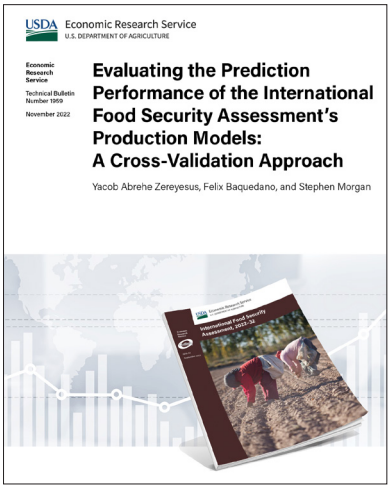
USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

A report summary from the Economic Research Service *November 2022*

Evaluating the Prediction Performance of the International Food Security Assessment's Production Models: A Cross-Validation Approach

Yacob Abrehe Zereyesus, Felix Baquedano, and Stephen Morgan

What Is the Issue?



Food insecurity exists when people do not have physical, social, and economic access to sufficient, safe, and nutritious food that meets their food preferences and dietary needs for an active and healthy life. To evaluate global food security status, the U.S. Department of Agriculture (USDA) Economic Research Service (ERS) developed the International Food Security Assessment (IFSA) model, which evaluates the food security status of 76 low- and middle-income countries. Each country's food security status is estimated for the current year and projected up to 10 years (Baquedano et al., 2020). The gap between domestic food production and food needs is used to estimate the implied additional supply required for each of the 76 countries. The gap is an indication of potential import needs, including food aid. The methodology used to report production figures only focuses on regional and subregional estimates. This research systematically selects the best performing model specification. Having better production estimates is generally important as shortfalls in output may produce global trade repercussions (Jagermeyr et al., 2020). In addition to robust domestic and regional markets, the importance of international markets for the supply of stable and affordable agricultural commodities will be increasingly critical as sources of nutrition for the world population (Hertel et al., 2020; Smith and Glauber, 2020).

What Did the Study Find?

This report examines the forecasting capabilities of the IFSA model. In terms of the new results compared with the previous approach, the findings are:

- The subregional model specification improves the yield prediction performance by 15 percent relative to the pooled IFSA model approach used in the past. In particular, the model improves the absolute difference between the observed and estimated yield (0.159 tons per hectare and 0.188 tons per hectare for the subregional model and pooled IFSA model, respectively).
- When the data are aggregated in alternative ways for forecasting, the subregional model performs better than the crop model. However, both of these model specifications perform better than a regional model in predicting yield.

How Was the Study Conducted?

This research evaluates the production side of the International Food Security Assessment (IFSA) model to determine the best performing prediction model specification. To do so, a "leave-one-out-cross-validation" (LOOCV) approach is used to simulate the out-of-sample model prediction performance. The best performing prediction model is chosen using mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) criteria. The data for the analyses come from the IFSA 2020 model dataset covering the years 2011 to 2019 for the estimations. Four model specifications are developed in this exercise. The first specification estimates the yield model using pooled data for all 76 countries, which does not allow for regional or subregional heterogeneity in the model's parameters. Three other specifications are estimated using regional, subregional, and crop level disaggregated datasets allowing for heterogeneity among the diverse countries included in the IFSA model.