USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Technical Bulletin Number 1959

November 2022

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

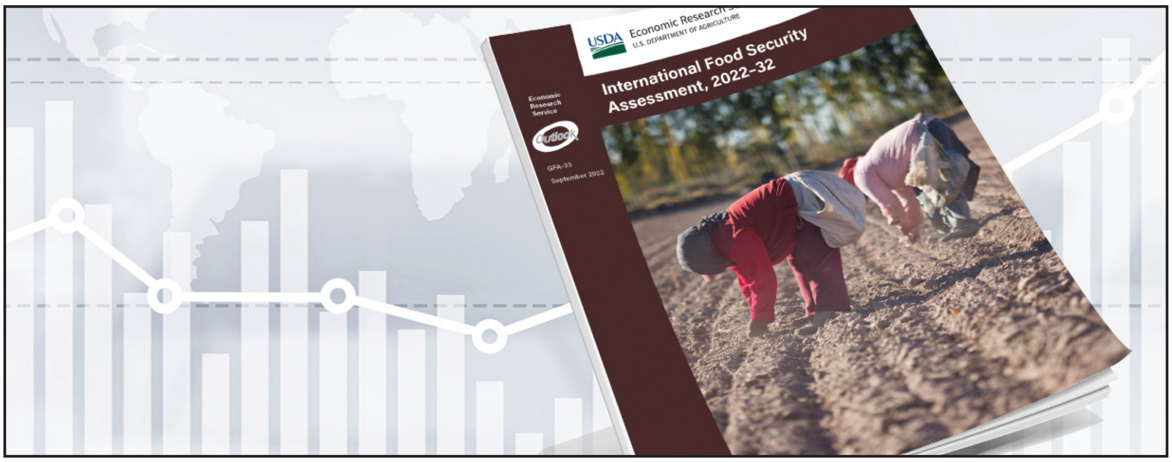
Yacob Abrehe Zereyesus, Felix Baquedano, and Stephen Morgan

USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service
www.ers.usda.gov

Recommended citation format for this publication:

Zereyesus, Yacob Abrehe, Felix Baquedano, and Stephen Morgan. November 2022. *Evaluating the Prediction Performance of the International Food Security Assessment's Production Models: A Cross-Validation Approach,* TB-1959, U.S. Department of Agriculture, Economic Research Service.



Cover image from AdobeStock and Getty Images to include front cover of previously released IFSA report.

USDA Economic Research Service
U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Technical Bulletin Number 1959

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

Abstract

The U.S. Department of Agriculture (USDA), Economic Research Service (ERS) International Food Security Assessment (IFSA) model was developed to help USDA and its stakeholders evaluate the food security status of 76 low- and middle-income countries. The IFSA model provides an estimate of total food demand and food production, both elements in measuring food security. The demand side of the IFSA model is used to estimate the prevalence of country-level food insecurity based on an aggregate food consumption threshold of 2,100 calories per capita per day. The gap between aggregate domestic food production and food demand is used to estimate the implied additional supply required for each of the 76 countries in the IFSA, which is an indication of potential import needs, including food aid. The primary objective of the IFSA's supply-side modeling work is to project production. This research evaluates the production model to determine the best performing prediction model specification. This report advances previous research by using a data-driven approach to select the best performing model specification.

Keywords: cross-validation, global food security, production, prediction, mean squared error, root mean squared error, mean absolute error

Acknowledgments

The authors thank Fred Gale and Jayson Beckman in the Market and Trade Economics Division (MTED) of USDA, ERS, for their valuable comments and review of the paper. We also thank Jeff Chaltas and Jana Goldman for editorial assistance. The findings and conclusions in this paper are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

About the Authors

Yacob Abrehe Zereyesus and Stephen Morgan are research agricultural economists, and Felix Baquedano is chief of the ITD branch of MTED at USDA, ERS.

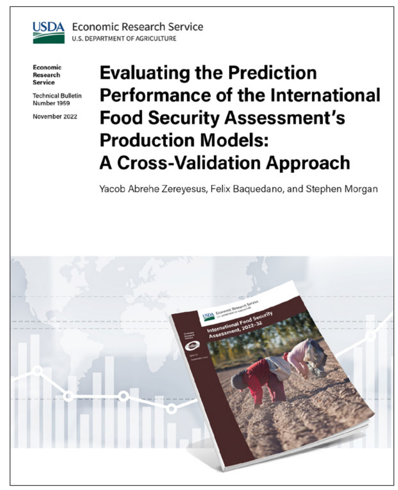
Contents

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.

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

Introduction

The International Food Security Assessment (IFSA) model of the U.S. Department of Agriculture's (USDA) Economic Research Service (ERS) was developed to help USDA and its stakeholders evaluate the food security status of low- and middle-income countries. Recent reports by Baquedano et al. (2020) and Thome et al. (2019) describe how the model projects per capita food demand and compares it with an average caloric level necessary to sustain a healthy and active lifestyle for 76 countries of Sub-Saharan Africa (SSA), Asia, Latin America and the Caribbean (LAC), and North Africa. Each country's food security status is estimated for the current year and projected up to 10 years. The IFSA model only provides an estimate of total food demand, but food production is also a critical factor in food security. An independent module estimates production for the current year and projects it over the subsequent 10-year period to obtain full closure of the model. This research describes the production module and recommends approaches for improving the prediction performance of the cereal yield equation.

The current production module of the IFSA model aggregates a panel of agricultural production data for 76 countries to provide model-based estimation and projection of yield, and area dynamics. This research aims to achieve two objectives:

- Improve the model's performance by disaggregating yield estimates using regional or subregional specifications in place of the aggregate IFSA yield model, and;
- Develop a model-based performance evaluation criterion to select the best performing model specification.

The first objective seeks to achieve further disaggregation based on the major crop of each country, including wheat, rice, and coarse grains (maize, sorghum, and millet). A disaggregated model could improve the accuracy of estimates by accounting for the diverse economic and geographic features of each region, subregion, or a particular major grain. The second objective is addressed using a cross-validation approach called a leave-one-out-cross-validation (LOOCV) to simulate model out-of-sample prediction performance (James et al., 2017). Cross-validation is a method for estimating prediction performance metrics using observed yield data. The best performing model is chosen by comparing three different performance metrics: mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). Because the cross-validation approach is applied using observed yield data, prediction performance for more than 1 year in the future becomes challenging without data. The authors assume that the best performing model specification based on the LOOCV approach will also result in the best prediction performance in each of the next 10 years covered by the IFSA.

Modeling Staple Cereal Production

The availability of cereal grains is critical to food security in low- and middle-income countries.[1] Thus, factors influencing cereal grain output are a key component of the production module.

The IFSA projects production using crop yield and area equations that incorporate parameters estimated from a pooled panel dataset of the 76 countries. The yield and area projections rely on the producer price projections in local currency units, world price projections, and macroeconomic trends from the *USDA Agricultural Projections to 2029* (USDA, Agricultural Outlook Board, and Interagency Agricultural Projections Committee, 2020).

