STATISTICAL ANALYSIS AND FORECASTING OF MAIN AGRICULTURE OUTPUT OF SRI LANKA: RULE-BASED APPROACH

T.M.J.A. COORAY
Department of Mathematics
University of Moratuwa
Moratuwa,
SRI LANKA.
STATISTICAL ANALYSIS AND FORECASTING OF MAIN AGRICULTURE OUTPUT OF SRI LANKA: RULE-BASED APPROACH
T.M.J.A. COORAY
Department of Mathematics
University of Moratuwa
Moratuwa, SRI LANKA.
email: cooray@math.mrt.ac.lk http://www.mrt.ac.lk

Abstract: - This paper examines the feasibility of rule-based forecasting, a procedure that applies forecasting expertise and domain knowledge to produce forecasts according to features of the data. Rule-based is developed to make annual extrapolation forecasts for econometric time series. The development of the rule-based drew upon protocol analyses of four experts on forecasting methods. Combined forecasts from four extrapolation methods (Decomposition method, Exponential smoothing, Winter’s Seasonal smoothing and ARIMA methodology). We looked at evidence from comparative empirical studies to identify methods that can be useful for predicting demand in various situations and to warn against methods that should not be used. In general, use structured methods and avoid intuition, unstructured meetings, focus groups and data mining. In such situations where there are sufficient data, use quantitative methods including extrapolation, quantitative analogies, rule-based forecasting and causal methods. Otherwise, use methods that structure judgment including surveys of intentions and expectations and structured analogies. Managers’ domain knowledge should be incorporated into statistical forecasts. Methods for combining forecasts improve accuracy. We provide guidelines for the effective use of forecasts, including such procedures as scenarios. Few organizations use many of the methods described in this paper. Thus, there are opportunities to improve efficiency by adopting these forecasting practices.

Key Words: Exponential Smoothing, Winter’s Seasonal, ARIMA, Rule-based, AIC

1 Introduction

In empirical studies of forecasting accuracy, no single extrapolation method has performed well across all types of data and all forecast horizons. In this paper, we describe an approach where the use of extrapolation methods depends upon the features of a time series. This approach incorporates judgment into the extrapolation process. A survey of experts [2] identified a need to integrate judgment and extrapolation. In reviewing the literature, [5] concluded that judgment and statistical forecasting methods should be integrated. Furthermore, [16] and [21] have provided empirical support.

Rule-based are developed and applied for extrapolations of annual economic and demographic time series and compared the accuracy of rule-based forecasting with other procedures for making extrapolations, in particular with equally-weighted combined forecasts.

1.1 Developing a Rule Base

Our intention for rule-based forecasting was to apply a validated, fully disclosed, and understandable set of conditional actions to make forecasts. Although rule development begins by specifying rules that seem reasonable to experts, the goal is to go beyond such specifications and to validate rules by prior research and empirical tests. By fully disclosed, means that the forecaster (or decision maker) has access to all of the rules. Understandable is judged from the perspective of forecasters and decision makers. We have developed a rule base that helps to realize our intention.

Our rule base integrates several strategies for extrapolation. Among these are (1) using features of the series to establish weights for combining forecasts, (2) using heuristics to establish parameters for an exponential smoothing model, (3) using separate models for long-range and short-range forecasts, (4) damping the trend under certain conditions and (5) incorporating domain knowledge in extrapolation. Most of these ideas have been
discussed in the forecasting literature and some have been supported by empirical research. In this study, we focused primarily upon combining. Of the many approaches proposed for combining forecasts, an equal weight has been shown to be particularly robust. However, our rule-based uses features of the series to produce differential weights.

1.2 Obtaining Information for Developing Rules

The forecasting literature provided some research streams that influenced our development of rules. Among these research streams were findings that favor decomposing time series, placing more weight on recent data, using simple methods, and making conservative trend estimates when uncertainty is high. Guidelines from the literature, however, typically lacked a precise statement of the conditions under which particular actions were likely to be useful. Also, such guidelines rarely specified precise actions needed to obtain forecasts.

The experts thought that the presence of an abrupt change was as important as the presence of trends. They also judged both the short-term and long-term trends to be important in the selection of an extrapolation method. The experts cautioned against the use of complex methods and they suggested combining forecasts for many situations. [14] Concluded that operational rules were most likely to be discovered when experts performed familiar, realistic tasks. The experts could view a variety of transformations, statistics, and forecasts.

