1991

A Note on Evaluation of Turning Point Accuracy

Rhonda K. Skaggs

Donald L. Snyder
Utah State University

November 1991 ERI Study Paper #91-01

A NOTE ON EVALUATION OF TURNING POINT ACCURACY

By

Rhonda K. Skaggs

Donald L. Snyder

A NOTE ON THE EVALUATION OF TURNING POINT ACCURACY

By

Rhonda K. Skaggs and Donald L. Snyder*

ABSTRACT

This paper develops a method for use in qualitative evaluation of forecasts. The method is based upon the traditional contingency table method, but is an improvement over both the 2 x 2 and the 4 x 4 matrices previously used in assessments of forecast turning point accuracy. The 2 x 2 matrix did not account for directional accuracy in turning points; the 4 x 4 matrix overcame that weakness, but it fails to account for no change points. An expanded 9 x 9 contingency matrix allows for complete turning point accuracy evaluation; therefore, it is better suited for forecasting applications using higher-order time-aggregated data.

KEY WORDS forecast evaluation, forecast accuracy, turning points

September 1990

---

*Assistant Professor, New Mexico State University Department of Agricultural Economics and Agricultural Business; and Professor and Department Head, Utah State University Department of Economics.

Journal article #1552 of the New Mexico Agricultural Experiment Station.

A NOTE ON THE EVALUATION OF TURNING POINT ACCURACY

This paper develops a method for use in qualitative evaluation of forecasts. The method is based upon the traditional contingency table method, but is an improvement over both the 2 x 2 and the 4 x 4 matrices previously used in assessments of forecast turning point accuracy. The 2 x 2 matrix did not account for directional accuracy in turning points; the 4 x 4 matrix overcame that weakness, but it fails to account for no change points. An expanded 9 x 9 contingency matrix allows for complete turning point accuracy evaluation; therefore, it is better suited for forecasting applications using higher-order time-aggregated data.

The technique of turning point evaluation is a popular method for assessing a forecasting model's directional accuracy. This qualitative measure of predictive ability can be compared and contrasted with quantitative measures of accuracy (such as root mean squared error and mean absolute error) for a comprehensive examination of forecasting model performance. Turning point accuracy is traditionally evaluated using an inventory of a forecasting model's hits and misses in accurately predicting the direction of turning points. The typical turning point inventory used for model evaluation is, however, incomplete. The purpose of this note is to illustrate the deficiencies of the turning point evaluation procedures as they are commonly applied, and offer an alternative procedure for use in assessing turning point accuracy.

Background

Theil suggested the use of turning point accuracy as a measure of forecasting ability was supported by the presence of significant positive serial correlation in most economic time series. Because these series tend to exhibit rather stable patterns of expansions and contractions, it is relatively easy to predict continuations of such trends (Theil, p. 28). A turning point is characterized by a change in the direction of movement of the variable being tracked, and exists if $P_t > P_{t-1} < P_{t-2}$ or $P_t < P_{t-1} > P_{t-2}$. Theil asserted a model was truly successful if it could predict the end or beginning of one-sided expansions or contractions. He used the 2 x 2 contingency matrix shown in table 1 for conducting the turning point inventory.

Based on table 1, the four possibilities with respect to the prediction of turning points are (Theil, pp. 28-30): 1) a turning point is correctly predicted - the model predicted a turning point, and there is an actual turning point (f_{11}); 2) a turning point is incorrectly predicted - the model predicted a turning point and there is no actual turning point (f_{21}); 3) a turning point is incorrectly not predicted - the model does not predict a turning point, and there is an actual turning point (f_{12}); 4) a turning point is correctly not predicted - the model does not predict a turning point and no turning point occurs (f_{22}).

Cases 2 and 3 are failures of the model, while cases 1 and 4 represent successful prediction. With this simple matrix, a quantitative measure of turning point accuracy is derived by dividing the number of misses or failures by the number of forecast periods examined. The result is a percentage of the forecast periods during which there were turning point errors. This turning point method has been frequently used in applied forecasting research (Bessler and Brandt 1979, 1981; Brandt and Bessler 1981, 1984; Bourke; Gellatly; Harris and Leuthold; Hudson and Capps; Kulshreshtha and Rosaasen; MacGregor and Kulshreshtha).

