U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Economic Research Report Number 285

February 2021

Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area

David Boussios, Sharon Raszap Skorbiansky, and Matthew MacLachlan

U.S. DEPARTMENT OF AGRICULTURE

Recommended citation format for this publication:

David Boussios, Sharon Raszap Skorbiansky, and Matthew MacLachlan. *Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area* ERR-285, U.S. Department of Agriculture, Economic Research Service, February 2021.

U.S. DEPARTMENT OF AGRICULTURE

Economic Research Service

Economic Research Report Number 285

February 2021

Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area

David Boussios, Sharon Raszap Skorbiansky, and Matthew MacLachlan

Abstract

This report examines the potential for statistical forecast models to improve the performance of the U.S. Department of Agriculture's (USDA) long-term agricultural baseline projections for the harvested area for U.S. corn, soybeans, and wheat. After-the-fact analysis for years 1997 to 2017 reveals the baseline projections have, historically, consistently overestimated the harvested area of wheat and underestimated soybean area. The baseline projections also tend to underestimate the corn area, though to a lesser degree. Part of the difference between the projections and realized values is likely attributable to policy, program, weather, and other unforeseen changes when USDA developed the projections. Still, the results of quantitative forecast models show there may be substantial potential for improvement on the existing methodology. Forecasts generated using 3 econometric time-series models did not improve performance relative to the current baseline approach for nearer forecast horizons but improved performance for projection horizon lengths of 8-10, 2-10, and 4-10 years for harvested area of corn, soybeans, and wheat, respectively, when using 1 of our statistical measures. The forecasts generated using the econometric models produce predictions with an average absolute forecasting error 10 years out that is between 26 percent to 60 percent smaller than those provided by baseline projections. The results suggest that econometric models offer the potential to improve the performance of forecasting long-term trends in agricultural markets. As of 2020, USDA begun using statistical forecast models such as these when developing its long-term agricultural projections as complements to the existing process. USDA is also in the process of testing these models for additional commodities to improve the long-term projections for all commodities.

Keywords: Agricultural baseline projections, harvested area, corn, soybeans, wheat, forecast, real-time dataset.

Acknowledgments

The authors would like to thank Todd Kuethe of Purdue University; Seth Meyers of USDA, Office of the Chief Economist; and four anonymous reviewers for their helpful comments on earlier drafts of this report. They also thank Carol Ready and Andres Guerrero, of USDA, Economic Research Service (ERS), for editorial and design services.

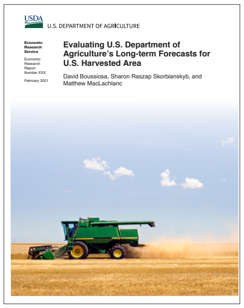
About the Authors

David Boussios is an economist with USDA, World Agricultural Outlook Board. Sharon Raszap Skorbiansky is an economist with USDA, Farm Production and Conservation Business Center. Matthew MacLachlan is an economist with USDA, ERS.

Contents

Summary i
Introduction 1
Forecast Errors of the Projections 6
Time-Series Analysis 10
Findings 12
Forecast Comparisons 16
Discussion 19
References 20
Appendix: model selection of time-series models 22

A report summary from the Economic Research Service February 2021



Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area

David Boussios, Sharon Raszap Skorbiansky, and Matthew MacLachlan

What Is the Issue?

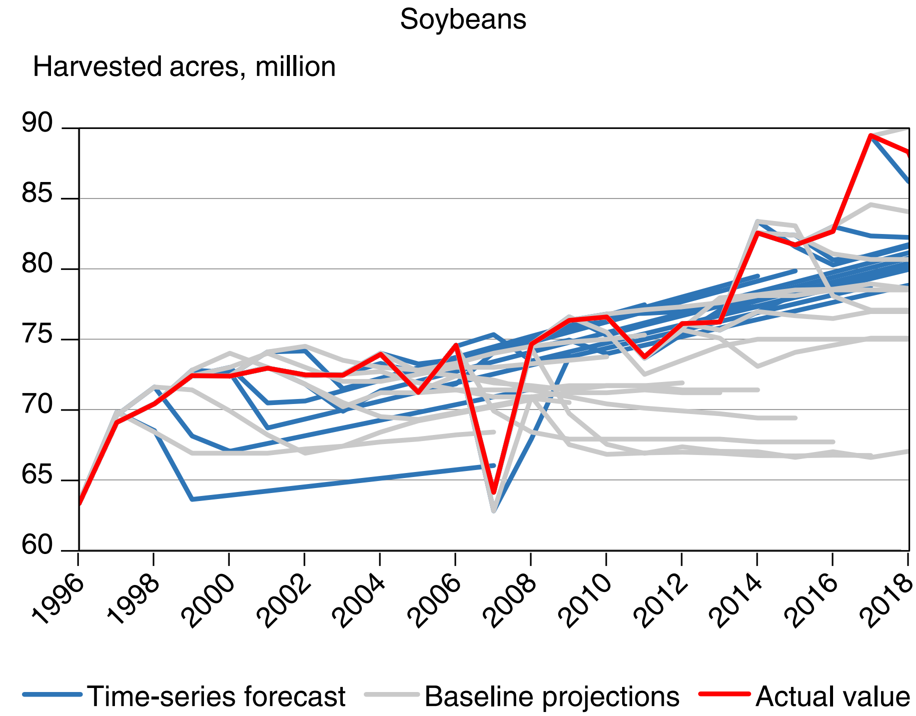
The long-term baseline projections of the U.S. Department of Agriculture (USDA) are a departmental consensus that provides commodity projections for the trajectory of global agricultural markets 10 years into the future. The baseline model is used in policy and budgetary matters and by USDA to evaluate the effect of shocks on agricultural markets. For example, the baseline projections are a key component for forecasting USDA outlays in the President's budget. These 10-year projections, which span many commodities and countries, are based on a combination of modeling and expert opinion. The baseline projections are based on incomplete knowledge about the future, so it is to be expected that some divergence between the projections and actual values will occur. Several modeling techniques are available, and USDA engages in continuous efforts to improve the accuracy of projections. This report analyzes the historical baseline projections for U.S. corn, soybean, and wheat harvested area for the period from 1997 to 2017 and examines the potential for two types of modeling approaches to improve the performance of USDA's baseline projections. While these crops are only one piece of the entire set of baseline projections, the three crops analyzed in this report make up the lion's share of U.S. planted acreage and, as such, these estimates are widely reviewed by USDA stakeholders due to their importance for policy and the public. As of the release of this report, USDA has taken steps to introduce statistical forecast models to the baseline process. First, in 2020, USDA estimated time-series models alongside the current baseline process for corn, soybeans, and wheat as part of the 2020 baseline projections. Second, USDA is in the process of testing these models with additional structure and for additional commodities.

What Did the Study Find?

This study found that the baseline projections consistently overestimated the harvested area of wheat area harvested while simultaneously underestimating soybean area. For example, when predicting seven periods ahead, the baseline projections had a mean percentage error of 10.4 percent and -7.3 percent for wheat and soybeans, respectively. The mean percentage error is a measure of forecast bias, i.e. the tendency to over-forecast or under-forecast repeatedly. Harvested corn area was also statistically underestimated, but to a lesser degree than the other two crops studied, with a mean percentage error of -3.3 percent when projecting seven periods ahead.

Naïve and statistical forecasts were used to examine whether the performance of the baseline projections could be improved. A naïve forecast uses only the most recent observation and assumes it will continue at that level for the duration of the forecast horizon. The statistical or time-series econometric models used here use the historical relationship between outcomes over time to forecast 10 periods into the future (or 10 forecast horizons). The study found that generally, both naïve and statistical forecasts outperformed the baseline projections over longer horizons but were less accurate for projecting some less distant outcomes. Results varied by crop.

Economic forecasts result in smaller forecast errors relative to baseline



Note: The red line is the final realized outcome. Gray lines are the baseline's projected values. Blue lines are time-series projections from econometric results.

Source: USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats. Time-series projections calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017

Forecasts generated using time-series models did not improve performance relative to the current baseline for projections in the first 3 years but did for more distant ones. The statistical forecasts outperformed baseline projections for soybean and wheat areas in 9 and 7 of 10 forecast horizons considered, respectively. The forecasts outperformed the baseline projections in 3 out of 10 horizon lengths for corn. The statistical forecasts reduced the mean absolute percentage error, a measure of forecast accuracy, by 60 percent relative to baseline projections for soybean forecasts of horizon lengths of 10 years. The reduction in forecast error (the difference between the realized and predicted values) was equivalent to improving the performance of the forecast by 6 million acres in 2018 values. The results indicate that statistical forecasting methods are a useful addition to the baseline process, particularly in forecasting more distant outcomes for corn, soybeans, and wheat. USDA has begun to incorporate these models into the baseline process for corn, soybeans, and wheat and evaluate their performance for additional countries and commodities.

How Was the Study Conducted?