The experts suggested rules for irregularities and changes in patterns. One rule was “If the series has a recent pattern that looks suspicious, and then increases the weight given to the random walk forecast. That is, rely more on an extrapolation of the last observation with no trend.” We found protocols especially useful in linking actions with conditions. On the other hand, the experts overlooked several conditions that were not represented in the test series. [2] Provide further details about some of these protocols.

1.3 Specifying the Rules

Our rules combine forecasts using weights that vary according to the features of the series. Given the desire for an understandable system, we chose to combine four simple, widely understood extrapolation methods:

1. The decomposition method of time series emphasizes the short-range perspective, it sets based on the assumption that there is a trend.
2. Exponential smoothing [15] captures information about short-range trends, and we call this the recent trend.
3. Winter’s Seasonal method, with trend [3] also measures the short-range trend.
4. Box-Jenkins ARIMA mythology, which fits a least squares line to the historical data (or transformed historical data), represents the long range.

We introduced rules only if they made sense to us. An initial set of time series was used to calibrate the rules. First, we made judgmental (eyeball) extrapolations. We then compared these with the rule-based forecasts to see if the latter looked reasonable to us. When the judgmental forecasts and the rule-based forecasts differed substantially, we searched for explanations. We also examined changes in the forecasts that resulted from adding a rule. These analyses led to many refinements in the rules. Because rules interact with one another, we examined the effects of removing a rule (or group of rules) from the rule base. We then put it back before analyzing other rules. We also examined different formulations for important rules. For example, we varied damping factors to determine the effects on accuracy. Experimentation was constrained because the number of series affected by some rules was small. These searches involved hundreds of runs and we recorded the results in a log.

1.4 Description of a Rule Base

One objective of the rule base is to provide more accurate forecasts. A second objective is to provide a systematic summary of knowledge. By expressing the knowledge explicitly and uniformly, we gain benefits generally associated with rule bases. These benefits include automating some tasks associated with maintaining a complex body of knowledge and providing knowledge in an accessible and modifiable form. Besides its usefulness in forecasting, knowledge in this form aids reasoning about forecasting, that is, the knowledge is useful both procedurally and declaratively. Such explicit representation also
can help in developing and testing theories [4]. Our goal in this study was to test the feasibility of rule-based forecasting as a procedure. To do this, we had to establish a set of reasonable rules. We do not presume that they are the best possible set of rules. Figure 1 shows the elements of the rule base. First, features of the series are identified. Rules are then applied to produce short and long-range forecasting models. To formulate these models we had to select smoothing factors for the exponential smoothing methods and make estimates of levels and trends for each model. For the long-range model, we formulated rules to damp the trend over the forecast horizon. Finally, there are rules for blending the forecasts from the short and long-range models.

Features of time series condition forecasting actions. We used 18 features to describe time series. All of the features are binary except the coefficient of variation (which is continuous). The analyst identifies certain features (e.g., irrelevant early data, unstable recent trend) by examining plots of the series. The identification of ‘last observation unusual’ is of particular concern given that the last observation typically plays an important part in the estimation of both level and trend. We expect that eventually the identification of such features will be performed by rules or analytical procedures. Domain knowledge (knowledge about the underlying process) is used to determine some features. Based upon knowledge of the underlying process, the analyst specifies the functional form as either additive or multiplicative. For example, many sales series follow a multiplicative process, that is, the trend can be thought of in percentage terms. Domain, knowledge also aids in identifying economic cycles, unusual events, and causal forces. Causal forces represent the net directional effect of the factors that influence a specific series at a given point in time. Causal forces play a role in many of the rules here. For example, sales of a product are often affected by forces that promote growth, such as an increasing number of customers, their greater wealth, their growing awareness of the product, and improvements in product quality. If a sales
trend is downward, a firm might try, to reverse it by reducing prices or increasing advertising. Consequently, the rules would extrapolate the estimated upward trends for this series, but they would damp downward trends. (For a discussion and evaluation of causal forces, see [2].