In the current model, agricultural production is decomposed into yield (production per hectare) and area for grains. Production (PR) for a given country *c* in year *t* is obtained by multiplying projected yield (YL) and area (AR) (Baquedano et al., 2020).

$$PR_{ct} = AR_{ct} \times YL_{ct} \tag{1}$$

Below is a description of the modeling and estimation of yield and area for the 76 countries covered by the IFSA model.

Modeling Yield

Yield parameters are estimated econometrically using panel data consisting of observations for each country and are calibrated to observed yields for the most recent 3 years (e.g., 2017–19 for the 2020 report). The calibration procedure involves in-sample prediction using observed yield data and consensus estimates for the expected return ratio—an indicator of the relative profitability of fertilizer use (Baquedano et al., 2020; Beghin et al., 2015). The production model used in the IFSA estimation posits that yields vary with the return ratio (RR) and technical change (T) (represented by a time trend):

$$YL_{ct} = f(RR_{ct}, T_t) \tag{2}$$

The return ratios are the ratio of the return per hectare—revenue from yield divided by the price of fertilizer, RRct = $((yp_{ct} * Y_{ct}) / fp_{ct})$, where yp_{ct} and fp_{ct} are prices of yield (output) and fertilizer, respectively. The expected return ratios include current year and long-term expectation components, and the ratios are expressed in a real local currency unit (rlcu).

The domestic price for each grain is linked to its world reference price, which is expressed in rlcu through the following equation:

$$P^{domestic} = 0.7 \times P^{world} + 0.3 \times I$$

The 0.7 is assumed to be the slope parameter that captures the price transmission from world to domestic prices. The expected domestic price for each grain is a weighted average of 70 percent of the current year world price and 30 percent of the mean domestic price *I* over the preceding 3 years.

[1] For more on the definition and coverage of cereals and a description of the International Food Security Assessment model, refer to Baquedano et al. (2020).

Modeling Area

Crop area, AR_{ct}, is modeled with the widely used Nerlovian specification, in which lagged area (AR_{ct-1}), expected crop price (yp_{ct}), fertilizer prices (fp_{ct}), and a time trend (T_t) are expressed as:

$$AR_{ct} = f(yp_{ct}, fp_{ct}, AR_{ct-1}, T_t) \quad (3)$$

The expected prices are averages of contemporaneous and lagged relative prices. A time trend is included in the area equation to capture non-price factors in area and a country fixed effect. Country fixed effects are included to capture heterogeneity across countries. The area equation is numerically calibrated to the base year average of the preceding 3 years of the report (e.g., 2017–19 for the 2020 report) using a predefined linear relationship for the price and lagged acreage responses and the ratio of domestic grain price to fertilizer price (Thome et al., 2019). The focus of the current report is on yield projections; however, the area estimations are used for computing the production estimates (i.e., $PR_{ct} = AR_{ct} * YL_{ct}$).

Yield Projection and Model Specifications

Following Beghin et al. (2015), the base specification for equation 2 is estimated using a generalized linear model:

$$YL_{ct} = \beta_1 MA_{2,ct} + \beta_2 MA_{5,ct} + \beta_3 \lambda_t + \theta_c \delta_c + e_{ct} \quad (4)$$

The dependent variable, YL_{ct}, measures yield in tons per hectare (tons/ha). Also included are a 2-year moving average ($MA_{2,ct}$) and a 5-year moving average ($MA_{5,ct}$) to provide the short- to medium-term dynamics of return to yield after accounting for input costs. The λ_t is a time trend that represents possible technical change during the analysis period, and δ_c refers to country fixed effect. The e_{ct} is the error term. Finally, the β_1, β_2, β_3, and θ_c are parameters to be estimated.

Four model specifications are developed as follows:

- The first specification estimates the yield model using pooled data for all 76 IFSA countries, which does not allow for regional, subregional, or crop-level heterogeneity in the model's parameters. This is referred to as the aggregate model.
- The second specification is based on a regionally disaggregated dataset. The regional specification disaggregates the estimation of yield by the four regional classifications of the IFSA countries (for more on the list of countries, refer to appendix 1): Sub-Saharan Africa (SSA), Asia, Latin America and the Caribbean (LAC), and North Africa (Baquedano et al., 2020; Thome et al., 2019).
- The third specification disaggregates the dataset based on the subregions included in the IFSA analysis. The subregional specification disaggregates the model to 10 subregions of the IFSA countries: Central Africa, East Africa, Southern Africa, West Africa, North Africa, Latin America and the Caribbean, Commonwealth of Independent States, Central and South Asia, Southeast Asia, and Other Asia.
- The fourth specification is based on crop-level disaggregation. The crop-level specification estimates the yield equation by disaggregating the data by wheat, rice, and coarse grains (maize, sorghum, and millet). Twenty-two countries have wheat as their major grain, 28 countries have rice as their major grain, and 22 countries have maize as their major grain. Three countries (Somalia, Sudan, and Chad) have sorghum as their major grain, while only one country (Niger in West Africa) has millet as its major grain.

Out-of-Sample Prediction Performance and Cross-Validation Approach

Model-based projection performance is assessed in terms of how well the specified model can be expected to perform on an independent out-of-sample dataset, often assessed by the actual estimate of the out-of-sample criteria (e.g., MAE, MSE, RMSE). When an independent out-of-sample dataset is not available, a cross-validation (CV) approach can be used to choose the best model by estimating out-of-sample performance criteria using an in-sample dataset. The out-of-sample error—often referred to as the test-error—is the average error that results from using the regression method to predict or project the response on a new observation that was not used in regression estimation. Given an in-sample dataset, the choice of a specification (the aggregate, regional, subregional, and crop-level model specifications in this report) is warranted if the model causes a low-test error (James et al., 2017). In the case of the IFSA model, model performance is evaluated across three different geographic specifications: aggregate (global), regional, and subregional. Given the broad coverage of the IFSA model with limited variables available across countries, a CV approach provides an objective and data-driven approach to compare the predictive value of the different geographic specifications. Furthermore, cross-validation has been an important part of the literature related to crop yield distribution modeling and insurance rates (e.g., Norwood et al., 2004; Lanoue et al., 2010; Woodard and Sherrick, 2011).[2]

Cross-Validation

The leave-one-out-cross-validation (LOOCV) approach is used where a single observation (x1, y1) contains the validation set, while the remaining observations ((x2, y2), . . ., (xn, yn)) contain the training set. The regression is fit on the n - 1 training observations, and a prediction ($\hat{y}_1$) is made for the excluded observation using its value (x1).

Three different performance metrics are used when evaluating yield estimates in this analysis. The first, MAE, evaluates the absolute distance between the observations and predictions of the model. In this case, a predicted value ($\hat{y}_1$) from x1 using parameters from a regression fit on the n -1 in-sample observations and the $MAE_1 = |y_1 - \hat{y}_1|$. This procedure is iteratively repeated, using each observation as the validation set such that the LOOCV estimate for the out-of-sample MAE is the average of the n out-of-sample error estimates: $MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$.