This study examined the accuracy of the USDA's long-term (10-year) projections of the harvested area of U.S. corn, soybeans, and wheat compared to the realized (actual) harvested area from 1997 to 2017. The study used the harvested area for corn, soybeans, and wheat as known at the time of each baseline projection. USDA researchers compared the relative performance of forecasts generated by three econometric models and one naïve specification to the baseline projections for these three crops using the same harvested acreage data as known at the time of each baseline projection. The econometric models were selected based on their suitability to reproducing the data generating process, as well as simplicity (parsimony). The quality of the estimates was statistically evaluated based on how frequently they over- or underestimated the actual values, as well as the size of the difference between the projected and actual values.

Evaluating U.S. Department of Agriculture's Long-Term Forecasts for U.S. Harvested Area

Introduction

The U.S. Department of Agriculture (USDA) Interagency Agricultural Projections Committee generates annual long-run, 10-year agricultural projections of 44 major exporting and importing countries and regions. These projections—called the long-term agricultural projections by USDA, baseline forecast by the General Accountability Office (GAO 1991, 1998), or often simply the baseline projections—are released every February.[1] The projections inform market participants and policymakers on a scenario for the agricultural sector and USDA's expectations for the future trajectory of agricultural markets given the scenario assumptions. The baseline projections are a critical component for projecting USDA expenditures as part of the President's budget each year and conducting program analysis. The results of the baseline projections are also widely used in scenario analyses to evaluate the long-term impact of policy changes on agricultural markets (e.g., Nigatu et al., 2017). To distinguish between baseline results and results from our statistical models later introduced, we reserve the word "forecast" for predictions from statistical models and projection for predictions from the annual baseline report.

Forecasting often involves combining statistical and econometric methods with the subjective beliefs or judgments of experts (Marris, 1954; Wallis and Whitley, 1991; Pole et al., 2018). The forecasters' expertise is a valuable component for ensuring that projected values do not stray from what is considered practical, as well as for incorporating information not reflected in the data, such as shifts in policy. However, forecasters' judgments can also introduce bias. Bias, a type of inaccuracy in forecasting, refers to the tendency to over-forecast or under-forecast repeatedly. This may occur if the forecaster is overly optimistic or pessimistic on the expected trajectory of the forecasted variable. For example, Batchelor (2001) showed systematic bias in International Monetary Fund and Organisation for Economic Co-operation and Development forecasts, and Batchelor (2007) showed systematic bias in real gross domestic product (GDP) and inflation forecasts by private-sector forecasters.

Several studies evaluated accuracy in other USDA market-related projections (e.g., Egelkraut et al., 2003; Isengildina et al., 2004; Good and Irwin, 2006). Irwin and Good (2015) compared baseline price errors with errors from a forecast created using futures prices, finding that the baseline performed better in shorter time horizons.[2] Thirty years ago, a study requested by the U.S. Senate, and undertaken by the Government Accounting Office (GAO; now the Government Accountability Office), reviewed the performance of USDA's projections in 1988 and 1991 for the season-average price of corn, soybeans, wheat, cotton, and production for dairy products. The GAO found the baseline projections outperformed naïve forecasts over the first two projection years but underperformed over more distant horizons (GAO, 1991). A second report by the GAO found that the consistent bias observed in the projections led to misallocation in the Federal Government's budget.[3] The GAO study found that USDA's budget esti-

[1] Early-release projections for the United States are made public in November of the preceding calendar year.

[2] World Agricultural Supply and Demand Estimates project into the short- and mid-term, up to 16 months forward for crops and 20 months forward for livestock, while baseline involves annual forecasts 10 years into the future.

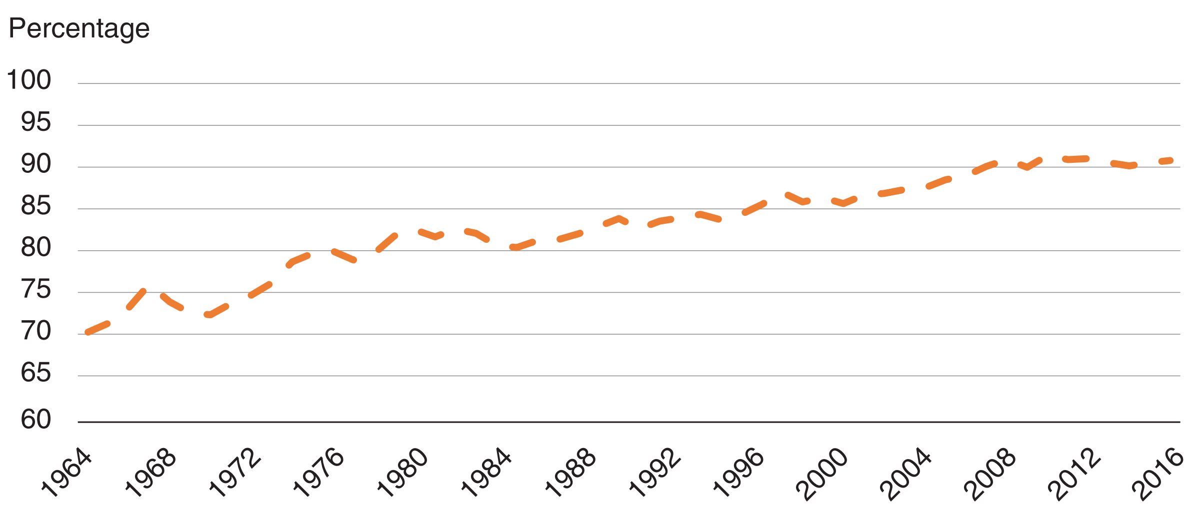
[3] Inaccurate estimation of projected program costs could affect funding decisions for multiyear mandatory spending and consequently, allocation of annual discretionary funds.

mates underestimated actual levels by an average of $3.1 billion per year from 1972 to 1986, stemming from forecast errors in the commodities and variables studied (GAO, 1988). Since then, USDA has taken several steps to improve the baseline process, including improvements in data and record management and providing further documentation of the baseline process (including the recent report by Hjort et al. [2018]).

Forecasting errors in the baseline projections are inevitable because the forecasts are made using incomplete information on future events. However, this does not mean that forecasts cannot be improved. Improving the accuracy of baseline projections is valuable for policymakers, farmers, agricultural industry members, and others who must consider long timelines in their decision making.

In this report, we examined the accuracy and bias of the long-run baseline projections for the harvested acreage (area) of three commodities in the United States: corn, soybeans, and wheat. Since 2007, the combined harvested area of these commodities has exceeded 200 million acres, representing 89 percent of the total harvested area of the commodities forecasted by the baseline projections (figure 1). While the baseline projections predict the entire supply-and-use balance sheet for agricultural commodities, we focused solely on area harvested for the three commodities.[4] We compared the baseline projections with forecasts generated using three time-series econometric models. Evaluating and comparing the forecasting approaches aid in improving the performance and reducing the bias of future projections. The analysis showed baseline projections tend to predict outcomes more accurately than time-series approaches over short horizons but predict outcomes less accurately compared to time-series for more distant projections. Across nearly all horizons and commodities, the time-series forecasts exhibited less bias than the baseline projections. These results suggest that USDA can improve baseline forecasts for corn, soybeans, and wheat by incorporating time-series projections in the baseline process. Additional testing is needed before these results can be extrapolated to additional commodities within the baseline model.

Figure 1
Percentage of major U.S. crop-harvested area in corn, soybeans, and wheat



Note: Major U.S. crops are barley, corn, cotton, oats, rice, sorghum, soybeans, and wheat.

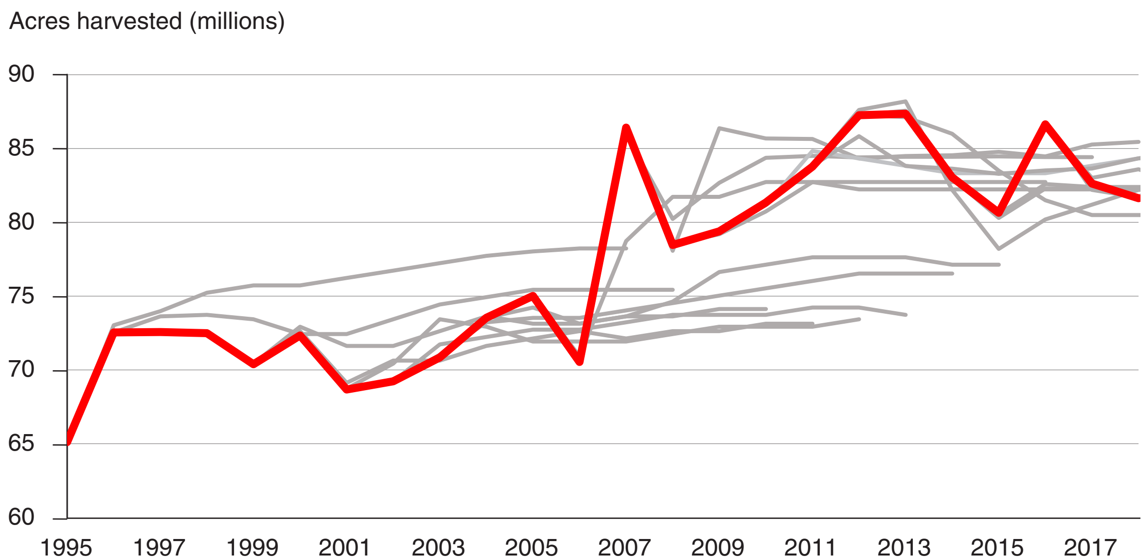
Source: Calculated by USDA, Economic Research Service using data from USDA, Foreign Agricultural Service's Production, Supply, and Distribution database, 2020.

