

Compiling Food Balance Sheet using detailed FAO Supply Utilization Accounting data to improve the understanding and modeling of the food system

Relevance of the Topic

Sound economic modeling hinges significantly upon the accessibility and quality of data (Bruckner et al., 2019; Calvin et al., 2022; Chepeliev, 2022). The Food Balance Sheet (FBS)ⁱ constructed by the Food and Agriculture Organization (FAO) of the United Nations is widely used as a primary source of data in research related to global agriculture development, food security, nutrient, and dietary health^{1,2}. It provides a comprehensive overview of a country's supply (e.g., production, import, and stock variation) and utilization (e.g., food, feed, export, processing, seed, loss, etc.) patterns and macronutrient (fat, protein, and carbohydrate) for ~100 aggregated food commodities on a **primary commodity equivalent (PCE) basis**. Compilation of the FBS at the aggregate commodity level, combining together primary and processed commodities (e.g. wheat and flour), creates major data challenges for the agro-economic modeling community, including that for the Global Trade Analysis Project (GTAP) nutritional Data Base construction process (Chepeliev, 2022). It is noteworthy that each agro-economic modeling team typically develops its own assumptions and methods to prepare and process FAOSTAT data (Bond-Lamberty et al., 2019). While largely overlooked, the uncertainty in the base data calibration data approach likely contributed to the disparities in model outcomes (von Lampe et al., 2014; Zhao et al., 2021a).

FAO has recently improved the methodology of compiling FBS³ and released the Supply Utilization Accountingⁱⁱ (SUA) data, which includes over 500 primary and processed commodities and serves as the primary data source for constructing FBS. Despite the significant improvement in the quality and documentation of the FBS and SUA datasets in recent years, a transparent, traceable, and consistent method for aggregating the 500+ SUA commodities into ~100 FBS commodities is not yet available. It is described that a PCE approachⁱⁱⁱ was used to aggregate commodities along the food processing chain. The approach uses extraction rates to scale the SUA balance of secondary commodities to their primary equivalent, ensuring consistency in commodity aggregation. However, the practical implementation of the PCE approach is considerably challenging, as it requires maintaining data balance simultaneously across different dimensions, including (1) supply and demand, (2) space (trade balance), (3) time (storage carryover), (4) along the processing chain (reasonable extraction rate).

Ensuring consistency between the FBS and SUA is crucial. The SUA data supplements the FBS data by providing detailed commodity information in original units, which is essential for understanding the FBS data presented in PCE. This more detailed commodity information also improves the portrayal of the food system. Additionally, SUA data complements the FBS by providing additional information such as (1) macronutrient and dietary energy supply at the detailed commodity level, (2) a broader commodity coverage, including the balance of non-food commodities (e.g., oilseed cake, fiber crops, rubber, etc.), and (3) additional category detail, such as opening and closing stock data. Carefully bridging the gap between the FBS and SUA also facilitates the compiling of the FBS and enhances its scientific robustness.

Thus, we are motivated to improve the PCE method to provide a transparent, reproducible, consistent, and more flexible approach to aggregating agricultural commodities across different dimensions.

Method

A robust PCE approach requires a thorough understanding of the commodity supply chain for all agricultural and food commodities. To accomplish this, we constructed a nested commodity mapping to understand the relationship across all 500+ SUA commodities (shown in **Fig. 1**, left panel) and employed a data-based approach to calculate the extraction rate (a ratio between the production of secondary products and processed use of the primary product). As discussed, this approach requires maintaining data balance across different

ⁱ <https://www.fao.org/faostat/en/#data>

ⁱⁱ Including the Commodity Balance dataset which provides the SUA data for non-food agricultural commodities.

ⁱⁱⁱ The approach is documented with only simple examples.

dimensions, such as supply and demand, space, time, and along the processing chain. In particular, we calculated commodity-processing tier-region-year-specific extraction rates (allowing output-specific rates when coproducts exist) to ensure that the PCE-scaled SUA data is balanced across the interactions of dimensions. For example, an international extraction rate is applied to imported commodities, and a lagged domestic extraction rate is applied to the opening stock. In addition, regional macronutrients and dietary energy are also traced in the aggregation process. The data processing is made in an extension of an open-source R package (*gcamfaostat*^{iv}), to ensure traceability, transparency, and reproducibility and to document the detailed mapping, parameter, and processing assumptions. The updated package further bridges the gap between FAOSTAT data and global economic modeling.