The second criteria, MSE, uses the squared value of the distance between predicted values of the model and observed data points. Using the iterative LOOCV process, $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$.

The third criteria, RMSE, is the square root of MSE where $RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$.

The three performance criteria operate differently in evaluating the yield models. For example, MSE and RMSE penalize large actual and predicted discrepancies by more than MAE. In other words, MSE and RMSE are more sensitive to outliers in the data. Additionally, compared with MSE, MAE and RMSE measure the out-of-sample error in the same units as the original estimates. Each of these values can be used to compare across models; however, comparing one performance criteria to another is not informative.[3]

[2] For example, Norwood et al. (2004) used a grouped cross-validation approach to evaluate the out-of-sample performance of six different yield densities for corn, soybeans, and wheat. The authors found that the flexible semiparametric model developed by Goodwin and Ker (1998) performed best using out-of-sample criteria.

[3] For example, MAE is always less than or equal to RMSE.

Data and Analysis

The data for the analyses come from the IFSA 2020 model dataset covering 2011 to 2019. Observed domestic prices are from the United Nations' Food and Agriculture Organization's (FAO) Global Information and Early Warning System on Food and Agriculture (GIEWS) Food Price Monitoring and Analysis (FPMA) tool. Tariff data are from the World Bank's World Integrated Trade Solution (WITS). Exchange rates and Consumer Price Indexes (CPIs) are from the USDA's ERS International Macroeconomic Data Set. World grain prices are from USDA's Agricultural Projections. The area harvested, production, and yield data used in the estimation come mainly from USDA's Production, Supply, and Distribution (PSD) database. When and if data from USDA's PSD data are not available for a country, FAO data are used.

For each country, the observed yield, area, and production data are examined for outliers to help purge the impact of influential observations. Observations that are three standard deviations away from the country mean of the respective variable are identified as outliers and consequently replaced with the median observation of the corresponding variable for each of the countries. For each of these variables, the number of the outliers were less than 1 percent of the respective sample sizes. All estimation and analysis of results are completed using the R statistical software (R Core Team, 2021), including glmnet, a regularization package for generalized linear models (Friedman et al., 2010) and boot, a bootstrapping package for the in- and out-of-sample cross-validation applications (Canty and Ripley, 2021).

Results and Discussion

Descriptive Statistics

The summary statistics of the area, yield, and production metrics are presented in tables 1 to 3. The average area under cultivation for all IFSA countries is 3.6 million hectares (table 1). The heterogeneity in terms of area allocated for staple grains is evident in the significant variability in the average area under cultivation in each of the regions. The Asia region has the highest average area under cultivation (8.6 million hectares) relative to the other regions. In terms of subregional summaries, Central and Southern Asia has the highest average area under cultivation (21.8 million hectares), mainly driven by countries such as India, Bangladesh, and Indonesia. Crop-level disaggregation also reveals that, on average, rice cultivation (6.5 million hectares) is the highest relative to the other commodities. The median values are also reported in the table, indicating that, while the general trends revealed by the mean values remain equal, the distribution of area estimates across all disaggregation levels is skewed to the right (table 1).

Summary statistics for yield (tons/ha)—disaggregated by regional, subregional, and crop level—are presented in table *2*. In terms of yield (tons/ha), the productivity rate also varies widely with the mean yield ranging from 0.9 to 2.4 tons/ha. The overall aggregated average yield across all IFSA countries is 1.5 tons/ha. The corresponding median value is 1.3 tons per hectare. Regionally, North Africa and Asia have the highest mean values of yield with 2.4 tons/ha and 2.1 tons/ha, respectively. Crop-level disaggregated estimates show that the average productivity of wheat yield is 1.8 tons/ha, followed by rice with an average value of 1.6 tons/ha. The average yield for coarse grain is the lowest (1.2 tons/ha) due to the lower per hectare productivity of sorghum and millet crops.

Summary statistics for the overall production values (thousand tons) are presented in table 3. The mean and median production estimates for the IFSA countries are 6.5 million tons and 1.4 million tons, respectively. Asia, driven by country-level estimates from the Central and Southern Asia subregion and Southeast Asia subregion, has the highest average production estimate compared with the other regions. At a subregion level, Central and Southern Asia has the highest average production (42.5 million tons), mainly because of the higher-than-average area allocated in both India and Bangladesh. Average total production is the highest for rice (12.7 million tons) due to the higher-than-average area allocated to this crop.

Table 1

Aggregated, regional, subregional, and crop-level area summary statistics, 1991–2019

	N*	Mean (Thousand hectares)	Median (Thousand hectares)	Standard deviation (Thousand hectares)
International Food Security Assessment aggregate	2,726	3,608	946	11,988
Regional disaggregation				
Asia	749	8,558	1,171	21,681
Latin America and the Caribbean	418	558	483	358
North Africa	152	2,878	2,688	1,507
Sub-Saharan Africa	1,169	2,240	1,151	3,309
Subregional disaggregation				
Commonwealth of Independent States	264	644	587	425
Central and Southern Asia	228	21,848	7,015	35,440
Southeast Asia	190	6,448	6,566	4,855
Other Asia	67	501	475	226
Central Africa	76	734	462	691
East Africa	341	2,669	1,864	2,718
Southern Africa	258	1,094	1,174	687
West Africa	494	2,775	1,113	4,368
Crop-level disaggregation				
Wheat	781	1,629	590	2,776
Rice	1,036	6,560	1,146	18,808
Coarse grains	909	1,943	1,023	2,321

Notes: N* = number of observations in each sample. The International Food Security Assessment (IFSA) refers to the overall aggregate model specification. The regional disaggregation refers to: Sub-Saharan Africa, Asia, Latin America and the Caribbean, and North Africa. The subregional specifications are: Central Africa, East Africa, Southern Africa, West Africa, North Africa, Latin America and the Caribbean, Commonwealth of Independent States, Central and South Asia, Southeast Asia, and Other Asia. Coarse grains refer to maize, sorghum, and millet.

Source: USDA, Economic Research Service using data from the USDA, Foreign Agricultural Service's Production, Supply, and Distribution database (2021).