[4] Area is typically used as the primary component to determine the rest of the commodity's supply-and-use balance sheet. We could have analyzed area planted instead; however, production is area harvested multiplied by the quantity produced on area harvested. Aside from area harvested being a function of area planted, the area planted has no direct relation to the supply and use of each commodity.

Historical Projections

This report assesses the historical annual USDA baseline projections from 1997 to 2017 for the harvested area of corn, soybeans, and wheat in the United States. Hjort et al. (2018) provides background on the baseline projection process. We first evaluated these historical projections by visualizations of the predictions and the realized outcomes. Figure 2 presents the annual baseline projections and realized outcomes for harvested corn area from 1997 to 2018. Observationally, there appears to be a change in the structure of harvested corn acreage after 2007 (figure 2). This structural shift is visible not only in the realized levels that spiked in 2007 and remained high in the years that followed but is visible, too, in the projections, none of which forecast harvested area above 80 million acres before 2007.

Figure 2
Harvested corn area, actual and baseline projections

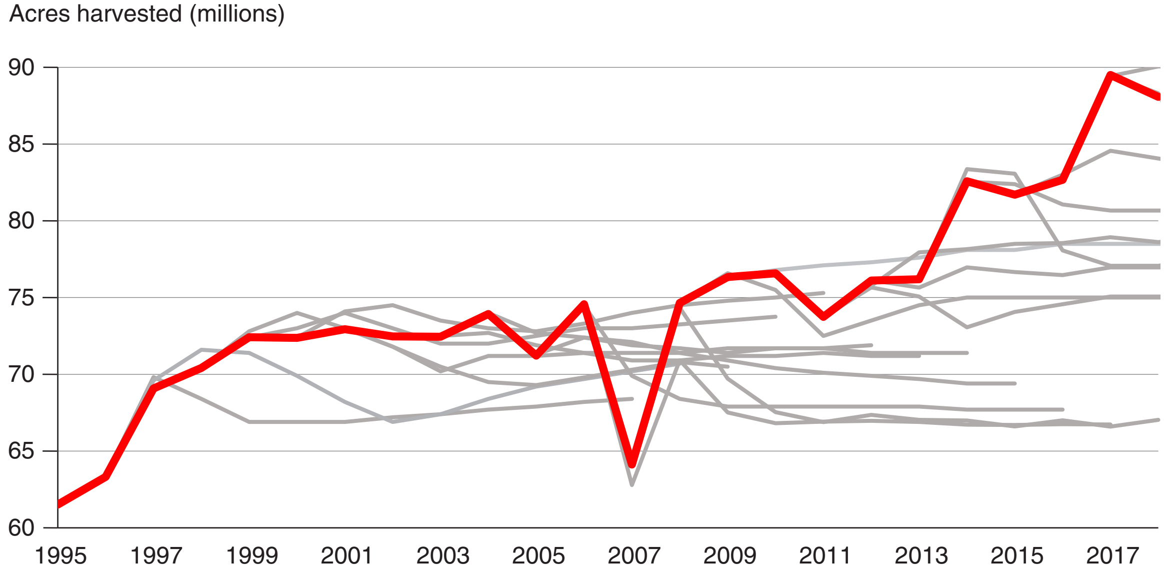


Note: The red line is the final realized outcome. Gray lines are the baseline's projected values.

Source: USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

The Energy Policy Act of 2005 and the Energy Independence and Security Act of 2007, which expanded the Renewable Fuel Standard (RFS) program (U.S. Environmental Protection Agency, 2015), generated a substantial increase in the demand for corn (Roberts and Schlenker, 2013; Condon et al., 2015), which likely caused the structural shift (Motamed et al., 2016). Demand also probably expanded at the time because the North American Free Trade Agreement (NAFTA) in 2008 lifted all tariffs and quotas on corn trade between the United States and Mexico (Zahniser et al., 2019). As with all projections, the baseline can incorporate only information available at the time of analysis. To ensure that predictions remain neutral to policy proposals, the baseline projections must consider only those existing policies.

Figure 3
Harvested soybean area, actual and baseline projections

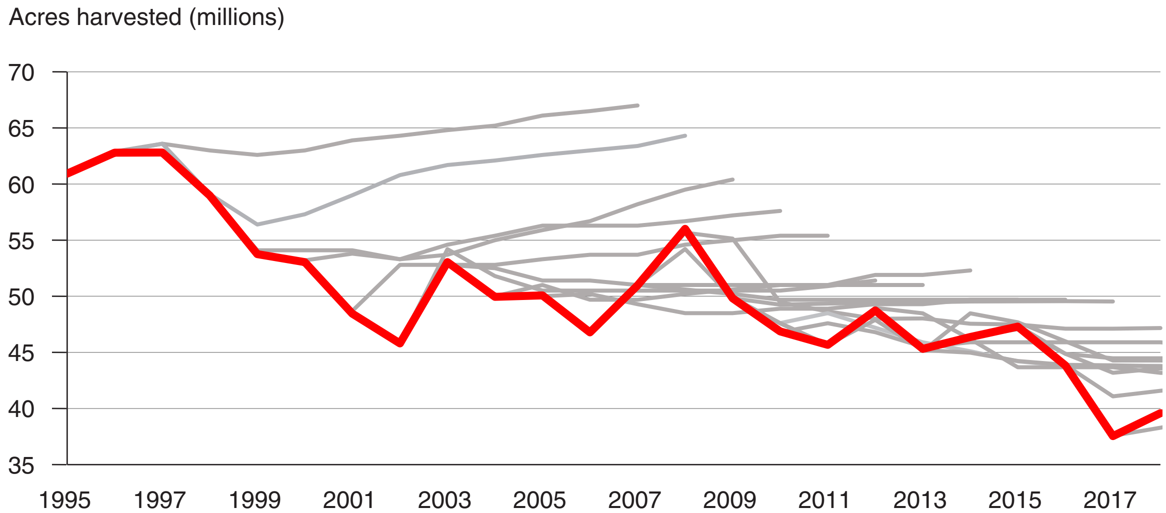


Notes: The red line is the final realized outcome. Gray lines are the baseline's projected values.

Source: USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

Baseline projections for soybeans appear to have frequently underestimated the area ultimately harvested (figure 3). In the late 1990s, soybean area increased, then flattened out until the mid-2000s. In 2007, soybean acreage dropped to a 12-year low following a record 2006 crop and near-record yields. Simultaneously, the price of corn was on the rise as demand for corn ethanol increased, making planting corn more attractive than soybeans (Ash and Dohlman, 2007). Increased global demand for soybeans led to expanded acreage in the late 2000s (Ash and Dohlman, 2006), as did rising demand for livestock feed (Lee et al., 2016). The baseline underestimated soybean harvested area during the periods of relatively flat harvested acreage and as area expanded past 85 million acres. The underestimation indicates a persistent downward bias in the long-run baseline projections.

Figure 4
Harvested wheat area, actual and baseline projections



Notes: The red line is the final realized outcome. Gray lines are the baseline's projected values.

Source: USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

In contrast to the baseline projections of soybean area, overestimation characterized the relationship between wheat projections and actual values (figure 4). Similar to soybean area, wheat area projections repeatedly forecasted flat trends. The realized values, however, show a significant decline in harvested area over time. Higher input prices and weakened demand for U.S. wheat (Ali and Vocke, 2009), along with changes in policies (Bonnen and Schweikhardt, 1998), likely contributed to the decline in wheat area. The frequency of the projected lines above the actual outcomes indicates a positive bias.

Forecast Errors of the Projections

While the visualizations (figures 2-4) display the general accuracy of the baseline, measures of forecasting error provide objective metrics to test for bias and compare the relative accuracy of different approaches. The statistical term "forecast error" refers to the difference between the realized value and the predicted value (whether projected or forecasted). In practice, the forecast error seldom equals zero. The goal in model selection is to select the model that provides the lowest forecast errors, and hence, the highest accuracy. The baseline projections are an example of a multi-step forecast because the projection includes estimates for multiple years in the future. To clarify the number of years between the projected value and when the projection occurred, we define the following notation, $F_{t,h}$, for a forecasted value for year t over forecast horizon length h. For example, the 2005 baseline forecast for 2006 outcomes is indexed $F_{2006,1}$, since it is a projection of a 2006 outcome one period away. As another example, $F_{2018,6}$ is the forecasted outcome in 2018, forecasted 6 periods before 2018, or the 2012 baseline projection of the year 2018. This notation tracks the forecasts across a given horizon length (e.g., all forecasts for outcomes 10 years out: $F_{t,10}$). The notation allows for evaluating the accuracy of a rolling forecast, which moves forward one period with each forecast. The rolling forecast approach follows the approach used by USDA and the Federal Government for estimating future budgets. Four commonly used measurements to assess the difference between actual and estimated values are: root mean squared error (RMSE), mean squared error (MSE), mean absolute percentage error (MAPE), and mean percentage error (MPE). The four statistics are calculated over each horizon length as follows:

Equation 1

$$RMSE_h = \sqrt{\frac{\sum_{\mathrm{t}=1997}^{2018}\left(F_{t,h} - \mathrm{A_t}\right)^2}{N}}$$

Equation 2

$$MSE_h = \frac{\sum_{\mathrm{t}=1997}^{2018}\left(F_{t,h} - \mathrm{A_t}\right)^2}{N}$$

Equation 3

$$MAPE_h = \frac{\sum_{t=1997}^{2018}\left|F_{t,h} - A_t\right|/A_t}{N}$$

Equation 4

$$MPE_h = \frac{\sum_{t=1997}^{2018}\left(F_{t,h} - A_t\right)/A_t}{N}$$

Equations 1-4 denote the realized or actual value as A_t. The difference between the forecast and the actual value is the forecast error ($e_{t,h}$). N represents the number of times each h horizon length is forecasted. Forecasts are evaluated over the 1997 to 2018 range, with 2017 being the latest baseline projection evaluated.