Rules and analytic procedures determine some features. Currently, these include features such as the direction of the basic trend, the statistical significance of the basic trend, and the amount of variation about the trend.

1.5 Estimating the Short-Range and Long-Range Models

The factors affecting short-range forecasts for a time series often differ from those affecting long-range forecasts. For example, in the short term, sales of automobiles could be affected by price reductions or by a new advertising campaign. Such things as design, service, restrictions on foreign trade, and product reliability influence long-range forecasts. Short and long-range models reflect different underlying processes. Separate models for short- and long-range extrapolations improved accuracy in a study by [17].

Smoothing Factors. A grid-search procedure provides estimates for smoothing factors that produce the best fit to the historical data for Holt's exponential smoothing. We used the grid search described in [17]. Smoothing factors for Brown's linear exponential smoothing were based largely upon rules from [1] and [13]. For example, \( \alpha \) (for levels) is decreased when the amount of historical variation is large, \( \beta \) (for trends) for the short-range model is increased if the causal forces support the recent trend direction.

Level. The rules apply weights to the estimated levels from the four methods-random walk, linear regression, Holt's, and Brown's. Short- and long-range models use different weights. For example, the short-range model puts more weight on the random walk. Weights for levels also vary depending upon the presence of discontinuities in the historical data, the proximity of the last value to a previous extreme in a cyclical series, the occurrence of an unstable recent trend or a changing basic trend, or the presence of a suspicious pattern. The assumption here is that features found in the historical data may occur again in the future. For example, when discontinuities are observed, rules put more weight on the latest observation (the random walk)

The rules adjust the level based upon how well the rule base was able to forecast the latest observation \( (t_0) \) given data through the preceding period \( (t-1) \). Rules adjust the estimate of the Level at \( t_0 \) in the direction, implied by the causal forces.

Trend. When the directions of the recent trend (estimated by Holt's exponential smoothing) and the basic trend (estimated by linear regression) different rules make a more conservative trend extrapolation; they do this by increasing the weight on the random walk. Similarly, when the causal forces are not in the same direction as the estimated trend, rules make a more conservative trend estimate.

The short- and long-range models employ different trend weights. The recent trend is more heavily weighted in the short-range model. The basic trend gets more weight in the long-range model than in the short-range model.

Other features also affect the weighting of extrapolation methods. For example, if there has been a long recent run (movements in the same direction for at least six consecutive years), a rule puts more weight on the recent trend and less on the basic. For series with causal forces that are regressing (that is, tending to return to a mean), a trend line is extended from the latest observation to the overall historical mean of the series. Rules then blend the previously estimated long-range model trend with this regressing trend.

Damping the Long-Range Model Trend. The damping factor represents the percentage of the long-range model trend that is ignored in the forecast. Rules establish a damping factor based upon the consistency of the recent and basic trends, and on the relationship of the direction of these trends to the causal forces. Rules then modify the factor by the amount of variation in the series and for suspicious patterns and unstable recent trends. The damping factor increases (i.e., the trend is further discounted)
Fig: 1. Original series for paddy data. The total values for the paddy production for "YALA" and "MAHA" seasons in Sri Lanka,

Source: Values are taken from annual bulletin published by Central Bank of Sri Lanka
Graph of paddy data series of Sri Lanka

As an illustration of the impact of the rules, Figure 1 presents one-year-ahead forecasts for paddy production of Sri Lanka. Figure 2 shows the forecast using an additive model, note the drop in the level. Upon examining this series but before looking at the holdout sample, we judged that an unusual event must have occurred in the year, 1982. (It did. We learned that a strike in 1982, 10% of its annual production.) Therefore, we adjusted the period, identified the last point as unusual, and applied the rules to produce a forecast. The rule for handling unusual last observations had a substantial impact. Of lesser importance was our adjustment of the last observation based on domain knowledge about the strike.

2.7 Hypotheses on Conditions Favoring Rule-Based Forecasting

The performance of rule-based forecasting depends not only on the rule base, but also on the conditions of the series. By conditions, we mean the set of features that describes a series. We expected both equal-weights combining and rule-based forecasting to be more accurate when
(1) Historical data have significant trends, and
(2) Uncertainty is low.

