functional form best fits the data. Once this form is known, yields in missing years should be estimated using this type of function.

**Principle 3: Include other relevant explanatory variables in the estimating regressions**

As noted above, several factors may influence yields for a given commodity in a given year. Nevertheless, in their estimations, country-level analysts should include some primary variables that have been shown to have a good explanatory relationship with the final crop yield. These may include, for example, average diurnal temperature, planting date, the occurrence of natural disasters, or the amount of rainfall during grain filling<sup>16</sup>.

Aside from following these three yield modelling principles, analysts should also take into account any relevant qualitative information on yields. For example, if it is known that a substantial pest infestation reduced output in the current year compared to the previous year, then estimated yields should reflect this decline, even if available imputation modules fail to capture a reduction.

Once country-level analysts have measured or imputed a crop yield, they can move to Step 2.

<sup>16</sup> It is likely that, for whatever crop is being considered, a simple Internet search will yield various scholarly articles related to yield estimation or projection that can help analysts to pinpoint the explanatory factors upon which they should focus.

## Step 2: Harvested area

If an estimate of harvested area has been established from surveys, then FBS compilers can use that estimate, along with the yield estimate discussed in step 1, and proceed to step 3.

If no harvested area estimate has been derived from surveys, then country-level compilers must estimate a likely harvested area. If some form of survey of sown area has taken place, then estimating harvested area should be straightforward, as harvested area is by definition the quantity of sown area minus any land that has not been harvested (also referred to as “abandoned” area). In this case, all that is needed to calculate a harvested area is the estimate of sown area, and an estimate of the percentage of land that was abandoned (represented by *abd*), as below in equation (3-3).

$$\text{Harvested area}_t = (1 - abd)\text{Sown area}_t \quad (3-3)$$

It may also be the case that rather than estimating a percentage of abandoned area, countries may possess information as to the actual area of land abandoned (most likely due to a particular phenomenon, such as a hurricane, flood or pest infestation). In this case, the harvested area may be estimated directly simply by subtracting the quantity of land abandoned from the quantity of area sown, as in equation (3-4) below.

$$\text{Harvested area}_t = \text{Sown area}_t - \text{Abandoned area}_t \quad (3-4)$$

If an abandonment rate or a quantity of abandoned area is unknown, and the sown area is known, FBS compilers may – as a last resort – use sown area to proxy for harvested area. They should note, however, that this approach will lead to overinflated production estimates, since crop yields are by definition estimated with respect to harvested area.

If sown area is not known, then FBS compilers must devise an alternative strategy for estimating harvested area. This may involve either first estimating sown area (and then deriving harvested area using either equation (3-3) or (3-4)), or, otherwise, estimating harvested area directly. It is recommended to first estimate sown area and then derive harvested area, as sown area is more likely to be correlated to other observable data, and can thus be more easily imputed. Sown area is commonly modelled as a function of either the previous year’s sown area or of farmer price expectations at the time of planting (or as a combination of the two). Country-level FBS compilers are encouraged to consult the relevant literature for the commodity in question to formalize their strategy for estimating sown area.

If modelling sown area does not appear to be feasible, the last approach that can be adopted is that of basing an area estimate on qualitative reports (as discussed above for yields). For example, early warning system reports often publish estimates of harvested or sown area, either in absolute terms or relative to the previous year. These reports can then be used to estimate harvested area.

Once compilers have an estimate of the harvested area (using any of the strategies elaborated above), they can proceed to step 3.

**Step 3: Derive production estimate by multiplying estimates for harvested area and yield**

With estimates of both harvested area and yield in hand, FBS compilers need only multiply the two together using equation (3-1). Compilers are reminded that in cases where either the yield or area are imputed or estimated, the quality flag assigned to the production estimate should reflect the quality of the yield and harvested area used to derive the production number.

**Processed products derived from crops**

In the context of the FBS, the “production” of processed products is directly linked to the amount of a primary, secondary or even tertiary good allocated to food processing (covered below in section 3.5.5). As such, imputation of the derived product production differs from imputation of primary product production, in that the only two pieces of information necessary for imputing production quantities for derived goods are (1) the amount of the primary good that is being processed (that is, quantities of primary goods assigned to the food processing variable); and (2) the extraction rate. For most products, the extraction rates will fluctuate very little over time, such that assuming fixed extraction rates is a reasonable approach. However, estimating the quantity of a given primary commodity destined for processing may be slightly more complicated, and is likely to require input from a panel of experts. In addition, if multiple derived goods stem from the same primary commodity, analysts will have to make assumptions on the share of processed use of the primary commodity that is being diverted into producing each of the derived goods. A sample exercise along these lines is provided in box 2-1, section 2.3.2.

**Livestock and livestock product imputation**

Imputing data for livestock, derived livestock products such as meat, and live animal products (such as dairy and honey) should follow a slightly different rationale than the imputation of missing crop production data. In this process, the objective of FBS compilers should be to synchronize production of all derived products, using livestock commodity trees as their guide. This is achieved by working backwards from any official data on production of a given derived commodity (by dividing by extraction rates, as in equation (2 8)) to impute values for higher levels of derived products. In this way, analysts can ultimately work backward to the number of functional animal units required to maintain reasonable yields of the product in question (for example, the number of animals slaughtered in the case of meat products, or the number of milking animals in the case of dairy products). In addition, compilers should be mindful of synchronizing production of co-products as outlined in commodity trees, using this synchronization to impute missing production data. For example, if official data is provided for cheese, country-level FBS compilers can use the inverse of known milk-to-cheese conversion factors to calculate a likely value of the amount of milk used to produce said cheese. This same amount can then be used to impute a value for production of the cheese co-product whey.

For missing production data on derived processed animal products (such as meat and skins), the lynchpin of the synchronization process is an estimation of the number of animals slaughtered. Using this estimate of animals slaughtered, and applying the appropriate yield conversion factor for the product in question, estimates for production of meat, offals, fat and hides or skins can be derived as in equation (3 5):

$Production (MT) = Yield \left( \frac{MT}{Animal} \right) * Animals Slaughtered$	(3-5)
--	-------

Note: MT = metric tonnes

If FBS compilers know the number of animals slaughtered, then imputing production of derived goods is simply a matter of applying the appropriate yield factors. If the number of animals slaughtered is not known, and production of at least one derived product is known, FBS compilers should start from that number and work backwards to first derive an estimate of the number of animals slaughtered, and then use that number of animals to impute production of the remaining slaughtered animal products using the yields applied in the previous year.

For example, many countries do produce an official estimate of meat production if the legal framework mandates that all meat must be inspected by a government entity. Using this estimate of meat production, FBS compilers can use the previous year’s carcass weight yields (as carcass weights do not vary greatly over time)<sup>17</sup> to work backwards to derive an estimated number of animals slaughtered, as below in equation (3-6).

$\text{Animals Slaughtered} = \frac{\text{Production (MT)}}{\text{Carcass Yield} \left( \frac{\text{MT}}{\text{Animal}} \right)}$	(3-6)
---	-------

Note: MT = metric tonnes

FBS compilers should note that carcass yields for various animals may be expressed in smaller units, although the overall national production estimate will likely be expressed in tonnes, such that additional unit conversions may be required. For example, carcass weight yields of chickens are likely to be expressed in grams or decigrams. In these cases, a production value can be calculated by either (1) first multiplying the carcass yield expressed in grams by the number of animals slaughtered, and then converting that number from grams to tonnes; or (2) first converting the carcass yield from grams per animal to tonne per animal, and then multiplying that value by the number of animals slaughtered.

All of these cases assume that some official data on livestock or livestock products exists and is complete. However, the reality for many countries is that while data may be available for officially registered or slaughtered animals, a substantial portion of livestock and production of livestock-derived goods may not be registered through official channels. In these cases, FBS compilers are advised to combine official data with an estimate of non-registered animals or production of livestock-derived goods outside official channels to derive a “total” figure from which the imputations described above can be carried out.

**3.5.2 Trade (imports and exports)**

For a given product, data on international trade in agricultural goods includes three dimensions: quantity (typically expressed in tonnes), value (expressed either in local currency or in US\$), and a unit value (the quantity divided by the value). While unit values may not be directly reported in international trade data, they are easily calculated by the identity noted above, and can be used to check the consistency of trade data across time – that is, while import quantities and import values may fluctuate greatly, unit values are more likely to remain at similar levels (or at least at the same order of magnitude) year after year.

Of all the variables involved in FBS compilation, data for the international trade variables of imports and exports are the most likely to be reported reliably by official sources (usually, the national customs office). This is because most countries mandate the collection of data on all cross-border goods transactions for tax purposes (as well as to comply with World Trade Organization – WTO – and WCO guidelines), most commonly through customs declarations.

<sup>17</sup> Compilers should note that, when domestic animal carcass weights differ substantially from imported animal weights, it may be preferable to use a weighted average carcass yield that takes these differences into account.

At the same time, particularly in the trade of agricultural goods, official reported trade flows may not encompass all cross-border transactions. First, food aid transactions are sometimes excluded from official trade flows, despite potentially being a source of a significant proportion of a country's food supply. Additionally, in some countries, agricultural goods are traded outside formal customs procedures. These trade flows, referred to here as unrecorded trade<sup>18</sup>, can be important contributors to both household income and localized food security (Afrika and Ajumbo, 2012). In particular, given the potential contribution of unrecorded trade to food security, it is vital that, where relevant, these transactions are accounted for in an FBS setting. These flows may be particularly important for the accurate estimation of livestock populations, especially for countries with large nomadic herder populations who may frequently cross national borders with their herds.

For these reasons, while official trade data is likely to be available in most cases, it may have to be supplemented with data from other sources to provide more accurate aggregate import and export estimates in an FBS setting. In these cases, several additional data resources are available. An additional valuable trade data cross-checking tool is the use of mirror statistics (trading partner data).

### 3.5.2.1 Official data sources

As described above, most of the world's countries collect official data on both imports and exports of goods via customs declarations. Customs declarations may require a wide variety of information to be provided on a given cross-border transaction; however, they must include a commodity code for the product to aid in classification (almost always an HS-based code), as well as the weight of the shipment<sup>19</sup>. Aggregated data from customs declarations for use in national FBS compilation can typically be accessed directly from the national administrative body charged with reporting on trade data, which may be the national customs office, the ministry of trade or NSO. Although other international data sources exist (covered below in section 3.5.2.2), using data sourced from domestic agencies will in some cases allow countries to produce more timely estimates, as this data tends to be updated and disseminated frequently.

Aside from customs declarations, country-level FBS compilers may wish to consult additional official administrative records if potential issues with official customs data are identified. Other sources that can be consulted include shipping manifests, ship registers, port administration reports and enterprise surveys.

In some cases, countries also produce official data estimating otherwise unrecorded trade flows. Uganda, for example, conducts an annual "Informal Cross Border Trade (ICBT)" survey to collect information on the scope of these trade flows<sup>20</sup>. The country's 2014 survey suggested that ICBT exports were nearly one-fifth the size of formal exports, and agricultural goods represented almost half of those unreported trade flows. These cases underscore the importance of monitoring or surveying unrecorded trade flows for statistical purposes in countries where these flows are recognized as occurring. For the purposes of compiling an FBS, it is crucial to combine an estimate of unrecorded trade flows with the official trade data to obtain a more realistic total trade estimate.

---

18 Depending upon the source publication, various terms are applied to these trade flows, including "informal cross-border trade," "informal trade," "grey trade," or "shadow trade."

19 For more information on the types of information typically included in a customs declaration, see UNSD (2004).

20 See Bank of Uganda, 2017, Informal Cross Border Trade (ICBT) Survey, available at: [https://www.bou.or.ug/bou/publications\\_research/icbt.html](https://www.bou.or.ug/bou/publications_research/icbt.html). Accessed on 19 January 2017.

### 3.5.2.2 Alternative data sources

Various additional trade data resources are available to FBS compilers who, for some reason, cannot access national customs data, or for compilers who wish to cross-check the data or consult other sources: international trade databases that cover mostly formal trade flows; mirror data from trading partners accessed through international trade databases; resources for estimating unrecorded trade; and data sources that document food aid shipments. Each of these is described below.

#### International trade databases

For the majority of the world's countries, official customs data is forwarded to the UNSD, where it is checked, organized and published in the UN's signature trade database, Comtrade<sup>21</sup>. Comtrade is a comprehensive database that publishes trade statistics by year, reporting country, partner country, trade flow type (import, export, re-export) and HS code. Countries may choose to utilize this data in place of national customs data when accessing the latter is difficult. It should be emphasized that the data published by Comtrade are official trade statistics, reported by national statistical authorities. As such, although data from Comtrade are not directly sourced from a national statistical agency, they can still be considered "official data" for the purposes of FBS compilation.

FAO also produces a data set on agricultural trade that FBS compilers may find useful. This dataset is elaborated on the basis of Comtrade data, but undergoes additional validation steps to identify outliers and, if necessary, adjust them. FAO's data set seeks to account for food aid and unrecorded trade flows, thus rendering it a more complete data set for FBS purposes. The final advantage of the FAO data set is that it attempts to fill gaps in Comtrade data by using partner trade flows to help document trade in countries that do not report to Comtrade, or for which data has not been updated in Comtrade. As such, the FAO data set contains imputations and estimations and is therefore not entirely an official data source.

#### "Mirror" data from trading partners

Within the Comtrade and FAO data sets, countries are encouraged to check the trade flows reported by their trading partners – referred to as "mirror" data – when no official trade data are available, or as an additional validation of their own national data. This is partly due to the greater overall reliability of import data. Because countries often tax imports through tariffs, there is a tendency for import data to be more closely and accurately monitored. As most exports are not taxed, those trade flows may not always be adequately captured, even in official data. In these cases, it may make sense for compilers to consult partner data and, if necessary, further investigate the reason behind the discrepancy, or even override their official trade figures. For countries that do not report trade data, consulting trade partner mirror data is necessary to formulate a detailed picture of a country's agricultural trading pattern.

#### Unrecorded trade resources

In countries where unrecorded trade is reported to be substantial, compilers should at least attempt to quantify those values. Some data sources are available for certain regions of the world, and could be helpful in this process. For example, the FEWS NET network produces periodic cross-border trade reports for both East and Southern Africa<sup>22</sup>. These reports typically estimate quantities of cross-border trade by commodity, and document the underlying dynamics driving changes in these trade flows.

---

21 This database is publicly available at <http://comtrade.un.org/>. Accessed on 10 June 2017.

22 These reports can be accessed via the FEWS NET Markets & Trade portal, at <https://www.fews.net/sectors/markets-trade>. Accessed on 19 January 2017.

The Permanent Interstates Committee for Drought Control in the Sahel (known by its French acronym, CILSS, for Comité permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel) also collects data on unrecorded trade flows between eight West African countries for both major cereal grains and livestock<sup>23</sup>. Compilers from countries in that region should consult the CILSS reports on intraregional agricultural trade to ensure the completeness of export and import data for FBS purposes.

### **Food aid data**

For some countries, food aid shipments are not reported in commercial shipment data, which can lead to a drastic underestimation of total imports, with consequences for food availability. If country-level compilers suspect that food aid shipments are excluded from trade data, efforts should be made to add food aid quantities to commercial quantities to obtain an estimate of total trade for FBS compilation purposes.

The World Food Programme (WFP) is the UN agency that oversees most deliveries of food aid. At the time of writing, however, the WFP is adjusting how it reports food aid shipments. Through 2016, the WFP published data on food aid shipments (through calendar year 2014) on their INTERFAIS website<sup>24</sup>. Beginning in 2017, this platform will be sunsetted and WFP will instead report its food aid shipments through the International Aid Transparency Initiative (IATI)<sup>25</sup>. In addition, the FAO GIEWS team will begin to collect and publish data on non-WFP food aid shipments.

Country-level FBS compilers should consult these resources for information on food aid flows. Compilers should be aware, however, that over the past decade there has been a dramatic shift in the way that food aid is delivered to countries: while donor countries previously gave physical quantities of food, today, they are much more likely to donate money, either for local procurement of food aid, or to assist the hungry in purchasing food directly from the market. For this reason, the actual quantity of physical food deliveries has declined over the years. As an example, the quantity of wheat food aid distributed by WFP has declined from more than 7.8 million tonnes in 1988 to just under 1 million tonnes in 2012 (although this decline was most likely the result of a combination of factors, and not solely due to a shift to more cash-based food aid) (WFP, 2017).

### **3.5.2.3 Imputation and estimation**

No methodology is suggested for imputation and estimation of trade data at the country level, as various data sets already exist that should cover the trade data needs of most countries. However, as described above, country-level compilers may wish to adjust official trade data based on unrecorded trade data or on mirror data from trading partners. These options are both detailed in section 3.5.2.2 above.

## **3.5.3 Stocks and stock changes**

The systematic long-term holding of stocks is typically limited to a handful of products that are non-perishable and most likely important to domestic food security needs: mainly grains, but also sugar, pulses and some oilseeds. At the same time, countries may hold short-term stocks of various other products from one marketing year to the next, such as horticultural goods (apples or potatoes), processed horticultural products (frozen concentrated orange juice or canned tomatoes) or processed dairy products (butter or cheese).

23 More information on this program is posted on the CILSS website, at: <http://www.cilss.int/index.php/flux-transfrontalier/>. Accessed on 11 August 2017.

24 The INTERFAIS data is available through this portal: <http://www.wfp.org/fais/>. Accessed on 19 January 2017.

25 The IATI website is available at: <http://www.aidtransparency.net/>. Accessed on 19 January 2017.

Given the effect that stock levels can have on food prices and their strategic use as a safeguard for domestic food security, the accurate measurement of total stockholdings among all actors (at least for the primary food commodities) should be a policy priority for countries. However, currently, data coverage of total stocks estimates is extremely limited. This is partially due to the complexity of measuring stock levels, as these can be held anywhere along the supply chain. For this reason, to accurately collect data on or estimate stock levels, it is strongly recommended that country FBS compilers first assess the stock situation in their country by speaking with industry experts and relevant government officials to determine which commodities are being stocked and how those stocks are organized (including which stakeholders are keeping those stocks, and the size of stockholdings among the different stakeholders). There are some efforts underway to improve data collection on stock levels; however, many country-level FBS compilers may find that stock changes will have to be imputed or estimated.

### 3.5.3.1 Official data sources

Official government agricultural surveys are the preferred mechanism through which to collect data on stock levels, as surveys can target the supply chain actors that are most likely to hold stocks: farm surveys can produce estimates of on-farm stocks, while surveys of processors, manufacturers, exporters or distributors can target stockholding elsewhere in the supply chain. Governments themselves may also be large stockholders of certain food commodities. If countries are able to collect data on stocks held at the farm level, in the private sector and in the public sector, then an overall picture of the country's stocks situation should be mostly complete and provide a solid estimate for FBS compilation purposes.

Because collection of stocks data is so critical to removing sources of error from the balance sheet, it is highly recommended that countries make explicit efforts to measure stock levels of major commodities rather than rely upon an imputation or estimation approach<sup>26</sup>. In this regard, two particular efforts are suggested. The first of these is the addition of a stocks module to periodic agricultural production surveys. This action would greatly improve the availability of data for on-farm stock levels for primary food commodities. The second action is the reporting of government-held stock levels. Particularly in countries where governments hold large inventories of important food staples, the absence of administrative data on stock levels will severely limit the usefulness of compiling FBS to assess overall food supply and demand.

As mentioned above, there are several global efforts underway to improve the measurement of stock levels within a general push for improved information related to agricultural statistics. In fact, the Global Strategy stresses the importance of information on stocks for developing countries by including this variable in the minimum set of core data that should be measured and disseminated annually (FAO et al., 2012). The most prominent efforts in this respect have been led by the global Agricultural Market Information System (AMIS). Subsequent to the Expert Meeting on Stocks Measurement, held in November 2014, AMIS has recently drafted *Guidelines for Designing and Implementing Grain Stock Surveys* (forthcoming), and is also planning various other activities designed to assist countries in improving their stocks measurement<sup>27</sup>. In the same vein, in November 2016, FAO and the Indian Ministry of Agriculture & Farmers Welfare (MAFW) held a joint seminar on Approaches and Methodologies for Private Food Grain Stock Measurement, and all of the presentations on the various approaches are available online as a resource to country-level FBS compilers<sup>28</sup>.

26 In fact, this is the recommendation of AMIS for AMIS member countries. In addition, they recommend that both on-farm stocks and stocks held by commercial actors both be surveyed and combined to produce an overall stocks estimate. For more information, see AMIS (2015).

27 For more information on the content of the discussions held during the Expert Meeting on Stocks Measurement, see AMIS (2015). The AMIS *Guidelines for Designing and Implementing Grain Stock Surveys* will be published in 2017. For more details on the planned program of work related to stocks measurement, see AMIS (2016).

28 Related documents are available at <http://www.fao.org/asiapacific/events/detail-events/en/c/1363/>. Accessed on 10 June 2017.

### 3.5.3.2 Alternative data sources

Outside of official sources, data on stocks is likely to be limited to a single aspect of the supply chain (processors, for example), and thus provide an incomplete picture of a country's total stock levels. For this reason, countries are encouraged to develop strategies to survey overall stock levels in an official capacity, rather than rely on incomplete estimations from one segment of the supply chain. For some supply chains, however, reports of stock levels from processors or industry could account for the majority of stockholdings, and thus be invaluable to estimating total stock levels.

Compilers may also wish to consult the AMIS database, which estimates closing stock levels for maize, wheat, rice and soybeans for over 20 of the world's largest producers and consumers of those commodities<sup>29</sup>. Similarly, estimates on global sugar stocks can be accessed from F.O. Licht, and stocks estimates for numerous oils and fats can be sourced from Oil World<sup>30</sup>.

### 3.5.3.3 Imputation and estimation

After all possible data collection opportunities have been exhausted, country FBS compilers may choose from a number of different approaches to impute or estimate stock changes, subject to some cumulative constraints on stock levels. The choice of approach may vary depending upon the commodity in question.

#### Suggested approach

From a purely mathematical point of view based on the supply = utilization identity, stocks represent the mismatch between supply and utilization in a given year. Because most domestic utilizations tend to change little from one year to another, changes in stock positions tend to be correlated with changes in production net of trade (that is, production plus imports, minus exports). As such, changes in stocks can be modelled as a function of changes in production net of trade, as follows.

$\Delta Stocks_t = f(\Delta ProdNT_t) + \varepsilon_t$	
<p>Where:  <math>\Delta Stocks_t</math> is equivalent to <math>Closing Stocks_t - Closing Stocks_{t-1}</math>,  <math>\Delta ProdNT_t</math> is equivalent to <math>[Production + Imports - Exports]_t - [Production + Imports - Exports]_{t-1}</math>, and  <math>\varepsilon_t</math> is an error term.</p>	(3-7)

FBS compilers can estimate this relationship using regression analysis and choose the functional form that is most appropriate to their situations. Compilers may wish to add additional variables in their regressions, although the basic approach should remain the same. It must be emphasized here that imputation of stock changes through this suggested approach relies on countries having historically measured stock levels for the commodity in question<sup>31</sup>.

29 See the AMIS database, available at <http://statistics.amis-outlook.org/data/index.html#HOME>. Accessed on 19 January 2017.

30 Both the F.O. Licht data and the Oil World data are behind a paywall. Further information is available on their respective websites. For F.O. Licht, see: <https://www.agra-net.com/agra/international-sugar-and-sweetener-report/>; For Oil World, see: <https://www.oilworld.biz/t/publications/data-base>. Both links accessed on 19 January 2017.