Both the RMSE and MSE are frequently used statistics for measuring the accuracy of projections. Squaring transforms forecast errors to positive values and places a higher penalty on larger deviations from actual values. The RMSE is often preferred to the MSE because after squaring the errors, values are normalized by a square root, which helps make the errors comparable to projected values. The MAPE similarly measures the accuracy of projections, though it can be interpreted more easily because values are a percentage of the realized value. The absolute value operator in the MAPE formula normalizes all forecast errors to positive values, similar to the RMSE, though it places equal weights on all errors. The RMSE, MSE, and MAPE measure how precisely the projections forecasted the actual values, though none offers information on the bias.

The MPE measures bias in forecasts because the MPE does not transform the sign of the forecast errors. An unbiased forecast would possess projection errors centered on zero. For that reason, the MAPE and the MPE together illuminate the proportion of the projection error that is one-sided. Forecasts with an absolute value of MPE close to the MAPE value indicate the forecast errors occur mostly in one direction.

While these statistics are informative in understanding the projections and their accuracy, statistical tests have been developed to more rigorously evaluate their performance. A straightforward method is to test whether the projection contains bias (i.e., whether they consistently overestimate or underestimate future outcomes).[5] We tested for forecast bias using the zero-mean test with the following regression:

Equation 5

$$\mathrm{A_t} - F_{t,h} = a + v_t .$$

The right-hand side of the equation was estimated using linear regression with just a constant (a) and a residual (v_t). The coefficient of the constant term, a, provides an average value for the forecast error. The null hypothesis that $a = 0$, statistically tests for forecast bias. If the null hypothesis was rejected, implying $a \neq 0$, the forecast was considered biased.

When analyzing forecasts of different horizon lengths, the number of projections that are available to evaluate differs. Shorter forecasts have a larger N, which determines the statistical power of the test defined in equation 5. For example, for the baseline projections from 1997 to 2017, there were 21 observations of forecasted and actual values for a 1-year horizon length.

This same timeframe had only 12 forecasts and actual values for a forecast horizon length of 10 years.

The baseline projections for corn area harvested exhibited the lowest RMSE and MAPE throughout the 1997 to 2017 timeframe, indicating corn is the most accurate baseline forecast of the three projected commodities (table 1).[6] An MPE calculated using the corn baseline projections was also close to zero for the first 2 years. The negative sign of this MPE value, thereafter, showed that the baseline was more likely to underestimate than overestimate area from 1997 to 2017. Projections of

[5] While it is common to use the Mincer-Zarnowitz (1969) test when looking at biases of forecasts, Mankiw and Shapiro (1986) showed the test may be inaccurate with small samples.

[6] This is a casual comparison, as we cannot statistically compare the size of each value across commodities.

wheat area typically had the largest MAPE and MPE for all projection horizons, reaching as high as 17.7 percent for forecast horizons of 10 years. The large positive MPE values for wheat projections indicated that the baseline generally overestimated harvested acreage from 1997 to 2017. Soybeans had lower values of MAPE and MPE across shorter horizon lengths, but error sizes grew closer to those observed in wheat. For wheat and soybeans, the absolute value of MPE was close to the MAPE, suggesting most projection errors arose from a persistent overestimation of growing area for wheat and a persistent underestimation of soybean area.

Table 1

Baseline projection statistics for harvested corn, soybean, and wheat area, 1997–2017

	Corn				Soybean				Wheat		
Forecast horizon	RMSE	MAPE	MPE		RMSE	MAPE	MPE		RMSE	MAPE	MPE
1 (N=21)	3,037	3.1%	0.8%		3,052	3.3%	-0.8%		3,028	5.0%	4.2%
2 (N=20)	3,898	3.4%	0.0%		5,245	5.4%	-3.5%		3,745	5.8%	4.0%
3 (N=19)	4,002	3.8%	-1.0%		5,996	6.3%	-4.9%		4,918	8.1%	4.9%
4 (N=18)	4,557	4.1%	-1.1%		6,444	7.2%	-5.8%		6,037	8.9%	6.2%
5 (N=17)	4,764	4.4%	-1.8%		6,972	7.9%	-6.3%		6,197	9.5%	7.5%
6 (N=16)	5,270	5.1%	-2.5%		7,801	8.7%	-6.7%		5,870	9.7%	8.0%
7 (N=15)	6,008	6.1%	-3.3%		8,388	9.1%	-7.3%		6,606	10.8%	10.4%
8 (N=14)	6,995	7.0%	-4.3%		9,121	9.4%	-8.0%		7,471	12.8%	12.8%
9 (N=13)	7,632	7.9%	-5.3%		10,645	10.6%	-9.1%		8,503	15.6%	15.6%
10 (N=12)	7,868	8.1%	-7.0%		11,484	11.1%	-9.7%		8,960	17.7%	17.7%

Notes: N reflects the number of forecast projections used in calculating the values in the table. The projected value for 10 years out has 12 observations, while the projection across a 1-year horizon has 21.
RMSE = root mean squared error; MAPE = mean absolute percentage error; MPE = mean percentage error.

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

The statistical tests for bias from equation 5 confirmed the prior results that the forecast errors statistically differ from zero (table 2). The null hypothesis of mean zero was rejected at a 10-percent confidence interval for horizons higher than 5 years. Baseline projections for wheat area were statistically biased positively across all horizon lengths. Soybeans area projections were biased negatively for all horizons longer than 1 year.

Table 2
Baseline projection tests for bias for harvested corn, soybean, and wheat area, 1997–2017

Forecast horizon (Years)	Corn	Soybean	Wheat
1 (N=21)	544	-717	2,012***
2 (N=20)	-152	-2,851***	1,840**
3 (N=19)	-940	-3,962***	2,166**
4 (N=18)	-1,082	-4,665***	2,750**
5 (N=17)	-1,664	-5,135***	3,379***
6 (N=16)	-2,239*	-5,540***	3,736***
7 (N=15)	-2,851**	-5,996***	4,837***
8 (N=14)	-3,718**	-6,643***	5,921***
9 (N=13)	-4,569***	-7,645***	7,153***
10 (N=12)	-5,910***	-8,187***	8,123***

Notes: Values for each commodity are estimated coefficients (a) from equation 1. Asterisks denote statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.10$. N reflects the number of forecast projections for each forecast horizon estimate. The projected value for 10 years out has 12 observations, while the 1-year projection has 21.

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017 and USDA, National Agricultural Statistics Service, QuickStats,

The results of the mean zero tests above showed that baseline projections were biased during the period of analysis. Several other statistical tests can also be used to analyze the efficiency and bias of the baseline projections. For example, Nordhaus (1987) and Patton and Timmermann (2012) developed various additional statistical tests that use prior forecasts as explanatory variables in future forecasts to assess forecast efficiency and optimality.[7]

As opposed to using only statistical tests of baseline projections, we instead investigated whether the projections could be improved using econometric procedures. The current baseline figures rely on the use of partial equilibrium models of supply and demand to project variables of interest (Hjort et al., 2018). Because a partial equilibrium approach uses parameter estimates from historical observations for each equation, the baseline projection process must either assume markets in the future possess parameter values that are equal to the average of the prior years or use parameters that have not been estimated. Relying on analysts to devise artificial parameters can introduce bias into the projection process.

[7] One disadvantage to the approach is that it requires using multiple years in the forecast evaluation, which due to smaller samples reduces the statistical power.

Time-Series Analysis

While forecast error statistics detail the projection's accuracy and bias, they are uninformative on the source or cause. For example, bias and accuracy errors could be triggered by large, unpredictable changes to the system, such as policy or faults in the estimation process. The statistics alone do not indicate the cause of the error. Statistical and econometric approaches, specifically time-series, provide avenues for comparing the projections with naïve forecasts. The statistical methods rely strictly on historical data available at the time of each historical projection, called real-time data, to forecast expected outcomes. Comparing the forecast errors of the statistical procedures with past baseline projections allows for an assessment of the baseline projections' forecast errors. The baseline's forecast errors are compared against econometric time-series models, as well as a naïve no-change forecast.

