Interval Forecast Comparison

Olga Isengildina-Massa and Julia L. Sharp[1]

Selected Paper prepared for presentation at the

Agricultural and Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting

Washington, DC, August 4-6, 2013

[1] Olga Isengildina is an Associate Clinical Professor at the University of Texas at Arlington and Julia Sharp is an Associate Professor in the Department of Mathematical Sciences, Clemson University. Research assistance of Maria Bagwell is gratefully acknowledged.

Interval Forecast Comparison

The need for probability and interval forecasting has been repeatedly expressed in the agricultural economics literature (e.g., Teigen and Bell; Timm; Bessler and Kling; Bessler). However, application and analysis of this type of forecasts is still relatively rare. Interval forecasts in agricultural economics can be found in price forecasting. For example, United States Department of Agriculture (USDA) provides interval price forecasts for major field crops in their World Agricultural Supply and Demand Estimates (WASDE) report. Interval price forecasts for livestock (hogs) are available from USDA as well as Iowa State University (ISU). The advantage of interval forecasts is that instead of a single value they provide a range of values that will likely contain a future outcome, thus communicating information about the uncertainty associated with the forecast. While interval forecasts normally specify the probability of the final value being contained within the interval (confidence level), agricultural interval forecasts usually omit this information. Furthermore, since interval forecasts in agricultural economics are so uncommon, the additional uncertainty information that they provide is typically lost in their evaluation and comparison since most studies analyze them as point forecasts using the midpoints of the published intervals (e.g. Kastens, Schroeder, and Plain, 1998; Egelkraut et al., 2003).

To the best of our knowledge, only three previous studies evaluated USDA price forecasts as intervals rather than reducing them to a point estimate. Sanders and Manfredo (2003) examined one-quarter-ahead WASDE interval forecasts of livestock prices from 1982 to 2002. Evaluation of hit rates, the proportion of time actual market prices fall in the forecasted ranges, revealed relatively low hit levels for livestock price forecasts, about 48% of the time for broilers, 41% of the time for cattle, and only 35% of the time for hogs. The authors did not conduct any formal tests of interval forecast accuracy other than showing that based on z-scores for testing equality in the proportion of hits, these forecasts were not significantly better than a proposed naïve alternative. Isengildina, Irwin, and Good (2004) showed that monthly WASDE interval forecasts of corn and soybean prices during the 1980/81 through 2001/02 marketing years also had relatively low hit rates ranging from 36 to 82% for corn and from 59 to 89% for soybeans depending on the forecast month. The authors applied unconditional and conditional tests of interval forecast accuracy developed by Christoffersen (1998) to test whether WASDE price forecast intervals were calibrated at two benchmark confidence levels. Isengildina and Sharp (2012) evaluated the implications of asymmetry on accuracy of USDA interval forecasts of corn, soybean and wheat prices. Although forecast intervals published by the USDA for corn, soybean and wheat prices are reportedly symmetric, they have shown that these intervals should not always be interpreted as symmetric. Their findings demonstrate that due to the uneven distribution of forecast misses around the interval, calibration of several corn, soybean and wheat price forecasts over 1980/81 through 2009/10 marketing years was rejected by basic coverage tests (suitable for symmetric intervals) but not rejected by the tests adjusted for asymmetry. In other words, these forecasts were asymmetric but accurate. However, in order to make the evaluation of these interval forecasts possible, the authors had to collect additional information and make some assumptions about the confidence level associated with these forecasts. In this study we relax these assumptions and extend the analysis of interval forecasts to their comparison rather than just accuracy evaluation. Thus the goal of this study is to develop a framework for interval forecast comparison and apply it to comparing interval forecasts of hog prices provided by USDA and ISU.

Methods

As mentioned before, previous studies of agricultural interval forecasts (e.g., Egelkraut et al. 2003; Sanders and Manfredo, 2003; and Colino et al. 2008) used the midpoint of USDA interval forecasts to compare them to other forecasts that were usually published as point estimates. These studies used the Modified Diebold and Mariano (MDM) test to analyze whether various accuracy measures, such as mean absolute percent error (MAPE), root mean squared percent error (RMSPE), mean absolute error (MAE), and root mean squared error (RMSE) are significantly different across alternative forecasts. Thus, the criterion that is used for point forecast comparison is the size of forecast error defined as the difference between the forecasted value and the final outcome. Recent studies of interval forecast comparison (Christoffersen, Hahn and Inoue, 2001; and Corradi and Swanson, 2011) develop tests for comparison of model-based forecasts which cannot be applied to "model-free" agricultural interval forecasts but provide guidance for criteria that should be used for interval forecast comparison. The following criteria are used in this study for interval forecast comparison:

1. ***Accuracy*** describes how often the intervals contain the final value;
2. ***Informativeness*** refers to how specific the intervals are based on their width;
3. ***Precision*** measures the size of error (the distance between the interval bound and the final value) for forecast "misses."