Table 2

Aggregated, regional, subregional, and crop-level yield summary statistics, 1991–2019

	N*	Mean (Tons per hectare)	Median (Tons per hectare)	Standard deviation (Tons per hectare)
International Food Security Assessment aggregate	2,726	1.5	1.3	0.9
Regional disaggregation				
Asia	749	2.1	2.0	0.8
Latin America and the Caribbean	418	1.9	1.7	0.7
North Africa	152	2.4	1.5	2.1
Sub-Saharan Africa	1,169	1.1	1.0	0.5
Subregional disaggregation				
Commonwealth of Independent States	264	2.4	2.2	0.8
Central and Southern Asia	228	2.0	1.9	0.6
Southeast Asia	190	2.1	2.2	0.8
Other Asia	67	1.0	0.9	0.3
Central Africa	76	1.1	0.9	0.4
East Africa	341	1.1	1.1	0.5
Southern Africa	258	1.2	1.1	0.6
West Africa	494	0.9	0.9	0.3
Crop-level disaggregation				
Wheat	781	1.8	1.5	1.2
Rice	1,036	1.6	1.4	0.8
Coarse grains	909	1.2	1.1	0.7

Notes: N* = number of observations in each sample. The International Food Security Assessment (IFSA) refers to the overall aggregate model specification. The regional disaggregation refers to: Sub-Saharan Africa, Asia, Latin America and the Caribbean, and North Africa. The subregional specifications are: Central Africa, East Africa, Southern Africa, West Africa, North Africa, Latin America and the Caribbean, Commonwealth of Independent States, Central and South Asia, Southeast Asia, and Other Asia. Coarse grains refer to maize, sorghum, and millet.

Source: USDA, Economic Research Service using data from the USDA, Foreign Agricultural Service's Production, Supply, and Distribution database (2021).

Table 3

Aggregated, regional, subregional, and crop-level production summary statistics, 1991–2019

	N*	Mean (Thousand tons)	Median (Thousand tons)	Standard deviation (Thousand tons)
International Food Security Assessment aggregate	2,726	6,492	1,353	23,525
Regional disaggregation				
Asia	749	17,698	2,769	42,474
Latin America and the Caribbean	418	1,162	780	1,031
North Africa	152	6,618	4,106	6,043
Sub-Saharan Africa	1,169	2,391	1,205	4,007
Subregional disaggregation				
Commonwealth of Independent States	264	1,690	1,309	1,672
Central and Southern Asia	228	42,492	11,855	68,940
Southeast Asia	190	16,254	11,504	14,861
Other Asia	67	490	467	245
Central Africa	76	998	452	1,145
East Africa	341	3,050	2,135	3,716
Southern Africa	258	1,308	1,261	1,041
West Africa	494	2,715	997	5,142
Crop-level disaggregation				
Wheat	781	3,468	1,000	6,696
Rice	1,036	12,667	1,706	36,811
Coarse grains	909	2,052	1,353	2,580

Notes: N* = number of observations in each sample. The International Food Security Assessment (IFSA) refers to the overall aggregate model specification. The regional disaggregation refers to: Sub-Saharan Africa, Asia, Latin America and the Caribbean, and North Africa. The subregional specifications are: Central Africa, East Africa, Southern Africa, West Africa, North Africa, Latin America and the Caribbean, Commonwealth of Independent States, Central and South Asia, Southeast Asia, and Other Asia. Coarse grains refer to maize, sorghum, and millet.

Source: USDA, Economic Research Service using data from the USDA, Foreign Agricultural Service's Production, Supply, and Distribution database (2021).

Estimation Results

The econometric estimation results of the four different specifications of the yield equation (i.e., equation 4) are summarized in table 4. The year variable is strongly statistically significant for all of the model specifications, indicating the technological progress during the analysis period. Yield responses to the trends captured by the 2-year and 5-year moving averages provide the short- to medium-term dynamics of return to yield after accounting for input costs. In most cases, these two terms are positive and significant at the 5- and 1-percent levels, with varying magnitudes depending on the model specifications.

Table 4

Estimation results of four yield equation specifications: aggregated, regional, subregional, and crop-level disaggregated specifications, 2011–2019

Model specifications	Year		2-year moving average		5-year moving average				
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error	Observation	Log likelihood	Akaike information criterion
International Food Security Assessment aggregate	0.020***	0.001	0.300***	0.030	0.300***	0.030	2,422	-57.4	272.9
Regional model									
Asia	0.040***	0.001	0.010	0.030	0.200***	0.040	671	150.6	-275.2
Latin America and the Caribbean	0.020***	0.001	0.200***	0.100	0.400***	0.100	374	-18.6	65.3
North Africa	0.030***	0.003	0.200**	0.100	0.300**	0.100	136	-47	108.1
Sub-Saharan Africa	0.010***	0.001	0.600***	0.040	0.100**	0.050	1,241	373.4	-662.9
Subregional model									
Commonwealth of Independent States	0.030***	0.003	0.800***	0.100	0.100	0.100	232	-60.8	143.6
Central and South Asia	0.040***	0.001	0.040	0.040	0.100***	0.100	204	106.2	-194.4
Southeast Asia	0.040***	0.001	-0.010	0.040	0.200***	0.100	170	74.1	-132.2
Other Asia	0.010**	0.003	1.100***	0.200	-0.200	0.200	65	9.7	-7.5
Central Africa	0.010***	0.001	0.800***	0.100	0.200	0.200	136	106	-197.9
East Africa	0.010***	0.001	1.000***	0.100	0.200**	0.100	301	133.2	-240.4
Southern Africa	0.010***	0.002	0.900***	0.100	0.200	0.100	260	-1.9	27.9
West Africa	0.010***	0.001	0.400***	0.040	0.100***	0.040	544	333.4	-628.7
Crop model									
Wheat	0.020***	0.001	0.600***	0.100	0.300***	0.100	675	-86.2	222.4
Rice	0.020***	0.001	0.100***	0.030	0.300***	0.030	924	174.6	-287.1
Coarse grains	0.010***	0.001	1.000***	0.100	0.100***	0.100	823	145.3	-232.7

Notes: Country fixed effect included in all model specifications. The International Food Security Assessment (IFSA) refers to the overall aggregate model specification. *, **, and *** denote statistical significance levels at 10, 5, and 1 percent, respectively. Coarse grains refer to maize, sorghum, and millet.