31 It is also noted that this approach does not consider prices or price changes, which have been shown to be correlated to stocks (see, for example, the elaboration of the "Supply of Storage" model in Bobenrieth et al. (2004) and a follow-up analysis on the relationship between prices and stocks-to-use ratios in Bobenrieth et al. (2013). In fact, in the past, the USDA utilized the relationship between prices and the stocks-to-use ratio to inform forecasts of domestic season average farm prices – see, for example, Westcott and Hoffman (1999). Knowledge of this relationship could assist countries in estimating stock changes, or could help to validate estimates derived through the approach proposed above.

At the same time, estimates of stock changes derived from regressions must be checked against a constraint for cumulative stock levels. That is, a negative stock change in any given year cannot exceed the previous level of stocks, as violating this constraint would mean that countries are using more stocks than they possess.

The last concept to keep in mind when modelling stock changes is that over several successive years, cumulative stock changes should sum to approximately zero. To understand why this should be the case, consider the following two scenarios: Country A adds to stocks every year, while Country B takes away from stocks every year. In the case of Country A, their constant accumulation of stocks has two primary problems. First, within every country, there are physical limits in storage infrastructure that by definition indicate that a country cannot stock a commodity every year. Second, the large amount of product held in stocks would eventually dominate a country's domestic supply, most likely depressing prices to the point that there would no longer be any incentive to produce more of the good. Thus, at some point, Country A will have to remove some supply of the good from its stocks. In the case of Country B, the limits are much more evident, in line with the constraint outlined above: at some point, there will be no more of the good in stock available for removal. This illustration should help country FBS compilers to recognize that, over a given period of time, cumulative changes to stocks should sum to zero.

### **Alternative approach**

In the absence of historical data on stock levels for grains, pulses, sugar and oilseeds, compilers can preliminarily use stocks to “balance” the supply and demand equation; however, this approach should only be utilized when there exists some measured data to derive estimates for food and any other relevant utilizations. Otherwise, compilers are dealing with an equation with two unknowns, and it is not possible to properly account for error. Even using this approach, compilers must check cumulative changes against a running estimate of stock levels to ensure that the estimated changes are feasible.

For some perishable products, stock changes can be used to smooth supply fluctuations from year-to-year. In these cases, compilers should be aware that the stocks accumulated in one year should, in most cases, be entirely or almost entirely used in the following year. Compilers should also consider adjusting loss to account for any stocks that have not been allocated to some other utilization in the following year. However, before this approach is followed, compilers should have a solid understanding of the supply chain for the respective product – in particular, whether it is feasible to hold stocks of that good through the following year, and if so, the feasible quantity of stocks.

### **3.5.4 Food availability**

As mentioned in section 2.2.1 above, “food availability” as defined in the FBS setting refers to the quantities of food available for human consumption at the retail level by the country's resident population. This resident population should include refugees and long-term guest workers, and exclude tourists or temporary visitors. Food availability also includes any loss or waste at the retail or consumer level. For this reason, total food availability estimates derived from the FBS are likely to be higher than actual average food consumption.

Directly measured data on food availability (as defined in the FBS setting) may be difficult to obtain. However, FBS compilers can derive estimates of food availability by making certain adjustments to other existing data sets that measure food production or consumption. The key to this process is understanding exactly how the measured quantities differ from FBS definitions, and ensuring that each one of these differences is accounted for in the adjustment process.

While adjusting certain underlying data to ensure its consistency with the FBS definition of food availability is the preferred course of action, some countries may instead choose to impute or estimate food use values. This process is facilitated by the fact that food availability will likely vary little from year to year – particularly for staple foods

that comprise the bulk of consumer diets – because countries are much more likely to see changes in trade or stock levels to maintain consumption of staple goods at fairly consistent levels.

#### 3.5.4.1 Official data sources

Two primary types of official data sources may provide information useful to the estimation of a country's food availability: industrial output surveys and household consumption or expenditure surveys. Both of these sources include certain caveats, however, that compilers should take note of when assessing the value of such data for FBS purposes.

##### **Industrial output surveys**

The first potential source of data is industrial output surveys from food processors, including flour mills, oilseed crushers, dairy processors or breweries. These data are useful for food estimates because they represent so-called “bottleneck” industries, through which all quantities of the primary commodity that will be used as food must first pass before they become edible. This phenomenon can be illustrated by the wheat industry. For many countries, wheat in its primary form is mostly fed to livestock, and for the bulk of the quantity of wheat consumed by people, it must first undergo a transformation to become wheat flour (although this may not be the case in all countries). In addition, wheat flour is not consumed by animals. Therefore, all production of flour (after accounting for net trade) is likely to be consumed as food.

While data from industrial output surveys can be useful for deriving food use estimates, FBS compilers should keep the following in mind when using these data:

- Data must represent a large proportion of total production. For this reason, these sources are only useful in countries where **most processed food production occurs at the industrial level and not at the farm or artisanal level**. As such, they are likely to be most applicable to either developed countries, or to developing countries with more industrialized food processing sectors and only very limited artisanal manufacturing. This point bears repeating: in countries where home processing is common, using industrial output surveys to estimate production of derived goods will result in an underestimation of the production quantity for the derived good in question.
- These data sources will only be available to facilitate the estimation of foods that are processed. This leaves out several commodity segments, principally fresh fruits and vegetables.
- In some cases, estimates of industrial output for food manufacturers may only be available in value terms. Quantities may be determined by dividing these values by current prices.
- Output data for food processors are technically production quantities for those SUA-level items. Therefore, in order to use this data to obtain an estimate of food availability at the SUA level, other uses (imports, exports, stock changes and tourist food) must first be netted out.

These observations aside, the advantages of using manufacturing output is that these data will cover processing use for all consumption occurring within a country, including food away from home and institutional consumption (including in schools, hospitals, jails or military installations).

### Household surveys

The second source of useful data on food availability is household surveys. While such surveys do provide a detailed portrait of consumption at the household level, food consumed outside the home may not be fully captured. As such, household surveys can, in most cases, be considered to provide a conceptual “lower bound” for food availability. Using consumption figures derived solely from household surveys is likely to underestimate total food availability within a country – potentially by large margins in countries where a large portion of calories are consumed away from home and are therefore not accounted for in the survey. At the same time, the trends in consumption evident in household surveys should also manifest themselves in overall FBS food availability levels. In fact, previous work has found that calorie estimates can vary widely between household surveys and FBS; however, shares of individual food groups in overall consumption (in the case of household surveys) or availability (in the case of FBS) tend to remain consistent<sup>32</sup>. They may therefore prove very useful in estimating or imputing food availability, provided that FBS compilers take note of, and adjust for, the other limitations of household surveys. These may include the following issues:

- Data are typically collected only for a brief period of time. However, strictly annualizing the data may be problematic for countries where consumption varies between seasons. Care should be taken to ensure that seasonality is accounted for where relevant.
- Data may be collected only infrequently every four or five years, such that quantities may need to be adjusted for subsequent years to account for changes in income or population, for example.
- Surveys may miss some underrepresented subgroups, thus biasing consumption estimates when extrapolated to the total population.
- Household surveys will entirely miss consumption occurring in schools, prisons, hospitals and military installations.
- Although it is increasingly less common, household surveys sometimes collect data only on expenditures and not quantities. In these cases, expenditures should be converted to quantities using consumer prices.
- Surveys do not include any accounting for food waste at the retail level and may not include food waste at the household level either, potentially underestimating total food availability.

Bearing these caveats in mind, food consumption estimates from household surveys can serve as a benchmark for estimations of FBS food availability, and can even in some cases be scaled up to better fit the FBS definition.

#### 3.5.4.2 Alternative data sources

Even if data is collected by other actors outside of official surveys, the same two sources mentioned above – food processor statistics and household surveys – provide the best snapshots of overall data on food use within a country. At the same time, additional scrutiny may be necessary if these data are collected outside of official channels.

For food processor statistics, FBS compilers may wish to either consult industry groups, processor associations or even a small number of firms (provided that they collectively account for a large share of the total market) to assess the availability of data at the first-line processor level. In each of these cases, compilers should note the representativeness of the data and make adjustments as necessary. For example, if a wheat flour millers’ association represents approximately 80 percent of the total market, then data on output from the association could be used to derive a total production of flour used for food simply by dividing by 0.8.

For household surveys, compilers should consider all of the caveats outlined above, paying additional attention to the representativeness of the survey.

---

32 For more information, as well as a methodology reconciling FBS and household estimates, see Grünberger (2014).

### 3.5.4.3 Imputation and estimation

In the absence of data on food availability from the sources described above, food availability can be imputed. Two possible approaches are outlined below.

#### Suggested approach

Recall from above that per capita food availability is likely to fluctuate little from year to year, as actors within countries use trade or stocks to smooth consumption. The basic approach to imputing food use relies on this assumption – modelling food availability in the current year based on availability levels in the previous year, while adjusting for changes in income and the overall trend in consumption. Imputation of food availability should also account for changes in population: even if each person in a country eats the same quantity of a certain food product from one year to another, adding additional people to a country’s population (assuming that dietary patterns remain unchanged) necessarily increases the amount of that product that is available for consumption as food.

The foundational linear equation for food use, which uses only population, trend and food use in the previous period, can then be defined as follows for a given commodity in a given country:

$$Food_t = \frac{Population_t}{Population_{t-1}} * Food_{t-1} * (1 + \phi) \quad (3-8)$$

where food availability in the current period  $t$  ( $Food_t$ ) is estimated as a function of the change in population (expressed here as the ratio of population in the current period to population in the previous period, or  $\frac{Population_t}{Population_{t-1}}$ <sup>33</sup>) multiplied by food availability in the previous period ( $Food_{t-1}$ ), multiplied by 1 plus the historical trend (for example, the growth rate) in food consumption ( $\phi$ ). In this specification,  $\phi$  should be estimated from a regression on the historical food availability data series.

The basic specification outlined in equation (3-8) provides a good foundation for a basic estimate of food availability. However, most country-level FBS compilers will have access to additional information that should provide a better estimate of food availability. Specifically, country-level compilers can consider introducing both income (in the absence of specific data on income, this may be proxied by either expenditure data sourced from national accounts or Gross Domestic Product (GDP), depending upon data availability and the preferences of the country)<sup>34</sup> and product-specific income elasticities of demand<sup>35</sup> into the equation. In doing so, the specification will depend on how income elasticities for the product in question have been estimated. The semi-log specification (indicated for income elasticities that have been estimated using an underlying semi-log functional form) is set out below, as it is very similar to the linear specification, with the only addition being the income elasticity  $\epsilon$  for the commodity in question multiplied by the log of the change in the income proxy (in this example, the ratio of household consumption expenditure in the current period to household consumption expenditure in the previous period, or  $\frac{Household\ consumption\ expenditure_t}{Household\ consumption\ expenditure_{t-1}}$ )<sup>36</sup> in the additive term at the end of the equation, as below in equation (3-9).

33 This expression is also equal to 1 plus the percent change in population from period  $t-1$  to period  $t$ .

34 There are several potential proxies for this exercise. Some suggestions include either final consumption expenditure, household consumption expenditure or GDP. All of these data are published by the UNSD in their National Accounts data sets, available as either “National Accounts Estimates of Main Aggregates” or “National Accounts Official Country Data”. The three referenced categories can be found under “GDP by Type of Expenditure” at both current and constant prices, for both data sets. This data is publicly available at <http://data.un.org/Explorer.aspx?d=SNAAM>. Accessed on 26 April 2017.

35 Income elasticities of demand measure the responsiveness of demand for a certain good to a change in income. Mathematically, this can be expressed as  $income\ elasticity\ of\ demand = \frac{\% \Delta in\ demand}{\% \Delta in\ income}$ . For example, an income elasticity of demand of 0.1 for a given good indicates that for every 10 percent increase in income, demand for the product rises by 1 percent. Almost all food products are normal goods; that is, an increase in income is associated with an increase in demand for the good.

36 This term is equal to 1 plus the percent change in household consumption expenditure from period  $t-1$  to period  $t$ .

$$Food_t = \frac{Population_t}{Population_{t-1}} * Food_{t-1} * \left[ 1 + \epsilon \log \left( \frac{Household\ consumption\ expenditure_t}{Household\ consumption\ expenditure_{t-1}} \right) + \phi \right] \quad (3-9)$$

Prior to the specification of a model using income elasticities, however, FBS compilers must first identify appropriate elasticities for each of the food products or food product groups. These are likely to have been estimated by academics and then used as inputs in computable general equilibrium models. If no domestically recommended database or source data can be identified, compilers can consult the database of country-specific income elasticities for food categories produced by USDA in 2005 as a next best option<sup>37</sup>.

### Alternative approach

For products where food use is the sole or the overwhelming utilization, countries can employ a balancer approach (similar to that described above for wheat flour), where food availability is calculated as the balance of production minus net trade (and any other small utilization elements). Because this approach will result in food use accumulating all of the error from the other utilization elements, this approach is most appropriate for products that have no or few other utilizations: principally, items that cannot be stocked for extended periods of time and are not used for feed, such as meat, eggs and certain fruits and vegetables or dairy products.

It should be noted, however, that in the final validation and balancing process, the food availability estimated using either approach may be adjusted.

## 3.5.5 Food processing

Food processing refers to quantities of a commodity that enter a manufacturing process for the production of a derived food product. As noted in the “Production” section, food processing quantities are linked to the production of derived commodities through extraction rates. That is, food processing is unique in that it can either be directly measured, or can be calculated by applying the extraction rate to the quantities of production of derived commodities. Thus, if data on either production of a processed commodity or input into a transformation process is present, the other quantity is easily calculated.

### 3.5.5.1 Official data sources

Two official data sources on food processing should be noted. The first are agricultural production surveys. For some commodities (such as fruits or milk), production surveys may include questions on whether the product is destined for the fresh market or whether it was sold to be further processed. Quantities reported as destined for further processing are then, by definition, food processing quantities.

<sup>37</sup> This database is publicly available at <http://www.ers.usda.gov/data-products/international-food-consumption-patterns.aspx>. Accessed on 19 January 2017.

Industrial output surveys are another potential data source for food processing data, albeit indirectly: if the production of derived goods is reported in an industrial output survey, compilers need only divide by the extraction rate to calculate the primary commodity equivalent used as input for that particular transformation process. As noted in section 3.5.4 on food availability, these official data sources are only useful if a majority of processing is covered by industrial output surveys. Where home processing is common, these data should be combined with an estimate of total production of the derived product at the household level to obtain an estimate of total production of the derived good, from which a food processing quantity can then be imputed.

**3.5.5.2 Alternative data sources**

Where official data is not available, data from commodity organizations, manufacturers’ associations or even specific food processing facilities may also be useful in the calculation of food processing quantities. In such instances, however, FBS compilers should make some effort to take into account the representativeness of said data. For example, if the members of a hypothetical “Orange Juice Producers’ Association” are thought to cover 90 percent of all production, then orange juice production data from that association can be utilized and scaled up to obtain an estimate of the country’s total orange juice production.

**3.5.5.3 Imputation and estimation**

Given that estimated quantities destined for food processing are linked to production quantities of derived commodities through extraction rates, the imputation of food processing can be fairly simple in cases where data on production of derived goods exist. As described in section 2.3.1 on commodity trees, this calculation can be illustrated by equation (2-9).

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

It should be emphasized here that “food processing” quantities must cover the inputs of all derived products. As such, the application of the above equation will only result in the total quantity of food processing if only a single derived product stems from the primary good. Of course, the equation can be applied multiple times, and the values of the primary commodity equivalents can be added together to derive the overall quantity of the primary good that entered all transformation processes.

If no data on derived commodity production are available, then it is recommended that total quantities destined to food processing be estimated by a panel of experts. This panel should also determine the share of the food processing quantity that is destined to different transformation processes.

In particular cases, food processing can also be used as a balancing item at the SUA level. Recalling the olive example from box 2-1, all olives are processed before they are consumed. As such, after accounting for net trade, loss and any other utilization, all remaining olives can be assumed to be destined for food processing.

### 3.5.6 Feed

One of the more dramatic shifts in the global food system over the past several decades is the increasing dietary demand for animal products (meat and dairy) as incomes rise in developing countries. With rising numbers of livestock globally, demand for animal feed has also risen, and now accounts for a large proportion of the global production of some crops that are also used for food (such as maize and soybean). In addition, feed can be obtained from a variety of sources – including own production, feed compounders, or even common pasture resources – depending upon relative prices and (to some extent) the intensity of a given country's livestock production system. At the same time, the composition of livestock rations can also shift depending upon changes in these relative prices of feed products, such that for an individual cereal grain, the quantities allocated to feed may fluctuate from year to year. However, aggregate nutrient availability from all feed sources should remain relatively stable on a per livestock unit basis. These trends may be evident in both official and unofficial data sources, and any developed imputation approach should also take this into account.

To improve the accuracy of feed estimations, FBS compilers should first research the characteristics of livestock rearing in their country. Both official data collection approaches and imputation strategies should take into account the structure of livestock production systems to more accurately estimate feed needs.

#### 3.5.6.1 Official data sources

Official data collection efforts on feed are subject to a similar limitation as stock data, in that feed can be sourced from a variety of actors. Therefore, obtaining an accurate picture of aggregate feed production requires various types of surveys. Questions can be added to farm-level surveys on own production reserved for feed; feed compounders can be surveyed as to their output; and pasture resources can be estimated using a variety of methods. If feed compounders are not surveyed, it may also be possible to derive an estimate of their output by consulting administrative records. If costs are an issue, then ad hoc surveys on feed use could be utilized to measure feed demand periodically, which could help to parametrize a country-specific module of feed demand and utilization.

However, it is important that this official data on feed production be cross-checked against actual livestock feed demands, in terms of both total energy and total protein requirements.

#### 3.5.6.2 Alternative data sources

If official data is incomplete or unavailable, unofficial data sources may also exist to help countries estimate feed use for certain commodities. First, commodity interest groups are likely to either have some measurement or estimate of the quantity or proportion of their particular commodity that is being used as animal feed. At the same time, livestock associations may publish data on feed usage, or may be able to at least provide some indication as to the composition of feed rations for certain animal groups. Regardless of whether these groups publish data on feed use, it is advised that FBS compilers consult them to gain a better understanding of the feed market and, ultimately, better inform their balance sheets.

Extension workers could also be consulted to glean information on herd sizes, most commonly utilized feedstuffs, and local animal production systems. While this information may be an approximation, it could still be useful for estimating overall feed use and the feed use of a given individual commodity.

### 3.5.6.3 Imputation and estimation

Imputation of missing feed data can pose a challenge to country-level FBS compilers, as feed sources tend to be substitutable in animal feed rations, and overall feed demand will shift based on livestock populations and productivity intensity levels. Although different actors have used several approaches to estimating feed use for individual commodities<sup>38</sup>, the general approach suggested here is based on the reconciliation of total feed demand and available feed supply. Three basic steps should be followed in this process:

#### Step 1: Understanding and estimating total feed demand

At its most basic level, total feed demand is merely a function of the total number of animals and the nutritional needs of those animals, in terms of both total energy and protein. With respect to energy needs, total required feed demand,  $FD$ , is merely the product of the number of animals belonging to species  $i$ ,  $N_i$ , multiplied by the amount of energy required per animal for that species,  $e_i$ , summed across all species,  $i$ , raised in the country, as in equation (3-10) below (total protein demand could be estimated using the same equation, substituting per animal protein requirement,  $p$ , for the energy requirement,  $e$ ).

$$FD = \sum_i N_i * e_i \quad (3-10)$$

Of course, the amount of energy required per animal can vary widely even within species, depending upon both the characteristics of the animal (for example, a lactating dairy cow's energy needs are much greater than those of a yearling calf) and the type of the production system in which the animal is raised, such as grassland-based systems versus more intensive industrialized systems.

It is important to note that FBS themselves cover only non-forage commodities. Therefore, this imputation method estimates feed demand only for non-forage commodities (in fact, it is redundant to attempt to estimate feed demand for forage, as feed is by definition the sole utilization of forage). To underscore this point, consider that livestock raised in pastoralist production systems are almost universally fed a diet solely of forage, such that their feed needs would be excluded from FBS.

At the same time, the supply of forage crops is necessary for the calculation of total feed supply under this method. As such, countries are encouraged to compile data on production of forage crops (in addition to production of non-forage crops) to ensure that total feed supplies from all sources are adequately measured.

Taking all these factors into consideration, the following suggestions will assist country-level FBS compilers in more accurately estimating total feed demand:

- Understand the different livestock production systems being used in the country for each species, including industrial, grassland-based, mixed, or backyard
- With the assistance of farm surveys and censuses, attempt to estimate the number of animals for each animal species raised in the country (including farmed fish and poultry), and the number of animals raised under each identified production system. These may include counts of animals raised under nomadic or transhumant systems, although the feed demand for such animals is typically satisfied by forage and not by grain crops and crop derivatives
- Determine the “average” animal's feed requirements for each production system
- If data on the number of animals raised under different production systems are of good quality, consider estimating the feed needs of animals raised under separate production systems as if they were a different species

<sup>38</sup> For a more extensive treatment of this topic, see AMIS (2014).

(for example, estimate feed needs of backyard chickens separate from industrial broiler production) to ensure as accurate of an estimate of total feed demand as possible

- Add up the feed needs of all animal species, both in terms of energy and protein

Once an estimate of total feed demand has been reached, FBS analysts can move on to accounting for the supply side.

### **Step 2: Understanding feed supply**

Understanding feed supply begins with an inventory of all of the products that are potentially used for feed in the country. This should be accompanied by an analysis or ranking of which commodities are most likely to be used to meet feed demand. Analysts should keep in mind that many byproducts from food processing (such as bran or fruit pulp) can be assumed to be utilized exclusively as feed in the FBS setting. In addition, as mentioned above, any production of forage can be assumed to be used for feed.

Once all the commodities used as feed have been identified (forage, cereals, root crops and processing byproducts, to name a few), their dietary values per unit should be recorded. These values are publicly available for almost 1 400 feed products via the Feedipedia online resource library<sup>39</sup>.

After all commodities used as feed have been identified, the work of allocating supplies to match with the feed demand calculated in Step 1 can begin. First, any official data on feed use should be recorded appropriately, converted to their total energy/protein equivalents (by multiplying quantities reported in tonnes by their unit dietary values), and subtracted from the total feed demand estimated in Step 1. After official feed use data have been accounted for, residual use (that is, production net of trade) of commodities only used for feed (including forage, bran and pulps) should be considered. As above, these volumes should be converted to energy/protein equivalents and subtracted from the remaining total feed demand.

### **Step 3: Allocating feed supply**

The final step is to allocate the remaining feed demand to the available commodities. The recommended approach is to assemble a TWG to discuss the most likely feed commodities, and distribute feed demand among them accordingly. Alternatively, FBS compilers can use whatever information is gathered in the initial feed inventory at the beginning of step 2, and allocate feed demand accordingly based on the ranking of which commodities are most likely to be used as feed. As an additional option, analysts can consult the available literature on feed demand in their country for assistance in estimating feed utilization. It should be emphasized, however, that regardless of how estimates of feed demand are derived, they should be validated by the TWG once FBS have been compiled for all commodities.

---

<sup>39</sup> This database is a collaboration between France's National Institute for Agricultural Research (*Institut National de la Recherche Agronomique, or INRA*), Agricultural Research Centre for International Development (*Centre de Coopération Internationale en Recherche Agronomique pour le Développement, or CIRAD*) and Association for Animal Production (*Association Française de Zootechnie, or AFZ*), and FAO, and can be accessed at <http://www.feedipedia.org/>. Accessed on 19 January 2017.