Traditional measures of interval forecast accuracy are hit rates and forecast coverage. Hit rates describe the proportion of times forecast intervals contain the final or "true" value (y_t) and may be defined as $E(I_t^k)$. Forecast coverage examines whether the proportion of times the forecast interval includes the true value corresponds to a stated confidence level, or, in other words, if the interval hit rate is equal to the coverage probability. The challenge with agricultural interval forecasts is that the confidence level is not stated. In this case the hit rates can be calculated and coverage can be evaluated for each forecast at several benchmark levels (in our study we evaluate coverage at 10 % increments, i.e., 10%, 20%, 30%, 40%, 50%). This approach allows us to avoid having to make an assumption about the intended confidence level and whether the confidence level is the same for two alternative forecasts. Forecast coverage is evaluated using the unconditional, independence and conditional coverage tests developed by Christoffersen (1998).

Unconditional coverage test examines whether the interval hit rate is equal to the coverage probability by testing the hypothesis H_0: $E(I_t^k)=\alpha$ against H_1: $E(I_t^k)\neq\alpha$. If H_0 is not rejected, forecasts are said to be calibrated. The likelihood function for the indicator variable I_t^k, which has a binomial distribution, is

(1) $$L(\alpha)=(1-\alpha)^{n_0}\alpha^{n_1}$$

under the null hypothesis and

(2) $$L(p)=(1-p)^{n_0}p^{n_1}$$

under the alternative hypothesis, where n_1 and n_0 are the number of times an interval was "hit" (1) or "missed" (0) in the indicator sequence I_t^k, and L is a likelihood function. Then, forecast coverage may be tested via the likelihood ratio test,

(3) $$LR_c=-2\ln\left(\frac{L(\alpha)}{L(\hat{p})}\right)\xrightarrow{asy}\chi^2(1)$$ (unconditional coverage test)

where $\hat{p} = n_1/(n_0 + n_1)$ is the maximum likelihood estimator of *p*. Because this test is termed unconditional coverage because it does not imply anything about the underlying information set.

Christoffersen argued, however, that in addition to coverage, interval forecasts should be dynamic, in the sense of being "narrow in tranquil times and wide in volatile times, so that the occurrences of observations outside the interval forecast would be spread out over the sample and not come in clusters" (p. 842). Christoffersen proposed testing independence of the indicator sequence I_t^k against an explicit first-order Markov alternative. First, define the transition probability of the first-order Markov chain for a given forecast date *k* as $\pi_{ij} = \Pr(I_t^k = j / I_{t-1}^k = i)$, where *j*=1,0 and *i*=1,0.[3] Then, the likelihood ratio test of independence is given by,

(4) $$LR_i = -2\ln\left[\frac{L(\hat{p})}{L(\hat{\pi}_{01}, \hat{\pi}_{11})}\right] \overset{asy}{\rightarrow} \chi^2(1)$$ (independence test)

where $L(\hat{\pi}_{01}, \hat{\pi}_{11}) = (1-\hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1-\hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}$, where n_{ij} is the number of observations with value *i* followed by *j*, $\hat{\pi}_{01} = n_{01}/(n_{00} + n_{01})$, and $\hat{\pi}_{11} = n_{11}/(n_{10} + n_{11})$.

The conditional coverage test combines an unconditional coverage test (equation 3) with a test of forecast independence (equation 4) to account for higher-order dynamics of time-series forecasts:

(5) $$LR_{cc} = LR_c + LR_i \overset{asy}{\rightarrow} \chi^2(2).$$ (conditional coverage test)

Thus, the conditional coverage test combines unconditional coverage and independence while retaining the individual hypotheses as subcomponents. According to Christoffersen, this test allows determination of whether "a given interval forecast deserves the label "good" (p.842).

Informativeness refers to the specificity of the forecast reflected in the width of the interval. Between two alternative forecasts calibrated at the same coverage level, a narrower interval would be preferred to a wider one. This characteristic can be examined by evaluating whether interval widths are significantly different from each other using MDM test. The MDM test is preferred over alternative test, such as a t-test because it provides a more accurate test for small sample sizes (Harvey et al. 1997). Also, the MDM test accounts for the auto-correlation in data sets that contain time-dependent entries. If the interval widths of the forecasts released at time *t* from two alternative sources are defined as w_{1t} and w_{2t} for $t = 1,...,n$, the difference between them, is $\mathrm{d_t} = |w_{1t}| - |w_{2t}|$, with the sample mean of $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_t$. It was shown by Diebold and Mariano (1995) that under the null hypothesis, the test statistic, S_1, follows an asymptotic standard normal distribution:

(6) $$S_1 = \frac{\bar{d}}{\sqrt{\frac{1}{n}\left(\hat{\gamma}_0 + 2\sum_{k=1}^{h-1}\hat{\gamma}_k\right)}} \xrightarrow{d} N(0,1)$$

where $\hat{\gamma}_0$ is the variance and $\hat{\gamma}_k$ is the auto-covariance function of d_t. The MDM test statistic uses an approximately unbiased estimator of the variance of $\bar{d}$. For a generalized h-step ahead forecast the test statistic is

(7) $$S_1^* = \left(\frac{n + 1 - 2h + n^{-1}h(h-1)}{n}\right)^{1/2} S_1,$$

which follows a student t distribution with $n-1$ degrees of freedom. This test will allow us to conclude whether or not the expected value of the difference between the interval widths is zero, where a difference of zero implies the two alternative interval forecasts are of similar informativeness.

Interval forecast precision focuses on forecast "misses" and evaluates the size of forecast errors by measuring the absolute distance between the final outcome and the closest interval bound. In this case, even if both intervals miss the final outcome, the interval that is closer to the final value is preferred to the one that is further away. This characteristic can be examined by evaluating whether these distances from the final value are significantly different from each other using MDM test. For precision evaluation the distance between the final value and the closest bound of the first forecast interval is denoted e_{1t} and the distance for the alternative interval forecast is e_{2t}. Let g(e) be the error function for the specified forecast error, where n is the sample size. Then the difference in the distances, d_t, is defined as

$\mathrm{d}_t = g(e_{1t}) - g(e_{2t}) = |e_{1t}| - |e_{2t}|, \quad t = 1,...,n$. Further description of the MDM test follows the one provided for testing informativeness in equations 6 and 7. Particular care should be applied for using this test to evaluate precision since a lot of zero values (illustrating "hits") are expected. The MDM test statistic is dependent on the variance of the forecast errors. Thus, the small values for the variance, specifically those values that approach zero, will increase the size of the test statistic. Therefore a bootstrapped estimate of the variance should also be considered. Ashley (1998) proposed a new post-sample inference procedure using bootstrapped parameters that avoids pre-test bias and allows for contemporaneously cross-correlation and serial dependence that is oftentimes found in post-sample time series data. All analyses are conducted using the statistical computing program R (2008).

Data

Interval forecasts for live hog prices from January 1990 through November 2011 are obtained from Livestock, Dairy and Poultry Outlook (USDA) and Iowa Farm outlook – Hog and Pig

Report Summary (ISU). Both agencies release forecasts quarterly. USDA forecasts are obtained from the reports released in February, May, August, and November of each year. The release of ISU forecasts follows the same periodicity, but the forecasts are published one month earlier that USDA, in January, April, June and October. For example, the USDA report in February of 1990 published the average price forecasts for the first, second, third and fourth quarter of 1990, then in the subsequent report in May, the forecasts for the second, third and fourth quarter of 1990 and the first quarter of 1991 are released and so on. Thus, for an objective price of a given quarter, four forecasts at different time points are obtained, one-quarter ahead, two-quarters ahead, three-quarters ahead and four-quarters ahead forecasts. As described in the methodology section, the time lag from the final value is denoted as h for h-step ahead forecasts. The final values of real average prices for each quarter are usually published in the report of the next quarter. After deleting occasional point forecasts, 84 observations for each agency were available for one-quarter-ahead comparison, 85 for two-quarter-ahead comparison, 83 for three-quarter-ahead comparison, and 69 for four-quarter-ahead comparison.

The data for the one, two, thee, and four step ahead forecasts from both sources and the final values are shown in figures 1 through 4, respectively. Figure 1 shows that 1-step ahead forecasts from both sources fit the final values very closely with similar interval width, but USDA appears to have a higher hit rate. In most cases, the final value tends to lie above the intervals, rather than below. Considering the 2, 3, and 4-steps ahead forecasts shown in figures 2 through 4, we can conclude that as h increases, the accuracy of interval forecasts decreases. While the width of the USDA's intervals tends to get larger as h increases, the width of the ISU's intervals appears to remain the same.

Empirical Results

Comparison of accuracy of USDA and ISU live hog forecasts focuses on the ability of these intervals to contain the final value. Table 1 demonstrates that the hit rates for 1-step ahead forecasts were higher for USDA at 36% vs 23%. Furthermore, unconditional coverage tests failed to reject calibration of 1-step ahead ISU forecasts at 20% -30%, while USDA's 1-step ahead forecasts were calibrated at 30% to 45%. Two-step ahead forecasts were calibrated at 15% to 30% for ISU and 20% to 30% for USDA indicating that these forecasts were very similar. Three-step ahead forecasts were calibrated at 20% to 30% for ISU and 15% to 25% for USDA, suggesting that USDA forecasts were slightly worse. Finally, four-step ahead forecasts were calibrated at 15%-25% for ISU but only at 10% to 20% for USDA, suggesting again lower accuracy for USDA forecasts. Thus, USDA's short-term forecasts appear better than ISU's, while long-term forecasts are worse.