Implications

As an illustration, we present examples of wheat product PCE in **Table 1**. Compared with the original FAO approach, our results show higher trade and feed consumption and lower food and remaining processed use of wheat products at the world level. However, the difference could be larger in results at the regional scale or for other commodities. Our approach is able to provide more detailed storage data and traceable aggregation information linking to SUA commodities. More results will be discussed at the conference. One important application of the PCA approach is to provide more flexible and consistent sector aggregation in global economic modeling. This data development is an ongoing process that would benefit from the continued efforts of a broader community with regard to updating, maintaining and improving the processing.

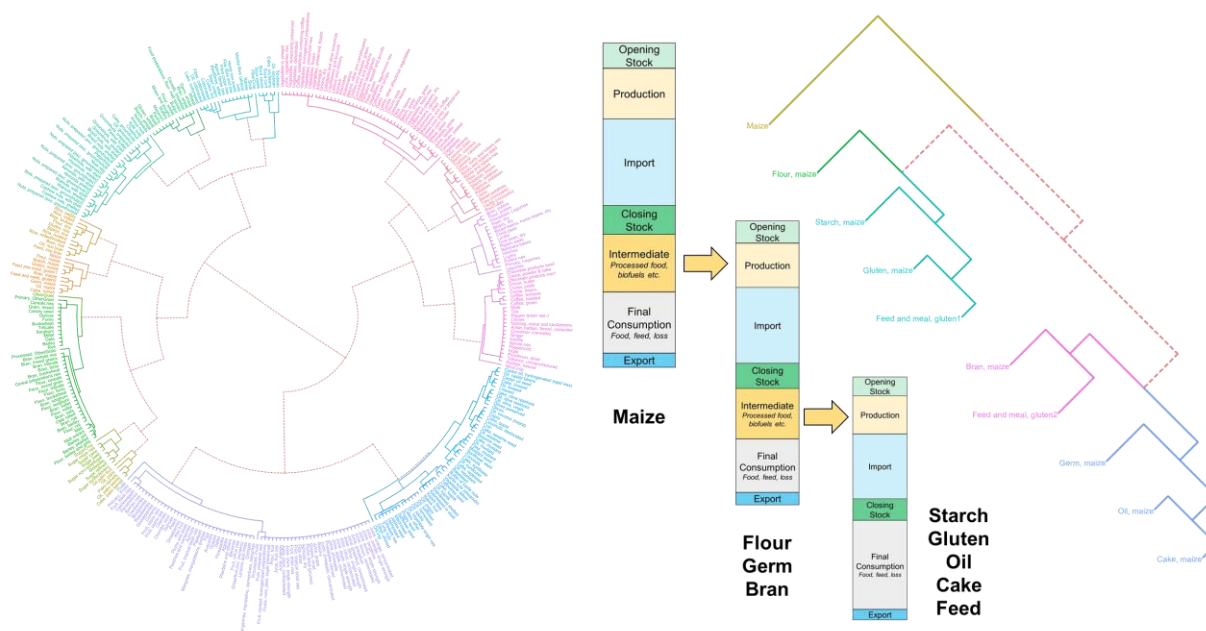


Fig. 1 Illustration of the nested mapping for all FAO crop commodities (left panel) and primary commodity equivalent (PCE) approach (maize example in the right panel). For example, maize is first processed into flour, germ, and bran. Flour, if processed, is used to produce starch (and then gluten & meal). Maize germ, if processed, produces oil and cake. There is an SUA balance for all primary and processed products along the supply chain.

^{iv} It is worth noting that we have extended the package framework of *gcamdata* as it provides an established and traceable structure for model and data development. *gcamdata* is used for processing raw data inputs into structured and balanced base data (in XML format) for GCAM (Global Change Analysis Model). The branch is https://github.com/JGCRI/gcamfaostat/tree/AgSUA_PCE.