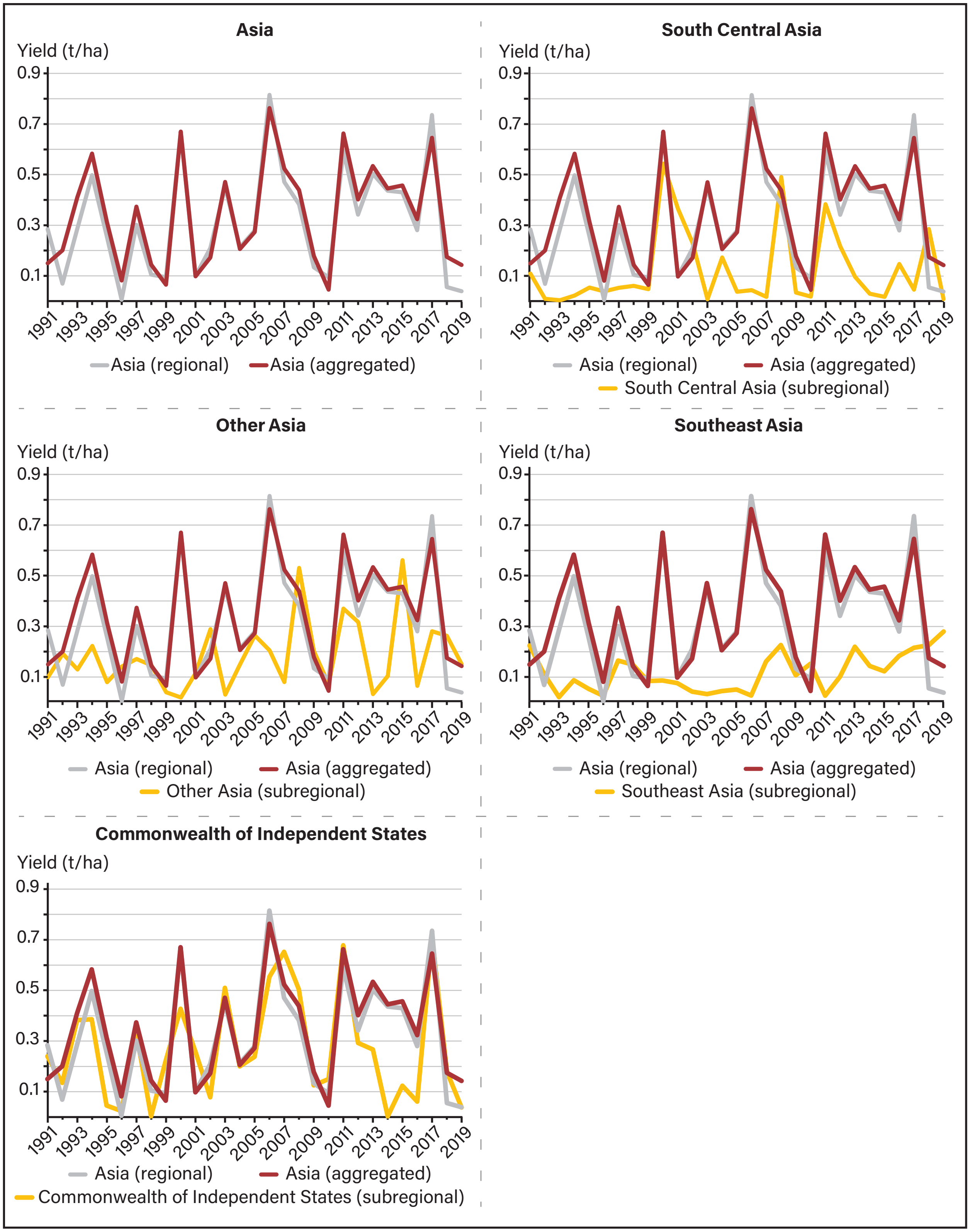
Source: USDA, Economic Research Service calculations based on the International Food Security Assessment 2020 model dataset covering 2011 to 2019.

Prediction Evaluations

Comparisons between the actual yield (i.e., observed yield value) and predicted yield (i.e., the estimated yield value by the model) are presented in figure 1 to figure 4 for each of the four regions. It is considered better (i.e., smaller prediction error) when the absolute yield difference between the actual and the predicted yield value is smaller. The absolute yield difference between the actual and predicted yield is plotted from 1991 to 2019. These figures display heterogeneity within and across the regions. Figure 1, for example, shows that the aggregated model and the regional model display higher absolute yield differences than the individual subregional models in most years. The Southeast Asia subregion particularly has the lowest deviations in projected yield compared with the other model specifications. The historical projections in the Sub-Saharan Africa region do not have any clear pattern apart from their support for the existing heterogeneity in the data. Because the number of countries in both the North Africa and Latin America and the Caribbean regions is small, the regional and subregional models are the same. Therefore, the comparison of these models is with respect to the IFSA aggregate model. The results show a similar trend for the Latin America and the Caribbean dataset. Whereas for the North Africa dataset, the world model is heavily weighted by the rest of the other regions (subregions) and hence shows lower absolute historical deviations compared with the North Africa region and subregion.

Figure 1

Historical comparison of absolute yield difference between actual and predicted values in the Asia region, 1991–2019