Test of independence shown in table 1 examines the ability of forecast intervals to adjust dynamically to the volatility of the forecasted series, with wider intervals published in more uncertain times. Independence was rejected for two-step ahead USDA forecasts and three-step ahead ISU forecasts, suggesting that interval width in these cases did not dynamically adjust to uncertainty in forecasted prices. Since the conditional coverage test is a combination of the unconditional and independence tests, none of the ISU's three-step ahead and USDA's two-step ahead forecasts are calibrated, while the results for other forecasts are similar to the ones for unconditional coverage tests. Therefore, based on conditional coverage test results (not shown here but available from authors upon request) we can only argue that one-step ahead USDA forecasts were better than ISU's and four-step ahead ISU forecasts were better than USDA's, but we cannot say anything about 2 and 3 step ahead forecasts.

Analysis of asymmetry illustrated in table 1 reveals that the probability of misses above was much greater than the probability of misses below for USDA's 1, 3, and 4-step ahead forecasts and ISU's 4-step ahead forecasts and to a less degree in USDA's 2-step ahead forecasts and ISU's 3-step ahead forecasts. This suggests that both agencies tend to be conservative in these forecasts, or that under-estimation of price is more common than over-estimation. Thus, these forecasts should not be interpreted as symmetric. At the same time, ISU's 1 and 2-step ahead forecasts appear fairly symmetric.

Informativeness test results reported in table 2 indicate that ISU's intervals were significantly wider than USDA's for 1-step ahead forecasts and significantly narrower for 3 and 4 step ahead forecasts. ISU commonly published 3$/cwt intervals for their hog price forecasts, but they were as narrow as 2$/cwt and as wide as 15$/cwt (January 1999). The size of these intervals appears to vary with the uncertainty of the underlying prices (with the exception of 3-step ahead forecasts as discussed above), but not with h (time lag to the final realization of price). On the other hand, USDA 1 through 3-step ahead forecasts appear to narrow as they get closer to the quarter they are predicting. Our findings of superiority for 4-step ahead ISU forecasts and 1-step ahead USDA forecasts are consistent across both accuracy and informativeness criteria.

Interval forecast precision is measured in table 3. While table 1 reports that the probability of misses above is higher than the probability of misses below in many cases, table 3 shows that the size of misses is fairly equal for USDA and may be a little larger for misses below for ISU. The test of precision indicates that the overall size of error, measured as the average distance from the closest interval boundary to the final value, is smaller for USDA's 1-step ahead forecasts relative to ISU's 1-step ahead forecasts. The differences in the size of the error are not statistically significant in all other cases. After repeated re-sampling with replacement from d_t using bootstrapping, we compare the calculated sample variance of d_t to the bootstrapped estimate of the variance of d_t. The results showed that the bootstrapped estimate of the variance was smaller than the computed sample variance of d_t for 2, 3, and 4-step ahead predictions (results not shown here but available from the authors upon request). Therefore, we can conclude that the MDM test statistic in equation (7) uses a conservative estimate for the variance of d_t, which means that we may fail to reject the null hypothesis more often than if the bootstrapped estimate was employed. This finding contributes evidence of superiority of USDA's 1-step ahead forecasts in terms of precision in addition to accuracy and informativeness.

Summary and Conclusions

This study sought to develop a framework for agricultural interval forecast comparison and apply it to comparing forecasts of live hog prices provided by USDA and ISU. Challenges in evaluation and comparison of agricultural interval forecasts stem from the fact that these forecasts are model-free and the confidence levels associated with intervals are not revealed. We propose to use three criteria for interval forecast comparison: accuracy, informativeness and precision. Accuracy describes how often the intervals contain the final value and may be evaluated using hit rates and Christoffersen's tests of forecast coverage. Informativeness refers to how specific the intervals are based on their width. We demonstrate how modified Diebold Mariano (MDM) test can be used to evaluate the difference in interval widths. Precision measures the size of error (the distance between the interval bound and the final value) for forecast "misses." The difference in the size of the error between alternative forecasts is also examined using MDM test.

This framework was applied to comparison of interval forecasts of live hog prices from USDA and ISU over 1990 through 2011. Our results show that 1-step ahead forecasts from USDA are more accurate, more informative and more precise than 1-step ahead ISU forecasts. On the other hand, 4-step ahead ISU forecasts are more accurate and more informative than 4-step ahead forecasts from USDA. At the same time 2 and 3-step ahead forecasts are not significantly different across the two agencies in terms of these criteria.

References

Ashley, R. 1998. "A new Technique for post sample model selection and validation." *Journal of Economic Dynamics and Control*. 22: 647-665.

Bessler, D.A. “Subjective Probability.” in P.J. Barry (ed.) *Risk Management in Agriculture*. Iowa State University Press: Ames, Iowa, 1984; pp. 43-52.