### 3.5.7 Seed

Although official estimates for seed may not exist in all countries, for any country with reliable crop area estimates, the process of imputing missing values is relatively straightforward: seed use is merely the product of an average seeding rate (the amount of seed required for a given sown area) and the sown area in the following year (since seed use in year  $t$  is actually merely set aside in year  $t$  to use for sowing in year  $t+1$ ). For most crops, the seeding rate does not vary substantially from year to year within a given country. However, more gradual changes in the seeding rate may be expected due to the adoption of new technologies, different planting methods, or even in instances where production shifts to new areas within a country.

After production and trade, seed estimates are the balance sheet variable for which official data is most common, likely due to the aforementioned ease of imputing missing values.

#### 3.5.7.1 Official data sources

Most official measurements of seed use data are sourced from agricultural surveys. Most surveys will include questions on both purchases of improved seed and quantities of own reserved seed; however, if the survey excludes purchases of improved seed, it may be possible to access the sales records of commercial seed companies to obtain a full estimate of total seed use. Trade data may also provide some indications on seed quantities if most seed is imported, as seeds typically have separate HS codes.

#### 3.5.7.2 Alternative data sources

If an estimate of total seed use is not available and a historical seeding rate cannot be calculated from the data, compilers should also investigate the possibility of whether data is available on seeding rates only. Information on either optimal or effective seeding rates may be available from a variety of sources. First, compilers can contact commercial seed companies to inquire about recommended seeding rates for major commodities for the varieties most commonly sold in their country. Additionally, agricultural research institutions and extension specialists may be able to provide estimates of the common seeding rates in certain production regions. In cases where governments have programs providing subsidized seed to growers, government administrative records are likely to contain information on average seeding rates.

Information on seeding rates for various commodities can also be found in the publication entitled *Technical Conversion Factors for Agricultural Commodities*<sup>40</sup>. However, compilers are advised that this publication is not recent, such that the published average rates may not reflect current production systems or technologies.

Regardless of the source, if the typical seed rate is known, then imputing total seed use is a simple calculation, according to the methodology described below in section 3.5.7.3.

---

40 This document is available here: <http://www.fao.org/fileadmin/templates/ess/documents/methodology/tcf.pdf>. Accessed on 19 January 2017.

### 3.5.7.3 Imputation and estimation

As stated above, seed use quantities in the FBS context represent the amount of seed set aside in the current year that will be used to produce a crop in the following year. As such, seed use in a given year  $t$  is a function of a seeding rate and a sown area in the following year,  $t+1$ , as expressed in the following identity:

$$\text{Seed use (MT)}_t = \text{Seeding rate} \left( \frac{\text{MT}}{\text{HA}} \right) * \text{Sown area (HA)}_{t+1} \quad (3-11)$$

Note: MT = metric tonnes

Given this identity, the process of deriving an imputed value for seed quantity is as follows:

**Step 1: Calculate or estimate a seeding rate.**

**Step 2: If missing, impute a value for sown area in the following year.**

**Step 3: Multiply the two values for an estimate of total seed use.**

Further details of each of these steps are laid out below:

#### Step 1: Seeding rate

If the country has previously planted the commodity in question, country-level FBS compilers are recommended to simply calculate the seeding rate using data from previous years. This can be done by rearranging equation (3-11) to solve for the seeding rate.

$$\text{Seeding rate} \left( \frac{\text{MT}}{\text{HA}} \right) = \frac{\text{Seed use (MT)}_t}{\text{Sown area (HA)}_{t+1}} \quad (3-12)$$

Note: MT = metric tonnes

This equation is then utilized on previous time periods for which seed use in one year and sown area in the following year have been reported.

If seed use is being estimated for the first time or if FBS compilers wish to ensure that utilized seed rates are current (in other words, that they take into account any changes in technology or shifts in production area that could affect the overall seeding rate), compilers should consult agricultural experts who are capable of providing an estimate of an average seeding rate (bearing in mind that this average will have to reflect any differences in seeding rates by production system)<sup>41</sup>. These experts could be extension agents, research scientists at public institutions or even persons within seed companies, provided that most utilized seed is purchased every year and does not come from household reserves. Seeding rate estimates from these sources at the country level are likely to be closer to reality than estimates produced by any general global model.

As a last resort, country-level FBS compilers can consider using seed rates from products in the same commodity group, or even use seed rates for the same product from similar (typically neighboring) countries.

41 Where seeding rates vary greatly by production system, the average seeding rate should be calculated as a weighted average of seeding rates in individual production systems. For example, the seeding rates for direct-seeded paddy rice are higher than those for transplanted rice and system-of-rice-intensification (SRI) techniques. As such, the “average” national seeding rate in these circumstances should be calculated by weighting the average seeding rate of each different technique by its share of planted area.

## Step 2: Area imputation

The statistical programs of many developing countries collect data only on sown area, and not harvested area. If an estimate for sown area in the following year is available, FBS compilers should go to step 3, as only sown area and seeding rate are necessary to derive an estimate of seed use.

If no estimate for sown area in the following year is available, then this area must be imputed. This can be done through one of three approaches, depending upon the data available to the country-level compiler: the first approach is to be preferred (however, it requires other input data); the second and third approaches are proposed as alternatives.

### Approach 1: The ratio approach

For countries where historical data is available on both sown and harvested area, then the average ratio of sown area to harvested area over the historical series,  $\overline{RatioSH}$ , can be used to impute a quantity estimate for sown area in the following year, provided that an estimate of the area harvested in the following year is available<sup>42</sup>. To derive the  $\overline{RatioSH}$ , country-level FBS compilers need simply calculate  $RatioSH_t$ , the ratio of  $\frac{\text{Sown area}}{\text{Harvested area}}$ , in each year for which there is a value for both variables, and then average those annual ratios. Once  $\overline{RatioSH}$  has been calculated, that value can be multiplied by harvested area in the following period,  $t+1$ , to obtain an estimate for sown area in that same year, as below:

$$Sown\ area_{t+1} = (\overline{RatioSH}) * Harvested\ area_{t+1} \quad (3-13)$$

The following scenario will help to illuminate this calculation. Imagine that Country A's FBS compilers have estimated harvested area of sunflowers for 2014 at 385 ha, but they have no estimate of sown area for that year. They do, however, have historic data on sown area and harvested area for sunflowers for 2010–2013. To impute sown area for 2014, the first step is to calculate  $RatioSH_t$  for each year (step one in table 3-4). Then, each of these annual  $RatioSH_t$  values are averaged over the series to derive  $\overline{RatioSH}$  (step 2).

**TABLE 3-4. HYPOTHETICAL SOWN AREA, HARVESTED AREA AND RATIO SH FOR SUNFLOWERS IN COUNTRY A.**

Year	Sown area (HA) (A)	Harvested area (HA) (B)	Ratio SH <sub>t</sub> (C=A/B)
2010	400	388	400/388 = 1.03
2011	425	405	425/405 = 1.05
2012	420	395	420/395 = 1.06
2013	390	370	390/370 = 1.05
<b>2014</b>	?	<b>385</b>	$\overline{RatioSH} = \frac{1.03+1.05+1.06+1.05}{4} = 1.05$

<sup>42</sup> Where sown or harvested areas vary widely from year to year, compilers may instead wish to calculate  $(RatioSH)^{-}$  as a geometric mean, since the geometric mean is less susceptible to extreme values than the arithmetic mean suggested above.

Now that Country A FBS compilers have both  $\overline{\text{RatioSH}}$  and harvested area in 2014, they can calculate sown area in 2014 using equation (3-13).

$\text{Sown area}_{t+1} = (\overline{\text{RatioSH}}) * \text{Harvested area}_{t+1}$	(3-13)
--	--------

$\text{Sown area}_{2014} = (\overline{\text{RatioSH}}) * \text{Harvested area}_{2014}$ $\text{Sown area}_{2014} = 1.05 * 385$ $\text{Sown area}_{2014} = 404$	
---	--

If no estimate of harvested area in the following year is available, country compilers can substitute the current year's harvested area as a stand-in until data on harvested area in the following year is available.

### Approach 2: Abandonment adjustment

If there is no historical data from which  $\overline{\text{RatioSH}}$  can be calculated but data on harvested area is available, sown area can be estimated using the harvested area data and an approximation of the amount of area that is sown but not harvested, which is referred to as the abandonment rate. The first step is to use the identity that harvested area is equal to the sown area multiplied by 1 minus whatever percentage of land is abandoned, *abd*.

$\text{Harvested area}_{t+1} = (1 - abd)\text{Sown area}_{t+1}$	(3-14)
---	--------

Rearranged to solve for sown area, the formula becomes:

$\text{Sown area}_{t+1} = \frac{\text{Harvested area}_{t+1}}{(1 - abd)}$	(3-15)
--	--------

For example, if Country A harvested 95 ha of wheat in 2015, and 5 percent of sown area is commonly abandoned before harvest in an average year, equation (3-15) can be utilized to calculate an estimated sown area, which in this case calculates to 100 ha.

$\text{Sown area}_{t+1} = \frac{\text{Harvested area}_{t+1}}{(1 - abd)}$	(3-15)
--	--------

$\text{Sown area}_{2015} = \frac{95}{(1-0.05)}$ $\text{Sown area}_{2015} = 100$	(3-15)
---	--------

### **Approach 3: Using harvested area as an approximation for sown area**

If it is not possible to calculate a historical ratio of sown to harvested area or an estimate of land abandonment, but data on harvested area is available, then as a final option, country-level FBS compilers can use harvested area in the following year to approximate sown area in the following year. This approach is equivalent to calculating sown area using either a  $\overline{\text{RatioSH}}$  of 1 in Approach 1, or an *abd* rate of 0 in Approach 2. It should be emphasized that this approach should only be used when either of the two options above are not possible, as no accounting for land abandonment will lead to chronic underestimating of seed use in the previous year.

### **Step 3: Multiply the two values**

Once a seeding rate and sown area in  $t+1$  have been estimated for the product in question, the two values are multiplied to obtain the quantity of seed needed in year  $t$ .

## **3.5.8 Tourist food**

Historically, food available for consumption by tourists and other visitors was not included as a separate category in most FBS exercises. Instead, these quantities were assumed to be covered by the catchall “other utilizations” category. However, estimating food available for consumption by visitors independently is encouraged for two reasons. First, data on visitor arrivals is widely accessible, such that it is possible for all countries to more specifically account for tourist food in their FBS. Second, for some countries (particularly Small Island Developing States), large numbers of visitors relative to the resident population have the potential to substantially alter the balance sheet landscape. For example, the UN Population Division reports that in 2013, the population of the Caribbean nation of Saint Lucia was of 182 000 people. In that same year, the UN World Tourism Organization (UNWTO) reported that the country had 921 000 visitors, including 602 000 same-day visitors and 319 000 overnights visitors, who stayed an average of 8.9 days (UNWTO, 2016). This concept of large visiting populations also applies to countries with large migratory labour forces. That is, if the country compiling the FBS has a large seasonal immigrant workforce that is not counted in population estimates as residents, then the food available to those migrant labour forces must be accounted for somehow. In these cases, it is evident that the failure to specifically account for food available for consumption by non-resident visitors (regardless of the duration of their stay) would lead to overestimation of the food available for consumption by local populations.

Likewise, the days that a country’s residents spend abroad should not be counted in domestic food availability, given that those persons are not at home to consume food, and food consumed abroad will be counted in the tourist food figures of other countries. For this reason, estimating tourist food should be done in net terms. That is, net tourist food should be calculated by subtracting the food that would otherwise be available to a country’s outbound travellers from the amount of food available to inbound visitors. As a consequence, it is likely that this information in the balance sheet will have to be populated through imputation, derived using numbers of visitors, visit lengths and the amount of calories historically available in the home and destination countries. This input data can be drawn from a mixture of both official and semi-official sources, as detailed below.

It should be emphasized that country-level compilers should ensure that all persons consuming food within a given country are accounted for, either as resident population (thus accounted for in food availability) or as visitors (accounted for in tourist food). This concept is straightforward when it comes to tourists, but may not be so evident when it comes to temporary migrants, who may spend months away from home. How these populations are classified may depend upon the country. However, the following should always apply: if an FBS compiling country’s outgoing migrants are counted as part of the resident population in their country of origin, then the number of visitor-days spent outside the country should be subtracted from tourist food; if outgoing migrants are counted as non-residents in their country of origin, any days that they spend back in their country of origin should be added to tourist food.

Of course, cases such as that of Small Island Developing States described above affect only a small subset of countries, and migrant worker populations may constitute a minor proportion too, such that many country-level FBS compilers may not find it worthwhile to estimate tourist food separately. Nevertheless, the general approach is outlined here, and full details on the necessary calculations are available in appendix 1.

### **3.5.8.1 Official data sources**

While data on arrivals and departures may be collected by immigration authorities, it is likely that national tourism offices in each country are the entities that publish the most detailed information available on visitor arrivals and departures. This data should be differentiated by country of origin and include numbers of both day visitors and overnight visitors, as well as the average length of stay for overnight visitors.

Through surveys, tourism boards may also publish figures on tourist food consumption patterns, which would certainly aid FBS compilers in estimating tourist food within the balance sheets.

### **3.5.8.2 Alternative data sources**

If FBS compilers do not have ready access to their country's data on visitor arrivals, they may instead consult reports from the UNWTO<sup>43</sup>. This organization compiles and publishes member country-provided data on the number of visitors<sup>44</sup>, average length of stay, and country of origin, as well as estimates on outbound tourism. Although no data on tourist food consumption is included, as above, the number of arrivals is a useful starting point for estimating tourist food.

Industry groups may have more detailed data about actual visitor food consumption, possibly including information on how tourist consumption patterns differ from the local population, or even the quantities of certain foods consumed by tourists. If such information is not available from industry groups, it is possible that for countries where tourism is clustered mostly into resort areas, sales or tax records from those establishments could be used as a first-level approximation of tourist food and then extrapolated to the entire tourist population using the appropriate weights.

### **3.5.8.3 Imputation and estimation**

The approach to imputing tourist food is merely a calculation and not an econometric model. However, given the number of steps involved in the calculation and the reality that many countries may decide not to estimate tourist food separately, for the sake of brevity, only the basic approach is outlined here. Instead, a step-by-step guide on how to calculate net tourist food is included in appendix 1.

Net tourist food is simply the amount of food available to incoming visitors minus the amount of food that would have been available to residents had they been present in the country. For each commodity, this amount can be calculated by first multiplying the number of incoming visitor days by the average daily food availability of that commodity, and then subtracting from this value the product of the number of outgoing traveller days and the average daily food availability for that commodity (3-16).

---

43 Although access to all data and reports requires a subscription, basic data on number of arrivals and country of origin for overnight visitors is available free of charge. See <http://www.e-unwto.org/toc/unwtotfb/current>.

44 Although the data is produced by the UNWTO, data on both personal and business travellers are included. For this reason, it is more accurate to refer to arrivals as "visitors" rather than "tourists".

$  \begin{aligned}  NetTF = & \text{[#Incoming visitor days * Daily food availability for visitors]} \\  & - \text{[#Outgoing traveller days} \\  & \text{* Daily food availability for residents]}  \end{aligned}  $	(3-16)
---	--------

In calculating the daily food available for consumption by tourists, the imputation process also assumes that incoming visitors follow the consumption pattern of the local population (that is, they experience the same dietary food availability shares of given products as in the countries they are visiting), but rather that they continue to expect the same overall amount of calorie availability as in their home country. This is done simply by scaling quantities of an individual commodity by the ratio of overall food availability in the two countries. For example, if food availability in Country A is 30 percent greater than in Country B, all quantities of food available for consumption for visitors from Country A to Country B would be scaled up by 30 percent when compared to the food available for the local population. The specifics of this approach are outlined in appendix 1.

If country-level compilers cannot access data on visitor country-of-origin, but still wish to account for “Tourist Food” in their balance sheet, a simplified calculation is suggested: compilers can also simply assume that visitors experience the same food availability for each commodity as does a resident. This approach may underestimate total tourist food; however, it is preferable to relegating food available for visitors to a residual component.

Again, more guidance on the recommended calculation is provided in appendix 1, although this brief overview should be sufficient to inform country-level compilers of the basic idea, should they consider it relevant to introduce net tourist food as a separate variable in their balance sheets.

### 3.5.9 Industrial use

As defined in section 2.2.1, “industrial use” refers to the utilization of any food items in any non-food industry. Industrial uses of agricultural products have been growing over the past few decades, to a large extent driven by the expansion of the biofuels market. For example, in certain countries, corn, rapeseed, soybean and sugarcane may all be used for this application. However, industrial applications are also growing for other commodities, such as palm oil and coconut oil, which are used in many cosmetics. In addition, many crop byproducts may have industrial applications. As two examples, wheat starch is commonly used in the paper industry, and limonene – produced from orange peels as a byproduct of the juicing process – is a common ingredient in cleaning products.

Because industrial uses of agricultural products are highly context-specific, it is not possible to provide universally applicable advice on data sources or imputation methodologies for this balance sheet variable. Instead, compilers are encouraged to first seek out industry and commodity experts (both in the public and private sector) to investigate which products are utilized for industrial purposes in their respective countries, and how their use can be modelled in cases of missing data. Nevertheless, some guidance on potential data sources is provided.

#### 3.5.9.1 Official data sources

Country FBS compilers are first encouraged to consult any official data sources about the possibility of industrial use of the commodities. Countries with large industrial utilizations of certain products may collect data on the quantity or share of production that is destined for such uses in either an annual statistical yearbook or in industry-specific input-output tables. If, during the process of the data assessment, it is discovered that there is a large amount of industrial use of a certain product that is not captured in current official surveys or input-output data, countries are encouraged to consider collecting official data on those uses, which will better inform markets and facilitate FBS compilation.

### 3.5.9.2 Alternative data sources

For countries where no official data collection on industrial uses is currently taking place, compilers do have some alternatives. In some countries, it may be possible to obtain estimates of industrial uses by accessing purchase or sales records from private agro-industrial companies. Particularly in countries where the processing of a given commodity for industrial uses is concentrated in the hands of a few processors, consulting those companies could provide valuable information for populating the FBS. Some estimates on industrial uses may also be obtained directly from commodity associations, who most likely already consult with or get information directly from agroprocessors.

Where industrial uses are almost entirely biofuel-related, countries may be able to use the current policy framework to assist in estimating industrial use data. For example, if a country has implemented a biofuels mandate, those thresholds may be useful in inferring industrial utilizations.

Where none of these strategies appear to be feasible, countries can also consult two additional data sources, which largely cover biofuel uses. The first of these is the OECD-FAO medium-term outlook, which provides estimates of ethanol production, biodiesel production and biofuel use for a selection of the world's countries<sup>45</sup>. Compilers may also wish to consult the USDA's Production, Supply and Distribution (PS&D) database estimates for the "Industrial Domestic Consumption" of oil crops. These estimates are typically derived from reports of agricultural attachés of the United States of America, and may provide a useful starting point for FBS compilation<sup>46</sup>.

### 3.5.9.3 Imputation and estimation

At present, there is no suggested imputation methodology for industrial uses, partly because industrial uses tend to be strongly related to the contexts of specific commodities and countries. To ensure that industrial uses are properly accounted for in the balance sheet framework, compilers are encouraged to focus their efforts on consulting with commodity experts, and advocating for official data collection if industrial uses are found to be large.

## 3.5.10 Loss

Recall that for the purposes of the FBS, "loss" most closely aligns with "post-harvest/post-slaughter loss", representing those quantities of food that leave the production/supply chain at any stage from post-harvest up to the retail level (the level of the supply chain at which "food availability" is defined). The accurate measurement or imputation of loss is important both because of its effect on the overall balance (failure to estimate loss could result in dramatically higher estimates of food availability, or of any of the other utilizations, for that matter) and as a means to help countries identify problems in production or in particular supply chains to underpin policy efforts seeking to maximize resource efficiency. Food loss is problematic because it is both a squandering of resources (for example, the land used to produce food that goes uneaten could have been dedicated to carbon sequestration) and an environmental problem in its own right (because rotting food emits methane gas), not to mention the untapped potential of that lost food to feed the world's almost 800 million hungry people. However, to most effectively target food loss interventions, loss must first be measured.

<sup>45</sup> This database can be accessed at <http://www.agri-outlook.org/database/>. Accessed on 19 January 2017.

<sup>46</sup> The PS&D database is available at <https://apps.fas.usda.gov/psdonline/app/index.html#/app/home>. Accessed on 19 January 2017.

### **3.5.10.1 Official data sources**

For the reasons elaborated above, countries are increasingly attempting to measure or estimate loss as part of their overall agricultural statistical programs. It is recommended that countries use targeted surveys to measure loss. This may include surveying loss in on-farm operations and storage, loss in warehouses or collection points, loss in transportation and loss in public storage. While surveying for information on loss can be expensive, countries can follow certain recommendations to reduce these costs, such as including a module on loss in annual production surveys at the farm level. The Global Strategy has already produced a methodological report on the measurement of post-harvest loss of grains and is planning to publish guidelines on the topic in 2017.

For countries holding large public food stocks (particularly of cereals) access to data on the loss both in public storage facilities and loss occurring during the transportation of these publicly held stocks is essential to accurate estimation of overall loss. Without such data, loss is likely to be severely underestimated.

### **3.5.10.2 Alternative data sources**

For most countries, at least some data on loss in specific segments of the supply chain is most likely available outside of official sources, as loss incurs real-world economic costs for supply chain actors. At a minimum, country-level FBS compilers are encouraged to consult warehouse managers and transportation firms or associations for a basic understanding of the scale of loss for the most important commodities.

In addition, country-level FBS compilers are advised to seek out case study investigations of that may contain loss estimates for particular sectors. Compilers are, however, encouraged to consider the statistical validity of the data, particularly its representativeness of the target population, before adopting an estimate published in a sector case study.

### **3.5.10.3 Imputation and estimation**

In the absence of data from official sources or information from alternative data sources at the country level, loss should be imputed for each primary commodity. The approach followed by a given country will be highly dependent upon the historic availability of data on post-harvest or post-slaughter loss for that country.

#### **Suggested approach**

In cases where some historic data is available, it may be optimal to estimate loss through a regression approach, such that loss is modelled as a function of certain other variables (potentially including covariates such as maximum temperature in the harvest areas, average moisture level of grain, miles of paved roads per square kilometre, refrigerated storage capacity or distance of main producing areas from the main terminal markets). Countries are encouraged to assess their particular situation on loss, including identifying critical loss segments in each supply chain, to determine whether their imputation would improve by using such an approach.

Where no historical data on loss exists, country-level FBS compilers are advised to scour any relevant information available that might inform an estimate of loss. This may include scaling up estimates from case studies, convening focus groups of supply chain experts, consulting industry organizations, or conducting controlled experiments or pilot studies to form some idea of the share of production (or share of supply) that ends up as loss. Then, this percentage can be applied to production (or supply) in subsequent years to impute a value for loss, as in equation (3-17).

$$\text{Quantity of Loss} = \text{Quantity of Production} * \text{Estimated \% Loss}$$

(3-17)

### **Alternative approach**

If no local information is available, country-level FBS compilers may consider imputing loss by relying on the available pool of global loss information. In essence, country-level compilers can estimate their loss using official data from countries that do report loss. This is done by estimating the relationship between officially reported production levels (or officially reported supply levels, but because official data on stock levels is relatively rare, it is recommended that countries use only production data in their estimations) and officially reported loss (for either the product in question or for similar goods) using what is called a hierarchical linear model<sup>47</sup>. This type of model uses nested data, where groups of units are clustered together in an organized fashion, such as “commodities” within “geographic areas” which in turn fall within “overarching commodity groups.” This data is then organized into levels, where the first level is the most specific and the highest level is the broadest. Conceptually, the approach is as follows:

1. If data on loss for a particular commodity in a particular country is reported, then no imputation is necessary. If not:
2. Loss of that commodity in that country is imputed by estimating the relationship between production and all other independent variables and loss of that commodity in all other countries of the world that reported official data on loss, and then using this relationship to calculate likely loss in the country in question. If no official data are reported for that commodity for any country in the world, then:
3. Loss of the commodity in that country is imputed by estimating the relationship between production and all other independent variables and loss of all commodities from the same commodity group as the commodity in question for all other countries of the world that reported official data on loss, and then using that relationship to calculate likely loss in the country in question.

