## Seed use

### A brief review of the previous approach

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

Seed data are collected through a special section in the FAO production questionnaire; where no information is provided by countries directly, seed use information is collected through the websites of the relevant national authorities, i.e. the National Statistical Offices or the Ministries of Agriculture. Where neither of these sources provide official data, seed estimates are calculated or estimated either as a percentage of production (e.g. eggs for hatching) or by multiplying the area sown/harvested with a seed rate. Ideally, seed estimates are based on the area sown rather than the area harvested; in practice, however, sparse data available for area sown make it inevitable that seed use estimates are based on the area harvested.

This can be formulated in the simple identity:

(Equation 28)

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

#### Seed rates

The seed rates currently used in the FBS/SUA system are based on an ad-hoc publication entitled “Technical conversion factors for agricultural commodities”. It provides, inter alia, seed rates for every country and primary commodity/item of FCL classification. It brings together information on seed rates by country and commodity, reflecting current production practices under different growing conditions. The compilation of seed rates benefited from information provided through the questionnaire replies, as well as from FAO expert advice.

The publication also provides information about the share of eggs that is typically used for hatching[[1]](#footnote-2). Additional information is provided to better assess the reliability of hatching rates. All hatching and seed rates have been reviewed for the new FBS system and have been changed where necessary.

### The new imputation approach

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

1. Impute area sown, when missing.
2. Estimate seed rates using a hierarchical linear model, but only if questionnaire results are unavailable.
3. Impute seed as the product of the previously imputed area sown with the estimated seed rate.

#### Imputation of Area Sown

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

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

#### Imputation of Seed Rates

The estimation of seed rates is performed via a hierarchical linear model (mentioned previously in the loss imputation). The rationale for this model is that it is capable of capturing and modelling complicated trends when data is available. Moreover, the hierarchy of the model allows accurate imputation on countries with very sparse data by pooling together global data. The mathematical model can be written as follows:

(Equation 29 a-c)

where are coefficients to estimate from the data, are error estimates, is the average annual temperature of country i (provided by the World Bank), and is measured in years and is included to capture linear trends over time. The i indices run over all countries, the j indices over all CPC groups, and the k indices over all unique country/CPC code combinations.

Thus, the model estimates seed use proportional to the area sown. The model also accounts for changes over time and differences across countries; the latter are captured by the annual temperature variable, assuming essentially that seed rates need to be higher where production conditions are difficult, with a potential of late and frequent frosts (Russia), and can be lower where production conditions are more favourable (UK).

If data for a particular country and commodity are sparse, then and will likely be estimated to be close to 0. Thus, will be close to its mean value, and the model will only account for availability within commodity groups. However, if data are available for a country or commodity, the estimates of and/or will be far from 0, and thus the model can adapt to the individual characteristics of a particular scenario.

1. Only 11 countries have provided official data in recent years. [↑](#footnote-ref-2)