Bessler, D. A. and J.L. Kling. “The Forecast and Policy Analysis.” *American Journal of Agricultural Economics*, 71(1989):503-506.

Christoffersen, P.F. 1998.“Evaluating interval forecasts.” *International Economic Review*, 39: 841–862.

Christoffersen, P.F., Hahn, and A. Inoue. 2001. “Testing and Comparing Value-at-Risk Measures.” *Journal of Empirical Finance* 8: 325-342.

Colino, E. V., S. H. Irwin, P. Garcia and Etienne. 2012. “Composite and Outlook Forecast Accuracy.” *Journal of Agricultural and Resource Economics*, 37(2):228-246.

Corradi, V. and N. R. Swanson. 2011. “A survey of recent advances in forecast accuracy comparison testing, with an extension to stochastic dominance.” Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions, San Diego, CA.

Diebold, F. X. and R. S. Mariano. 1995. “Comparing predictive accuracy.” *Journal of Business and Economic Statistics*, 13(3): 253–263.

Egelkraut, T. 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.” *American Journal of Agricultural and Applied Economics*, 35: 79–95.

Harvey, D., S. Leybourne, and P. Newbold. 1997. “Testing the equality of prediction mean squared errors.” *International Journal of Forecasting*, 13: 281–291.

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.

Isengildina, O. and J. Sharp. 2012. "Evaluation of USDA Interval Forecasts Revisited: Asymmetry and Accuracy of Corn, Soybean, and Wheat Prices" *Agribusiness: An International Journal*, 28 (3): 310–323.

Kastens, T.L., T.C.Schroeder and R. Plain. "Evaluation of Extension and USDA Price and Production Forecasts." *Journal of Agricultural and Resource Economics*, 23(1998):244-261.

Sanders, D. R. and M. R. Manfredo. 2003. "USDA livestock price forecasts: A comprehensive evaluation." *Journal of Agricultural and Resource Economics*, 28(2): 316–334.

Teigen, L.D. and T.M. Bell. "Confidence Intervals for Corn Price and Utilization Forecasts." *Agricultural Economic Research*, 30(1978):23-29.

Table 1. Accuracy Comparison of ISU and USDA Live Hog Price forecasts, 1990-2011.

Forecast Source		Unconditional Coverage Test for Various *p* levels									Independence Test	Hit Rate	N	Asymmetry	
		10%	15%	20%	25%	30%	35%	40%	45%	50%				% Miss Below	% Miss Above
ISU	1-Step Ahead	12.73 ***	4.05 **	0.55	0.14	1.96	5.57 **	10.80 ***	17.57 ***	25.94 ***	0.03	23.26	86	36.04	40.70
	2-Step Ahead	8.98 ***	2.16	0.05	0.79	3.61	8.14 ***	14.22 ***	21.81 ***	30.98 ***	0.57	20.93	86	37.21	41.86
	3-Step Ahead	15.51 ***	5.68 **	1.24	0.00	1.04	3.90 **	8.38 ***	14.39 ***	21.98 ***	4.29 **	25.00	84	33.33	41.67
	4-Step Ahead	7.34 ***	1.57	0.00	0.99	3.85 **	8.26 ***	14.07 ***	21.27 ***	29.90 ***	0.24	20.25	79	30.38	49.37
USDA	1-Step Ahead	41.92 ***	23.07 ***	11.90 ***	5.16 **	1.45	0.04	0.57	2.84	6.79 ***	0.08	36.04	86	23.26	40.70
	2-Step Ahead	14.80 ***	5.20 **	1.00	0.02	1.33	4.49 **	9.28 ***	15.65 ***	23.61 ***	7.25 **	24.42	86	31.39	44.19
	3-Step Ahead	7.55 ***	1.54	0.00	1.19	4.37 **	9.21 ***	15.56 ***	23.39 ***	32.77 ***	1.04	20.00	85	29.41	50.59
	4-Step Ahead	0.59	0.26	2.48	6.34 **	11.47 ***	17.74 ***	25.10 ***	33.60 ***	43.33 ***	0.00	12.86	70	34.29	52.85

Notes: Christoffersen's unconditional coverage test results are reported for *p* confidence levels. The chi-squared critical value is 3.84. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1%, respectively.

Table 2. Informativeness Comparison of ISU and USDA Live Hog Price forecasts, 1990-2011.

	ISU Intervals					USDA Intervals					N joint	MDM test
	N	mean	std. dev.	min	max	N	mean	std. dev.	min	max		
			$/cwt					$/cwt				
1-Step Ahead	86	3.186	1.739	2	15	86	2.361	0.853	0	4	84	3.927 ***
2-Step Ahead	86	2.988	0.360	2	4	87	3.103	1.578	2	6	85	-0.25
3-Step Ahead	85	3.118	0.420	2	5	87	4.138	1.173	0	6	83	-5.216 ***
4-Step Ahead	81	3.185	0.635	2	7	73	3.767	1.048	1	6	69	-4.123 ***

Notes: MDM test is the Modified Diebold Mariano test of the difference in the mean interval width. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1%, respectively.