A defining feature of time-series data is the relationships among and dependence on past, current, and future observations. For reasons such as long-term contractual relationships, rigidity in decision-making, specialized capital investment, policies, and risk aversion, prior data hold predictive power for projecting future outcomes. These factors mean that even after controlling for relevant covariates, deviations in prior years typically correlate with deviations in future years. Time-series econometric models replicate these tendencies and incorporate them into a forecast.

We evaluated three different time-series forecasts and their results; the univariate autoregressive (AR) forecast, a univariate moving-average (MA) forecast, and multivariate vector-autoregressive (VAR) forecast. The AR forecast uses a linear combination of lagged values of the dependent variable. By contrast, the MA forecast uses a linear combination of lagged values of the errors.[8] Similar to the AR model, the VAR model uses linear combinations of lagged values, but instead of just one variable, it includes the lags of multiple variables. We used models of order one for all three forecast models. See the appendix for more details on these models and why the length was selected.

Including forecasts from a VAR model with the univariate AR and MA models examines whether the inclusion of lagged area of other crops improves the accuracy of forecasts. Farmers evaluate many crops as potential alternatives, causing a natural relationship between the crops at the farm level. As such, the VAR model could improve on the projections if the farm-level relationship between crops translates to consistent behavior at the aggregate level. In addition to the aggregation, for the additional information used in a VAR model to improve on the univariate forecasts, the correlations of the crops must not change across the historical and forecasted periods.

All the time-series forecasts used real-time data (Croushore, 2011). Real-time refers to the fact that often historical data is revised after having been first reported. Consider a forecast that used last year's harvested data, only to find out a revision to the prior year's data transpired after the forecast occurred. A researcher looking back 20 years likely would see only the final revised value, as opposed to the actual information that was available at the time of each forecast.

While typically, it is advantageous to use the biggest sample available for building a mathematical model, this is not always true of a time-series model. It is often the case that forecast accuracy can be improved by conditioning the sample on a smaller subset of the data. Changing relationships of

[8] Note that an AR process can be transformed into an infinite-order MA model, so either model could be appropriate for forecasting a particular series.

outcomes over time explain why this might occur. For example, 50 years ago, farmers nationwide were more limited than they are now in the crops they could plant and grow. Including the correlation between crops 50 years ago may ignore the technological advances that influence planting decisions today. Statistical tests, called structural break tests, have been developed to determine whether the data series used for forecasting should be truncated to prevent introducing historical correlations that are no longer related to current market outcomes. Structural breaks can occur for a variety of reasons, including shifts in supply (e.g., policy changes, technological change, new persistent weather patterns) or demand (e.g., changes in consumer preferences, biofuel policy).

We examined for breaks in the time-series by finding the maximum of the sample Wald statistics. Rejecting the null hypothesis from a Wald statistic implies a structural break occurred in the sample. The results of structural break tests for the three crops from the dataset for the earliest baseline projections in the study, 1997, found the largest Wald statistic for the year 1983. Interestingly, the largest Wald statistic corresponds to 1983 for the most recent sample forecast, 2017, as well. Data before 1983 were not used in this projection exercise. While further steps could be used to test the already restricted sample for additional structural breaks, and thus change the sample used for each forecast, we chose to keep the forecasts simpler by holding the start date of the data constant across all models.

We examined the results of a no-change forecast, where the most recent observation was the same forecasted value 10 years forward. No-change forecasts are used widely in the forecast evaluation literature to assess the accuracy and benchmark the results against the most straightforward approach.[9] The no-change forecast is beneficial for assessing whether the values follow a random walk, which would imply any forecast model would be at best just as accurate as a no-change forecast. A defining feature of both the no-change forecast and the time-series forecasts is that they are naïve forecasts, only using historical data for forecasting forward. These forecasts make no assumptions on future policy and do not incorporate any expectations of future outcomes, such as those that might be gleaned from commodities futures markets.

[9] No-change models are simple and widely used in both the literature and the field. For example, the Federal Reserve Bank of Philadelphia used the no-change forecast as its first benchmark for forecasting variables such as gross domestic product or unemployment rate (Stark, 2010).

Findings

Table 3 presents the forecast error statistics for a no-change forecast for corn, soybean, and wheat harvested area. The no-change forecast uses the most recent harvested area value at the time of the baseline projection and extends the value forward with no change (in other words, for the 2007 projection, the no-change forecast assumes all forecasted values are equal to the 2006 area). The results showed that the no-change forecast for corn area performed worse in each of the three statistics compared with the baseline projections. The MPE was close to the MAPE in absolute levels. Because the MPE was not calculated in absolutes, the positive and negative values offset each other. When the MPE was not zero, it highlighted the direction in which the series was biased, in this case, highlighting an upward trend in corn area. For soybeans, the no-change forecast error statistics for RMSE and MAPE were smaller than the baseline errors for all horizons except for the 1-year forecast. The negative MPE emphasized that the projections were downward biased. The no-change forecast for wheat area exhibited larger forecast error statistics over the shorter horizons (years 1-3) but outperformed baseline according to the RMSE statistic for forecast horizon lengths 4, 5, and 7-10. The MAPE for the wheat 10-year forecast horizon was 26 percent smaller under the no-change forecast than the baseline projections. Interestingly, the MPE for wheat area 1 year out was 38 percent smaller for the no-change forecast relative to baseline. Also, baseline projections for wheat area had a lower RMSE and MAPE but a higher MPE than no-change for years 1-3. It was not uncommon for these measures to disagree, given that each had a different measurement objective.

Table 3

No-change forecast for harvested corn, soybean, and wheat area, 1997–2017

Forecast horizon (Years)	Corn				Soybean				Wheat		
	RMSE	MAPE	MPE		RMSE	MAPE	MPE		RMSE	MAPE	MPE
1 (N=21)	4,644	4.1%	-0.1%		4,295	3.7%	-0.8%		3,882	7.1%	2.6%
2 (N=20)	5,000	5.5%	-0.7%		4,654	4.2%	-2.0%		5,070	8.9%	4.7%
3 (N=19)	5,372	5.3%	-1.8%		5,390	4.9%	-2.8%		5,465	9.3%	6.0%
4 (N=18)	6,028	5.7%	-2.5%		5,834	5.7%	-3.7%		5,981	9.3%	7.0%
5 (N =17)	6,707	6.6%	-3.7%		6,765	6.4%	-4.6%		6,048	9.5%	7.6%
6 (N=16)	7,805	7.4%	-5.3%		7,369	7.3%	-5.1%		6,186	11.5%	8.1%
7 (N=15)	8,453	8.0%	-7.0%		8,226	7.6%	-5.6%		6,583	12.1%	9.2%
8 (N =14)	9,575	9.9%	-8.3%		8,164	8.0%	-6.1%		6,916	11.9%	10.6%
9 (N=13)	9,766	10.3%	-9.6%		9,017	8.8%	-7.0%		7,166	12.4%	11.7%
10 (N=12)	11,072	11.9%	-11.2%		10,206	9.3%	-7.7%		6,885	13.1%	12.1%

Notes: N reflects the number of forecast projections used in calculating the values in the table. The projected value for 10 years out has 12 observations, while the 1-year projection has 21.
RMSE = root mean squared error; MAPE = mean absolute percentage error; MPE = mean percentage error.

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

While a no-change forecast was optimal if the projections followed a random walk, time-series models often outperformed the no-change forecast if the data followed a trend. Tables 4, 5, and 6 present the forecast error statistics for harvested corn, soybeans, and wheat area, respectively, using the autoregressive (AR [1]), moving average (MA [1]), and multivariate vector autoregressive (VAR [1])

time-series models.[10] These models were estimated with a starting point of 1983 and used the year of the baseline projection as the last data point. For example, the forecast for 2007 was estimated using data from 1983 to 2007. Our last time-series forecast was the 2017 baseline forecast.
Table 4 presents the forecast evaluation statistics of the time-series models for corn area. The AR and MA forecasts outperformed the VAR across all forecast horizon lengths using the RMSE and MAPE criteria. The MA forecast performed slightly better than the AR forecasts by the same criteria. The similarity of results for AR and MA implies the linear trend in each model explains much of the forecast. By contrast, the VAR forecasts were less accurate. However, at forecast horizon lengths 4-10, VAR estimation exhibited less bias than the other 2 time-series models. Across all three forecasts, the absolute value of the MPE was similar to the MAPE—an indication of bias. The biases in the time-series forecast and the baseline projections are indicative of the substantial policy shifts in the mid-2000s, which caused forecasts to underestimate future harvested area.