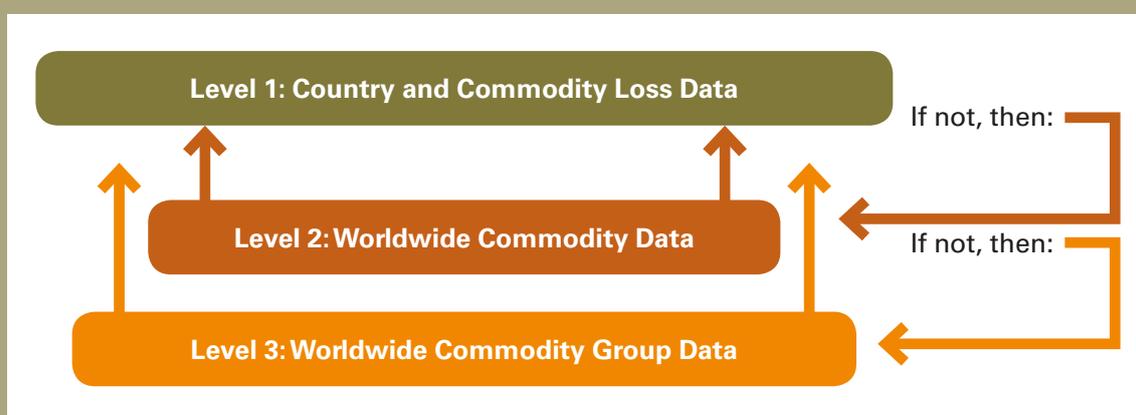
Estimating loss in this manner then requires that countries consult a data resource where other countries have reported official estimates for both loss and production. This data can be accessed in the FAO-compiled FBS.

<sup>47</sup> For more information on this approach, see Gelman and Hill (2007).

### BOX 3-1. EXAMPLE OF IMPUTATION OF LOSS USING A HIERARCHICAL APPROACH.

A practical example and graphical representation of the application of the hierarchical linear model for the imputation of loss should facilitate understanding of this structure. Say that an FBS compiler from Country A is attempting to impute loss for wheat. Country A does not collect data on wheat loss (level 1), thus making an imputation necessary. However, many other countries do measure loss of wheat. Therefore, Country A can estimate the average relationship between loss of wheat and production on a global level (level 2), and then use that relationship to calculate an imputed value for their loss of wheat. In this case, the only information used from the hierarchy would be that at level 2: there is no data for level 1 and because data on level 2 was available, the compiler did not need to move further up the hierarchy (box figure 1).

BOX FIGURE 1



Suppose that in addition to wheat, Country A must also impute loss of oats. Because Country A does not collect data on oats, they automatically proceed to level 2, the relationship between production and loss of oats at the global level. However (at least for the purposes of this hypothetical illustration), they discover that no other countries report official data on loss of oats. In this case, the compilers would proceed to level 3. Because oats are a cereal grain, the compilers would first estimate the relationship between loss and production for all countries for all cereal grains, and then use that relationship to calculate an imputed value for Country A's loss of oats.

### 3.5.11 Residual and other uses

Residual and other uses is a unique balance sheet element, in that its purpose and calculation varies depending upon the needs of the country in question (indeed, countries may wish to not utilize this category at all, entirely dropping it from the balance sheet).

First, this category can be calculated *ex post* as a balancing item at the SUA level. As such, it would be estimated in a manner similar to that of the “imbalance” in the supply = utilization identity, after quantities have been estimated for each of the other variables. However, as elaborated in earlier sections, this strategy should only be utilized when the imbalances in the equation are small.

There may also be cases where “residual and other uses” is utilized to capture a category that the country itself deems important to include in the FBS. While these guidelines have striven to cover all possible commodity utilizations with the previously identified categories, there may be additional uses for certain commodities in certain countries that country-level FBS compilers wish to account for separately (such as food available for consumption by refugees). In this case, it is not possible to recommend specific data sources or imputation methods; users are encouraged to consult experts on the commodity supply chain in question to determine whether said variable is or can be measured, and, if not, how it can be reliably estimated.

### 3.5.12 Additional parameters

#### Population

In the calculation of country-level FBS, compilers may use either domestically produced or internationally standardized population estimates from the UN Population Division. Domestically produced estimates and UN Population Division estimates may diverge if certain groups are counted in domestically produced figures but are excluded from the figures provided by the UN Population Division (or vice versa). As such, it is up to the compiling country to decide which population estimate they should use for the purposes of FBS compilation. It is important to recognize that the choice of population data will directly impact the final DES figures, such that the rationale for choosing one figure over the other should be discussed and documented by an FBS TWG. This decision should be based upon which figure better captures the full picture of a country’s resident population (counting undocumented workers, refugees, etc.). In many instances, national population data may better account for these resident groups.

Countries may find that using internationally produced population data has specific advantages. Such data are produced according to a standardized, internationally recognized methodology, which facilitates comparison between countries. Furthermore, the estimates from the UN Population Division are used as a common denominator for deriving per capita estimates under global development initiatives, including the recently concluded Millennium Development Goals (MDGs) and the ongoing SDGs. If compilers choose to use this data, they may note that it is publicly available and can be easily accessed from the following website: <https://esa.un.org/unpd/wpp/>.

If there are large differences between nationally produced population estimates and those provided by the UN Population Division, countries are encouraged to clarify the underlying explanation behind the discrepancies to hopefully better reconcile the two estimates.

#### Nutrient estimates

As described in section 2.2.2, the relevant nutrient estimates that are currently commonly covered in FBS composition are energy (expressed in kcal), protein (expressed in g) and fat (also expressed in g). Data on nutrient content are available from various sources; however, two international sources should be noted for their relevance to FBS compilation:

- FAO and INFOODS publishes an international food composition table and database directory, available at <http://www.fao.org/infoods/infoods/tables-and-databases/en/><sup>48</sup>. This website includes links to various regional and country-specific food composition tables. In addition, the FAO/INFOODS Analytical Food Composition Database (AnFood1.1) contains analytical food composition data on specific foods from all over the world.
- FAO publishes the conversion table that its Statistics Division uses in the composition of FBS at [http://www.fao.org/fileadmin/templates/ess/ess\\_test\\_folder/Food\\_security/Excel\\_sheets/Nutritive\\_Factors.xls](http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Food_security/Excel_sheets/Nutritive_Factors.xls)<sup>49</sup>.

Despite the existence of these internationally compiled resources, it is recommended that country-level FBS compilers first verify whether an equivalent nationally produced conversion table exists before using another source table. This is because the nutrient content of products may vary slightly depending upon the characteristics of the variety of the product that is consumed. As such, if no national estimates are available, it is recommended that countries first consult tables from their own region or neighbouring countries before using an international reference table.

It should also be emphasized that countries should note whether their reference nutrient tables are done on the basis of gross or net product weight – that is, whether the per ton nutrient conversions take into account inedible parts, or whether inedible parts should first be subtracted from food quantities through the appropriate application of a refuse factor.

### **Extraction rates**

Extraction rates for processed products vary by country according to the technologies utilized, and in some cases may vary based on the specific properties of the input products used in that country. These rates may increase over time as improved technologies are adopted; however, rates tend to change little from one year to another. Keeping this in mind, compilers are encouraged to hold extraction rates fixed in their FBS calculations. This simplification is recommended because only in a few rare instances in developed countries are both the input quantity of raw material and the output quantity of processed product actually measured, which would be the necessary condition for the endogenous calculation of an extraction rate. Additionally, very few countries separately publish and/or update their average extraction rates annually. In contrast, it is more common for the output of a processed product to be known, but not the input. For example, countries may publish an estimate of their production of apple juice, but not the quantity of apples used as input in that transformation process.

The second reason why fixed extraction rates are recommended is the observed tendency to overestimate quantities assigned to food processing while simultaneously underestimating (or leaving out altogether) other uses of a product. To continue with the above example of apples, this would be equivalent to assuming that all apples not consumed in the fresh market are processed into juice with no accounting for loss on the utilization side of the equation. In this case, an overestimation of the amount of input due to missed accounting for loss would lead to the deriving of a lower-than-expected extraction rate. The last reason why the usage of fixed extraction rates is recommended is to aid the standardization process, such that the parameter used to convert derived products back to primary equivalents remains unchanged from year to year. Even so, countries are encouraged to periodically review their extraction rates to ensure their accuracy.

In sourcing extraction rates, countries are encouraged to speak directly to domestic processors to either collect data on, or more accurately estimate, national extraction rates. More structured supply chain studies (potentially including research student theses or dissertations) may also be a useful source of information, provided that they are sufficiently representative. Focus group discussions including industry experts and researchers could also be a useful source of extraction rate data.

<sup>48</sup> Accessed on 19 January 2017.

<sup>49</sup> Accessed on 19 January 2017.

As a last resort, FBS compilers can consult the FAO handbook on Technical Conversion Factors for Agricultural Commodities, publicly available on FAO's website: <http://www.fao.org/fileadmin/templates/ess/documents/methodology/tcf.pdf>. However, FBS compilers should note that this document was produced in 1999, in some cases based on data that may already have been more than a decade old, and therefore that it does not capture any recent innovations in food processing technologies.

### **Processing shares**

Processing shares are extremely dependent upon the structure of each country's particular product supply chains (as laid out in country-specific commodity trees). As such, there are no international sources that can be consulted in the search for processing shares data. Instead, it is recommended that country-level FBS compilers consult any available supply chain analysis investigations as possible data sources. Industry experts may also be consulted, as they may be able to provide certain insights into the supply chain that would assist FBS compilers in estimating likely processing shares.

## **3.6 SUMMARY**

This chapter has focused intensively on the data necessary to compile FBS. The process of compiling data should begin with the establishment of a TWG that defines the scope of the work and assigns data responsibilities among the TWG participants. Then, the data search process may begin, keeping in mind that data should be comparable by product, unit and reference period. The topic of data quality considerations, including the application of flags and assigning of a priori tolerance intervals was discussed, and advice on performing the data search and assessment was provided. The chapter offered suggested data sources and imputation approaches for all variables in the FBS. Once all of the data is in hand, FBS compilers can move on to fitting the data together in the FBS framework, which will be covered in detail in chapter 4.

---

<sup>50</sup> Accessed on 19 January 2017.

# 4

## Step-by-step guide to FBS construction

### 4.1 INTRODUCTION

As emphasized in previous chapters, the bulk of the work in compiling country-level FBS takes place in the preparatory stages – identifying data sources and making adjustments to ensure that data are comparable (including the filling in of a data assessment grid), followed by the imputation of values for any missing data in the balance sheet. Once this has been done, the task of compiling the actual supply utilization accounts for commodities should proceed more quickly. This requires first filling in the empty SUA tables, balancing SUA-level commodities, standardizing and aggregating products in the SUA to primary commodity equivalent level, balancing the identity at the primary commodity level, converting food values from volumes to dietary equivalents, and finally deriving per capita dietary estimates. This chapter will outline this process by tracing the development of an FBS account for one commodity tree: oats.

## 4.2 FILLING IN AN SUA TABLE

Once the first two steps in the FBS compilation process have been completed (data assessment and imputation of any missing data, covered in the previous chapter), FBS compilers can proceed with filling in the blank SUA table for each commodity in question. This process begins with consulting the relevant commodity tree to ensure the accounting for of the primary commodity and all derived products. For oats, the commodity tree includes the primary product and two derived co-products that result from the milling process: rolled oats and bran of oats<sup>1</sup>. Once this information has been verified, the blank SUA table can be set up, such that the columns account for all utilizations, each product in the commodity tree has its own row, and the first row is reserved for the primary commodity (table 4 1). Note that in this example, the compilers chose to not include a “Residual and other uses” variable in their calculations. As such, it is excluded from the tables.

**TABLE 4-1. BLANK OATS SUA TABLE.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	-	-	-	-	-	-	-	-	-	-	-
Rolled oats	-	-	-	-	-	-	-	-	-	-	-
Bran of oats	-	-	-	-	-	-	-	-	-	-	-

From this blank table, FBS compilers should start by filling in the different supply and utilization variables, beginning from the left-hand side with production. At this stage, compilers should keep in mind that for some derived products, production will be filled only after quantities for amounts sent to processing have been estimated.

In our sample case, the filling in of the SUA table begins with official figures for production for both oats and rolled oats (one coming from an agricultural survey and one from an industrial output survey) (table 4-2). Since it is known from the commodity tree that bran of oats and rolled oats are co-products, the production of rolled oats implies that there must also be production of oat bran. However, there is no official figure for oat bran production; therefore, it is not included at this time. Instead, as alluded to above, a figure for oat bran production will be derived after a number for food processing has been entered.

<sup>1</sup> It should be noted that other products can be derived from oats (including breakfast cereals, cereal preparations and distilled alcoholic beverages); however, in this example, oats are only processed into the above goods.

**TABLE 4-2. OATS SUA TABLE WITH OFFICIAL PRODUCTION DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	-	-	-	-	-	-	-	-	-	-
Rolled oats	54 789	-	-	-	-	-	-	-	-	-	-
Bran of oats	-	-	-	-	-	-	-	-	-	-	-

Next, official trade data is added (table 4-3). At this stage, analysts should note that there is no harmonized HS6 category for rolled oats or bran of oats. Instead, countries may either have a more detailed breakout at the 8- or 10-digit level for these products, or alternatively, trade data for these two products would be included under another HS6 basket category (most likely covering bran or worked grains for other cereals not elsewhere defined in the HS classification). In the second case, country-level analysts would have to estimate the proportion of the relevant 6-digit basket category that these products represent, and then estimate or impute some value for trade. In this example, however, the hypothetical country in question has a more detailed trade breakout at the HS 8-digit level, such that trade data for rolled oats and bran of oats are collected under specific individual 8-digit codes that can be directly added to the table.

**TABLE 4-3. OATS SUA TABLE WITH OFFICIAL TRADE DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	-	-	-	-	-	-	-	-
Rolled oats	54 789	14 074	2 130	-	-	-	-	-	-	-	-
Bran of oats	-	688	1 436	-	-	-	-	-	-	-	-

Once trade has been added to the table, stock changes should be accounted for (table 4-4). Because oats are a grain, they can be stored from one year to the next. However, very few countries collect comprehensive data on total stock levels (which include public stocks, private commercial stocks, farm-level stocks, and, in some cases, consumer-held stocks) through surveys<sup>2</sup>. Instead, stock changes must be imputed, or stock levels must be estimated. In line with this situation, the country in our example does not collect data on stock levels. Therefore, a value of stock change is imputed using a model (refer to section 3.5.3.3 for an example).

Stocks are not kept (or accounted for) for many derived commodities. This is usually the case with derived grain products such as flour and bran, as the shelf life of flour can be considered to be only a few months, while whole, unprocessed cereal grains can be stored for more than one year. By contrast, stock-keeping is common for some derived commodities (such as cheese, juice or olive oil). FBS compilers should be aware of the form in which commodities are typically stored to ensure that they are correctly imputing stock levels. In the case of oats, however,

<sup>2</sup> AMIS (2015) provides more information on the status of the global measurement of agricultural stocks.

stocks are kept in their raw, unprocessed form. Reflecting that fact, zeros have been added in the “Stock change” column in table 4-4 for both rolled oats and bran of oats.

**TABLE 4-4. OATS SUA TABLE WITH IMPUTED STOCK CHANGE DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	-	-	-	-	-	-	-
Rolled oats	54 789	14 074	2 130	0	-	-	-	-	-	-	-
Bran of oats	-	688	1 436	0	-	-	-	-	-	-	-

Next, the variables of food and food processing should be addressed. As with most cereal grains, oats are typically not eaten in their primary form, such that there will be no “food” estimate for the primary good. Instead, the primary product (in this case, oats) is processed into a derived commodity (rolled oats in this example), which are then consumed as food: thus, in lieu of accounting for food at the primary level, there should be some accounting for “food processing” of oats. Analysts should note that the “food” value for the primary equivalent will come later, once the “food” value for rolled oats is aggregated up to the primary equivalent level. The value for “food processing” will be dropped in the same step.

To account for food processing, the production of all derived goods should be converted back to their primary equivalent (if data on their production is available). In this example, since rolled oats and bran of oats are co-products, all that is required is a conversion back of the production of rolled oats into the amount of oats sent to food processing, through an application of the extraction rate using equation (2-8). In this case, the extraction rate for rolled oats from unprocessed oats is 0.65. As such, the calculation proceeds as follows:

$$Quantity\ of\ input = \frac{Quantity\ of\ output}{Extraction\ rate} \tag{2-8}$$

$$Quantity\ of\ input = \frac{54\ 789\ MT\ rolled\ oats}{\frac{0.65\ MT\ rolled\ oats}{1\ MT\ oats}}$$

$$Quantity\ of\ input = 84\ 291\ MT\ oats$$

Note: MT = metric tonnes

Once the value of processed oats has been calculated, it can be added to the SUA table (table 4 5). Since rolled oats and bran of oats are not further processed, it is also possible to add zeros to the table under “food processing” for those products.

**TABLE 4-5. OATS SUA TABLE WITH CALCULATED FOOD PROCESSING DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	-	84 291	-	-	-	-	-
Rolled oats	54 789	14 074	2 130	0	-	0	-	-	-	-	-
Bran of oats	-	688	1 436	0	-	0	-	-	-	-	-

It is important to note that this particular example – in which there is official production data for a derived product – may not be characteristic of most developing country supply chains. As such, it would not be possible to calculate “food processing” by working backwards from the production of the derived products. Instead, a value for food processing would have to be imputed, and that value would then need to be directed towards particular transformation processes using processing shares, with the end goal of calculating the production of the derived commodities (see box 2-1). It should be recalled that the calculated production of the derived commodities would be subject to the constraint that the production of all derived goods is sufficient to cover net trade. That is, if the production of a derived commodity calculated from an imputed “food processing” value is insufficient to cover net trade, then the “food processing” value will have to be raised to a level that would cover production net of trade.

Because a value for the amount of oats sent to processing is now available, that extraction rate can be applied to calculate a number for the production of bran of oats. To do so, equation (2-7) may be used, as follows:

<i>Quantity of output = Quantity of input * Extraction rate</i>	(2-7)
---	-------

$Quantity\ of\ output = 84\ 291\ MT\ oats * \frac{0.20\ MT\ bran\ of\ oats}{1\ MT\ oats}$ $Quantity\ of\ output = 16\ 858\ MT\ bran\ of\ oats$	
--	--

Note: MT = metric tonnes

This value is then added to the table (table 4-6).

**TABLE 4-6. OATS SUA TABLE WITH CALCULATED BRAN OF OATS PRODUCTION DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	-	84 291	-	-	-	-	-
Rolled oats	54 789	14 074	2 130	0	-	0	-	-	-	-	-
Bran of oats	16 858	688	1 436	0	-	0	-	-	-	-	-

As described above, the only product in the oats tree in this example that is consumed by humans is rolled oats. As such, a value for rolled oats used as food must be filled in this step. For the sake of this illustration, it is assumed that there is no official data on the food availability of oats, and therefore the value must be imputed. There are two options as to how this can be accomplished. First, food could simply be treated as the balancer in the equation after accounting for all other uses. If this approach is adopted, then the value for food should be left unfilled until the other utilizations have been accounted for, given that this is a derived good produced solely for human consumption. If this approach is chosen, a value will have to be imputed for net tourist food before food can be used to balance the SUA level account for rolled oats. The second option is to independently impute a value for food utilization using a model that is based on historic food availability data, but that takes into account population and income growth. For the sake of this illustration, it is assumed that the value is imputed. The results of this imputation are then added to the SUA table (table 4-7). It is also possible to add a zero for both regular oats and bran of oats used for food, because neither product is utilized in that way.

**TABLE 4-7. OATS SUA TABLE WITH IMPUTED FOOD DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	-	-	-	-	-
Rolled oats	54 789	14 074	2 130	0	66 700	0	-	-	-	-	-
Bran of oats	16 858	688	1 436	0	0	0	-	-	-	-	-

Feed is the next component in the balance sheet. In this example, two of the products in the tree are used for feed. Therefore, both of these must be filled. However, the strategy for doing so will differ somewhat for each product. First, it is assumed that the sample country surveys both farmers and animal feeding operations to publish official data on the feed use of oats. This number can be added directly to the table on the line for the primary commodity (table 4 8). With respect to rolled oats, a zero can be added to the table, as this product is not used for feed. Finally, considering the feed use of bran of oats, although an estimate of this quantity is unavailable, it is known that the primary utilization of bran of oats is for feed. As such, feed will balance this product's SUA. For this reason, this cell may be left blank for the time being while estimates for the other utilizations are considered.

**TABLE 4-8. OATS SUA TABLE WITH OFFICIAL FEED DATA ON OATS ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	-	-	-	-
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	-	-	-	-
Bran of oats	16 858	688	1 436	0	0	0	-	-	-	-	-

From feed, the next step is seed. In this scenario, the country does not collect data on seed use. However, an average seeding rate and the value of sown area in the subsequent year are both available, such that imputing a value for seed use is possible. As seed is only relevant for the primary commodity, the rest of the column can be filled with zeros (table 4 9).

**TABLE 4-9. OATS SUA TABLE WITH IMPUTED SEED USE DATA ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	-	-	-
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	0	-	-	-
Bran of oats	16 858	688	1 436	0	0	0	-	0	-	-	-

Next, net tourist food is considered for every product that is consumed as food. For countries where net tourist food is likely to be negligible, country-level compilers may wish to either eliminate this variable entirely from their balance sheets, or keep the variable but estimate the value as “0” for the time being, as a consideration that it may be estimated in future. For countries wishing to consider net tourist food, this value can be imputed according to the approach laid out in section 3.5.8.3 and described in detail in appendix 1, which uses incoming and outgoing tourist flows and historical food availability data to estimate the likely quantity of food available for consumption by incoming tourists, net of food that would have been available for consumption to domestic residents that travelled abroad. This imputed value is then added to the table (table 4-10). Note that in this case, the positive value for net tourist food indicates that incoming tourists consume more rolled oats than outgoing residents. Again, as with food, since rolled oats are the only product that can be used for food by tourists, zeros are added for the other commodities in the tourist food column.

**TABLE 4-10. OATS SUA TABLE WITH IMPUTED TOURIST FOOD ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	0	-	-
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	0	750	-	-
Bran of oats	16 858	688	1 436	0	0	0	-	0	0	-	-

After tourist food, industrial use is addressed. In this example, FBS compilers in the country know from industry experts that some outgraded oats not used for feed are diverted to industry to be processed into soaps and lotions. This quantity is not known with certainty, and there is no methodology to impute the value; therefore, it is estimated according to the best available information, such as market data from industry organizations. This value is then added to the table, and zeros fill the remaining rows in the column (table 4-11).