Table 3. Precision Comparison of ISU and USDA Live Hog Price forecasts, 1990-2011.

	Avg. Miss Below	Avg. Miss Above	Average Absolute Error	Avg. Miss Below	Avg. Miss Above	Average Absolute Error	N	MDM Test	MDM Test with Bootstrapping
			$/cwt						
1-Step Ahead	5.18	3.36	8.54	1.89	1.85	3.74	84	4.353***	4.459***
2-Step Ahead	6.55	4.92	11.47	4.99	5.09	10.08	85	1.1	1.454
3-Step Ahead	7.52	5.59	13.11	5.89	5.37	11.26	83	0.242	0.312
4-Step Ahead	8.72	6.29	15.01	6.41	6.55	12.96	69	0.691	0.942

Notes: MDM test is the Modified Diebold Mariano test of the difference in the average absolute error. Single, double, and triple asterisks (*, **, ***) denote statistical significance at 10%, 5%, and 1%, respectively.

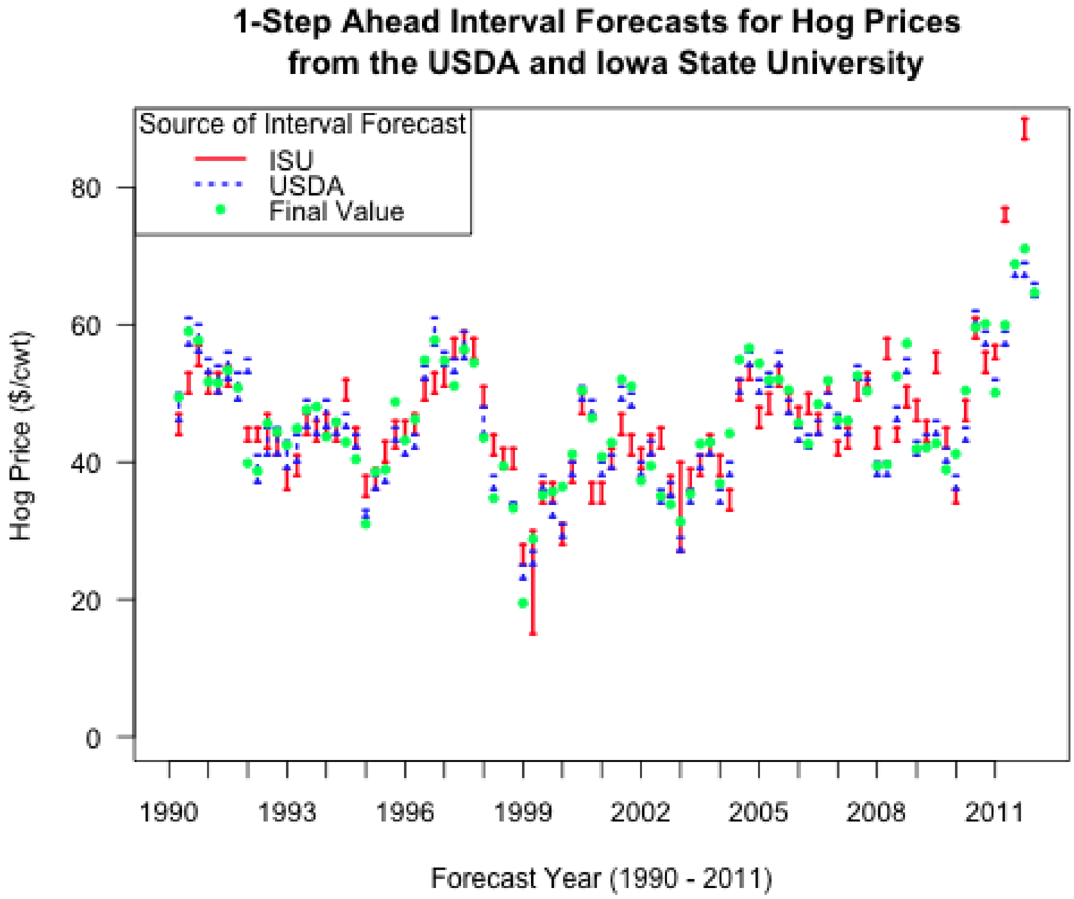

Figure 1. Final values and 1-step ahead interval forecasts for the USDA and ISU hog prices, 1990 - 2011.