Table 4
Time-series forecast results for harvested area of corn, 1997–2017

Forecast horizon (Years)	AR(1)				MA(1)				VAR(1)		
	RMSE	MAPE	MPE		RMSE	MAPE	MPE		RMSE	MAPE	MPE
1 (N=21)	3,777	3.7%	-0.3%		3,772	3.7%	-0.4%		4,583	4.4%	-1.3%
2 (N=20)	4,311	4.4%	-0.5%		4,287	4.4%	-0.5%		5,345	5.5%	-0.8%
3 (N=19)	4,662	4.7%	-0.8%		4,628	4.7%	-0.7%		5,680	6.1%	-0.8%
4 (N=18)	5,148	5.5%	-1.0%		5,129	5.4%	-0.9%		6,256	6.7%	-0.8%
5 (N =17)	5,403	5.7%	-1.6%		5,397	5.7%	-1.5%		6,606	7.2%	-1.4%
6 (N=16)	5,780	5.9%	-2.4%		5,694	5.9%	-2.3%		6,891	7.4%	-2.0%
7 (N=15)	5,952	6.1%	-3.2%		5,914	6.1%	-3.1%		7,045	7.4%	-2.7%
8 (N =14)	6,322	6.3%	-3.7%		6,314	6.3%	-3.6%		7,325	7.5%	-3.0%
9 (N=13)	6,303	6.5%	-4.3%		6,239	6.5%	-4.2%		7,303	6.9%	-3.3%
10 (N=12)	6,125	6.1%	-5.4%		5,994	6.0%	-5.3%		6,466	6.1%	-4.1%

Notes: N reflects the number of forecast projections that go into calculating the values in the table. The projected value for 10 years out has 12 observations, while the 1-year projection has 21.
AR = autoregressive; MA = moving average; VAR = multivariate vector autoregressive.
RMSE = root mean squared error; MAPE = mean absolute percentage error; MPE = mean percentage error.
Forecast errors from AR(1), MA(1), and VAR(1) models with a linear time trend.

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

[10] We present lower-order models only as they outperform higher-order models.

Table 5 presents the forecast evaluation statistics of the time-series forecasts for soybean area. Similar to the corn forecast, the AR and MA forecasts outperformed the VAR across most evaluation statistics and horizon lengths. The MA appeared more accurate than the AR for horizon lengths 6-10. Interestingly, the forecast statistics of the MA model were similar in accuracy across all forecast horizon lengths. One might expect the forecast errors to increase more substantially for more distant outcomes, but the results here indicate only marginal to negligible differences. This result highlighted the predictability of soybean acreage using a moving average of past outcomes and a linear trend. The substantially less accurate VAR forecasts highlighted the change in U.S. soybean planting over time. Each VAR forecast estimated the correlations of the three crops historically and projected outward. The results of the VAR forecasts indicate the historical correlations of soybeans with corn and wheat have changed meaningfully in more recent years.[11]

Table 5
Time-series results for harvested area of soybean, 1997–2017

Forecast horizon (Years)	AR(1)				MA(1)				VAR(1)		
	RMSE	MAPE	MPE		RMSE	MAPE	MPE		RMSE	MAPE	MPE
1 (N=21)	3,730	3.3%	-0.6%		3,939	3.9%	-1.2%		3,867	3.5%	0.2%
2 (N=20)	3,937	3.9%	-1.3%		4,541	4.8%	-2.5%		4,191	4.5%	0.8%
3 (N=19)	4,287	4.3%	-1.6%		4,587	4.6%	-2.4%		4,854	5.3%	1.7%
4 (N=18)	4,566	4.6%	-1.8%		4,736	4.7%	-2.3%		5,786	5.8%	2.7%
5 (N =17)	4,680	4.5%	-2.0%		4,784	4.5%	-2.3%		6,435	6.8%	3.6%
6 (N=16)	5,401	5.3%	-2.0%		5,102	4.7%	-2.2%		7,999	8.3%	4.7%
7 (N=15)	5,584	5.6%	-2.1%		5,256	5.3%	-2.1%		8,953	9.6%	5.9%
8 (N =14)	5,711	5.7%	-2.2%		5,000	5.4%	-2.0%		10,248	11.1%	7.3%
9 (N=13)	5,978	5.7%	-2.7%		5,326	5.3%	-2.3%		10,913	11.7%	8.2%
10 (N=12)	5,868	5.5%	-2.9%		4,736	4.4%	-2.4%		11,477	13.0%	9.6%

Notes: N= the number of forecast projections that go into calculating the values in the table. The projected value for 10 years out has 12 observations, while the 1-year projection has 21.
AR = autoregressive; MA = moving average; VAR = multivariate vector autoregressive.
RMSE = root mean squared error; MAPE = mean average percentage error; MPE = mean percentage error.
Forecast errors from AR(1), MA(1), and VAR(1) models with a linear time trend.

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

[11] This result echoes the analysis in Camp (2019).

Table 6 presents the forecast evaluation statistics of the time-series forecasts for harvested wheat area. The AR forecasts appeared the most accurate, according to the RMSE statistic. However, the VAR forecasts were more accurate according to the MAPE values. Higher MAPE but lower RMSE implies the AR is more likely to miss the correct value but less likely to miss by a more substantial proportion. These two differing statistics raise the question of whether it is better to miss closely but more often or miss widely but be accurate more often. Given the small samples and large forecast errors, the comparative gap in accuracy across the two models is not statistically different.[12] The VAR forecasts also resulted in the smallest MPE, meaning they were less biased than the other two forecasts.

Table 6
Time-series results for harvested area of wheat, 1997–2017

Forecast horizon (Years)	AR(1)				MA(1)				VAR(1)		
	RMSE	MAPE	MPE		RMSE	MAPE	MPE		RMSE	MAPE	MPE
1 (N=21)	3,752	6.6%	3.8%		3,680	6.4%	3.8%		3,578	6.2%	2.4%
2 (N=20)	4,917	7.9%	5.7%		5,001	8.1%	6.2%		4,903	7.4%	4.6%
3 (N=19)	5,402	8.8%	6.5%		5,446	8.9%	6.7%		5,355	8.2%	5.1%
4 (N=18)	5,823	9.0%	7.1%		5,886	9.0%	7.2%		6,135	8.4%	5.3%
5 (N =17)	5,790	9.0%	7.1%		5,863	9.1%	7.2%		6,189	8.6%	4.9%
6 (N=16)	5,110	8.7%	6.6%		5,144	8.7%	6.7%		5,322	8.2%	3.9%
7 (N=15)	5,525	8.7%	7.4%		5,590	8.8%	7.5%		6,028	7.9%	4.2%
8 (N =14)	5,683	9.2%	7.9%		5,746	9.3%	8.0%		6,396	9.1%	4.3%
9 (N=13)	6,143	9.7%	8.6%		6,194	9.9%	8.7%		7,132	9.7%	4.7%
10 (N=12)	5,561	9.6%	8.8%		5,601	9.6%	8.9%		6,118	8.9%	4.2%

Notes: N= reflects the number of forecast projections that go into calculating the values in the table. The projected value for 10 years out has 12 observations, while the 1-year projection has 21.
AR = autoregressive; MA = moving average; VAR = multivariate vector autoregressive.
RMSE = root mean squared error; MAPE = mean absolute percentage error; MPE = mean percentage error.
Forecast errors from AR (1), MA (1), and VAR (1) models with a linear time trend.

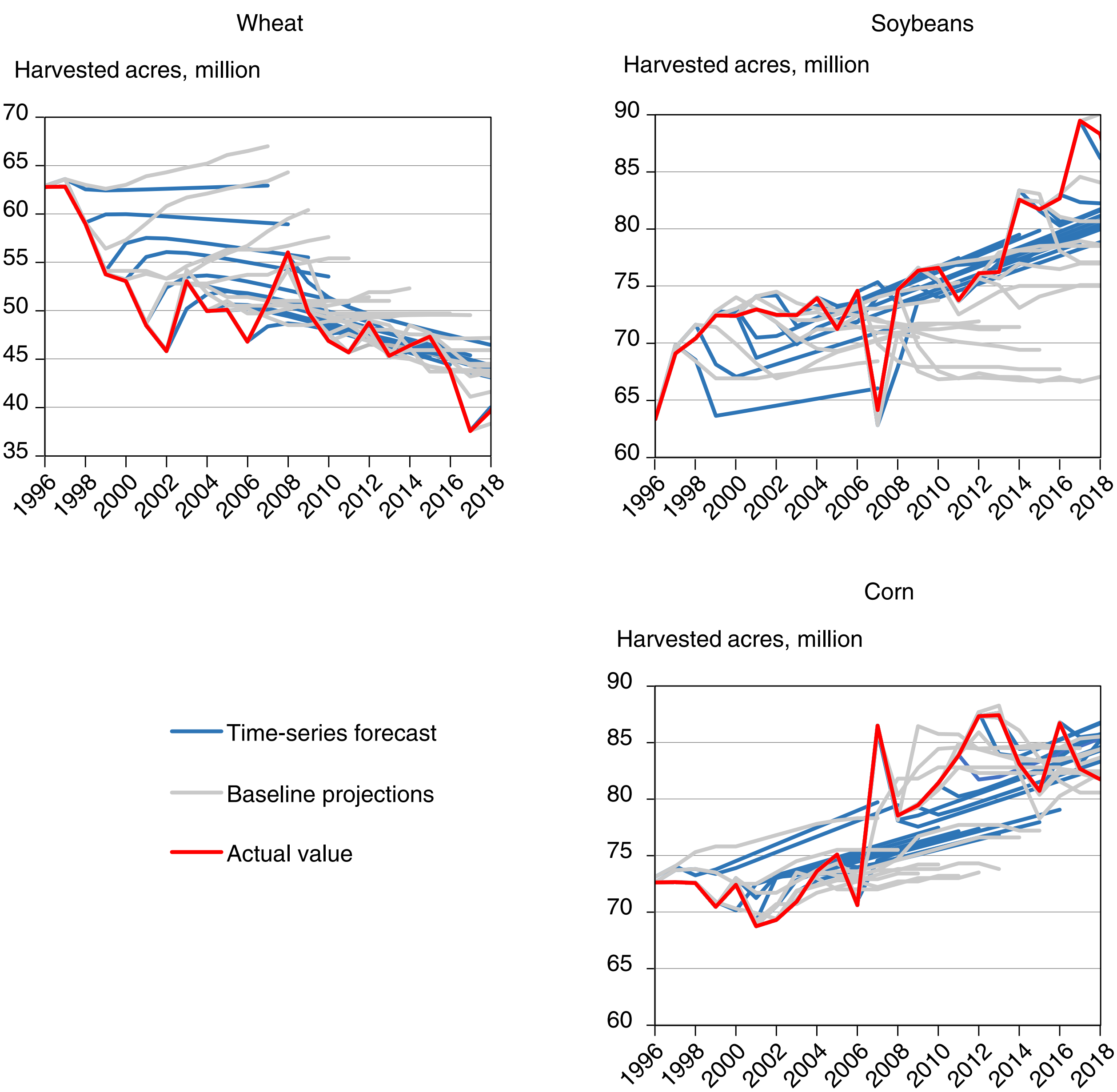
Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