**TABLE 4-11. OATS SUA WITH ESTIMATED INDUSTRIAL USE ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	0	2 500	-
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	0	750	0	-
Bran of oats	16 858	688	1 436	0	0	0	-	0	0	0	-

The last variable to be filled is loss. Many countries do not yet collect data on loss of all products. Partly due to this scarcity of data, countries may provisionally choose to either use estimations or country-specific econometric models. Failing these options, loss may be imputed at the primary-commodity level according to the approach described in section 3.5.10.3. Once this value has been imputed according to the chosen method, the value is added to the table, and zeros fill the remaining rows (table 4-12).

**TABLE 4-12. OATS SUA WITH IMPUTED LOSS ADDED.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial Use	Loss
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	0	2 500	3 940
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	0	750	0	0
Bran of oats	16 858	688	1 436	0	0	0	-	0	0	0	0

Once loss has been added to the table, all variables have been accounted for in each of the SUA accounts.

### 4.3 BALANCING SUA ACCOUNTS FOR DERIVED COMMODITIES

Before standardization and aggregation up to the primary-commodity level can begin, FBS analysts must first check that derived product accounts are balanced. To do so, a calculation is made at this step to check that the supply = utilization identity holds for the derived commodities (table 4-13).

**TABLE 4-13. BALANCE CHECK ON SUA ACCOUNTS FOR COMMODITIES DERIVED FROM OATS.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial use	Loss	Imbalance
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	0	2 500	3 940	
Rolled oats	54 789	14 074	2 130	0	66 700	0	0	0	750	0	0	-717
Bran of oats	16 858	688	1 436	0	0	0	-	0	0	0	0	16 110

In this case, balancing the account for bran of oats is straightforward, as it had already been noted that feed should be the primary utilization for this product. For this reason, the imbalance in this line is assigned to feed. However, the case for rolled oats is slightly different. Because production, imports and exports are all official data that should remain unchanged, the only way that this account can be balanced is if the imbalance is absorbed by either food or net tourist food, which were both imputed using models.

Because the accounts may be adjusted again after the standardization process, it is recommended that at this stage the imbalance be absorbed by the largest utilization. In this case, food is by far the largest non-official utilization, such that the adjustment should be made here. In table 4 14, it may be seen that food use has been reduced, and the equation now balances.

**TABLE 4-14. BALANCED SUAS FOR PRODUCTS DERIVED FROM OATS.**

Product	Production	Imports	Exports	Stock change	Food	Food processing	Feed	Seed	Net tourist food	Industrial use	Loss	Imbalance
Oats	131 259	188 219	3 439	12 350	0	84 291	182 950	12 300	0	2 500	3 940	
Rolled oats	54 789	14 074	2 130	0	65 983	0	0	0	750	0	0	0
Bran of oats	16 858	688	1 436	0	0	0	16 110	0	0	0	0	0

From here, the standardization process can begin.

## 4.4 STANDARDIZATION AND AGGREGATION

Recall from section 2.3 that to obtain a single FBS account at the FBS level, the SUA accounts must be standardized (that is, converted back to their primary commodity equivalent) before they can be added together. This process of standardization is accomplished by dividing by the extraction rate and then adding all the values together. However, only the values of certain variables are added together in the standardization process. In this section, we proceed variable by variable to outline this process precisely.

However, before beginning, it is imperative to note that in order to avoid double-counting, only one commodity from each transformation process is standardized and aggregated. Because the goal of this process is to produce FBS, this means that the product that is standardized will typically be the one that makes the largest contribution to food. That also means that commodities such as bran, which is a co-product of the flour milling process whose primary utilization is feed, are not standardized and aggregated in this framework. This point will be re-emphasized in the variable-by-variable run-down below in section 4.4.1.

### 4.4.1 Standardization rules by variable

#### Production

The rule for production is perhaps the easiest to remember. Production of derived commodities is never standardized to the “production” variable: the production value of the primary commodity remains in this cell for the primary equivalent level. This is because the only commodity that is truly being produced in the balance is the primary good; the “production” of the remaining goods can instead be conceptualized as the conversion of the primary good into other products.

At the same time, FBS compilers should still standardize and aggregate the production of derived goods (though not of co-products from the same transformation process) to ensure that the production of derived goods is equal to the “food processing” variable in the primary commodity’s SUA.

## Trade

In contrast to production, both imports and exports are always standardized, but only for the principal derived product (to avoid double-counting, as noted above) or for products that are not themselves part of another balance (in other words, wine trade would not be standardized back up into the grape balance). This standardization is achieved by dividing the import and export quantities for the derived commodities by their extraction rates, and then adding those standardized import and export values, respectively, together to obtain a standardized, aggregated value of imports and exports.

In this example, that means that only imports and exports of rolled oats will be standardized and aggregated to avoid double-counting the input quantity in the rolled oats production process, because bran of oats is a co-product of the transformation of oats into rolled oats.

This process begins by recalling equation (2-9). To convert rolled oats to its primary equivalent, the quantity of rolled oats must be divided by its extraction rate. For imports, the result is:

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

$\text{Primary commodity equivalent} = \frac{14\,074 \text{ MT rolled oats}}{\frac{0.65 \text{ MT rolled oats}}{1 \text{ MT oats}}}$	
$\text{Primary commodity equivalent} = 21\,652 \text{ MT oats equivalent}$	

Note: MT = metric tonnes

For exports, the same calculation can be performed:

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

$\text{Primary commodity equivalent} = \frac{2\,130 \text{ MT rolled oats}}{\frac{0.65 \text{ MT rolled oats}}{1 \text{ MT oats}}}$	
$\text{Primary commodity equivalent} = 3\,277 \text{ MT oats equivalent}$	

Note: MT = metric tonnes

It is suggested to then include all standardized values in a single table to aggregate the respective primary equivalent import and export totals, as below in table 4-15.

**TABLE 4-15. STANDARDIZATION OF COMMODITIES TO PRIMARY EQUIVALENT, AND SUBSEQUENT AGGREGATION.**

	Quantity	Extraction rate	Primary commodity equivalent
<b>Imports</b>			
Oats	188 219		188 219
Rolled oats	14 074	0.65	21 652
Oats equivalent total			<b>209 871</b>
<b>Exports</b>			
Oats	3 439		3 439
Rolled oats	2 130	0.65	3 277
Oats equivalent total			<b>6 716</b>

**Stocks**

In most cases, stocks are only estimated at the primary-commodity level, which would not require any standardization. However, in cases where stocks are reported for derived commodities (such as concentrated orange juice), they will have to be standardized. In this example for oats, because stocks are only estimated at the primary commodity level, no standardization is necessary.

**Food**

Values for food are always standardized and aggregated<sup>3</sup>. This is particularly important because, for many commodities, no food use occurs at the primary commodity level, such that without standardization and aggregation, food values in the final balance would appear as zero. In this example for oats, the food estimate is standardized as above, by dividing by the extraction rate.

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

$\text{Primary commodity equivalent} = \frac{65\,983 \text{ MT rolled oats}}{\frac{0.65 \text{ MT rolled oats}}{1 \text{ MT oats}}}$	
$\text{Primary commodity equivalent} = 101\,512 \text{ MT oats equivalent}$	

Note: MT = metric tonnes

**Food processing**

In most cases, food processing should not be standardized and aggregated into the final balance for the primary equivalent. Instead, it is dropped altogether to avoid double-counting, as the ultimate use of quantities processed into food commodities is as food. The only exception is in those cases where some processed product is standardized into another primary commodity equivalent of the FBS. For example, food processing of grapes into wine would not be standardized back up in to the grape balance, as there is a separate primary commodity balance for wine. In such cases, a quantity will remain in the grape balance under “food processing”.

<sup>3</sup> In the rare cases where a single transformation process produces more than one derived commodity used as food, compilers must choose which commodity to standardize and aggregate in order to avoid double-counting.

In the oats example, food processing can be dropped from the balance at this point, as there are no derived products that become separate commodity balances.

**Feed**

Feed quantities will not be standardized and aggregated for many commodities, as co-products of transformation processes (such as brans, germs or oilcakes) tend to be largely utilized as feed (if it is a useful exercise for countries, these quantities can be added together in a separate feed products balance). However, quantities estimated as feed for commodities that are the primary output of a transformation process should be standardized and aggregated.

With respect to the oats example, this means that the feed quantity for oat bran will not be standardized and aggregated. Rather, the feed quantity of the primary commodity (oats) will remain as the feed quantity in the primary commodity equivalent balance.

**Seed**

Because seed only comes from the primary commodity, this value will remain unchanged in the FBS primary equivalent balance.

**Net tourist food**

As with food, quantities of food available for consumption by tourists must be standardized by dividing by the extraction rate, as below.

$\text{Primary commodity equivalent} = \frac{\text{Quantity of derived product}}{\text{Extraction rate}}$	(2-9)
---	-------

$\text{Primary commodity equivalent} = \frac{750 \text{ MT rolled oats}}{\frac{0.65 \text{ MT rolled oats}}{1 \text{ MT oats}}}$	
$\text{Primary commodity equivalent} = 1\,154 \text{ MT oats equivalent}$	

Note: MT = metric tonnes

**Industrial use**

Any quantities of industrial use must also be standardized, provided that they are for the primary output product of a given transformation; the industrial use of co-products like brans or germs would not be standardized. In this example, industrial use is only recorded at the primary commodity level; therefore, no further calculation is necessary.

**Loss**

Finally, quantities allocated to loss need to be standardized. As with industrial use, in our example, loss has only been recorded at the primary level, such that no further calculation is necessary.

#### 4.4.2 Standardized and aggregated table

The rules for each of the variables in the balance sheet were reviewed above. In this example, the only quantities to be standardized and aggregated are those for imports, exports, food and tourist food. Food processing is dropped at this stage to avoid double-counting utilizations. For the remaining variables (stock change, feed, seed, industrial use, and loss), the quantities for the primary commodity (oats) will be entered into the (as yet unbalanced) primary commodity equivalent FBS table (table 4 16).

**TABLE 4-16. UNBALANCED FBS TABLE FOR OATS COMMODITY EQUIVALENT.**

Product	Production	Imports	Exports	Stock change	Food	Feed	Seed	Net tourist food	Industrial use	Loss
Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940

## 4.5 BALANCING

The balancing process is necessary to ensure that supply is equal to utilization in the FBS account after standardization and aggregation. As noted in the introduction to balancing given in section 2.4, the process of balancing primary-commodity-equivalent FBS balances comprises three steps: (1) calculate the imbalance from the supply = utilization identity; (2) distribute the imbalance; and (3) check that balanced quantities are within any set bounded values and rebalance, if necessary. These steps are outlined below.

### Step 1: Calculate the imbalance

This step is straightforward and should be carried out regardless of the distribution mechanism that country-level FBS compilers choose. To perform this calculation, it helps to recall equation (2-12).

$$\begin{aligned}
 \textit{Imbalance} = & \textit{Production} + \textit{Imports} - \textit{Exports} - \Delta\textit{Stocks} - \textit{Food} \\
 & - \textit{Food Processing} - \textit{Feed} - \textit{Seed} - \textit{Tourist Food} \\
 & - \textit{Industrial Use} - \textit{Loss} - \textit{Residual Use}
 \end{aligned}$$

This simple calculation is performed based on the values in the unbalanced FBS table (table 4-16 above in the example). Performing this calculation leads us to the imbalance noted on line B below in table 4-17.

**TABLE 4-17. CALCULATING THE IMBALANCE IN THE UNBALANCED FBS TABLE.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Stock change (4)	Food (5)	Feed (6)	Seed (7)	Net tourist food (8)	Industrial use (9)	Loss (10)	
A	Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6-A7-A8-A9-A10]											17 708

From here, country-level FBS compilers must choose which balancing approach they prefer.

**Step 2: Distribute the imbalance**

While different strategies for distributing the imbalance in the primary commodity equivalent are outlined in section 2.4.1, these guidelines will concentrate on the recommended “proportional error” approach detailed in section 2.4.1.1.

The first step in this process is using tolerance intervals and point estimates to quantify the error of each variable. Recall that tolerance intervals are assigned by variable a priori, based on the relative confidence in the data (sample tolerance intervals can be found in section 3.4.2.3, although countries should designate their own intervals after considering the quality of their available data). In this instance, official data were available for production, imports, exports and feed. As such, compilers in this country had a high degree of confidence in the data and assigned these variables tolerance intervals of zero. Data for food was imputed using a model; however, because food availability should change little from one year to the next, a low tolerance level of 10 percent was assigned. Similarly, seed use was imputed; however, as the imputation for seed implies a certain seeding rate, the confidence in the seed estimate should also be high. For this reason, a tolerance interval of 15 percent was assigned for seed. Data for stock changes, net tourist food, and loss were also imputed using models, although these quantities are much more likely to fluctuate. Consequently, compilers assigned each of these variables a tolerance interval of 30 percent. Finally, the data for industrial use were obtained using expert estimation. Country-level compilers had little confidence in this estimate and therefore assigned it a tolerance interval of 40 percent. These percentages were then added to the table in line C (table 4-18).

**TABLE 4-18. TOLERANCE INTERVAL PERCENTAGES ASSIGNED.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Stock change (4)	Food (5)	Feed (6)	Seed (7)	Net tourist food (8)	Industrial use (9)	Loss (10)	
A	Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6-A7-A8-A9-A10]											17 708
C	Tolerance interval (in %)	±0%	±0%	±0%	±30%	±10%	±0%	±15%	±30%	±40%	±30%	

Once the percentages have been added, the data in line A can be multiplied by the percentages in line C to obtain quantified error estimates in line D. These quantified errors are then summed to one another to obtain an aggregated error, calculated on line E (table 4 19).

**TABLE 4-19. ERRORS QUANTIFIED AND AGGREGATED ERROR CALCULATED.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Stock change (4)	Food (5)	Feed (6)	Seed (7)	Net tourist food (8)	Industrial use (9)	Loss (10)	
A	Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6-A7-A8-A9-A10]											17 708
C	Tolerance interval (in %)	±0%	±0%	±0%	±30%	±10%	±0%	±15%	±30%	±40%	±30%	
D	Error [D=A*C]	0.0	0.0	0.0	3 705.0	10 151.2	0.0	1 845.0	346.2	1 000.0	1 182.0	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6+D7+D8+D9+D10]											18 229.4

From here, the error is distributed proportionally. To do so, the proportion of error must first be calculated. This is achieved by dividing the individual variable errors in line D by the aggregated error estimate on line E (table 4 20).

**TABLE 4-20. PROPORTION OF AGGREGATED ERROR FOR EACH INDIVIDUAL VARIABLE CALCULATED.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Stock change (4)	Food (5)	Feed (6)	Seed (7)	Net tourist food (8)	Industrial use (9)	Loss (10)	
A	Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6-A7-A8-A9-A10]											17 708
C	Tolerance interval (in %)	±0%	±0%	±0%	±30%	±10%	±0%	±15%	±30%	±40%	±30%	
D	Error [D=A*C]	0.0	0.0	0.0	3 705.0	10 151.2	0.0	1 845.0	346.2	1 000.0	1 182.0	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6+D7+D8+D9+D10]											18 229.4
F	Proportion of aggregated error [F=D/E]	0.0%	0.0%	0.0%	20.3%	55.7%	0.0%	10.1%	1.9%	5.5%	6.5%	

Next, the quantity of adjustment for each variable is calculated by multiplying the proportion of aggregated error calculated in line F by the equation's imbalance in line B. Each of the unbalanced quantities from line A are then adjusted by the quantities in line G. It should be noted that the imbalance is positive. This means that there is more supply than there is demand. For this reason, the adjustments in all the demand variables (that is, everything except production and imports) will be positive. Had there been any adjustments in supply variables, the adjustments there would be negative.

The other point to note is that the imbalance in the equation is lower than the aggregated error. FBS compilers will find that this is not always the case: sometimes, the equation's imbalance may be much greater in magnitude than the aggregated error. This situation suggests that the bounds imposed through the tolerance intervals were excessively restrictive, and may indicate that one of the imputations (or even one of the official estimates) should have a larger tolerance interval. Unfortunately, there is no way of knowing exactly which variable is problematic. Such a situation should, however, be taken as a signal to FBS compilers that a closer examination of tolerance error percentages, and perhaps their reconsideration, may be warranted. Where the imbalance is greater than the aggregated error, FBS compilers should note that the adjustments will necessarily be larger than the estimated error.

In this case, however, the imbalance is lower than the aggregated error, such that the adjustments will be lower than the estimated errors (compare line G in table 4-21 to line D).

**TABLE 4-21. DISTRIBUTING THE IMBALANCE PROPORTIONALLY.**

Line	Product	Production (1)	Imports (2)	Exports (3)	Stock change (4)	Food (5)	Feed (6)	Seed (7)	Net tourist food (8)	Industrial use (9)	Loss (10)	
A	Oats primary equivalent	131 259	209 871	6 716	12 350	101 512	182 950	12 300	1 154	2 500	3 940	
B	Imbalance for A [A=A1+A2-A3-A4-A5-A6-A7-A8-A9-A10]											17 708
C	Tolerance interval (in %)	±0%	±0%	±0%	±30%	±10%	±0%	±15%	±30%	±40%	±30%	
D	Error [D=A*C]	0.0	0.0	0.0	3 705.0	10 151.2	0.0	1 845.0	346.2	1 000.0	1 182.0	
E	Aggregated error [E=D1+D2+D3+D4+D5+D6+D7+D8+D9+D10]											18 229.4
F	Proportion of aggregated error [F=D/E]	0.0%	0.0%	0.0%	20.3%	55.7%	0.0%	10.1%	1.9%	5.5%	6.5%	
G	Adjustment [G=B*F]	0.0	0.0	0.0	3 599.1	9 861.0	0.0	1 792.3	336.3	971.4	1 148.2	
H	Oats equivalent adjusted values [for (1) and (2), H=A-G, for remaining, H=A+G]	131 259.0	209 871.3	6 715.9	15 949.1	111 373.3	182 950.0	14 092.3	1 490.1	3 471.4	5 088.2	
I	Imbalance for H [I=H1+H2-H3-H4-H5-H6-H7-H8-H9-H10]											0.0

**Step 3: Ensure that balanced quantities are within any set bounded values**

This step serves as the final “reality check” on the now-balanced quantities. This step will depend upon the bounds that FBS compilers have set in their analysis of each commodity’s supply and demand situation. The most obvious check is that cumulative stocks are non-negative given the changes. Other hypothetical checks in our example situation could include things such as ensuring that the balanced estimate for seed implies a reasonable seeding rate. If any value is identified as outside of bounds, that value is set at the bound, assigned a zero tolerance interval, and the process is repeated.

## 4.6 NUTRIENT SUPPLIES AND CALORIE ESTIMATES

Nutrient supplies are then added, for each of the SUA level commodities, based on SUA-level food estimates before standardization. However, if the food quantity of the FBS primary-equivalent level was adjusted during balancing, then food quantities for all commodities in the account should first be scaled up by the same percentage. That is, a scaling factor should be calculated that compares the adjusted food value to the food value in the unbalanced identity. This scaling factor is simply the ratio of the adjusted food quantity to the food quantity before balancing, as below in equation (4-1).

$$\text{Scaling Factor} = \frac{\text{Adjusted FBS Food Value}}{\text{Beginning FBS Food Value}} \quad (4-1)$$

$$\text{Scaling Factor} = \frac{111\,373.3}{101\,512}$$

$$\text{Scaling Factor} = 1.09714$$

Note: MT = metric tonnes

Then, once this scaling factor has been calculated, all food quantities for each commodity in the SUA should be multiplied by this scaling factor. In the oats example, because the final balanced primary equivalent-level FBS food quantity rose compared to the unbalanced quantity, the food quantity for rolled oats must first be scaled up before the quantity can be converted into its nutrient equivalents (table 4-22).

**TABLE 4-22. FOOD QUANTITIES SCALED UP BASED ON BALANCED FBS.**

Product	SUA Food Quantity (A)	Scaling Factor (B)	Scaled Food Quantity (C) (C=A*B)
Rolled oats	65 983	1.09714	72 393

After the calculation of the scaled food quantity, the nutrient conversion can proceed. At this stage, analysts should pay careful attention to ensure that the units are correct. For example, nutrient tables may be published in kcal per 100 grams (as is the case with the FAO reference nutrient table), or even in kilojoules. Deriving a quantity for daily nutrient availability values is a simple conversion, whereby the volumes of food are multiplied by their nutritional unit values, and then divided by the number of days in a year.

In addition, analysts should pay careful attention as to whether their nutrient conversion factors take into account inedible parts. If their nutrient factors are on the basis of only edible product weights, then the food quantities (which, it is recalled, are expressed in gross product weight including both edible and inedible parts) must first be converted to the edible quantity of food through the application of a refuse factor. As the nutrient conversion factors published by the FAO Statistics Division (and utilized in this example) are calculated on the basis of gross weight (taking into account the portion of each commodity that is inedible), this step is not necessary here.

In the oats example, the nutrient table could be organized as below. Country-level FBS compilers may wish to compile it differently (using different units, for example). However, this organization is intended to stress the

importance of the units in the conversion process. First, the commodities for which food utilizations were recorded are added to the table. In this example, rolled oats was the sole commodity utilized for food; therefore, it alone is added to the table (table 4 23). Then the scaled-up food quantities (calculated above in table 4 22) are added. For ease of comparison, the quantities in tonnes are first converted to grams (column 1 to column 2). Again, this is because the nutrient values (included in the orange columns) are all indicated in grams. To calculate a country's daily total availability for the different nutrients derived from each commodity, the value in column 2 is then multiplied by the nutrient values in columns 3 through 5 and divided by 365, resulting in the data in columns 6 through 8. It should be recalled that these are aggregated, country-level estimates. Per capita estimates will be defined in the final step.

**TABLE 4-23. CONVERTING FOOD QUANTITIES INTO NUTRIENT VALUES.**

Line	Product	Quantity Food (tonnes) (1)	Quantity Food (g) (2)	kcal energy/g (3)	Protein/g (4)	Fat/g (5)	Energy (kcal/day) [A6=A2*A3/365] (6)	Protein (g/day) [A7=A2*A4/365] (7)	Fat (g/day) [A8=A2*A5/365] (8)
A	Rolled oats	72 393	72 393 000 000	3.84	0.16	0.063	761 614 027	31 733 918	12 495 230

For many commodity trees, it is likely that several commodities in the tree will be consumed as food. Thus, after food quantities are converted into nutrient values, the next step is to add up the total nutrients for the commodity tree. In this example, a food quantity was only recorded for rolled oats. Therefore, the total energy, protein and fat for the whole oats tree comes from rolled oats, as can be seen when line B (total nutrients from the oat tree) is compared to line A (nutrients from rolled oats alone) (table 4 24).

**TABLE 4-24. NUTRIENT VALUES ARE SUMMED TO ONE ANOTHER.**

Line	Product	Quantity Food (tonnes) (1)	Quantity Food (g) (2)	kcal energy/g (3)	Protein/g (4)	Fat/g (5)	Energy (kcal/day) [A6=A2*A3/365] (6)	Protein (g/day) [A7=A2*A4/365] (7)	Fat (g/day) [A8=A2*A5/365] (8)
A	Rolled oats	72 393	72 393 000 000	3.84	0.16	0.063	761 614 027	31 733 918	12 495 230
B	<i>Total national nutrient availability for oats and products [B=sum of nutrients for all products used for food in the account]</i>	-	-	-	-	-	761 614 027	31 733 918	12 495 230

## 4.7 DERIVING PER CAPITA ESTIMATES

The final step is to convert the national aggregate nutrient estimates into per capita equivalents. This is accomplished by dividing the national nutrient totals in line B by population in line C to obtain per capita estimates of nutrient availability in line D (table 4 25).