With respect to equal-weights combining, prior research provides much support for the first hypothesis and fair support for the second. We also hypothesized that rule-based forecasting would provide more accurate forecasts than the random walk or equal-weights combining when
(3) Historical data show stable patterns, and
(4) Good domain knowledge is available.

Hypotheses 3 and 4 represent logical extensions of prior research. We also expected rule-based forecasting to be more accurate for longer horizons. Extrapolating trends is risky because if the trend forecast is in the wrong direction, the resulting forecast will be less accurate than the random walk. Also, if the trend changes substantially, it can produce errors larger than those from the random walk.

We expected a trended forecast to be most useful when the basic trend in the historical data is statistically significant. When it is not, the random walk will tend to perform relatively well and the rule-based forecast will tend to move toward the random walk forecast. The rule base relies upon the patterns in the data to decide which extrapolation methods to use and how to weight them. When uncertainty is low, patterns are easier to detect, so rule-based forecasting is expected to be more accurate. Conversely, high uncertainty makes it difficult for rule-based forecasting to select the proper method. Here, equal weights combining should perform fairly well according to findings by [21] and [16]. When uncertainty is high, rule-based forecasting tends to approximate an equally weighted combined forecast. It also uses more damping, leading it to more closely approximate a random walk. Therefore, we expected rule-based forecasting to improve accuracy for series with low uncertainty, and to pose little risk for series with high uncertainty.
When series are unstable (with discontinuities, changing trends, etc.), the rule base moves toward a random walk forecast. Hence, rule-based forecasting can help in stable situations, and it is unlikely to harm the forecasts when the historical series show unstable patterns. Some techniques, such as Winter’s seasonal method, might encounter difficulties when historical series show unstable patterns. Knowledge about the subject area that is being forecast should be useful. Domain knowledge has improved accuracy in previous extrapolation studies (e.g., [18], [21]). Given its ability to incorporate domain knowledge, rule-based forecasting should be more accurate than the other extrapolation procedures when the forecaster has domain knowledge. Because rule-based forecasting can incorporate causal information, we expected that it would be more useful for long-range than for short-range forecasts, because causal factors have stronger effects in the long term. We also expected rule-based forecasting to be more appropriate for long-range forecasts because it uses information about the patterns in the data.

3 Analysis
The goal of this section is to provide analysis of certain data series by using different time series models. There are five methods were described, which are namely
(1) Decomposition of time series models.
(2) Exponential smoothing models.
(3) Winter’s Seasonal models.
(4) Box-Jenkins.

3.1 Data used for the study

ARIMA Methodology

We have mentioned alternative time series models, which are, used in practice by experts according as their experiences. In order to illustrate how it works in each model, we have to choose at least two data sets. Where one should visualizes seasonal pattern and other may not. To fulfill these requirement, Sri Lanka’s monthly total production of tea in Million kilograms in Sri Lanka for the period of January 1988 to September 2004 and total paddy production of Sri Lanka were used. According to the statistical analysis of log(paddy) data series, it is inclined to declare that tentative model is ARIMA(0,1,1)(0,1,1)2. In order to examine the situation further the possible model were evaluated and respective AIC values are also calculated in the Table 1

Table 1 Calculated Box-Pierce values, degrees of freedom, sample variance of error term and AIC values for all possible models applied to the Ln(paddy) data

<table>
<thead>
<tr>
<th>Model</th>
<th>Box-Pierce Q</th>
<th>df</th>
<th>Pvalue</th>
<th>$\sigma^2_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(0,1,1)(0,1,0)2</td>
<td>65.8</td>
<td>46</td>
<td>0.029</td>
<td>0.039</td>
</tr>
<tr>
<td>ARIMA(0,1,0)(0,1,1)2</td>
<td>96.3</td>
<td>46</td>
<td>0.000</td>
<td>0.036</td>
</tr>
<tr>
<td>ARIMA(0,1,1)(0,1,1)2</td>
<td>44.7</td>
<td>45</td>
<td>0.483</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Fig: 3. Graph of total production of Tea in Sri Lanka, data series
Table 2 Calculated Box-Pierce Q statistic and degrees of freedom, sample variance and AIC values for all possible models applied to the tea data