The 2 x 2 contingency table method has more recently been rejected in favor of a 4 x 4 matrix (Naik and Leuthold). These authors have proposed expansion of the matrix to account for the shape of the turning points. The analysis developed by Theil, and still prevalent in the literature of applied forecast evaluation, has a limitation that can bias the assessment of a model's turning point accuracy. According to the criteria of the 2 x 2 contingency table, a model that forecasts every turning point accurately can actually result in a 100% failure when prediction of turning point direction is considered. Using the 2 x 2 contingency table, a model that predicts a peak turning point (/\) when there is actually a trough turning point (\/) would have accurately predicted the directional change. The 4 x 4 contingency table suggested by Naik and Leuthold evaluates the model in terms of its ability to forecast types of turning points (peaks, troughs and no turns) and compares those forecasts against the actual peaks, troughs and no turns. The 4 x 4 contingency table is shown in table 2.

A peak turning point (PTP) is defined as $P_t < P_{t-1} > P_{t-2}$, a trough turning point (TTP) is $P_t > P_{t-1} < P_{t-2}$, an upward no turning point (UNTP) is $P_t > P_{t-1} > P_{t-2}$, and a downward no turning point (DNTP) is $P_t < P_{t-1} < P_{t-2}$. Accurate forecasts are found on the principal diagonal of the 4 x 4 matrix. A forecast classified as 'worst' is one that moves in the opposite direction to that of the actual values (as in the case of predicting a peak when the actual data indicated a trough for that period).

The 4 x 4 contingency table is an improvement over the simple 2 x 2 table, but remains incomplete. Higher-order, time-aggregated data (e.g., monthly, weekly and daily data) often exhibit no change from one period to the next. The 4 x 4 contingency table is simply incapable of classifying no change points in the forecast and actual values, thus providing an incomplete and inaccurate measure of forecast accuracy.

A Suggested Solution

To classify the no change points that frequently occur in forecasting applications using higher-order, time-aggregated data, five more cases must be defined in addition to the four listed above. These cases are: 1) flat downward turning point (FDT), defined as $P_t < P_{t-1} = P_{t-2}$; 2) a flat upward turning point (FUT), $P_t > P_{t-1} = P_{t-2}$; 3) an upward flat turning point (UFT), $P_t = P_{t-1} > P_{t-2}$; 4) downward flat turning point (DFT), $P_t = P_{t-1} < P_{t-2}$; and 5) a straight no change situation, $P_t = P_{t-1} = P_{t-2}$. The nine cases are summarized diagrammatically in table 3.

A 9 x 9 contingency table is used to inventory the 81 potential cases, as presented in table 4. Accurate forecasts are located on the principal diagonal of the matrix. The following ratios (continuing with the notation adopted by Naik and Leuthold) will quickly summarize the qualitative performance of a forecasting model:

1) Ratio of accurate forecasts (RAF)

$$RAF = (f_{11} + f_{22} + f_{33} + f_{44} + f_{55} + f_{66} + f_{77} + f_{88} + f_{99}) / \Sigma_i \Sigma_j f_{ij} ;$$

2) Ratio of worst forecasts (RWF)

$$RWF = (f_{12} + f_{21} + f_{34} + f_{43} + f_{56} + f_{65}) / \Sigma_i \Sigma_j f_{ij} ;$$

3) Ratio of accurate to worst forecasts (RAWF)

$$RAWF = (f_{11} + f_{22} + f_{33} + f_{44} + f_{55} + f_{66} + f_{77} + f_{88} + f_{99}) / (f_{12} + f_{21} + f_{34} + f_{43} + f_{56} + f_{65});$$

4) Ratio of inaccurate forecasts (RIF)

$$RIF = (f_{13} + f_{14} + f_{15} + f_{16} + f_{17} + f_{18} + f_{19} + f_{23} + f_{24} + f_{25} + f_{26} + f_{27} + f_{28} + f_{29} + f_{31} + f_{32} + f_{35} + f_{36} + f_{37} + f_{38} + f_{39} + f_{41} + f_{42} + f_{45} + f_{46} + f_{47} + f_{48} + f_{49} + f_{51} + f_{52} + f_{53} + f_{54} + f_{57} + f_{58} + f_{59} + f_{61} + f_{62} + f_{63} + f_{64} + f_{67} + f_{68} + f_{69} + f_{71} + f_{72} + f_{73} + f_{74} + f_{75} + f_{76} + f_{78} + f_{79} + f_{81} + f_{82} + f_{83} + f_{84} + f_{85} + f_{86} + f_{87} + f_{89} + f_{91} + f_{92} + f_{93} + f_{94} + f_{95} + f_{96} + f_{97} + f_{98}) / \Sigma_i \Sigma_j f_{ij};$$

5) Ratio of inaccurate and worst forecasts (RIWF)

$$RIWF = RWF + RIF.$$

Application of the Methodology

The 9 x 9 contingency table and the related ratios are applied here for the qualitative evaluation of the simulated actual and forecast data patterns shown in figure 1. Numerous no-change situations are apparent throughout both the forecast and actual series. To further illustrate the no-change ranges of points, figure 2 is an enlargement of observations 3 through 11 from figure 1.