**TABLE 4-25. PER CAPITA NUTRIENTS CALCULATED.**

Line	Product	Quantity Food (tonnes) (1)	Quantity Food (g) (2)	kcal energy/g (3)	Protein/g (4)	Fat/g (5)	Energy (kcal/day) [A6=A2*A3/365] (6)	Protein (g/day) [A7=A2*A4/365] (7)	Fat (g/day) [A8=A2*A5/365] (8)
A	Rolled oats	72 393	72 393 000 000	3.84	0.16	0.063	761 614 027	31 733 918	12 495 230
B	<i>Total national nutrient availability for oats and products [B=sum of nutrients for all products used for food in the account]</i>	-	-	-	-	-	761 614 027	31 733 918	12 495 230
C	Population						38 360 000	38 360 000	38 360 000
D	<i>Per capita average national nutri-ent availability for oats and prod-ucts [D=B/C]</i>						19.85	0.83	0.33

Once these nutrients are calculated, they can be added to the primary commodity equivalent supply and utilization table to obtain the final food balance for oats and products (table 4 26).

**TABLE 4-26. BALANCED FBS TABLE FOR OATS.**

Product	Domestic Supply				Domestic Utilization						Per capita supply		
	Production	Imports	Exports	Stock change	Food	Feed	Seed	Net tourist food	Industrial use	Loss	Energy (kcal/day)	Protein (g/day)	Fat (g/day)
Oats and products	131 259	209 871	6 716	15 949	111 373	182 950	14 092	1 490	3 471	5 088	19.85	0.83	0.33

It should be noted that this is the table for only one primary commodity equivalent. The process outlined above would then have to be repeated for all food commodities, including fisheries commodities, and then combined into one comprehensive national FBS<sup>4</sup>. When the lines for all commodities have been included, the per capita daily supplies of energy, protein and fat for the individual commodities can then be added up to calculate total daily supplies. The total obtained for energy is the daily dietary energy supply (DES).

## 4.8 VALIDATION AND TROUBLESHOOTING UNFEASIBLE SOLUTIONS

In the course of compiling FBS and once the accounts are completed, FBS compilers should embark on a process of validating their estimates. These processes should be undertaken with the goal of addressing the following: are these estimates feasible? If not, what is the best manner in which to correct this estimate to improve its feasibility? Validation processes will vary from country to country. However, some of the most common checks are described below.

### Extremely large changes in food availability/DES

As noted multiple times throughout these guidelines, overall food availability (as expressed through the DES) should be fairly stable from one year to another, not varying by more than 50 kcal, and very rarely by more than 100 kcal (which is most likely to occur in cases of war, famine or other crises). Countries should analyse their final DES total and consider the result in the context of a time series of data. Feasibility may be judged in terms of an absolute threshold change or a percentage change from the previous year.

In addition to the overall DES, per capita calories for particular products or product groups can be a useful measure of feasibility in and of itself. Analysts should use their understanding of their particular country circumstances in this process. However, some suggestions are made below.

- The overall makeup of the DES by commodity category should change little from one year to another. In other words, the percentage of calories from different commodity groups is unlikely to vary by more than a percentage point or two on an annual basis. If larger jumps are identified, analysts should check the annual changes in the most important products in that basket first, to pinpoint where revisions may be necessary. If, however, the annual jumps or declines are justified, analysts should make a note of the justification.

<sup>4</sup> It is noted that while these guidelines have not explicitly covered fishery commodities, the approach to compiling FBS for agriculture versus fishery products is the same. In fact, many of the same imputation and estimation approaches may also be applicable to fishery goods. However, even in cases where calories derived from fisheries are estimated in a different way, it should nonetheless be combined with the estimate from agricultural goods to attain a total DES.

- Calories may change significantly for foods that are highly substitutable; however, the overall calories level derived from substitute products should not change substantially. For instance, the total calories derived from vegetables should remain stable, and the calories derived from fresh green beans may fall by half while the calories sourced from green peas and carrots both rise.
- If the FBS suggest that there has been a great change in the food availability of a given good in a particular year, one way that analysts can verify the change is by examining changes in consumer prices: if prices have climbed dramatically, a large reduction in food availability is more feasible (and vice versa).

### **Large, improbable percentage changes in any other variable from year to year**

As with food availability, many of the other utilization variables are unlikely to experience huge changes from one year to the next, except in very particular circumstances. The magnitude of the feasible annual percentage change will depend upon the circumstances of the country and the product, however; it is therefore key that FBS compilers strive to understand their country's market dynamics as holistically as possible. Country-level FBS compilers are advised to calculate year-over-year percentage changes in all supply and utilization variables and then assess those changes on a case-by-case basis.

Analysts should be advised that large annual percentage changes in balance sheet values may be completely valid. For example, while a doubling of production of a staple food such as wheat is highly improbable from one year to the next, the doubling in production of a highly specialized product such as a minor fruit crop may be possible (particularly in a small country). In the same vein, a doubling of the industrial utilization of a given commodity from one year to the next may appear suspect, but it may also be feasible if a new utilization was found or a new processing facility was constructed.

The conclusion should be that annual percentage changes are a useful indicator that something in the balance sheet may be incorrect, but should not be interpreted as an absolute indicator that something is necessarily incorrect.

### **Identifying production of new commodities from trade discrepancies**

One of the balancing constraints noted in section 2.4.2 is the row constraint: production net of trade and stock changes must be positive. The rationale behind this constraint is obvious: a country cannot export more than it is producing, importing and removing from stocks. This intuitive concept can be a useful way for countries to identify growing production of commodities that may not be included in annual production surveys (typically, minor specialty crops). If no production is reported for a particular commodity and exports exceed imports, then that is an indication that production is occurring, and should either be surveyed or a production quantity should be imputed to cover the trade deficit.

FBS compilers should note that this constraint is commonly used to identify production of derived commodities in the absence of any information on the production of derived commodities. For example, if a country grows tomatoes and suddenly begins to export canned tomatoes, this is an indication that some of the fresh tomatoes are being processed into canned tomatoes.

The constraint can also be useful in identifying the likely production of new commodities, particularly with regard to new horticultural products that are grown exclusively for export markets.

### **No balanced solution possible, given all constraints**

Although this is improbable if the proportional imbalance balancing mechanism is used, the existence of constraints on specific variables opens the possibility that no solution is found. As an example, consider a country that measures all variables, save for stocks, and assumes that all the other values have zero tolerance intervals. After balancing, compilers note that the calculated withdrawals from stocks are greater than the estimated stock levels – an impossibility. How can such a situation be resolved?

Examples such as these are an indication that the bounded values are excessively stringent. In the above example, two scenarios are possible. First, it is possible that the estimated stock levels were incorrect, and there are sufficient stocks to permit the calculated withdrawal. Second, it is possible that the assumed tolerance intervals for the other variables was excessively stringent and must be relaxed for the equation to balance. This determination will have to be made at the discretion of the country-level FBS compilers.

## **4.9 VALIDATION THROUGH WORKING GROUPS**

In most countries, the majority of the work of compiling FBS is likely to fall to a particular person, or a specific group within a designated agency. However, because the supply and demand situation for commodities is dynamic, it is recommended that the balance sheets for the major commodities be validated by the TWG suggested in section 3.2. This process of validation through working groups has various advantages that will contribute to the long-term accuracy and utility of the FBS compilation process. First, group review provides FBS compilers with additional reviewers, to help spot either data inconsistencies or unfeasible solutions that may escape the notice of compilers who may have less expertise with the product in question. Second, the review of draft FBS by product experts ensures that FBS compilers are kept abreast of developments in the commodity, ensuring that any new uses are identified so that a strategy can be devised to properly account for these new uses within the FBS framework. Finally, the process of validating FBS through working groups can also facilitate greater buy-in to the overall FBS process from users and stakeholders. Allowing relevant groups to contribute to the process, voice their opinions and understand how the various estimates are devised or where the data comes from helps to foster a sense of ownership among stakeholders. While a sense of ownership is valuable in and of itself, it can also benefit the FBS process by convincing relevant parties of the value in contributing data to better inform the exercise. Outside stakeholders may also be more likely to promote the final product if they have been able to contribute to the process.

## **4.10 SUMMARY**

This chapter has educated users about the process of constructing an FBS through the step-by-step working of a practical example. This example has walked users through the filling in of the SUA table, the balancing of SUA accounts for derived products, the processes of standardizing and aggregating the accounts, the balancing mechanism for the FBS primary-equivalent product, the calculation of nutrient supplies, and the derivation of per capita estimates. The chapter closed by offering some suggestions for basic FBS validation, while also encouraging that final FBS validation be undertaken by a TWG. Now that users are comfortable with the process of producing FBS, chapter 5 will offer some final considerations on data quality, data dissemination and FBS interpretation.

# 5

## Data quality considerations, dissemination and FBS interpretation

### 5.1 INTRODUCTION

As emphasized throughout these guidelines, FBS are analytical data sets that strive to integrate, reconcile and cross-validate data from multiple sources. This is a challenging undertaking, and is frequently plagued by either an absolute lack of data, or the imperative to reconcile data of poor quality. Extensive parts of these guidelines have been devoted to the idea of data quality: that is, not every estimate is precise or reliable to the same degree. However, there are additional dimensions to data quality that FBS compilers should consider, both to ensure the integrity of their FBS preparation process, and to facilitate the dissemination and interpretation of the final FBS data set by users. This chapter briefly covers these data quality considerations before turning to recommendations on data dissemination, ending with some final notes intended to facilitate interpretation of FBS data.

## 5.2 QUALITY CONSIDERATIONS

The “garbage-in, garbage-out” philosophy is an oft-repeated caution for those working in the field of data analysis, and holds true for the process of FBS compilation as well. However, it is also true that good quality input data alone is not sufficient to ensure that a given statistical output will be successful. The concept of “data quality” has as much to do with the processes devoted to data production as it does with the final product itself. In this regard, FBS compilers should conduct their work keeping the following five quality dimensions in mind<sup>1</sup>:

- Relevance
- Accuracy and reliability
- Timeliness and punctuality
- Coherence and comparability
- Accessibility and clarity

The specific manner in which these five dimensions relate to the compilation of FBS is elaborated below.

### **Relevance**

FBS should not be compiled as an end in itself. National-level FBS compilers should begin the process of FBS compilation with a goal for how the data will be used and analysed by stakeholders. To determine the needs of stakeholders, initial consultations with users may be held, or user surveys may be conducted once the first year of data has been produced and distributed.

It then follows that FBS should be compiled with the intention of meeting the identified needs of current and future users. In the context of the FBS, this may mean covering only the commodities that are most relevant to stakeholders (for example, the top 25 most consumed products instead of all products). This may also mean computing certain indicators from the data to ease interpretation.

### **Accuracy and reliability**

In the realm of data quality, “accuracy” refers to how close an estimate is to its true value (which can be a consequence of either variance or bias), while “reliability” instead focuses on the magnitude of revisions, or how close initially reported values are to subsequent or final data. Both concepts are relevant to FBS data. However, the concept of accuracy deserves special emphasis. One of the primary motivations for updating the FBS methodology within these guidelines was a desire to improve the transparency of the accuracy of individual estimates, and to take advantage of available information on the perceived accuracy of a given point estimate to balance the supply = utilization identity in a way that eliminates accumulated error in a single balancer. As such, with respect to accuracy, the two most relevant suggestions of these guidelines are that countries (1) measure or approximate, document, and publish error estimates; and (2) develop and institute a data flag system that succinctly communicates information on the source of the data (and therefore some information on implied accuracy) to users. Recall from section 3.4.2.2 that data should be published with flags indicating the data source. Although it may not be feasible to also communicate the tolerance interval of the estimate in the table, this information should be published, either in footnotes to the table, or in the accompanying metadata that describes the processes of compiling the FBS. How this is communicated is at the discretion of country-level FBS compilers. In any case, two examples are provided below.

The first suggested layout is the addition of a column after each variable denoting the flag. The flag should then be explained, either within the table notes or as additional cells within the table itself. In the example below in table 5 1, the notes are included below the table.

---

1 These five dimensions of quality come from the FAO statistical quality assurance framework, and are equally relevant to country-level agricultural statistics systems. The full framework is available here: <http://www.fao.org/docrep/019/i3664e/i3664e.pdf>. Accessed on 19 January 2017.

**TABLE 5-1. SAMPLE PRESENTATION 1 OF DATA TABLE INCLUDING FLAGS.**

Year	Product	Production	Flag	Imports	Flag	Exports	Flag	Food	Flag	Loss	Flag	Energy (kcal/day)	Protein (g/day)	Fat (g/day)
2008	Onion	2 000	T	58 000		1 000		72 000	I	6 000	I	38	9	4

NOTE: Flags indicate the source of the data, as follows:  
 " " indicates that the data come from an official source.  
 "T" indicates that the data come from a semi-official source.  
 "I" indicates that the data are an imputed value.

An alternative presentation of the same data does not present the flag in columns. Rather, it refers the user to table footnotes that provide further explanations on the data source. In this presentation, it is also feasible to include information on the tolerance interval assigned to each variable (table 5 2).

**TABLE 5-2. SAMPLE PRESENTATION 2 OF DATA TABLE DETAILING SOURCES AND TOLERANCE INTERVALS IN TABLE FOOTNOTES.**

Year	Product	Production <sup>a</sup>	Imports <sup>b</sup>	Exports <sup>b</sup>	Food <sup>c</sup>	Loss <sup>d</sup>	Energy (kcal/day)	Protein (g/day)	Fat (g/day)
2008	Onion	2 000	58 000	1 000	72 000	6 000	38	9	4

a Data come from a semi-official source. This estimate has an assigned tolerance interval of  $\pm 10\%$ .  
 b Data come from an official source. This estimate has an assigned tolerance interval of  $\pm 0\%$ .  
 c Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .  
 d Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .

Of course, there are myriad other ways in which these data can be presented. The key element to note is that whatever presentation is chosen should include data flags or indications of data sources, and may also communicate assigned tolerance intervals to better convey data accuracy to users. It bears repeating that if tolerance intervals are not communicated to users in data tables, further information on them should be included in the table's accompanying metadata.

While accuracy refers to the closeness of an estimate to its true value, reliability instead refers to how close an initial estimate is to subsequent estimates. This is relevant to FBS because compilers should always be open to updating their FBS as new information becomes available – perhaps trade data has been revised, or a country may have begun a new program of measurement for loss and has found that estimates in previous years were much too high, necessitating back-year revisions. In all cases, FBS compilers should compare the initial estimates to subsequent ones and analyse their statistical processes to check whether initial estimates can be improved. In addition, it is important that data producers are transparent about revision policies and timelines. At a minimum, users should be notified of revisions (and, if relevant, the justification for the revisions).

### **Timeliness and punctuality**

In the context of data dissemination, “timeliness” indicates how long after the reference year a dataset is published, while “punctuality” refers to releasing the data on a previously indicated date.

When publishing FBS data, compilers typically face a trade-off between timeliness and accuracy: on the one hand, there is a desire to release data to users as soon as possible, although doing so may mean that not all data has had a chance to be properly validated. In the context of the FBS, much of the underlying data is produced on a lag. As an example, official production data for 2015 may not be released until 2017. Even then, that data may be revised in 2018 or 2019 as more information becomes available. In these instances, the choice is between whether to wait for the ostensibly more accurate final data at the risk of the data becoming irrelevant to users, or release more timely data that may be subject to large revisions.

To best meet data user needs, it is suggested that country-level FBS compilers aim for releasing their preliminary data sets the year after the FBS reference year. However, this does not preclude compilers from revising data sets should new information become available. In that case, preliminary data can be labelled as such, to differentiate it from finalized estimates.

With respect to punctuality, it is recommended that compilers publish an expected date of delivery on their agency’s external website, and then strive to deliver the product by that date. Workflow planning can assist in achieving this goal.

### **Coherence and comparability**

The concepts of coherence and comparability are highly relevant to FBS compilation, with “coherence” referring to the ability of statistical outputs to be combined with other data or utilized for various purposes, while “comparability” refers to the extent to which data is comparable between areas or across time.

First, focusing on coherence, the underlying structure of the FBS data makes it highly likely that this data will be coherent with other domains. In fact, the data contained within FBS will in most cases be sourced from a variety of domains in the first place. As long as data sources and characteristics are properly documented, users should find that FBS estimates are coherent with most other agricultural data sets. Of course, this will only be the case in countries that produce unique data sets for the different supply and utilization elements. Where a country produces multiple production estimates, for example, FBS production estimates will most likely be based on only one of the sources, and thus will not be coherent with the other. In these cases, FBS compilers should justify why one source was preferred over another.

With respect to comparability, as emphasized in these guidelines, the nature of FBS compilation is that data from various domains is identified and combined in such a way as to make it comparable. It should be recalled from section 3.4 that within the initial data assessment that precludes the process of putting the FBS together, analysts must first ensure that data is comparable in three key dimensions:

- **Product comparability:** data for each of the variables in the balance must refer to the same product. This process can be greatly aided by the use of numerical product classifications such as the CPC.
- **Unit comparability:** data for all variables should be elaborated in the same unit. For most countries, this is the tonne. Compilers should use internationally recognized conversion factors if any data are to be converted into the unit in which the balance is being compiled.
- **Temporal comparability:** all data should refer to the same reference period, whether a calendar year or a marketing year.

Even for the most diligent FBS compiler, full data comparability may not be possible. In these cases, at the very least, these differences should be documented so that users are aware of the data’s limitations.

### **Accessibility and clarity**

“Accessibility” can be defined as the ease with which users are able to obtain the data in question, while “clarity” refers to the availability of adequate documentation that explains the data set to the user. Additional guidance on data dissemination and accessibility is provided below in section 5.3; the focus of the present section will be on clarity.

Clarity covers all aspects that users must know to understand the data. This includes metadata, data quality indicators and advice on data interpretation. When disseminated to a general audience, it is recommended that the data be accompanied by the following:

- Metadata detailing various aspects of the data set, including units, time period covered, classifications and underlying data sources (for ease of reference, a separate discussion on metadata is given below in section 5.3.3).
- Quality reporting detailing how much of the data has been imputed, versus how much comes from official sources. Essentially, this reporting should allow users to understand the potential limitations of the data in a transparent way, as well as provide compilers themselves with an indication of how the data set can be improved in future.
- Public documentation of the general FBS compilation methodology utilized should be posted alongside the data set to facilitate user understanding of how the data set was compiled. While it is unnecessary to externally publish the cell-by-cell documentation of any edits or revisions, it is considered best practice to internally keep this type of detailed records. This document can be referred to if there are questions on a particular estimate, or if new information surfaces or an improved methodology is developed to better estimate a given value.

## **5.3 DISSEMINATION**

### **5.3.1 The case for wider dissemination of FBS data**

As noted elsewhere in these guidelines, FBS provide countries with a unique and valuable tool for analysis and policy-making. However, such analysis need not take place solely within the confines of the group that compiles the FBS. Indeed, the FBS data set should be considered a valuable public resource. For this reason, country-level compilers should go beyond merely compiling FBS, and also plan to disseminate the data to the public. FBS data sets should enter into the groups of data sets that are most likely already published on either NSO or ministry of agriculture websites (such as data sets on agricultural production and trade). If the country produces an annual agricultural statistical abstract, the FBS data tables should also be considered for inclusion in that publication.

While posting the data sets to an official website is an excellent first step, country-level FBS compilers are encouraged to go beyond publishing the data and provide some basic data analysis and context to better inform both policy-makers and laypeople. If the FBS are drafted by a purely statistical unit, this exercise may require first establishing a linkage or coordination mechanism between the statistical group and other researchers or analysts within the country’s ministry of agriculture, NSO or external research institutions.

Even the most basic accompanying explanatory text can be useful. For example, if wheat imports were shown to fall dramatically from one year to the next, some analysis on why that occurred might be helpful to users. To carry this example further, if calories from wheat-derived goods were to fall as a result of the contraction in imports, it may be useful to point out that calories from rice (another staple cereal good) increased in the same year, indicating that consumers substituted rice for wheat to maintain a stable overall cereal consumption level from year to year.

Particularly if carried out in conjunction with analytical units, the drafting of these types of analytical reports should be relatively straightforward. Moreover, the potential benefits of doing so are numerous, including increased

visibility of the data product, increased literacy of the data product among all users, and increased demand for the data product as users in both academia and the general public become accustomed to the types of analyses that can be undertaken using FBS data.

### 5.3.2 Suggested presentation

FBS data can be presented in any number of ways (depending upon the policy and data priorities of the country), because such a wealth of data is contained within the FBS. However, at least two presentations are suggested: data tables and data visualizations.

With respect to data tables, information can be presented in a number of layouts. However, two of the most common layouts are by year and by commodity. Because one of the primary objectives of FBS compilation is to illustrate the complete dietary picture of a country at a given period in time, the first data table layout that compilers should consider presenting is the single-year portrait, whereby each primary commodity equivalent is displayed as a separate line in a single comprehensive table (table 5-3). It is suggested that on the first line of the table, the sum of total nutrient levels from all foods be displayed. Compilers may also choose to display certain aggregates in their tables, such as cereals, tubers, fruits or meat. Compilers may also consider including an additional data flag or note to indicate where data has been standardized or aggregated.

**TABLE 5-3. SAMPLE TABLE LAYOUT OF FBS DATA ORGANIZED BY YEAR.**

Year	Product	Domestic supply				Domestic utilization				Per capita supply		
		Production <sup>a</sup>	Imports <sup>b</sup>	Exports <sup>b</sup>	Stock change <sup>c</sup>	Food <sup>d</sup>	Feed <sup>e</sup>	Seed <sup>f</sup>	Loss <sup>g</sup>	Energy (kcal/day)	Protein (g/day)	Fat (g/day)
2007	Total All Foods									2 863	65	72
2007	Maize and products	15 000	65 000	0	0	71 000	3 000	1 000	5 000	38	9	4
2007	Potatoes and products	0	23 000	1 000	-6 000	27 000	0	0	1 000	14	1	0
2007	Sunflower seed oil	0	17 000	0	3 000	14 000	0	0	0	7	0	20
2007	Tomatoes and products	0	11 000	0	0	11 000	0	0	0	3	0	0
2007	Bananas	0	6 000	0	0	6 000	0	0	0	5	0	0
2007	Milk excluding butter	113 000	86 000	0	-7 000	172 000	19 000	0	8 000	144	8	7
2007	(Etc.)	-	-	-	-	-	-	-	-	2 652	47	41

NOTE: Data for all commodity groups represents standardized and aggregated values of commodities derived from the indicated primary commodity. Further information on commodity trees can be found in the metadata.

a,b Data come from an official source. This estimate has an assigned tolerance interval of  $\pm 0\%$ .

c,d Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .

e Data come from a semi-official source. This estimate has an assigned tolerance interval of  $\pm 10\%$ .

f,g Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .

While the data tables by year inform users about the complete dietary picture (aside from supplying a snapshot of supply and utilization for each commodity in a single year), the table layout of data by commodity allows users to analyse the time series to compare how the different aspects of supply and demand have shifted over a given number of years. A sample table layout of data organized by commodity may be seen below in table 5 4.