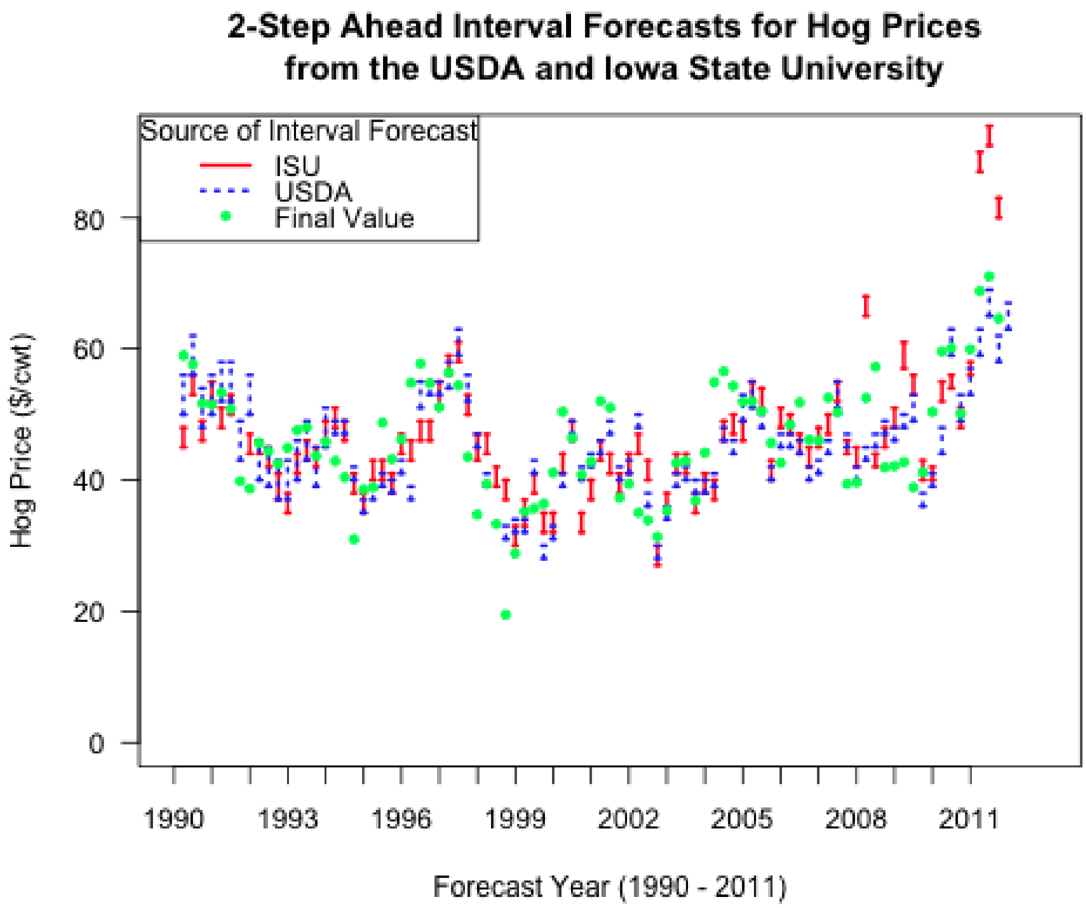

Figure 2. Final values and 2-step ahead interval forecasts for the USDA and ISU hog prices, 1990 - 2011.

