Key differences between new and old Food Balance Sheet (FBS) methodology

The key difference between the new and old food balances (FBS) methodologies is the absence of a balancer variable. In the past, one of the components of the FBS (often stocks, industrial utilization or feed) would take on the outstanding unbalanced amounts thus inheriting all the statistical errors. With the new methodology, the imputations for the FBS components (the data of which are often not provided by countries) are generated by dedicated modules; and a balancing mechanism will then proportionally spread the imbalances out among all the components.

The most important output of the FBS is the Dietary Energy Supply (DES) in kilocalories per person per day in a given country. This DES is based on the imputations generated by the dedicated new **food module**. This module, in year t, is heavily influenced by food in year t-1, and considering that the other variables in the module are changes in real GDP, population, and commodity demand elasticity – all of which do not usually exhibit large variations from one year to another - the food quantities can be assumed to be comparable between the two methodologies. However, the FBS balancing mechanism, in allocating all available quantities to the food component for 'food only' commodities (e.g. meat of cattle), may result in discrepancies with the old time series. Further analysis of this still needs to be carried out.

Revised **population** figures would have a more evident effect on the DES. The FBS for 2014-17 have been compiled using the 2019 UNPD population data; whereas the series up to 2013 has been compiled using the 2015 (or even earlier) version of the UNPD population data. Therefore, once the FBS preceding 2014 will be recompiled using the new methodology, it is the changes in population data that would have the most impact on the food availability per person (for example, the new revised population numbers for Cuba are some 25% higher than the previous time series).

Lastly, regarding the food basket per country, the new FBS methodology does not alter this between the old and new series. As the commodities consumed in a country are established by already validated production and import data that will not change during the recompilation using the new FBS methodology.

Other highlights:

Stocks, data for which are notoriously unavailable, incomplete or confidential, are now imputed using a new module that monitors stock levels vis-a-vis the supply of that commodity - thus averting unrealistically high stocks, and also calculating to avoid negative stocks – as would occur in the past (mainly, because stocks were often the 'balancer'). Furthermore, a more realistic reference file has been created for potentially stockable commodities (e.g. fresh meats can be expensively stocked only in certain rich countries). Lastly, much wider use is now made of USDA stock data, and from other specialized commodity institutions (such as OilWorld).

The **feed module** now generates feed requirements based on the actual animal numbers and species; and also on the typology of livestock farming, such as intensive using concentrated feeds, or pasture grazing using grasses and forage. More use is being made of feed and forage imports in assessing the availability of commodity specific amounts to be destined for animal feed. Feed only commodities (e.g. cereal cakes) are exhausted first to meet the calculated requirements before deducting further quantities from mixed food and feed commodities (e.g. maize).

The new **loss module** is a linear hierarchical algorithm that imputes for losses across the whole value chain up to and excluding the retail level. The hierarchy is based upon commodity and country groups. In addition, much more use is made of web scraping, text mining and academic/research articles and publications. Thus, the historical loss percentages in the food balances are consistently being revised based upon the new findings.

The **proportional balancing** mechanism is based upon a 3-year moving average of the share of each variable in the total utilizations. Several balancing iterations may be required to solve the imbalance. However, there are limitations put in place – such as a maximum of 10 iterations, and upper and lower boundaries for the different utilizations (based on the max/min over the time series of the share) - which may cause a residual to still remain unsolved. This quantity will be allocated to the 'residual' component and will indicate the amounts that just could not be allocated within the established criteria. Most likely, such residuals would find an allocation under stocks, feed, industrial use, tourist consumption, or a combination thereof, if more information were available.