[12] We tested for significance with the Diebold-Mariano (1995) test.

Forecast Comparisons

While tables 3-6 are informative for forecast evaluation, we expanded our analysis by more rigorously comparing all model results across all horizons and their associated statistics. Figure 5 visually compares one econometric forecast with the baseline projection for corn, soybeans, and wheat. The figure provides a visualization of the forecasts used to generate the results in the earlier tables. Most importantly, overestimation and underestimation associated with econometric forecasts were far less apparent than the bias associated with the baseline projections.

Figure 5
Harvested corn, soybean, and wheat area, time-series and baseline projections versus actual



Notes: AR = autoregressive; MA = moving average. Time-series projections from econometric results: Corn–AR(1); Soybeans–AR(1), and Wheat–MA(1).

Source: Calculated by USDA, Economic Research Service using data from USDA, Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

Table 7 presents the results of statistical comparisons of the baseline projections to the time-series forecasts. We statistically compared the projections to the time-series forecasts using the conventional Diebold-Mariano test's loss functions (Diebold and Mariano, 1995). A loss function is a statistical concept for comparing two forecasts according to specific properties of forecast error statistics. The RMSE, MSE, MPE, and MAPE are all forms of a loss function. We used a loss function such as MSE or MAPE to evaluate forecasts because it can reveal more than just comparing the average errors of two forecasts.[13] We chose to use the Diebold-Mariano test to compare forecasts according to the MSE and MAPE loss functions.[14] We also compared forecasts by testing the bias across forecasts using two-sample z-scores.[15] Asterisks compare the baseline projections to whichever time-series forecast possessed the smallest error value.

Table 7

Baseline versus time-series forecast errors 1997–2017, smallest absolute values for mean squared error (MSE), mean absolute percentage error (MAPE), and bias

Forecast horizon (Years)	Corn			Soybeans			Wheat		
	MSE	MAPE	Bias	MSE	MAPE	Bias	MSE	MAPE	Bias
1 (N=21)	Base.**	Base.	AR	Base.	Base.	AR	Base.	Base.**	AR
2 (N=20)	Base.	Base.***	Base.	AR**	AR***	AR**	Base.*	Base.***	Base.
3 (N=19)	Base.**	Base.***	MA	AR**	AR***	AR***	Base.	Base.	Base.
4 (N=18)	Base.	Base.***	MA	AR**	AR***	AR***	AR	VAR	Base.
5 (N=17)	Base.	Base.	MA	AR***	AR***	AR***	AR	VAR	AR
6 (N=16)	Base.	Base.	MA	MA*	MA***	AR***	AR	VAR*	AR
7 (N=15)	MA*	Base.	MA	MA*	MA**	MA***	AR**	VAR***	AR
8 (N=14)	MA	AR	MA	MA	MA*	MA***	AR***	VAR***	AR*
9 (N=13)	MA**	MA*	MA	MA	MA**	MA***	AR***	VAR***	AR**
10 (N=12)	MA**	MA***	MA	MA*	MA***	MA***	AR***	VAR***	AR***

Notes: AR = autoregressive; MA = moving average; VAR = multivariate vector autoregressive. Model name indicates the smallest absolute value for MSE, MAPE, or bias statistic. Asterisks denote statistical significance: *** p<0.01, ** p<0.05, * p<0.10 MSE and MAPE are tested using Diebold-Mariano tests, while Z tests for differences in coefficients are used to test bias. For the projection of corn area harvested for a horizon 1 year ahead, baseline projections have the smallest MSE of all forecasts. They are statistically smaller than the best time-series forecast at a 5-percent confidence interval. Baseline has the smallest MAPE of all corn area projections with a 1-year horizon, though it is not statistically smaller than a time-series model.

Source: Calculated by USDA, Economic Research Service using data from Agricultural Baseline Projections, 1997–2017, and USDA, National Agricultural Statistics Service, QuickStats.

---

[13] For example, two forecasts could have the same mean error (or MPE), but one of the errors from one forecast could be twice as large as the other. If both forecasts miss equally above and below the actual value, they would have the same average error. Accordingly, the Diebold-Mariano test was developed to statistically analyze other forecast error statistics, like MSE, MAPE, and MAE, to determine if one forecast was more accurate than the other or if the differences in values between the two forecasts are just due to unexplained variations

[14] Throughout the paper, we have used RMSE but used the MSE for the Diebold-Mariano tests. The switch in the convention is due to the statistical properties of comparing MSE. The ordering or ranking of forecast errors is equivalent using either the RMSE or the MSE.

[15] The z-score is estimated from the β from equation 5 for two different forecasts. $Z = \frac{\beta_1 - \beta_2}{\sqrt{SE_1^2 + SE_2^2}}$, where 1 and 2 indicate two different coefficients from two different bias tests. *SE* is the standard error from the β of each equation.

The results from Table 7 can be interpreted as follows. If the table shows a time-series model for a particular row and column, this means the statistical test compared that forecast with the baseline projections, and the time-series model had a lower loss function value than the baseline projections. In some instances, multiple time-series models (AR and MA) statistically outperformed the baseline projections. The forecast comparisons show the baseline projections for corn area produce smaller errors across shorter forecast horizons but larger errors for horizons longer than 6 years. According to the MSE, the moving-average forecast produced the smallest errors for forecasting corn area for horizons greater than 6 years. A time-series model forecasted the least amount of bias over 9 of the 10 horizons, though statistically, the errors are not differentiable from the baseline projections. The non-differentiability of the bias for forecasting corn area is likely because of the policy shift toward ethanol that expanded the demand for corn.

We found the time-series forecasts were statistically more accurate than the baseline projections across most horizon lengths. No soybean area forecast was statistically more accurate over a 1-year forecast horizon length, though the baseline did have smaller MSE and MAPE. Over every other forecast horizon length, either the AR or MA outperformed the baseline projections in predicting soybeans area, with the latter performing slightly better over more distant horizons.[16] The forecasting bias for soybeans was statistically smaller with the time-series models, implying even though the time-series models are likewise biased toward underpredicting area, the bias is reduced by enough to make it statistically different from the baseline projections.

Evaluating the forecasts for wheat area, we found results similar to those for soybean area. The baseline projection was more accurate statistically for shorter forecast horizons but produced significantly larger errors statistically for horizons 7 years or longer. As with soybeans, the bias was reduced with the AR forecast for more distant forecast horizons.

---

[16] The AR and MA forecasts for soybeans are not statistically differentiable over any horizon length. Both, however, statistically differ and are smaller than the baseline projections at most horizons.

Discussion

A comparison of the historical projection errors of baseline and time-series models is informative for improving future projections. While the baseline projection's forecast errors are the smallest and least biased for corn relative to soybeans and wheat, the forecast errors indicate a potential for improvement of the baseline process. In the short run, the current baseline projections typically outperform the naïve time-series approaches in terms of accuracy—that is, how close the projections are to realized harvested area for corn, soybean, and wheat. However, the forecast accuracy advantage of the baseline approach diminishes over time for the three commodities. Among forecast horizon lengths greater than or equal to 2 years for soybeans, 4 years for wheat, and 7 years for corn, the time-series models outperform the baseline approach in terms of accuracy. For projections with a time horizon over 7 years, the time-series forecast errors are meaningfully smaller. The RMSE from time-series projections at 10-year horizon lengths are 1.9, 6.7, and 3.4 million acres smaller than the baseline projections for corn, soybean, and wheat area, respectively. The improvement in forecast accuracy with time-series models has precedent, with similar research showing univariate projections often better-forecast more complicated models (Faust and Wright, 2009).