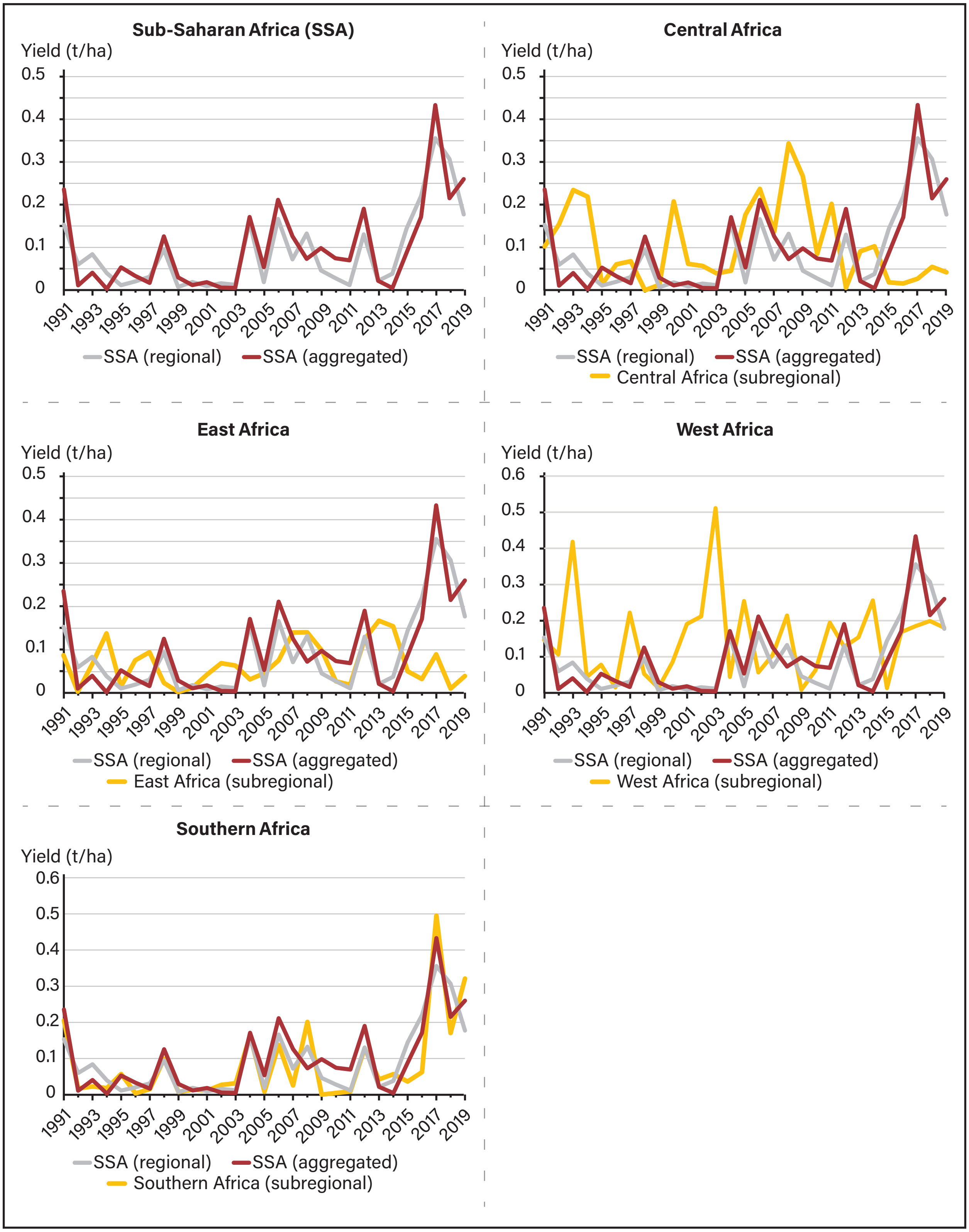


t/ha = metric tons per hectare.

Source: USDA, Economic Research Service calculations based on the International Food Security Assessment 2020 model.

Figure 2

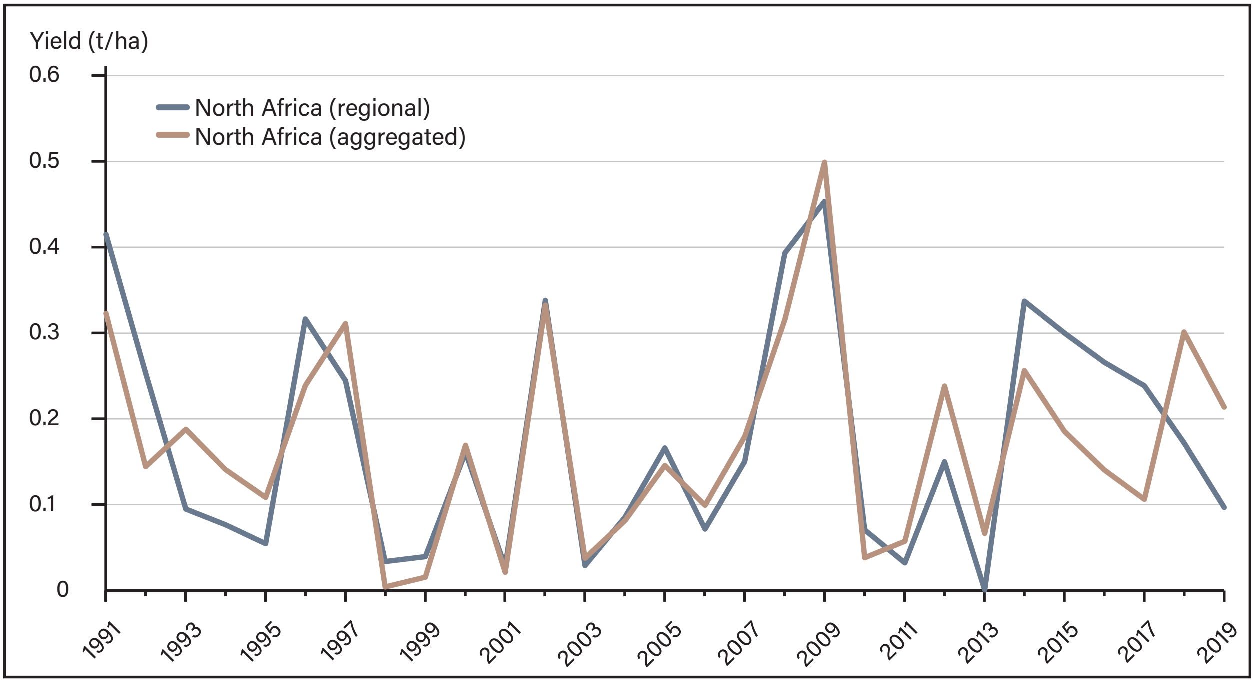
Historical comparison of absolute yield difference between actual and predicted values in the Sub-Saharan Africa region, 1991–2019



t/ha = metric tons per hectare.

Source: USDA, Economic Research Service calculations based on the International Food Security Assessment 2020 model.

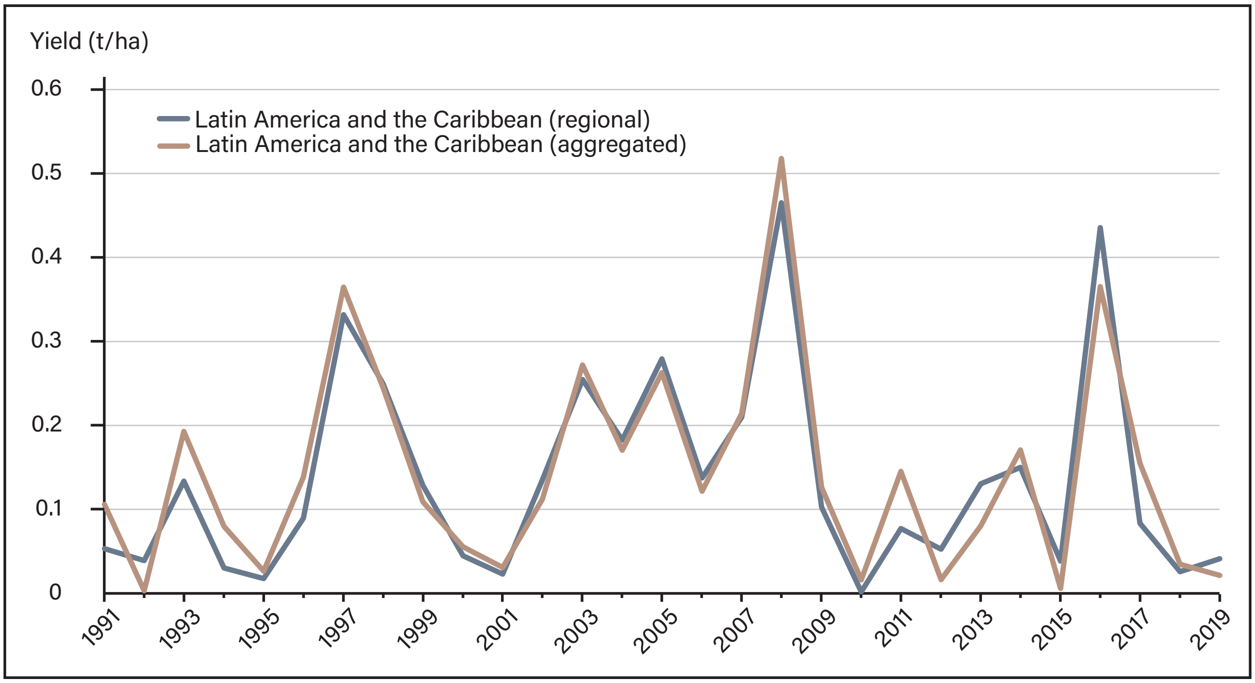
Figure 3
Historical comparison of absolute yield difference between actual and predicted values in the North Africa region, 1991–2019



t/ha = metric tons per hectare.