**TABLE 5-4. SAMPLE TABLE LAYOUT OF FBS DATA ORGANIZED BY COMMODITY.**

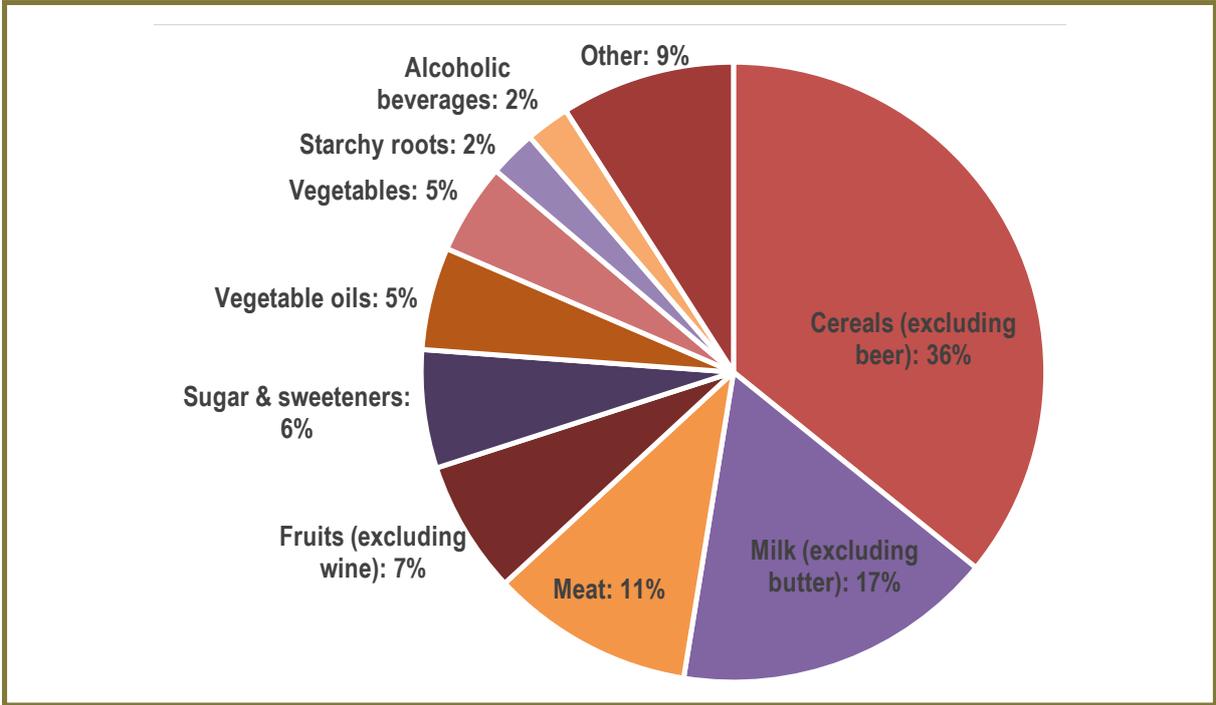
Year	Product	Domestic supply				Domestic utilization				Per capita supply		
		Production <sup>a</sup>	Imports <sup>b</sup>	Exports <sup>b</sup>	Stock change <sup>c</sup>	Food <sup>d</sup>	Feed <sup>e</sup>	Seed <sup>f</sup>	Loss <sup>g</sup>	Energy (kcal/day)	Protein (g/day)	Fat (g/day)
2008	Maize and products	2 000	58 000	1 000	-23 000	72 000	3 000	1 000	6 000	38	9	4
2009	Maize and products	9 000	76 000	0	5 000	69 000	4 000	2 000	5 000	36	8	3
2010	Maize and products	12 000	75 000	4 000	-3 000	74 000	3 000	2 000	7 000	38	9	4
2011	Maize and products	16 000	87 000	13 000	9 000	68 000	3 000	2 000	8 000	35	8	3

NOTE: Data for all commodity groups represents standardized and aggregated values of commodities derived from the indicated primary commodity. Further information on commodity trees can be found in the metadata.  
a,b Data come from an official source. This estimate has an assigned tolerance interval of  $\pm 0\%$ .  
c,d Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .  
e Data come from a semi-official source. This estimate has an assigned tolerance interval of  $\pm 10\%$ .  
f,g Data are an imputed value. This estimate has an assigned tolerance interval of  $\pm 25\%$ .

It should also be noted that the format suggestion here is to publish only the FBS primary-commodity-equivalent accounts. If countries wish to increase the level of detail available to users, they may consider publishing the more detailed SUA-level accounts as well. However, this should only be done if countries have explicitly outlined the methodology by which SUAs are standardized and aggregated to the primary-equivalent level to prevent confusion by users.

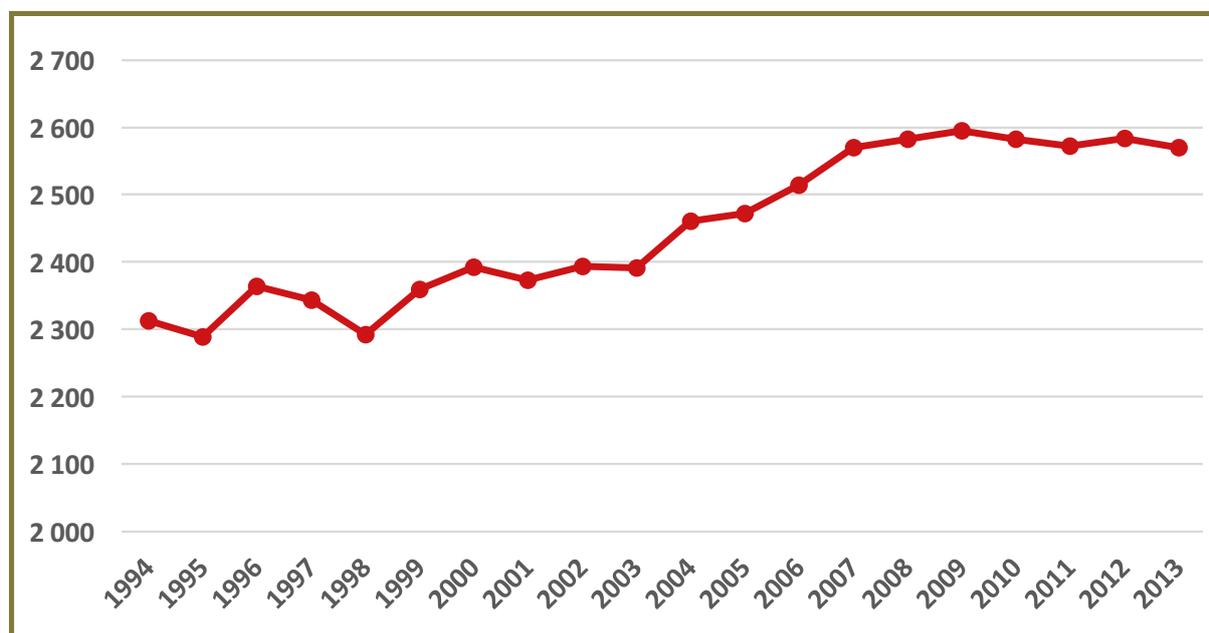
Data visualizations, in the form of graphs or charts, are another option for data presentation. The advantage of such visualizations is that they can communicate a great deal of information to users in an immediate, straightforward manner. As above, visualizations can be used to describe data at a given point in time, or can illustrate changes across time. For example, in addition to displaying a table that lists the per capita energy supply from each commodity group, a pie chart can be used to convey the information, as below in figure 5-1.

**FIGURE 5-1. SAMPLE VISUALIZATION OF FBS DATA, ILLUSTRATING DIETARY COMPOSITION BY COMMODITY GROUP IN A GIVEN YEAR.**



Illustrating changes in data across time is another possibility that data visualizations can accommodate, as they can allow users to better understand the overall trend behind any given variable. While these graphics can be published for any variable in the balance sheet, given the emphasis placed on the DES for policy formation and analysis, a line graph of this variable is an obvious choice, as below in figure 5 2.

**FIGURE 5-2. SAMPLE VISUALIZATION OF FBS DATA, ILLUSTRATING CHANGING DES FROM 1994 TO 2013.**



### 5.3.3 Metadata

Metadata is underlying information that helps users to better understand a given data set. For every data set produced, it is essential that it be accompanied by metadata so that users are aware of exactly which underlying concepts and processes went into the production of said data. This includes definitions of variables or commodities, clarification on units of measurement, definition of the reference period, a listing of the underlying data sources, an explanation of the methodology behind the data set's compilation, and some indication of data quality. The Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS) has produced a Common Metadata Framework (including publications on *Statistical Metadata in a Corporate Context* and *Metadata Concepts, Standards, Models and Registries*) that countries may find helpful as they seek to establish and implement their own metadata frameworks according to international best practices<sup>2</sup>.

In the context of FBS, this means indicating to users which commodity classification is being used, how the different supply and utilization variables are defined (for example, whether production excludes backyard production), a discussion of the underlying methodology, a disclosure of the sources for the different data components, a list of the assumptions behind extraction rates and processing shares, and an assessment of the quality of the data (which can be based on flags, for example). In some cases, it may be sufficient to provide a link to a background document detailing the methodology under which the data was produced.

All of the data sets available on FAOSTAT are published along with their metadata. These metadata can serve as a template for country-level FBS compilers who are not sure as to exactly what kind of detail to include when publishing their own metadata<sup>3</sup>.

<sup>2</sup> All of these documents are available from UNECE's website: <http://www1.unece.org/stat/platform/display/metis/The+Common+Metadata+Framework>. Accessed on 26 April 2017.

<sup>3</sup> See, for example, metadata for the crops production data set, available at: <http://www.fao.org/faostat/en/#data/QC/metadata>. Accessed on 13 February 2017.

## 5.4 INTERPRETATION

While FBS data can be used in numerous applications (including econometric studies, trend analyses, investigations of dietary content, and nutritional adequacy analyses), the data should be understood for what they are: an analytical data set that helps to capture a holistic snapshot of food supply and utilization for any one country at a given point in time. As such, to facilitate the accurate interpretation of FBS, several of their characteristics bear repeating.

### **Expected values, not point estimates**

FBS are useful for capturing general trends in food availability, both in the aggregate and at the individual commodity level. However, because input data for each variable in the commodity balance will have been estimated with a certain margin of error, it is incorrect to treat the derived estimates for calories per capita per day for an individual commodity (not to mention the total DES) as a point estimate. In reality, each of these derived estimates is an expected value as well, with some unknown error distribution. Therefore, while a country's FBS may estimate per capita availability of carrots at 9 calories, for example, the more accurate interpretation may be that availability is more likely somewhere between 7 and 11 calories per person. This does not make this estimate of 9 calories inaccurate, or useless; it remains useful when compared to likely values of food availability over the previous years (to track availability trends over time), or in comparison to other products within the country's balance sheet (that is, per capita availability of carrots is higher than per capita availability of green beans, which for the same country was estimated at 3 calories per capita). The same can be said of the overall DES: DES estimates should be understood as likely values, but not as exact values.

At the same time, if it is assumed that error distributions do not change significantly from one year to the next, then basing analyses on changes in availability (comparing one year to the previous year) should reduce the effect of any bias in measurement. In other words, looking at the change in variables can reduce the effect of systematic errors on interpretation of FBS estimates.

### **Measurement of apparent consumption**

In contrast to household surveys, which specifically measure the consumption of a good, FBS measure apparent consumption. In simple terms, this means the amount of food that was available for consumption, whether it was consumed or not. As such, DES estimates should be understood not as a precise value of per capita food intake, but rather as the amount of food available for human consumption within a country on a per capita basis. There are many reasons why the DES may not be equivalent to average levels of effective consumption. First, balance sheets do not take into account distributional differences in consumption (that is, the wealthy in a country are likely to consume more than those in poverty). Although a certain amount of food may be considered "available" for consumption by all in an FBS setting, in reality, only a small segment of a country may have economic access to that food. Second, balance sheets do not take into consideration the spatial distribution of food within a country. Where most production of a given food item takes place in one part of the country but the bulk of consumers are elsewhere and do not have physical access to this food item, balance sheets will still show that the food product is "available" to the average person. Last, FBS estimates for food availability are assumed to be calculated at the retail level. This means that food waste at both the retail and the consumer levels are contained within the estimate of food availability. In countries with substantial consumer food waste problems, using DES as a proxy for effective consumption would lead to a large overestimation in the amount of calories consumed by the average citizen.

### **Context**

Barring war, disease or natural disaster, overall consumption (as well as consumption of staple foods and dietary shares sourced from a given food group) is likely to change little from one year to another. However, that does not mean that large year-to-year swings will not be seen in particular food items. While the FBS framework may provide little information on what caused a change in the availability of a particular food, it does provide users with some clues to how both countries and consumers have compensated for that change. For example, consider

that Country A is a mid-level wheat producer, the wheat crop of which was devastated in one year by a late freeze. This fact in isolation appears to indicate an extremely dire situation, and would certainly make for an eye-catching headline (“Domestic wheat crop destroyed!”). However, because the FBS framework brings together additional information, the damage to the wheat crop is not the end of the story. The country could increase wheat imports to compensate for the crop loss, or perhaps withdraw from stocks to ensure that domestic supplies remained unchanged. Alternatively, it is possible that consumers switch to eating more rice in the face of higher wheat prices. In each of these scenarios, it is completely possible that overall dietary availability will remain unchanged for this country, even in the face of this poor harvest.

It is this context – the ability to compare different variables, years and commodities to draw some conclusions on the overall situation – that makes FBS such a powerful tool. While this context may not be so straightforward to the novice FBS user, with practice, FBS can be used to underpin the narrative for analysis of virtually any agrifood market.

## **5.5 SUMMARY**

This chapter has provided FBS compilers with additional content that may be relevant to countries once the FBS have been finalized – namely, considerations of data quality, dissemination and interpretation. In ensuring that quality data are produced, compilers should consider the five aspects of data quality: relevance, accuracy and reliability, timeliness and punctuality, coherence and comparability, and accessibility and clarity. It is also suggested that countries disseminate their data for wider public consumption after it has been produced. To that end, a few different data presentations have been suggested. The chapter closes with some final thoughts on the interpretation of FBS data that will help both compilers and users to appropriately utilize FBS data for policy analysis.

# References

- Afrika, J.K. & Ajumbo, G.** 2012. Informal Cross Border Trade in Africa: Implications and Policy Recommendations. *Africa Economic Brief*, 3(10). Available at <https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Economic%20Brief%20-%20Informal%20Cross%20Border%20Trade%20in%20Africa%20Implications%20and%20Policy%20Recommendations%20-%20Volume%203.pdf>. Accessed on 19 January 2017.
- Agricultural Market Information System (AMIS).** Forthcoming. *Guidelines for Designing and Implementing Grain Stock Surveys*. AMIS Research Report.
- AMIS.** 2016. *AMIS Programme of Work and Budget: 2016-2017*. Available at: [http://www.amis-outlook.org/fileadmin/user\\_upload/amis/docs/Rapid\\_Response\\_Forum\\_5/RRF\\_5\\_2016\\_3.pdf](http://www.amis-outlook.org/fileadmin/user_upload/amis/docs/Rapid_Response_Forum_5/RRF_5_2016_3.pdf). Accessed on 19 January 2017.
- AMIS.** 2015. *Report of the Expert Meeting on Stocks Measurement*. AMIS Rapid Response Forum Report. Available at: [http://www.amis-outlook.org/fileadmin/user\\_upload/amis/docs/Rapid\\_Response\\_Forum\\_4/RRF\\_4\\_2015\\_9.pdf](http://www.amis-outlook.org/fileadmin/user_upload/amis/docs/Rapid_Response_Forum_4/RRF_4_2015_9.pdf). Accessed on 19 January 2017.
- AMIS.** 2014. *Improving Feed Use Estimations: Data, Methodologies and Challenges*. AMIS Scoping Study: Rome. Available at: [http://www.amis-outlook.org/fileadmin/user\\_upload/amis/docs/resources/Improving\\_Feed\\_Use\\_Estimations\\_-\\_Data\\_Methodologies\\_and\\_Gaps.pdf](http://www.amis-outlook.org/fileadmin/user_upload/amis/docs/resources/Improving_Feed_Use_Estimations_-_Data_Methodologies_and_Gaps.pdf). Accessed on 19 January 2017.
- AMIS.** undated. *Market Database*. Available at: <http://statistics.amis-outlook.org/data/index.html>. Accessed on 19 January 2017.
- Bank of Uganda.** 2017. *Informal Cross Border Trade Survey*. Available at: [https://www.bou.or.ug/bou/publications\\_research/icbt.html](https://www.bou.or.ug/bou/publications_research/icbt.html). Accessed on 19 January 2017.
- Bobenrieth, E., Bobenrieth, J. & Wright, B.** 2004. *A Model of Supply of Storage*. *Economic Development and Cultural Change*, 52, no. 3 (April 2004): 605-616.
- Bobenrieth, E., Wright, B. & Zeng, D.** 2013. Stocks-to-use Ratios and Prices as Indicators of Vulnerability to Spikes in Global Cereal Markets. *Agricultural Economics*, 44, supplement: 43-52.
- Buzby, J., Wells, H. & Hyman, J.** 2014. *The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States*. USDA-ERS Economic Information Bulletin No. 121. Available at: [https://www.ers.usda.gov/webdocs/publications/eib121/43680\\_eib121.pdf](https://www.ers.usda.gov/webdocs/publications/eib121/43680_eib121.pdf). Accessed on 19 January 2017.
- Comité Permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel (CILSS).** undated. *Flux transfrontalier*. Available at: <http://www.cilss.int/index.php/flux-transfrontalier/>. Accessed on 11 August 2017.
- Food and Agriculture Organization of the United Nations (FAO).** 2017a. *Commodities which are Aggregated or Standardized to Their Primary Equivalent in Commodity Balances Demand and Food Balance Sheets*. Available at: [http://www.fao.org/fileadmin/templates/ess/ess\\_test\\_folder/Food\\_security/Excel\\_sheets/Commodities\\_which\\_are\\_aggregated\\_or\\_standardized.xls](http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Food_security/Excel_sheets/Commodities_which_are_aggregated_or_standardized.xls). Accessed on 22 February 2017.

**FAO.** 2017b. *CPC Ver. 2.1 Expanded for Agriculture (Crops, Livestock and Derived Products) and Correspondences to FAOSTAT Commodity List (FCL)*. Available at: [http://www.fao.org/fileadmin/templates/ess/classifications/Corr\\_11Jan2017.xlsx](http://www.fao.org/fileadmin/templates/ess/classifications/Corr_11Jan2017.xlsx). Accessed on 22 February 2017.

**FAO.** 2017c. *FAO/INFOODS Food Composition Databases*. Available at: <http://www.fao.org/infoods/infoods/tables-and-databases/faoinfoods-databases/en/>. Accessed on 19 January 2017.

**FAO.** 2017d. *FAOSTAT Commodity List*. Available at: <http://www.fao.org/economic/ess/ess-standards/commodity/en/>. Accessed on 25 January 2017.

**FAO.** 2017e. *FAOSTAT commodity definitions and correspondences: FCL – HS 2007*. Available at: <http://www.fao.org/economic/ess/ess-standards/commodity/item-hs/en/>. Accessed on 22 February 2017.

**FAO.** 2017f. *Food Balance Sheets: Definitions and Standards*. Available at: <http://www.fao.org/faostat/en/#data/FBS>. Accessed on 19 January 2017.

**FAO.** 2017g. *Food Security Methodology*. Available at: <http://www.fao.org/economic/ess/ess-fs/fs-methods/fs-methods1/en/>. Accessed on 19 January 2017.

**FAO.** 2017h. *Global Information and Early Warning System*. Available at: <http://www.fao.org/giews/en/>. Accessed on 19 January 2017.

**FAO.** 2017i. *FAOSTAT commodity definitions and correspondences: HS 2007 – FCL*. Available at: <http://www.fao.org/economic/ess/ess-standards/commodity/hs-item/en/>. Accessed on 22 February 2017.

**FAO.** 2016. *Crops Production Metadata*. Available at: <http://www.fao.org/faostat/en/#data/QC/metadata>. Accessed on 13 February 2017.

**FAO.** 2014. *The FAO Statistics Quality Assurance Framework*. Available at: <http://www.fao.org/docrep/019/i3664e/i3664e.pdf>. Accessed on 19 January 2017.

**FAO.** 2012. *FAO Statistical Yearbook 2012: World Food and Agriculture*. Available at: <http://www.fao.org/docrep/015/i2490e/i2490e00.htm>. Accessed on 27 April 2017.

**FAO.** 2011. *Nutritive Factors*. Available at: [http://www.fao.org/fileadmin/templates/ess/ess\\_test\\_folder/Food\\_security/Excel\\_sheets/Nutritive\\_Factors.xls](http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Food_security/Excel_sheets/Nutritive_Factors.xls). Accessed on 19 January 2017.

**FAO.** 2001. *Food Balance Sheets: A Handbook*. FAO Publication: Rome. Available at: <http://www.fao.org/docrep/003/X9892E/X9892E00.HTM>. Accessed on 19 January 2017.

**FAO.** 1999. *Technical Conversion Factors for Agricultural Commodities*. FAO Publication: Rome. Available at: <http://www.fao.org/fileadmin/templates/ess/documents/methodology/tcf.pdf>. Accessed on 19 January 2017.

**FAO, WB & UNSC.** 2012. *Action Plan of the Global Strategy to Improve Agricultural and Rural Statistics*. FAO Publication: Rome. Available at: <http://www.fao.org/docrep/016/i3082e/i3082e.pdf>. Accessed on 19 January 2017.

**Famine Early Warning Systems Network (FEWS NET).** undated. *Markets & Trade*. Available at: <https://www.fews.net/sectors/markets-trade>. Accessed on 19 January 2017.

**F.O. Licht.** undated. *International Sugar and Sweetener Report*. Available at: <https://www.agra-net.com/agra/international-sugar-and-sweetener-report/>. Accessed on 19 January 2017.

**FUSIONS.** 2016. *Estimates of European Food Waste Levels*. Available at: <https://www.eu-fusions.org/phocadownload/Publications/Estimates%20of%20European%20food%20waste%20levels.pdf>. Accessed on 19 January 2017.

**Gelman, A. & Hill, J.** 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press: Cambridge, UK.

**Global Strategy to improve Agricultural and Rural Statistics (Global Strategy).** 2015a. *Guidelines for the Integrated Survey Framework*. FAO Publication: Rome. Available at: [http://gsars.org/wp-content/uploads/2015/05/ISF-Guidelines\\_12\\_05\\_2015-WEB.pdf](http://gsars.org/wp-content/uploads/2015/05/ISF-Guidelines_12_05_2015-WEB.pdf). Accessed on 19 January 2017.

**Global Strategy.** 2015b. *Guidelines on International Classifications for Agricultural Statistics*. Joint FAO-UNSD Publication: Rome. Available at: <http://gsars.org/wp-content/uploads/2015/12/Guidelines-for-Int-Classifications-on-Agricultural-Statistics-web.pdf>. Accessed on 19 January 2017.

**Global Strategy.** 2014. *SPARS: Strategic Plans for Agricultural and Rural Statistics*. FAO Publication: Rome. Available at: <http://gsars.org/wp-content/uploads/2014/07/SPARS-final-3007.pdf>. Accessed on 20 February 2017.

**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. Available at: <http://www.fao.org/3/a-i4315e.pdf>. Accessed on 19 January 2017.

**International Aid Transparency Initiative (IATI).** undated. *International Aid Transparency Initiative*. Available at <http://www.aidtransparency.net/>. Accessed on 19 January 2017.

**INRA, CIRAD, AFZ & FAO.** 2016. *Feedipedia Animal Feed Resources Information System*. Available at: <http://www.feedipedia.org/>. Accessed on 19 January 2017.