<table>
<thead>
<tr>
<th>Model</th>
<th>Box-Pierce Q</th>
<th>df</th>
<th>P value</th>
<th>$\sigma^2$</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(1,0,1)</td>
<td>134.7</td>
<td>45</td>
<td>0.000</td>
<td>10.31</td>
<td>332.9</td>
</tr>
<tr>
<td>ARIMA(1,0,0)(\times) (1,0,0)_6</td>
<td>73.8</td>
<td>44</td>
<td>0.005</td>
<td>9.38</td>
<td>319.6</td>
</tr>
<tr>
<td>ARIMA(1,0,0)(\times) (1,0,1)_6</td>
<td>48.9</td>
<td>44</td>
<td>0.01</td>
<td>8.50</td>
<td>307.7</td>
</tr>
<tr>
<td>ARIMA(1,0,1)(\times) (1,0,1)_6</td>
<td>39.0</td>
<td>43</td>
<td>0.645</td>
<td>8.07</td>
<td>302.4</td>
</tr>
</tbody>
</table>

With reference to Table 1, Box-Pierce value minimizes in the last model and that value is $\chi^2_{45} = 44.7$ which is statistically insignificant and $\sigma^2$ minimizes in the same with value(0.026).

The modified Box-Pierce Q statistic for the model ARIMA(0,1,1)\(\times\) (0,1,1)\_2, is 39.3 with 45 degrees freedom which is statistically insignificance. That also verifies the results obtained. Estimated model:

\[
(1 - B)(1 - B^2) \ln x_t = (1 - 0.69B)(1 - 0.93B^2)z_t
\]

Similarly for the total production of Tea in Sri Lanka series shows tentative ARIMA model as in the Table 2. It is clear that minimum AIC and Box Pierce Q statistic showed at the model ARIMA(1,0,1)\(\times\) (1,0,1)\_6

The parameter estimates are $\alpha_1 = 0.86$, $\alpha_2 = 0.99$, $\beta_1 = 0.495$, $\beta_2 = 0.91$ with the estimated variance, $\sigma^2 = 8.07$. The MINITAB output includes the standard errors for the parameter estimates. The modified Box-Pierce Q statistic measuring the lack of fit is now very reasonable, as shown in the last row of Table 2. To decide which model is the most appropriate one, AIC values are computed and the minimum value of AIC corresponds to the ARIMA(1,0,1)\(\times\) (1,0,1)\_6 model which confirms the results. Estimated model:

\[
x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-6} - \beta_1 z_{t-1} - \beta_2 z_{t-6} + z_{t-6}
\]

Fitted for Tea data is:

\[
x_t = 0.860x_{t-1} + 0.9973x_{t-6} - \beta_1 0.4954z_{t-1} - 0.9151z_{t-6} + z_t
\]

Table 3. Forecast performances of data series By using different techniques

<table>
<thead>
<tr>
<th>Model</th>
<th>Decomposition Method</th>
<th>Exponential Smoothing</th>
<th>Winter’s Seasonal</th>
<th>ARIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paddy</td>
<td>4.66%</td>
<td>3.7%</td>
<td>2.45%</td>
<td>1.04%</td>
</tr>
<tr>
<td>Tea</td>
<td>6.89%</td>
<td>4.87%</td>
<td>Not applicable</td>
<td>1.24%</td>
</tr>
</tbody>
</table>

4 Conclusions

This study is about statistical analysis of time series, to investigate the effectiveness of time series analysis and forecasting performance for real data sets based on ruled-based procedure. It address the question of how to analyze time series data how to identify structure, how to explain behavior, how to model those structure and how to use insight gained from the analyze to make informed forecasts. For the purpose of the study total production of paddy and total Tea production of Sri Lanka were used. Those values were obtained from the Annual Bulletin, published by Central Bank of Sri Lanka.

Therefore this paper is devoted to test for empirical evaluation of data series in order to
investigate the effectiveness of rule based estimation of time series. According to the forecast performance of ARIMA time series models are highly accurate than the other models, that means, percentage error (MAPE) from each data series, those MAPE values are relatively small so that ARIMA estimation of time series models give higher degrees of accuracy. The ARIMA model estimation of time series models can play an important role of time series modeling. However, Hence without identifying the appropriate structure of parameter variation these methods could be implemented in contrast to “en-bloc” procedure s which could be used only after assuming the specific type of ruled-based procedure.

References