Source: USDA, Economic Research Service calculations based on the International Food Security Assessment 2020 model.

Figure 4
Historical comparison of absolute yield difference between actual and predicted values in the Latin American and the Caribbean region, 1991–2019



t/ha = metric tons per hectare.

Source: USDA, Economic Research Service calculations based on the International Food Security Assessment 2020 model.

The prediction performances of regional and subregional model specifications are assessed using their MSE scores. Tables 5 and 6 provide the in-sample and out-of-sample MSEs of the model specifications, respectively. The average overall MSE for the regional and subregional model specifications are obtained by taking the weighted average of the respective individual regional and subregional MSEs. The weight in the MSE average is defined as the ratio of the number of countries in each subregion divided by the total number of IFSA countries (i.e., 76). The in-sample MSE is normally lower than the out-of-sample MSE. The MSE for the aggregate IFSA model is 0.061, and the corresponding values for the regional, subregional, and crop model specifications are 0.055, 0.045, and 0.051, respectively. Similarly, the MAE for the aggregate IFSA model is 0.182, and the corresponding values for the regional, subregional, and crop model specifications are 0.168, 0.152, and 0.164, respectively. These results directly corroborate the historical yield deviations presented in figures 1 to 4.

Out-of-sample scores presented in table 6 show that the MSE for the aggregate IFSA model is 0.066. The average regional, subregional, and crop MSEs—0.060, 0.050 and 0.055, respectively—indicate that the regional model's prediction performance is 8.5 percent, the subregional model's performance is 23.9 percent, and the crop model is 16.7 percent better than the aggregate IFSA model. Thus, the subregional specification appears to perform best, based on its low MSE. Similarly, based on the MAE, the subregional model specification improves the prediction performance by 15.4 percent relative to the aggregate IFSA model. However, it is worth mentioning that despite the weighted average MSE of the subregional model being relatively the smallest of all other aggregations, the model also contains subregions with the highest individual MSE scores (e.g., the Commonwealth of Independent States and the Southern Africa subregions have 0.11 and 0.09 MSE scores, respectively, among the highest scores).

Table 6 also compares the prediction performance criteria for each of the disaggregated models. The Sub-Saharan Africa model has the lowest MSE (0.035), while the North Africa model has the highest (0.132). Of the subregional disaggregated models, Central Africa (0.014) and West Africa (0.014) have the lowest MSEs, while the Commonwealth of Independent States (0.109) and North Africa (0.132) have the highest.

Table 5

The estimated in-sample mean squared errors, root mean squared errors, and mean absolute errors for the various model specifications

Model specifications	MSE	MSE percent change relative to aggregate model	RMSE	RMSE percent change relative to aggregate model	MAE	MAE percent change relative to aggregate model
International Food Security Assessment aggregate	0.0613	NA	0.2477	NA	0.1820	NA
Regional	0.0553	-9.8	0.2288	-7.6	0.1675	-8.0
Subregional	0.0450	-26.6	0.1997	-19.4	0.1516	-16.7
Crop	0.0506	-17.5	0.2225	-10.2	0.1639	-9.9
Regional models						
Sub-Saharan Africa	0.0320	-47.8	0.1789	-27.8	0.1274	-30.0
Asia	0.0810	32.1	0.2846	14.9	0.2114	16.2
Latin America and the Caribbean	0.0643	4.9	0.2536	2.4	0.1878	3.2
North Africa	0.1152	87.9	0.3394	37.0	0.2610	43.4
Subregional models						
Commonwealth of Independent States	0.0981	60.0	0.3131	26.4	0.2342	28.7
Central and South Asia	0.0205	-66.6	0.1431	-42.2	0.1062	-41.6
Southeast Asia	0.0242	-60.5	0.1556	-37.2	0.1291	-29.1
Other Asia	0.0421	-31.3	0.2051	-17.2	0.1619	-11.0
Central Africa	0.0121	-80.3	0.1102	-55.5	0.0795	-56.3
East Africa	0.0240	-60.8	0.1549	-37.5	0.1185	-34.9
Southern Africa	0.0590	-3.8	0.2428	-2.0	0.1811	-0.5
West Africa	0.0171	-72.1	0.1309	-47.2	0.1022	-43.8
Crop models						
Wheat	0.0754	23.0	0.2745	10.8	0.2026	11.3
Rice	0.0400	-34.7	0.2001	-19.2	0.1552	-14.7
Coarse grains	0.0410	-33.1	0.2026	-18.2	0.1405	-22.8

NA = not applicable. MSA = Mean Squared Errors. RMSE = Root Mean Squared Errors. MEA = Mean Absolute Errors.

Note: The International Food Security Assessment (IFSA) refers to the overall aggregate model specification.

Source: USDA, Economic Research Service calculations using the International Food Security Assessment 2020 model dataset covering 2011–2019.

Table 6

The estimated out-of-sample mean squared Errors (MSEs), root mean squared errors (RMSEs), and mean absolute Errors (MEAs) for the various model specifications

Model specifications	MSE	MSE percent change relative to aggregate model	RMSE	RMSE percent change relative to aggregate model	MAE	MAE percent change relative to aggregate model
International Food Security Assessment aggregate	0.0658	NA	0.2565	NA	0.1884	NA
Regional	0.0602	-8.5	0.2386	-7.0	0.1743	-7.5
Subregional	0.0501	-23.9	0.2105	-17.9	0.1593	-15.4
Crop	0.0548	-16.7	0.2317	-9.7	0.1702	-9.7
Regional models						
Sub-Saharan Africa	0.0345	-47.6	0.1856	-27.6	0.1320	-29.9
Asia	0.0881	33.9	0.2969	15.8	0.2201	16.8
Latin America and the Caribbean	0.0702	6.7	0.2649	3.3	0.1956	3.8
North Africa	0.1316	100.0	0.3627	41.4	0.2768	46.9
Subregional models						
Commonwealth of Independent States	0.1088	65.3	0.3298	28.6	0.2463	30.7
Central and South Asia	0.0225	-65.8	0.1500	-41.5	0.1114	-40.9
Southeast Asia	0.0270	-59.0	0.1643	-35.9	0.1360	-27.8
Other Asia	0.0536	-18.5	0.2316	-9.7	0.1809	-4.0
Central Africa	0.0138	-79.0	0.1176	-54.2	0.0844	-55.2
East Africa	0.0264	-59.9	0.1624	-36.7	0.1240	-34.2
Southern Africa	0.0943	43.3	0.2567	0.1	0.1907	1.2
West Africa	0.0185	-71.9	0.1359	-47.0	0.1060	-43.7
Crop models						
Wheat	0.0819	24.5	0.2862	11.6	0.2108	11.9
Rice	0.0431	-34.5	0.2075	-19.1	0.1608	-14.6
Coarse grains	0.0447	-32.1	0.2115	-17.5	0.1459	-22.6

NA = not applicable. MSA = Mean Squared Errors. RMSE = Root Mean Squared Errors. MEA = Mean Absolute Errors.