Of the 50 simulated forecast and actual observations, 48 points were processed (the first two are dropped from the analysis), with the complete matrix and summary ratios presented in table 5. Based on this table, we can conclude the forecast accurately predicted the true direction of movement (relative to the actual pattern) 12.5% of the time. In 8.3% of the 48 forecast periods, the forecast predicted a direction of movement exactly opposite to that which occurred. These are the cases of trough versus peak (TTP vs. PTP), straight downward versus straight upward movements (DNTP vs. UNTP), and flat upward turning point versus flat downward turning point (FUT vs. FDT). The forecasts incorrectly predicted the direction of movement 79% of the time, although these were not the worst cases. In total, the model inaccurately forecast the pattern of movement 87.5% of the time. If the 4 x 4 matrix had been applied in this case, the turning point evaluation would have been unable to classify 38, or 79% of the forecast periods.

The 9 x 9 contingency table and related ratios were recently applied in the qualitative evaluation of several models used to generate monthly forecasts of alfalfa hay and feeder steer prices. It was essential to use the 81-cell matrix for assessing the directional accuracy of the monthly forecasts. Actual data with this level of aggregation commonly exhibit no change over two or three periods. The nine cases presented in table 3 were programmed with a series of conditional statements classifying the relationships between the actual and forecast series over the period of the forecasting competition. After reading the forecast and actual data series, the program produced the completed contingency matrix and summary ratios.

Results of the expanded turning point evaluation were subsequently compared with the following quantitative measures of forecast accuracy: mean error, root mean squared error, root mean squared percentage error, and Theil's Inequality Coefficient. Tradeoffs between point and turning point accuracy were apparent.

Forecasting procedures which accurately track and predict turning points are likely to be preferred by some forecast users. These users may favor an early warning of directional changes over a high degree of point accuracy (e.g., low forecast root mean squared error). They could be more risk averse than other forecast users, or may be maximizing profit by buying and selling speculatively, based on predicted market upturns and downturns. In these cases, emphasis should be placed on a careful assessment of directional forecasting precision. This evaluation can be accomplished using the expanded contingency matrix and related ratios presented above. This procedure is essential for forecasting applications using higher-order, time-aggregated data.

References

Bessler, David A. and Jon A. Brandt. Composite Forecasting of Livestock Prices: An Analysis of Combining Alternative Forecasting Methods. West Lafayette, Indiana: Dept. of Agricultural Economics, Agricultural Experiment Station, Purdue University, Station Bulletin No. 265, December 1979.

Bessler, David A. and Jon A. Brandt. "Forecasting Livestock Prices with Individual and Composite Methods." Applied Economics 13(1981):513-522.

Bourke, Ian James. "Comparing the Box-Jenkins and Econometric Techniques for Forecasting Beef Prices." Review of Marketing and Agricultural Economics 47(1979):95-106.

Brandt, Jon A. and David A. Bessler. "Composite Forecasting: An Application With U.S. Hog Prices." American Journal of Agricultural Economics 63(1981):135-140.

Brandt, Jon A. and David A. Bessler. "Forecasting With Vector Autoregressions Versus a Univariate ARIMA Process: An Empirical Example With U.S. Hog Prices." North Central Journal of Agricultural Economics 6(1984):29-35.

Gellatly, Colin. "Forecasting N.W.W. Beef Production: An Evaluation of Alternative Techniques." Review of Marketing and Agricultural Economics 47(1979):81-93.

Harris, Kim S. and Raymond M. Leuthold. "A Comparison of Alternative Forecasting for Livestock Prices: A Case Study." North Central Journal of Agricultural Economics 7(1985):40-50.

Hudson, Michael A. and Oral Capps, Jr. "Forecasting Ex-Vessel Prices for Hard Blue Crabs in the Chesapeake Bay Region: Individual and Composite Methods." Journal of the Northeastern Agricultural Council 13(1984):112-118.

Kulshreshtha, Surendra N. and K.A. Rosaasen. "A Monthly Price Forecasting Model for Cattle and Calves." Canadian Journal of Agricultural Economics 28(1980):41-62.

MacGregor, R.J. and S.N. Kulshreshtha. An Econometric Forecasting Model of the International Coarse Grains Market. Saskatoon, Saskatchewan: Dept. of Agricultural Economics, University of Saskatchewan, Economic Working Paper No. 11, May 1980.