Table 1 Comparison of primary commodity equivalent (PCE) between FAO and our new approach, world value in 2015 (million tonnes)

Element	FBS commodity			SUA commodity								
	PCE (Diff.)	PCE (FAO)	PCE (New)	Maize	Flour, maize	Bran, maize	Germ, maize	Starch, maize	Gluten, maize	Cake, maize	Oil, maize	Feed
Opening stocks		NA	352	348	2	0	0	0	0	0	1	2
Production	0	1052	1053	1053	159	21	11	35	5	3	3	10
Import	11	150	161	148	2	2	0	2	0	0	1	5
Regional supply			1565	1550	162	23	12	36	5	3	4	18
Export	9	152	161	148	2	2	0	2	0	0	1	5
Processed	-24	50	26	210	44	7	6	21	3	0	0	0
Food	-9	136	126	4	108	0	0	2	0	0	2	0
Feed	33	575	608	571	3	14	5	0	1	3	0	9
Seed	0	8	8	8	0	0	0	0	0	0	0	0
Other uses	-3	204	202	180	4	1	0	12	0	0	0	0
Loss	0	58	58	58	0	0	0	0	0	0	0	0
Closing stocks		NA	379	374	2	0	0	0	0	0	1	2
Residuals	4	-6	-2	-3	0	0	0	0	0	0	0	0
Regional demand			1567	1553	162	23	12	36	5	3	4	18
Stock Variation	1	26	27	26	0	0	0	0	0	0	0	1

Potential for generating discussion

Compared to the FAO's FBS, which report food and nutritional information for 98 composite categories, in many cases combining primary and processed commodities (e.g. wheat and flour) within a single category (e.g. wheat and products), the constructed SUA explicitly trace the transformations between primary and processed commodities, while reporting nutritional details at the highly disaggregated level (over 500 commodities). In doing so, the constructed dataset substantially simplifies the FAOSTAT data processing steps by explicitly distinguishing food and nutritional supply at the individual commodity level and facilitating a straightforward mapping of the corresponding data to the global agro-economic models.

The analyses provided in this paper would be of particular interest to the participants that are involved in supporting the decision-making process within the agricultural and food systems. The paper provides important insights into better understanding the food supply chains, developing a new more rigorous methodological basis for future modeling and analytical applications and has the potential to generate lively discussion. Our initiative also seeks to enhance the quality and accessibility of data for the global agro-economic modeling community, with the aim of fostering more robust and harmonized outcomes in a collaborative, efficient, and open-source framework.

References

- Bond-Lamberty, B. *et al.* gcamdata: An R Package for Preparation, Synthesis, and Tracking of Input Data for the GCAM Integrated Human-Earth Systems Model. *Journal of Open Research Software* **7**, (2019).
- Chepeliev, M. Incorporating Nutritional Accounts to the GTAP Data Base. *Journal of Global Economic Analysis* **7**, 1–43 (2022).
- Food and Agricultural Organization (FAO). Guidelines for the Compilation of Food Balance Sheets. *Global Strategy: Rome, Italy* (2017).
- Thar, C.-M., Jackson, R., Swinburn, B. & Mhurchu, C. N. A review of the uses and reliability of food balance sheets in health research. *Nutrition Reviews* **78**, 989–1000 (2020).
- von Lampe, M., Willenbockel, D., Ahammad, H., Blanc, E., Cai, Y., Calvin, K., Fujimori, S., Hasegawa, T., Havlik, P., Heyhoe, E., Kyle, P., Lotze-Campen, H., Mason d'Croz, D., Nelson, G.C., Sands, R.D., Schmitz, C., Tabeau, A., Valin, H., van der Mensbrugghe, D., van Meijl, H., 2014. Why do global long-term scenarios for agriculture differ? An overview of the AgMIP Global Economic Model Intercomparison. *Agricultural Economics* **45**, 3–20. <https://doi.org/10.1111/agec.12086>
- Zhao, X., Calvin, K.V., Wise, M.A., Iyer, G., 2021. The role of global agricultural market integration in multiregional economic modeling: Using hindcast experiments to validate an Armington model. *Economic Analysis and Policy* **72**, 1–17. <https://doi.org/10.1016/j.eap.2021.07.007>