Note: The International Food Security Assessment (IFSA) refers to the overall aggregate model specification.

Source: USDA, Economic Research Service calculations using the International Food Security Assessment 2020 model dataset covering 2011–2019.

Conclusion and Recommendation

Given that the primary objective of the IFSA supply-side exercise is to project production in the future, the analysis presented in this study provides alternative selection criteria among MSE, MRSE, or MAE (i.e., the best performing model) for projecting IFSA's agricultural production estimates. For this exercise, the subregional production model specification provided the best prediction result using the three criteria—MSE, MRSE, or MAE. For example, based on the MAE results presented in table 6, the subregional model specification provides a 15.4-percent improvement in the prediction performance relative to the aggregate IFSA model specification. The out-of-sample mean absolute difference between the actual and predicted yield is 0.188 tons per hectare based on the aggregate IFSA model. This figure declines to 0.159 tons per hectare based on the subregional model (table 6). The variability in the predication results, as shown in table 6, indicates the need for addressing the geographic and crop-level heterogeneity in the model.

Within the IFSA modeling framework, the domestic production projections—together with the total grain demand projections—are used to quantify the amount that is known as the Implied Additional Supply Required (IASR). The IASR quantifies the total grain demand in each country that is not projected to be met through domestic production (Thome et al., 2019). These figures show wide-ranging implications for country-level and regional demand and supply dynamics that are interrelated with each other. The total grain demand that is unmet through domestic production may either be supplied through trade arrangements or some sort of aid. For example, for all 76 IFSA countries, the total grain demand for 2020 and 2030 was estimated to be 912 million metric tons and 1,205 million metric tons, respectively, growing at an annual rate of 2.8 percent. Similarly, the estimated domestic production for 2020 and 2030 was 644.8 million metric tons and 788.7 million metric tons, respectively, with an annual growth rate of 2 percent. The total IASR for 2020 was then estimated at 267.4 million metric tons, and the corresponding figure for the year 2030 was 416.5 million metric tons, growing annually at a rate of 4.5 percent (Baquedano et al., 2020). This implies that, for 2030, 65 percent of the total grain demand for all IFSA countries is projected to be met through domestic production. Accurate estimations and projections of the domestic production that accounts for geographic and crop level heterogeneities will help increase the accuracy of yield projections that will ultimately enhance the supply and demand dynamics in the IFSA assessment. Having better production estimates, in general, is important as shortfalls in output may have 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). Better projections of output in the IFSA countries will also help provide more robust estimates that complement USDA's ERS baseline work.

References

Baquedano, F., C. Christensen, J. Beckman, and K. Ajewole. 2020. *International Food Security Assessment 2020–30 of the Economic Research Service of the USDA*, U.S. Department of Agriculture, Economic Research Service.

Beghin, J., B. Meade, and S. Rosen. 2015. *A Consistent Food Demand Framework for International Food Security Assessment*, U.S. Department of Agriculture, Economic Research Service.

Canty A., and B.D. Ripley. 2021. *Boot: Bootstrap R (S-Plus) Functions*, R package version 1.3–28.

Friedman, J., T. Hastie, and R. Tibshirani. 2010. "Regularization Paths for Generalized Linear Models Via Coordinate Descent," *Journal of Statistical Software* 33(1):1.

Hertel, B.T.W., U.L.C. Baldos, and K.O. Fuglie. 2020. "Trade in Technology: A Potential Solution to the Food Security Challenges of the 21st Century," *European Economic Review* 127:103479.

Jagermeyr, J., A. Robock, J. Elliott, C. Muller, L. Xia, N. Khabarov, C. Folberth, E. Schmid, W. Liu, F. Zabel, S.S. Rabin, M.J. Puma, A. Heslin, J. Franke, I. Foster, S. Asseng, C.G. Bardeen, O.B. Toon, and C. Rosenzweig. 2020. *A Regional Nuclear Conflict Would Compromise Global Food Security*, Proceedings of the National Academy of Sciences of the United States of America *117*(13):7071–7081.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2017. *An Intoduction to Statistical Learning:With Applications in R.* Springer Publishing, New York, N.Y.

R Core Team. 2021. R: *A Language and Environment for Statistical Computing* (software), R Foundation for Statistical Computing, Vienna, Austria.

Smith, V. H., and J.W. Glauber. 2020. "Trade, Policy, and Food Security," *Agricultural Economics (United Kingdom)*, *51*(1):159–171.

Thome, K., M.D. Smith, K. Daugherty, N. Rada, C. Christensen, and B. Meade. 2019. *International Food Security Assessment, 2019–2029*, *Electronic Outlook Report From the Economic Research Service*, GFA-30, U.S. Department of Agriculture, Economic Research Service.

U.S. Department of Agriculture, Foreign Agricultural Service, 2021. Production, Supply, and Distribution Database, 2021.

U.S. Department of Agriculture, World Agricultural Outlook Board, and Interagency Agricultural Projections Committee. 2020. *USDA Agricultural Projections to 2029*," *Long-Term Projections Report to OCE-2020-1* 1–124.

Appendix

International Food Security Assessment countries included in the study

Asia	**Latin America and the Caribbean**	**Sub-Saharan Africa**	
Commonwealth of Independent States	Bolivia	**Central Africa**	**Southern Africa**
Armenia	Colombia	Cameroon	Angola
Azerbaijan	Ecuador	Central African Republic	Lesotho
Georgia	Peru	Congo, Republic of the	Madagascar
Kyrgyzstan	Dominican Republic	Congo, Democratic Republic of the	Malawi
Moldova	El Salvador	**East Africa**	Mozambique
Tajikistan	Guatemala	Burundi	Namibia
Turkmenistan	Haiti	Chad	Eswatini
Uzbekistan	Honduras	Eritrea	Zambia
Central and Southern Asia	Jamaica	Ethiopia	Zimbabwe
Afghanistan	Nicaragua	Kenya	**West Africa**
Bangladesh	**North Africa**	Rwanda	Benin
India	Algeria	Somalia	Guinea-Bissau
Nepal	Egypt	Sudan	Burkina Faso
Pakistan	Morocco	Tanzania	Cabo Verde
Sri Lanka	Tunisia	Uganda	Côte d'Ivoire
Other Asia			Gambia
Korea, Democratic People's Republic of			Ghana
Mongolia			Guinea
Yemen			Liberia
Southeast Asia			Mali
Cambodia			Mauritania
Indonesia			Niger
Laos			Nigeria
Philippines			Senegal
Vietnam			Sierra Leone
			Togo