Naik, Gopal and Raymond M. Leuthold. "A Note on Qualitative Forecast Evaluation." American Journal of Agricultural Economics 68(1986):721-726.

Theil, H. Economic Forecasts and Policy. Amsterdam: North Holland Publishing Company, 1961.

Table 1. 2 x 2 Contingency Table for Evaluation of Turning Point Precision

		Forecast Values	
		Turning Point (TP)	No Turning Point (NTP)
Actual Values	Turning Point (TP)	f_{11}	f_{12}
	No Turning Point (NTP)	f_{21}	f_{22}

Table 2. 4 x 4 Contingency Table For Evaluation of Turning Point Precision (Naik and Leuthold, p.724)

Actual Values	Forecast Values: Peak Turning Point (PTP)	Trough Turning Point (TTP)	Upward No Turning Point (UNTP)	Downward No Turning Point (DNTP)
Peak Turning Point (PTP)	f_{11}	f_{12}	f_{13}	f_{14}
Trough Turning Point (TTP)	f_{21}	f_{22}	f_{23}	f_{24}
Upward Turning Point (UNTP)	f_{31}	f_{32}	f_{33}	f_{34}
Downward Turning Point (DNTP)	f_{41}	f_{42}	f_{43}	f_{44}

Table 3. Diagrammatic Representation of Nine Turning Point Cases

#	Case	Representation
1	PTP	
2	TTP	
3	UNTP	
4	DNTP	
5	FDT	
6	FUT	
7	UFT	
8	DFT	
9	NC	

Table 4. 9 x 9 Contingency Table for Evaluation of Turning Point Precision

		Forecast Values								
		Peak Turning Point (PTP)	Trough Turning Point (TTP)	Upward No Turning Point (UNTP)	Downward No Turning Point (DNTP)	Flat Downward Turning Point (FDT)	Flat Upward Turning Point (FUT)	Upward Flat Turning Point (UFT)	Downward Flat Turning Point (DFT)	No Change (NC)
Actual Values	Peak Turning Point (PTP)	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}	f_{17}	f_{18}	f_{19}
	Trough Turning Point (TTP)	f_{21}	f_{22}	f_{23}	f_{24}	f_{25}	f_{26}	f_{27}	f_{28}	f_{29}
	Upward No Turning Point (UNTP)	f_{31}	f_{32}	f_{33}	f_{34}	f_{35}	f_{36}	f_{37}	f_{38}	f_{39}
	Downward No Turning Point (DNTP)	f_{41}	f_{42}	f_{43}	f_{44}	f_{45}	f_{46}	f_{47}	f_{48}	f_{49}
	Flat Downward Turning Point (FDT)	f_{51}	f_{52}	f_{53}	f_{54}	f_{55}	f_{56}	f_{57}	f_{58}	f_{59}
	Flat Upward Turning Point (FUT)	f_{61}	f_{62}	f_{63}	f_{64}	f_{65}	f_{66}	f_{67}	f_{68}	f_{69}
	Upward Flat Turning Point (UFT)	f_{71}	f_{72}	f_{73}	f_{74}	f_{75}	f_{76}	f_{77}	f_{78}	f_{79}
	Downward Flat Turning Point (DFT)	f_{81}	f_{82}	f_{83}	f_{84}	f_{85}	f_{86}	f_{87}	f_{88}	f_{89}
	No Change (NC)	f_{91}	f_{92}	f_{93}	f_{94}	f_{95}	f_{96}	f_{97}	f_{98}	f_{99}

Table 5. Turning Point Evaluation of Simulated Forecast and Actual Data for Figure 1

		Forecast Values								
		PTP	TTP	UNTP	DNTP	FDT	FUT	UFT	DFT	NC
	PTP	1	0	1	0	0	0	0	1	2
	TTP	2	1	1	0	1	1	0	0	0
	UNTP	1	1	1	1	0	1	2	1	0
	DNTP	0	0	0	0	1	1	0	0	0
Actual	FDT	0	0	2	0	0	0	0	0	1
Values	FUT	1	1	1	0	1	1	2	0	1
	UFT	0	2	3	0	0	3	1	1	0
	DFT	1	0	0	0	0	0	0	1	0
	NC	1	1	1	0	0	0	1	0	0

Ratio of Accurate Forecasts (RAF)	.1250
Ratio of Worst Forecasts (RWF)	.0833
Ratio of Accurate to Worst Forecasts (RAWF)	1.5006
Ratio of Inaccurate Forecasts (RIF)	.7917
Ratio of Inaccurate + Worst Forecasts (RWF+RIF) ...	.8750

