New Food Balances

Description of utilization variables

I. 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).

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.

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[1]. 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

[1] 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).

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[2]. 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.[3]

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[4]. 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[5].

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.

$$\Delta Stocks_t = f\ (\Delta ProdNT_t) + \varepsilon_t$$

Where:
$\Delta Stocks_t$ is equivalent to $Closing\ Stocks_t - Closing\ Stocks_{t-1}$,

[2] 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).
[3] Related documents are available at http://www.fao.org/asiapacific/events/detail-events/en/c/1363/. Accessed on 10 June 2017.
[4] See the AMIS database, available at http://statistics.amis-outlook.org/data/index.html#HOME. Accessed on 19 January 2017.
[5] 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.

	$\Delta ProdNT_t$ is equivalent to $[Production + Imports - Exports]_t - [Production + Imports - Exports]_{t-1}$, and ε_t is an error term.	

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[6].

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.

[6] 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.

II. Food availability

"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.

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[7]. 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.

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

- 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.

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.

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:

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

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 $\frac{Population_t}{Population_{t-1}}$[8]), multiplied

[8] This expression is also equal to 1 plus the percent change in population from period *t-1* to period *t*.

by food availability in the previous period ($Food_{t-1}$), multiplied by 1 plus the historical trend (e.g. the growth rate) in food consumption (ϕ). In this specification, ϕ should be estimated from a regression on the historical food availability data series.

The basic specification outlined in equation 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)[9] and product-specific income elasticities of demand[10] 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 ϵ 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 $\frac{\text{Household consumption expenditure}_t}{\text{Household consumption expenditure}_{t-1}}$)[11] in the additive term at the end of the equation, as below in equation.

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

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.[12]

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.

[9] 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.

[10] 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.

[11] This term is equal to 1 plus the percent change in household consumption expenditure from period *t-1* to period *t*.

[12] This database is publicly available at http://www.ers.usda.gov/data-products/international-food-consumption-patterns.aspx. Accessed on 19 January 2017.

III. 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.

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.

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 II 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.

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.

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 on commodity trees, this calculation can be illustrated by equation.

$$Primary\ commodity\ equivalent = \frac{Quantity\ of\ derived\ product}{Extraction\ rate}$$

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.

IV. 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.

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.

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.

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[13], 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 below (total protein demand could be estimated using the same equation, substituting per animal protein requirement, p, for the energy requirement, e).

$$FD = \sum_i N_i * e_i$$

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:

[13] For a more extensive treatment of this topic, see AMIS (2014).

- 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 (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[14].

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.

[14] 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.

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.

V. 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.

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.

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*[15]. However, compilers are advised that this publication is not recent, such that the published average rates may not reflect current production systems or technologies.

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

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 0.

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:

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

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 to solve for the seeding rate.

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

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)[16]. 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.

[16] 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.

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.

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, $\overline{\text{RatioSH}}$, 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[17]. To derive the $\overline{\text{RatioSH}}$, 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 $\overline{\text{RatioSH}}$ 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:

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

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 1). Then, each of these annual $RatioSH_t$ values are averaged over the series to derive $\overline{\text{RatioSH}}$ (step 2).

Table 1. Hypothetical sown area, harvested area and RatioSH for sunflowers in Country A.

Year	*Sown area (HA) (A)*	*Harvested area (HA) (B)*	$RatioSH_t$ *(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
2014	*?*	***385***	$\overline{RatioSH} = \frac{1.03+1.05+1.06+1.05}{4} = 1.05$

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.

[17] Where sown or harvested areas vary widely from year to year, compilers may instead wish to calculate $\overline{\text{RatioSH}}$ as a geometric mean, since the geometric mean is less susceptible to extreme values than the arithmetic mean suggested above.

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

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

$$\text{Sown area}_{2014} = 1.05 * 385$$
$$\text{Sown area}_{2014} = 404$$

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*.

$$Harvested\ area_{t+1} = (1 - abd)Sown\ area_{t+1}$$

Rearranged to solve for sown area, the formula becomes:

$$Sown\ area_{t+1} = \frac{Harvested\ area_{t+1}}{(1 - abd)}$$

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 can be utilized to calculate an estimated sown area, which in this case calculates to 100 ha.

$$Sown\ area_{t+1} = \frac{Harvested\ area_{t+1}}{(1 - abd)}$$

$$Sown\ area_{2015} = \frac{95}{(1-0.05)}$$

$$Sown\ area_{2015} = 100$$

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.

<u>Step 3: Multiply the two values</u>

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.

VI. 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 travelers 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.

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.

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[18]. This organization compiles and publishes member country-provided data on the number of visitors[19], 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.

It is possible that 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.

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.

$$NetTF = [\#Incoming\ visitor\ days * Daily\ food\ availability\ for\ visitors] - [\#Outgoing\ traveller\ days * Daily\ food\ availability\ for\ residents]$$

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

[18] 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.
[19] Although the data is produced by the UNWTO, data on both personal and business travelers are included. For this reason, it is more accurate to refer to arrivals as "visitors" rather than "tourists".

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.

VII. Industrial use

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.

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.

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[20]. 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[21].

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.

VIII. 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.

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

[20] This database can be accessed at http://www.agri-outlook.org/database/. Accessed on 19 January 2017.
[21] The PS&D database is available at https://apps.fas.usda.gov/psdonline/app/index.html#/app/home. Accessed on 19 January 2017.

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.

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.

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.

$$Quantity\ of\ Loss = Quantity\ of\ Production * Estimated\ \%\ Loss$$

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[22]. 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:

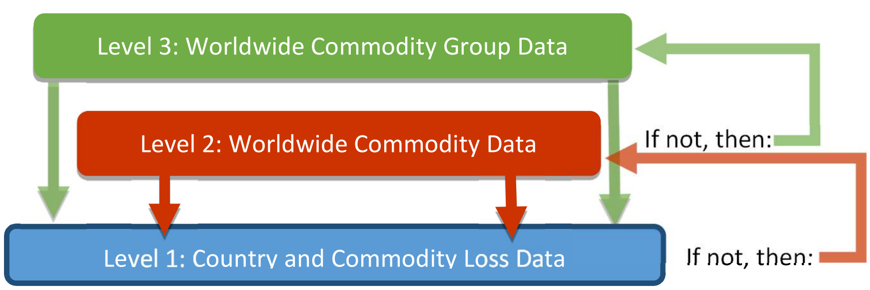
1) If data on loss for a particular commodity in a particular country is reported, then no imputation is necessary, if not:
2) Loss of that commodity in that country is imputed by estimating the relationship between production and all other independent variables and loss of that commodity in all other 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.

Box 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.

[22] For more information on this approach, see Gelman and Hill (2007).

IX. 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.