While this research highlights the role and effectiveness of using time-series models to forecast area for corn, soybeans, and wheat—especially in the case of more distant events—the approaches were strictly naïve, using only data on historical outcomes. Efforts toward incorporating additional information, such as that from commodity markets, which are often more focused on shorter horizons, would likely help to improve forecasts over nearer horizons. Additionally, the forecasts herein used only one methodology at a time, either AR, MA, or VAR, and efforts could be made to find combinations of these forecasts or other more complicated forecast methods to improve predictive accuracy.

The baseline projections are valuable to a variety of stakeholders, presenting a conditional, long-run scenario about what would be expected to happen to market outcomes under a continuation of current farm legislation and other specific assumptions. Multiple agencies across the USDA provide input into these projections, helping to ensure integrity. The results of our study highlight the value of using time-series approaches in assessing and improving the projection process for corn, soybeans, and wheat. While we conclude that no projection performs best across all timeframes and commodities, each method discussed provides insight into improving the accuracy of the baseline projections more broadly. The USDA has begun to incorporate the use of these models into the baseline process, particularly for initializing projections of more distant outcomes, alongside the current baseline model (Hjort et al., 2018).

References

Ali, M., and G. Vocke. 2009. *Consequences of Higher Input Costs and Wheat Prices for U.S. Wheat Producers*, WS-09c-01, U.S. Department of Agriculture, Economic Research Service.

Ash, M., and E. Dohlman. May 2007. "International Trade, Biofuel Initiatives Reshaping the Soybean Sector," *Amber Waves*, U.S. Department of Agriculture, Economic Research Service.

Ash, M., and E. Dohlman. July 2007. *Soybean Acreage Drops to 12-Year Low.* OCS-07f. OilCrops Outlook, U.S. Department of Agriculture, Economic Research Service.

Batchelor, R. 2001. "How Useful Are the Forecasts of Intergovernmental Agencies? The IMF and OECD Versus the Consensus," *Applied Economics* 33(2):225-35.

Batchelor, R. 2007. "Bias in Macroeconomic Forecasts," *International Journal of Forecasting* 23(2):189-203.

Bonnen, J.T., and D.B. Schweikhardt. 1998. "The Future of U.S. Agricultural Policy: Reflections on the Disappearance of the 'Farm Problem'," *Review of Agricultural Economics* 20(1):2-36.

Camp, K.M. 2019. "The Relationship Between Crude Oil Prices and Export Prices of Major Agricultural Commodities," *Beyond the Numbers*, U.S. Department of Labor, Bureau of Labor Statistics.

Condon, N., H. Klemic, and A. Wolverton. 2015. "Impacts of Ethanol Policy on Corn Prices: A Review and Meta-analysis of Recent Evidence," *Food Policy* 51:63-73.

Croushore, D. 2011. "Frontiers of Real-time Data Analysis," *Journal of Economic Literature* 49(1):72-100.

Diebold, F.X., and R.S. Mariano. 1995 "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics* 13(3):253-63.

Egelkraut, M., P. Garcia, S.H. Irwin, and D.L. Good. 2003. "An Evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agencies," *Journal of Agricultural and Applied Economics* 35(1):79-95.

Faust, J., and J.H. Wright. 2009. "Comparing Greenbook and Reduced Farm Forecasts Using a Large Realtime Dataset," *Journal of Business & Economic Statistics* 27(4):468-79.

General Accounting Office. 1988. "USDA's Commodity Program: The Accuracy of Budget Forecasts," United States General Accounting Office. PEMD-88-8. Washington, D.C.

General Accounting Office. 1991. "USDA Commodity Forecasts: Inaccuracies Found May Lead to Underestimates of Budget Outlays," United States General Accounting Office. PEMD-91-24. Washington, D.C.

Good, D.L., and S.H. Irwin. 2006. "Understanding USDA Corn and Soybean Production Forecast: Methods, Performance and Market Impacts over 1970–2005." University of Illinois at Urbana-Champaign, Department of Agriculture and Consumer Economics.

Hjort, K., D. Boussios, R. Seeley, and J. Hansen. 2018. "The ERS Country-Commodity Linked System: Documenting Its International Country and Regional Agricultural Baseline Models." TB-1951, U.S. Department of Agriculture, Economic Research Service.

Irwin, S., and D. Good. 2015. "Long-term Corn, Soybeans, and Wheat Forecasts and the Farm Bill Program Choice." farmdoc daily (5):20, University of Illinois at Urbana-Champaign.

Isengildina, O., S.H. Irwin, and D.L. Good. 2004. "Evaluation of USDA Interval Forecasts of Corn and Soybean Prices," *American Journal of Agricultural Economics* 86(4):990-1004.

Lee, T., A. Tran, J. Hansen, and M. Ash. 2016. "Major Factors Affecting Global Soybean and Products Trade Projections," *Amber Waves*, U.S. Department of Agriculture, Economic Research Service.

Mankiw, N.G., and M.D. Shapiro. 1986. "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models." *Economics Letters* 20(2):139-45.

Mincer, J., and V. Zarnowitz. 1969. *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, The National Bureau of Economic Research: 3-46.

Motamed, M., L. McPhail, and R. Williams. 2016. "Corn Area Response to Local Ethanol Markets in the United States: A Grid Cell Level Analysis," *American Journal of Agricultural Economics* 98(3):726-43.

Nigatu, G., J. Hansen, N. Childs, and R. Seeley. 2017. "Sub-Saharan Africa is Projected to Be the Leader in Global Rice Imports," Amber Waves, U.S. Department of Agriculture, Economic Research Service.

Nordhaus, W.D. 1987. "Forecasting Efficiency: Concepts and Applications," *The Review of Economics and Statistics* 69(4):667-74.

Patton, A.J., and A. Timmermann. 2012. "Forecast Rationality Tests Based on Multi-horizon Bounds," *Journal of Business & Economic Statistics* 30(1)1-17.

Pole, A., M. West, and J. Harrison. 2018. *Applied Bayesian Forecasting and Time Series Analysis.* Chapman and Hall/CRC Press.

Roberts, M.J., and W. Schlenker. 2013. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the U.S. Ethanol Mandate," *The American Economic Review* 103(6):2265-95.

Stark, T., 2010. "Realistic Evaluation of Real-time Forecasts in the Survey of Professional Forecasters," Federal Reserve Bank of Philadelphia.

U.S. Environmental Protection Agency, Office of Air and Radiation, "Overview for Renewable Fuel Standard." Overviews and Factsheets, accessed April 4, 2015.

Zahniser, S., and N.F. López López, M. Motamed, Z. Silva Vargas, and T. Capehart. 2019. *The Growing Corn Economies of Mexico and the United States*, OCS-19F-02, U.S. Department of Agriculture, Economic Research Service.

Appendix: model selection of time-series models

We compare the U.S. Department of Agriculture's agricultural baseline projections for U.S. corn, soybean, and wheat area harvested with forecasts generated using three econometric models. We restricted the presentation of the forecasting models to just these three to highlight the potential for these models to forecast forward and improve existing approaches. This section discusses explicitly why these models were chosen and how that could have impacted the results presented in the paper.

Despite appearing relatively similar, time-series econometric models differ dramatically in how they model a corresponding time-series data set. Evaluation of autocorrelation (AC) and partial autocorrelation (PAC) values, as well as unit-root Dickey-Fuller stationarity tests, can be helpful for testing and examining which time-series model would best fit the unique behavior of each model and accurately model the true data generating process (DGP) of the observations over time. While the tests are informative to uncovering the correct DGP, the tests are often not definitive, and ambiguity often remains as to which model to use. This is true if the forecaster is examining one data series or multiple ones.

In the forecasting exercise used in this report, 21 different 10-year projections are required for three crops. This means there are 63 different real-time data sets. We tested all these data sets for stationarity, looked at many of their AC and PAC plots. As one might expect, results of the tests were not definitive across commodity, lag-length, use of a time-trend, or timeframe. Some data sets were able to reject the null hypothesis of a random walk, while other series came back with inconclusive test results.

The advantage with model selection for forecasting is that the goal is to find the forecast model that exhibits the most accuracy or least bias. Because the forecasts are out-of-sample, the forecasts themselves test whether the forecast models explain the data or not. Thus, while the pre-estimation tests standard in time-series modeling is an important starting place for finding the appropriate model, the forecast results are also informative for selecting the optimal time-series model. The one concern with strictly using the forecast errors as a check for model selection would be that perhaps the true data generating process differs, and the accuracy of the forecasts was due to luck. However, with enough out-of-sample forecasts, this concern can be somewhat diminished.

As a result of the combination of inconclusive tests across commodities and the out-of-sample forecast errors, we chose to present the results of the forecasts with three different time-series models, each of order one. We believe each model is informative to understanding not only which forecast performs best over time but also how information should be used when building each projection. In addition to the models shown in this paper, we forecasted numerous combinations of model types. We tested multiple lag lengths, first differencing, various ARIMA models, and more. However, we found the forecast errors were smallest on average with the presented models of order one. Further research could expand on this list of models as well, incorporating additional variables such as forward-looking futures market prices to improve the forecast as well.