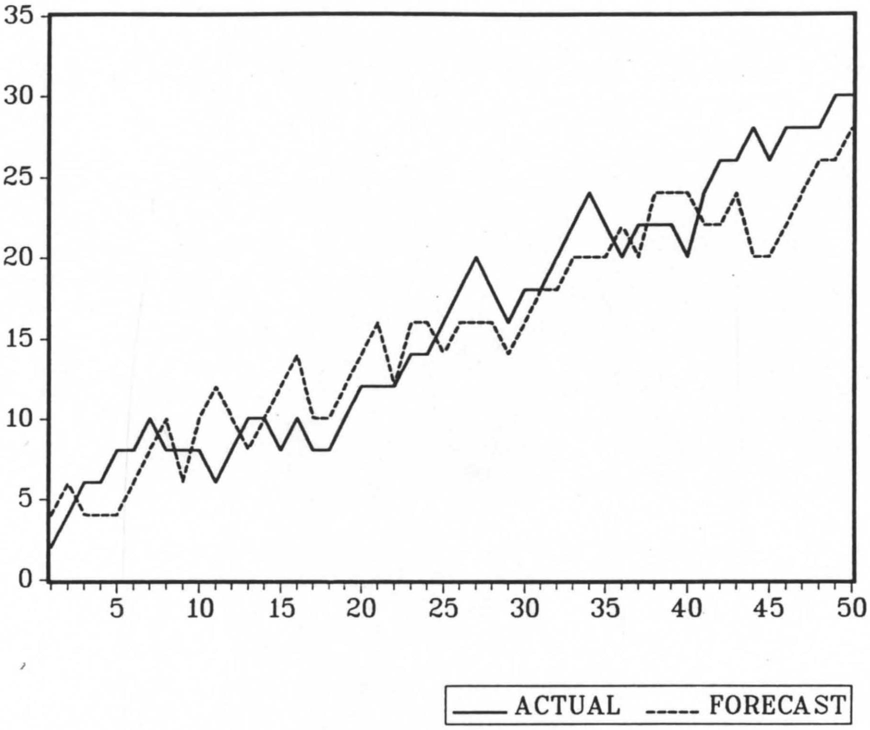

Figure 1. Simulated forecast and actual data patterns, observations 1-50

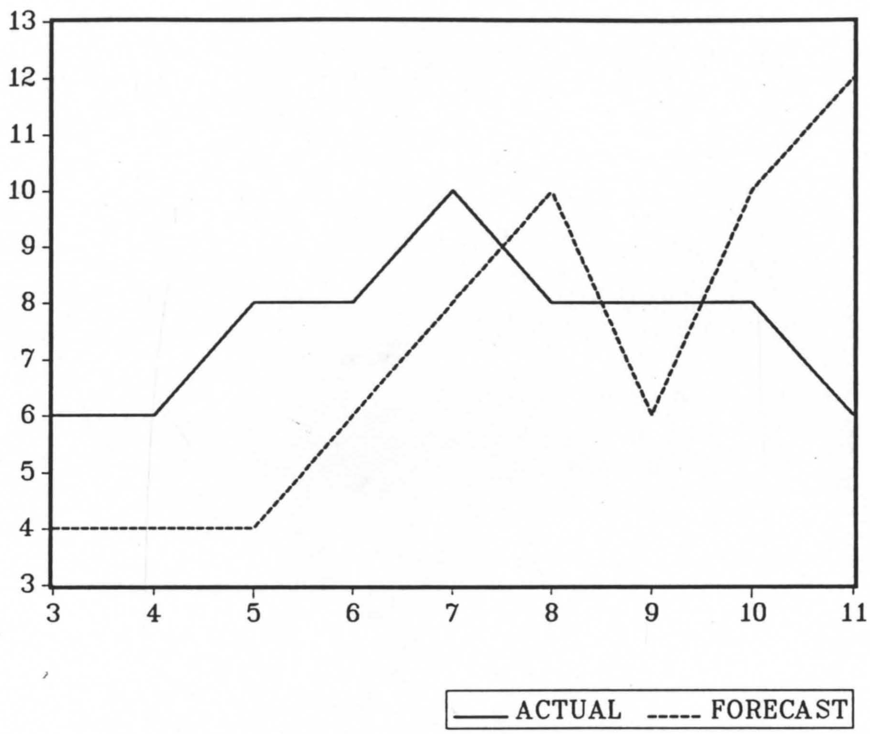

Figure 2. Simulated forecast and actual data patterns, observations 3-11