**Lin, B. & Mentzer Morrison, R.** 2016. A Closer Look at Declining Fruit and Vegetable Consumption Using Linked Data Sources. *Amber Waves*. Available at: <https://www.ers.usda.gov/amber-waves/2016/july/a-closer-look-at-declining-fruit-and-vegetable-consumption-using-linked-data-sources/>. Accessed on 19 January 2017.

**Moss, J., Morley, P., Baker, D., Al-Moadhen, H., & Downie, R. & University of New England.** 2016. *Improving Methods for Estimating Livestock Production and Productivity: Literature Review*. Global Strategy to Improve Agricultural and Rural Statistics Technical Report Series No. GO-11-2016. Available at: [http://gsars.org/wp-content/uploads/2016/04/LR\\_Improving-Methods-for-Estimating-Livestock-Production-and-Productivity-220416.pdf](http://gsars.org/wp-content/uploads/2016/04/LR_Improving-Methods-for-Estimating-Livestock-Production-and-Productivity-220416.pdf). Accessed on 19 January 2017.

**Organisation for Economic Co-operation and Development (OECD) & FAO.** undated. *OECD-FAO Agricultural Outlook Database*. Available at: <http://www.agri-outlook.org/database/>. Accessed on 19 January 2017.

**Schmidhuber, J.** 2016. *FAO Resource Book for the Compilation of Food Balance Sheets*. FAO Statistics Division draft internal document (Publication forthcoming).

**Statistical Data and Metadata eXchange (SDMX).** 2016. *Code List for Observation Status*. SDMX Statistical Working Group Statistical Guidelines, Version 2.1, 15 September 2016. Available at: [https://sdmx.org/wp-content/uploads/CL\\_OBS\\_STATUS\\_v2\\_1.docx](https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_1.docx). Accessed on 26 April 2017.

**SDMX.** 2015. *Guidelines for the Creation and Management of SDMX Code Lists*. SDMX Guidelines, Version 2.0, 15 January 2015. Available at: [http://sdmx.org/wp-content/uploads/SDMX\\_Guidelines\\_for\\_CDCL.doc](http://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_CDCL.doc). Accessed on 26 April 2017.

**Servicio de Información Agroalimentaria y Pesquera (SIAP).** 2017. *Balanza del Sorgo*. Available at: <http://www.numerosdelcampo.sagarpa.gob.mx/publicnew/productosAgricultolas/cargarPagina/5>. Accessed on 19 January 2017.

**SIAP.** undated. *Cosechando Números del Campo*. Available at: <http://www.numerosdelcampo.sagarpa.gob.mx/publicnew/>. Accessed on 19 January 2017.

**Sud, U.C., Ahmad, T., Gupta, V.K., Chandra, H., Misra Sahoo, P., Aditya, K., Singh, M., Biswas, A. & ICAR-Indian Agricultural Statistics Research Institute.** 2016. *Research on Improving Methods for Estimating Crop Area, Yield and Production under Mixed, Repeated and Continuous Cropping*. Global Strategy Working Paper No. 5. Available at: [http://gsars.org/wp-content/uploads/2016/01/WP\\_Synthesis-of-Literature-and-Framework\\_Improving-Methods-for-Estimation-of-Crop-Area-190116.pdf](http://gsars.org/wp-content/uploads/2016/01/WP_Synthesis-of-Literature-and-Framework_Improving-Methods-for-Estimation-of-Crop-Area-190116.pdf). Accessed on 19 January 2017.

**United Nations (UN).** undated. *Comtrade Database*. Available at: <https://comtrade.un.org/>. Accessed on 19 January 2017.

**United Nations Economic Commission for Europe (UNECE).** 2013. *The Common Metadata Framework*. Available at: <http://www1.unece.org/stat/platform/display/metis/The+Common+Metadata+Framework>. Accessed on 26 April 2017.

**UN Population Division.** 2015. *World Population Prospects, the 2015 Revision*. Available at: <https://esa.un.org/unpd/wpp/>. Accessed on 19 January 2017.

**UN Population Division.** undated. *Glossary of Demographic Terms*. Available at: <https://esa.un.org/unpd/wpp/General/GlossaryDemographicTerms.aspx>. Accessed on 19 January 2017.

**UN Statistics Division (UNSD).** 2017a. *Correspondence between HS 2012 and CPC Ver.2.1*. Available at: <http://unstats.un.org/unsd/cr/registry/regso.asp?Ci=81&Lg=1>. Accessed on 25 January 2017.

**UNSD.** 2017b. *The System of National Accounts (SNA)*. Available at: <http://unstats.un.org/unsd/nationalaccount/sna.asp>. Accessed on 19 January 2017.

**UNSD.** 2017c. *UN Data*. Available at: <http://data.un.org/Explorer.aspx?d=SNAAM>. Accessed on 26 April 2017.

**UNSD.** 2016. *Environment Glossary: "Apparent Consumption"*. Available at: <http://unstats.un.org/unsd/environmentgl/gesform.asp?getitem=116>. Accessed on 19 January 2017.

**UNSD.** 2004. *International Merchandise Trade Statistics: Compilers Manual*. UNSD Publication: New York. Available at: [http://unstats.un.org/unsd/publication/SeriesF/seriesf\\_87e.pdf](http://unstats.un.org/unsd/publication/SeriesF/seriesf_87e.pdf). Accessed on 19 January 2017.

**UN World Tourism Organization (UN WTO).** 2016. *Saint Lucia: Country-specific: Basic indicators (Compendium) 2011-2015*. Available at: <http://www.e-unwto.org/doi/abs/10.5555/unwtotfb0662010020112015201611>. Accessed on 19 January 2017.

**UNWTO.** 2015. *UNWTO eLibrary*. Available at: <http://www.e-unwto.org/toc/unwtotfb/current>. Accessed on 19 January 2017.

**United States Department of Agriculture (USDA).** undated. *Production, Supply and Distribution Online Database*. Available at: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/home>. Accessed on 19 January 2017.

**USDA Economic Research Service (USDA-ERS).** 2016a. *Food Availability Documentation of the Food Availability (Per Capita) Data System*. Available at: <https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/food-availability-documentation/>. Accessed on 19 January 2017.

**USDA-ERS.** 2016b. *International Food Consumption Patterns*. Available at: <https://www.ers.usda.gov/data-products/international-food-consumption-patterns.aspx>. Accessed on 19 January 2017.

**USDA Office of the Chief Economist (USDA-OCE).** undated. *How the WASDE is Prepared*. Available at: <https://www.usda.gov/oce/commodity/wasde/prepared.htm>. Accessed on 19 January 2017.

**Van der Vorst, J.G.A.J., da Silva, C.A. & Trienekens, J.H.** 2007. *Agro-industrial Supply Chain Management: Concepts and Applications*. FAO Agricultural Management, Marketing and Finance Occasional Paper No. 17. Available at: <http://www.fao.org/3/a-a1369e.pdf>. Accessed on 30 January 2017.

**Van Leeuwen, M., Salamon, P., Fellmann, T., Banse, M., von Ledebur, O. Salputra, G. & Nekhay, O.** 2012. *The Agri-food Sector in Ukraine: Current Situation and Market Outlook Until 2025*. Available at <http://ftp.jrc.es/EURdoc/JRC71776.pdf>. Accessed on 19 January 2017.

**Van Leeuwen, M., Salamon, P., Fellmann, T., Koç, A., Bölük, G., Tabeau, A., Esposti, R., Bonfiglio, A., Lobianco, A. & Hanrahan, K.** 2011. *Potential Impacts on Agricultural Commodity Markets of an EU Enlargement to Turkey*. Available at <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC60663/jrc60663.pdf>. Accessed on 19 January 2017.

**World Customs Organization (WCO).** 2017. *HS Nomenclature 2017 edition*. Available at: <http://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs-nomenclature-2017-edition/hs-nomenclature-2017-edition.aspx>. Accessed on 17 February 2017.

**Westcott, P. & Hoffman, L.** *Price Determination for Corn and Wheat: The Role of Market Factors and Government Programs*. USDA ERS, Technical Bulletin No. 1878, July 1999. Available at: <http://usda.mannlib.cornell.edu/usda/ers/cornwheatprices/tb1878.pdf>. Accessed on 24 April 2017.

**World Food Programme (WFP).** 2017. *International Food Aid Information System*. Available at: <http://www.wfp.org/fais/>. Accessed on 19 January 2017.

# Annex 1

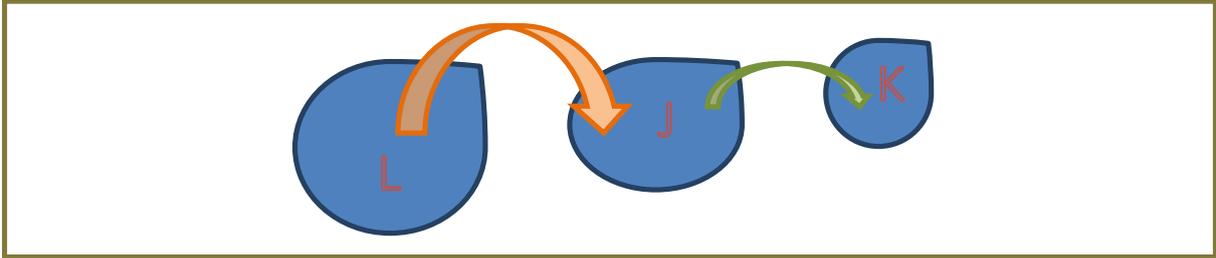
## Tourist food

The estimation of tourist food begins with a simple calculation of the number of visitor days,  $N$ , for travelers originating from country  $l$  and visiting country  $j$ . To represent this flow of visitors traveling from country  $l$  to country  $j$ , the notation  $N_{lj}$  is used – it should be noted that the first subscript indicates the country of origin and the second subscript indicates the destination country. Visitors to country  $j$  include both day visitors and overnight visitors, and these visits must be standardized to provide a total number of visitor days. To do so, the number of day visitors from  $l$  to  $j$ ,  $N_{Dlj}$ , is added to the number of overnight visitors from  $l$  to  $j$ ,  $N_{Olj}$ , multiplied by the average number of days an overnight visitor stayed,  $\bar{D}$ , as below in equation (6-1).

$$N_{lj} = N_{Dlj} + N_{Olj} * \bar{D} \quad (6-1)$$

Now that the convention of the notation has been introduced, a graphical representation of visitor flows better illustrates the conceptual framework of the imputation approach (figure 6-1). Consider that the goal is to estimate net tourist food for country  $j$ . Every year, thousands of visitors travel from country  $l$  visit country  $j$ . These flows are denoted as  $N_{lj}$  – that is, the number of visitor-days for those travelling from country  $l$  to country  $j$ . In figure 6 1, these flows are represented by the large orange arrow. At the same time, hundreds of residents of country  $j$  travel abroad as well, to visit country  $k$ . These flows can be denoted as  $N_{jk}$ , the number of visitor-days for residents of country  $j$  travelling to country  $k$ , with these flows represented by the small green arrow.

**FIGURE 6-1. SIMPLIFIED REPRESENTATION OF NET VISITORS TO COUNTRY J.**



In this three-country scenario, for any individual commodity  $i$ , net tourist food for country  $j$  ( $NetTF_{ij}$ ) can be represented as the amount of food (in tonnes) that is available for consumption in country  $j$  by incoming visitors from  $l$  to  $j$ ,  $Food_{ij}$ , minus the amount of food available for consumption by residents of country  $j$  that are instead visiting country  $k$ ,  $Food_{ijk}$  (equation (6-2)).

$NetTF_{ij} = Food_{ij} - Food_{ijk}$	(6-2)
---------------------------------------	-------

The amounts of food available for consumption by visitors in tonnes ( $Food_{ij}$  for incoming visitors and  $Food_{ijk}$  for outgoing travellers) is calculated as the product of the amount of visitor-days  $N$  (as defined above in equation (6-1)) and the number of kilocalories available for consumption daily, divided by the amount of kilocalories per tonne for commodity  $i$ . At this point, however, a number of intuitive assumptions about the quantities and dietary patterns of visitors should be introduced. Visitors are limited in their food choices to what is on offer locally. Although there are certainly instances where local restaurants cater to visitors by offering foods from the visitor’s country of origin, it is assumed to be more likely that visitors follow the local dietary pattern (as a stark example, Mexican visitors to Italy are more likely to consume pizza than tacos).

At the same time, visitors are likely to consume the same overall amount of food that they would eat at home. To see why this is the case, consider this exaggerated illustration. Imagine a tourist from a wealthy country (with a high daily DES of 3 500 kcal/day) visiting a poor country (with lower DES estimated at 2 000 kcal/day). Although this tourist may consume the foods of the local diet, they are not likely to reduce their consumption by nearly half to fully adapt to the local diet. Instead, we assume that they will eat local foods, and at a scale that accounts for their own typical daily caloric consumption. Keeping in mind the number of visitors, the dietary pattern in the destination country, the level of food availability in the origin country, and the amount of kcal per tonne, total food availability in tonnes of commodity  $i$  for visitors from country  $l$  to country  $j$ ,  $Food_{ij}$ , can be written as:

$Food_{ij} = \frac{\left[ N_{lj} * \left( f_{ij} * \frac{\sum_i f_{il}}{\sum_i f_{ij}} \right) \right]}{\frac{kcal_i}{MT}}$	(6-3)
---	-------

Note: MT = metric tonnes

where  $N_{lj}$  is the number of visitor-days for visitors traveling from country  $l$  to country  $j$ ,  $f_{ij}$  represents the amount of kcal of commodity  $i$  available for consumption in country  $j$ , the term  $\frac{\sum_i f_{il}}{\sum_i f_{ij}}$  is the relative amount of total historical

caloric availability<sup>1</sup> in country *l* compared to country *j*, and  $\frac{kcal_i}{MT}$  is the number of kcal contained in 1 tonne of commodity *i*. This adjustment of kcal per unit is required to convert the quantity from an estimate in kcal to an estimate in weight needed for the SUA account for commodity *i*.

Building on the above example of the relative consumption of a wealthy country and a poor country for illustrative purposes, consider the following scenario. FBS compilers in country *j* are estimating inbound tourist consumption of beer of barley. For the purposes of this exercise, we assume that country *j* is the relatively poorer country with daily DES of 2 000 kcal/cap/day (therefore, total daily caloric availability *f* for commodity *i*, summed over all commodities, or  $\sum_i f_{ij}$ ), and country *l* is the wealthy country where DES ( $\sum_i f_{il}$ ) is 3 500 kcal/cap/day. Tourists from country *l* spend 50 000 tourist-days in country *j* in the reference period. In addition, in country *j*, daily food availability of beer of barley, *f*, is 25 kcal/cap/day, and consulting a calorie conversion table, they find that there are approximately 430 kilocalories per kilogram of beer of barley, equivalent to 430 000 kilocalories per tonne of beer of barley.

Using this information, daily kilocalories of beer of barley available for consumption by visitors from country *l* to country *j* can be imputed by performing the following calculation, according to equation (6-3).

$$Food_{beerlj} = \frac{\left[ N_{lj} * \left( f_{beerj} * \frac{\sum_i f_{il}}{\sum_i f_{ij}} \right) \right]}{\frac{kcal_{beer}}{MT}} \quad (6-3)$$

$$Food_{beerlj} = \frac{\left[ 50\,000 * \left( 25 * \frac{3\,500}{2\,000} \right) \right] kcal}{430\,000 \frac{kcal}{MT}}$$

$$Food_{beerlj} = 5.09 MT$$

Note: MT = metric tonnes

If countries are unable to access information on total food availability for visitors differentiated by country to calculate food as above in (6-3), then a simplification can be substituted (equation (6-4)). In these cases, the issue is that kilocalories cannot be scaled appropriately based on the relative overall DES for visitors from each country. As such, the simplification here is that the scaling term can be dropped, and compilers may assume that visitors consume the same amount of the commodity as do residents. Depending upon the country and the relative country of origin of visitors, dropping this scaling may underestimate tourist consumption in some cases and overestimate it in others.

$$Food_{ij} = \frac{(N_{lj} * f_{ij})}{\frac{kcal_i}{MT}} \quad (6-4)$$

Note: MT = metric tonnes

For outbound travellers, no scaling is necessary either, as the purpose of accounting for outbound travellers is to subtract what they would have eaten at home, and not to detail exactly what they will eat while abroad. Therefore, similarly to the above simplification, for outbound travellers the amount of kilocalories of commodity *i* available for consumption per day in country *j*,  $f_{ij}$ , is simply multiplied by the number of visitor-days for travellers leaving country *j*,  $N_{jk}$ , and this number of total kilocalories is then converted into a quantity in tonnes by dividing by the relevant nutrient factor,  $\frac{kcal_i}{MT}$ , represented below in equation (6-5).

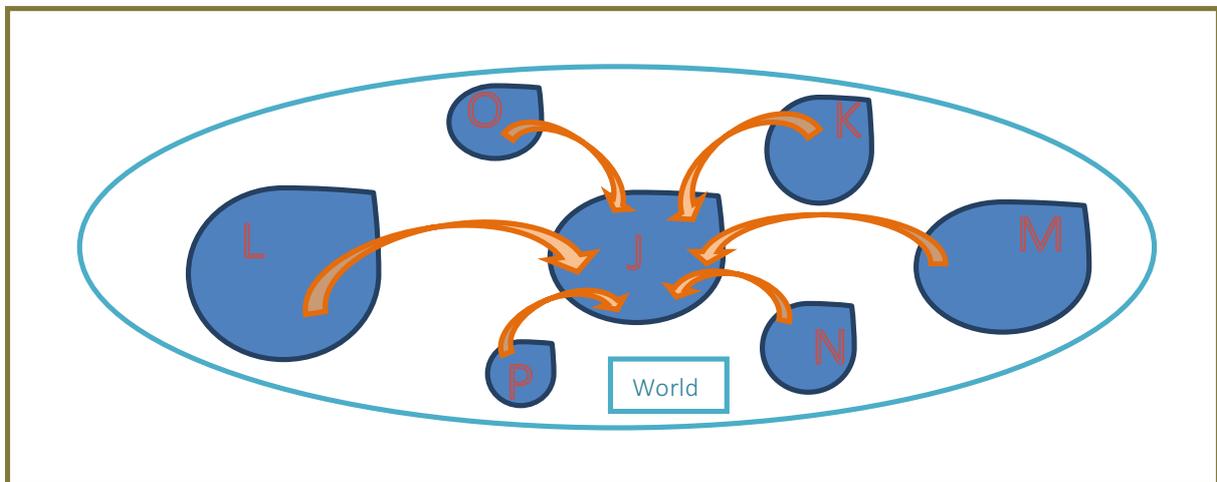
<sup>1</sup> Country-level analysts should define what “historical caloric availability” means; two suggestions are (1) level of caloric availability in the previous year, or (2) average caloric availability over the previous three years.

$$Food_{ijk} = \frac{(N_{jk} * f_{ij})}{\frac{kcal_i}{MT}} \quad (6-5)$$

Note: MT = metric tonnes

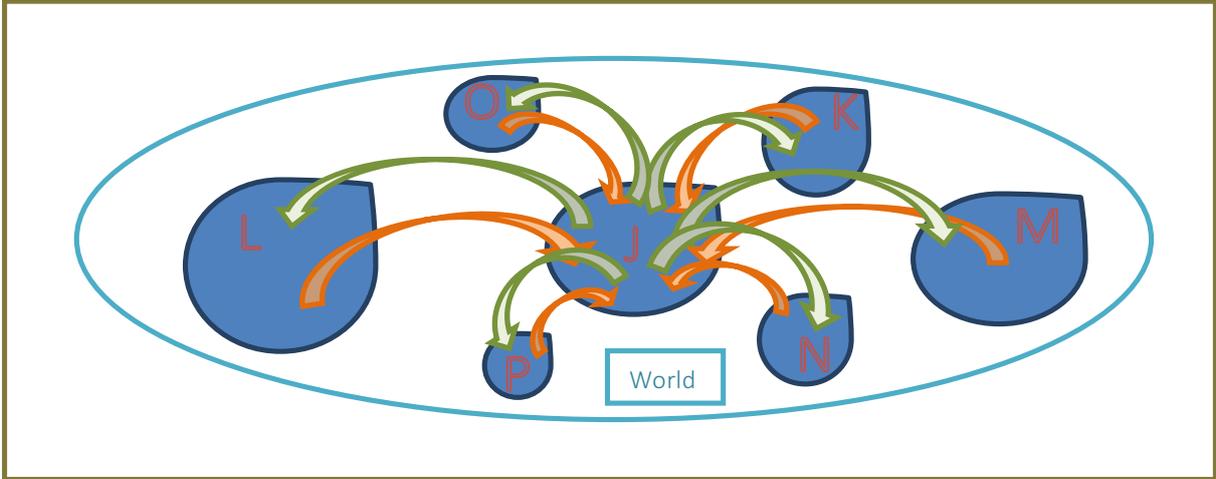
This simplified illustration hopefully aids in the understanding of how tourist food is calculated for visitors from a single country. However, in reality, it is known that visitors are flowing into country *j* from many countries. In fact, they may potentially come from every country in the world except country *j* itself. These multiple inbound visitor flows are represented by the orange arrows below, in figure 6-2.

**FIGURE 6-2. REPRESENTATION OF INBOUND VISITOR FLOWS TO COUNTRY J.**



At the same time that country *j* is receiving visitors from other countries, it is also potentially sending visitors out to these same countries. Therefore, to comprehensively represent net tourism, country-level FBS compilers must account for multiple inbound visitor flows (orange arrows), and multiple outbound traveller flows (green arrows), as represented below in figure 6-3.

FIGURE 6-3. COMPREHENSIVE REPRESENTATION OF NET TOURISM FLOWS FOR COUNTRY J.



The approach required, then, is to take the intuitive single country model presented in equation (6-2) and expand upon it by summing up all the individual food quantities from the unique bilateral visitor flows. This aggregation can be represented using summation notation, combining equations (6-2), (6-3), and (6-5) into a single model, described below in equation (6-6).

$$\text{NetTC}_{ij} = \underbrace{\frac{\sum_{l=1, l \neq j}^x \left[ N_{lj} * \left( f_{ij} * \frac{\sum_i f_{il}}{\sum_i f_{ij}} \right) \right]}{\frac{\text{kcal}_i}{\text{MT}}}}_{\text{Food consumed by incoming visitors}} - \underbrace{\frac{(\sum_{k=1, k \neq j}^x N_{jk}) * f_{ij}}{\frac{\text{kcal}_i}{\text{MT}}}}_{\text{Food consumed by outgoing visitors}} \tag{6-6}$$

Note: MT = metric tonnes

The only additions to this notation are the country summation terms,  $\sum_{l=1, l \neq j}^x [ ]$  and  $\sum_{k=1, k \neq j}^x [ ]$ , indicating that the calculation should be performed for all countries in set  $x$  (in this case, the world) beginning with country  $l$  but excluding country  $j$  in the first case, and beginning with country  $k$  and excluding country  $j$  in the second case. Calculations for the individual flows are done as described above, but then summed in the case of both the incoming and outgoing visitor flows.

While the process of imputing missing data on tourist food requires several calculations and the combining of data from different sources, it is highly recommended that countries make an effort to include it in their balance sheets. This is because specifically accounting for tourist food as a separate utilization variable is an easy way to remove some error from the estimation of the rest of the balance sheet, as previously tourist food was either not differentiated from food availability, accounted for in a basket “other utilizations” category, or simply assumed to be a component of the residual.