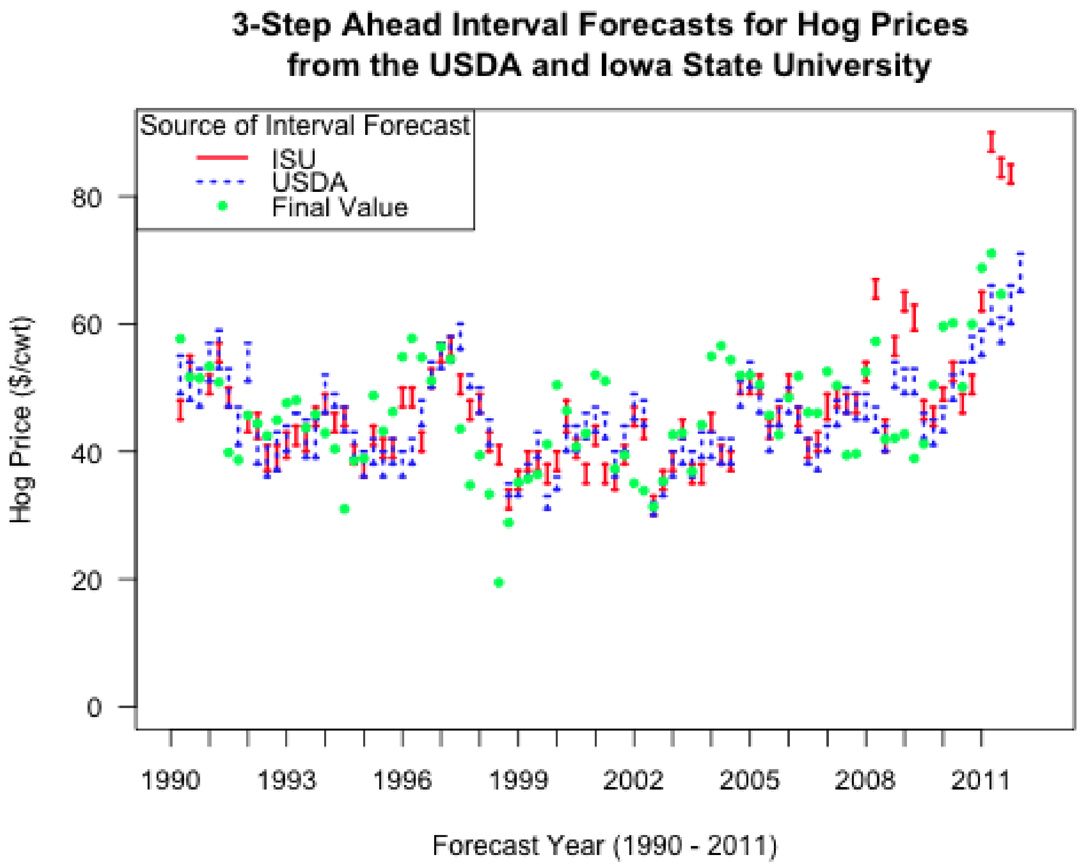

Figure 3. Final values and 3-step ahead interval forecasts for the USDA and ISU hog prices, 1990 - 2011.

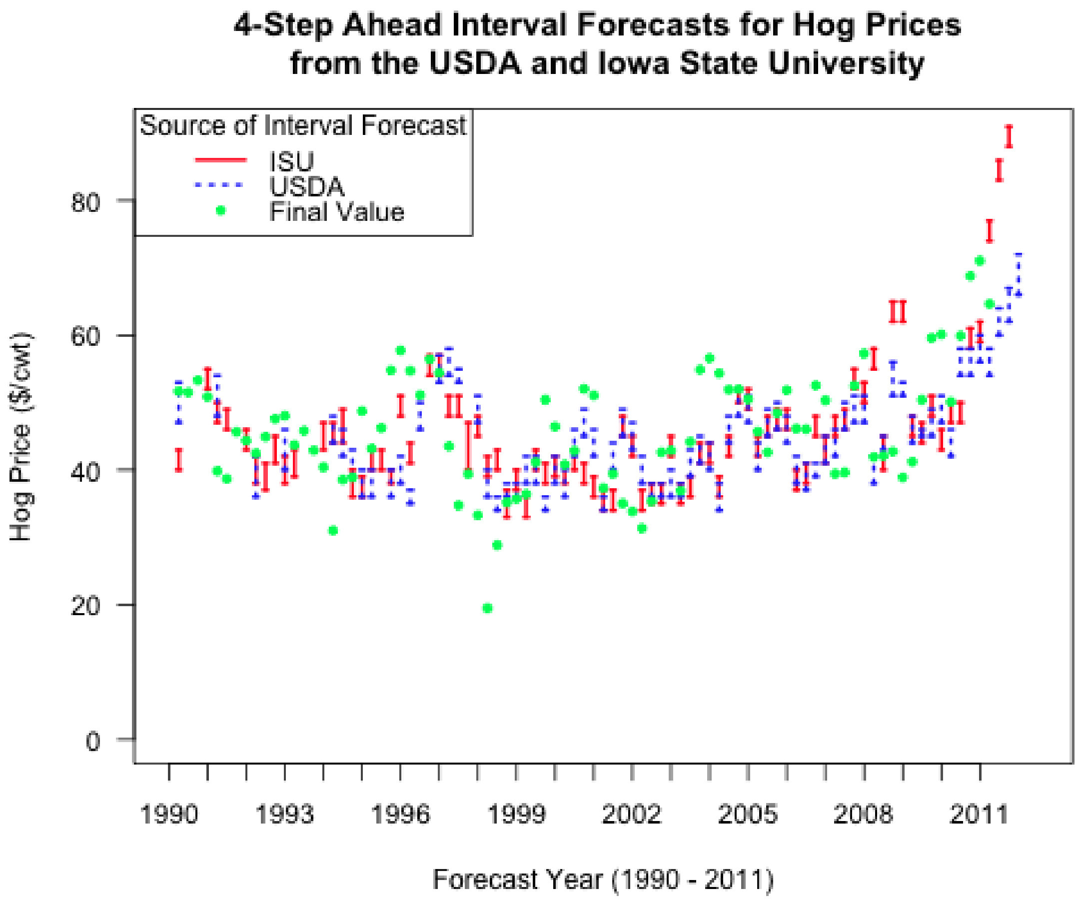

Figure 4. Final values and 4-step ahead interval forecasts for the USDA and ISU hog prices, 1990 - 2011.