An application of rule-based forecasting to a situation lacking domain knowledge

Monica Adyaa,[a,*] J. Scott Armstrong,[b] Fred Collopy,[c] Miles Kennedy [c]

[a] Department of Management, DePaul University, Chicago, IL 60604, USA
[b] The Wharton School, University of Pennsylvania, Philadelphia, PA, USA
[c] The Weatherhead School, Case Western Reserve University, Cleveland, OH 44106, USA

Abstract

Rule-based forecasting (RBF) uses rules to combine forecasts from simple extrapolation methods. Weights for combining the rules use statistical and domain-based features of time series. RBF was originally developed, tested, and validated only on annual data. For the M3-Competition, three major modifications were made to RBF. First, due to the absence of much in the way of domain knowledge, we prepared the forecasts under the assumption that no domain knowledge was available. This removes what we believe is one of RBF's primary advantages. We had to re-calibrate some of the rules relating to causal forces to allow for this lack of domain knowledge. Second, automatic identification procedures were used for six time-series features that had previously been identified using judgment. This was done to reduce cost and improve reliability. Third, we simplified the rule-base by removing one method from the four that were used in the original implementation. Although this resulted in some loss in accuracy, it reduced the number of rules in the rule-base from 99 to 64. This version of RBF still benefits from the use of prior findings on extrapolation, so we expected that it would be substantially more accurate than the random walk and somewhat more accurate than equal weights combining. Because most of the previous work on RBF was done using annual data, we especially expected it to perform well with annual data. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Rule-based forecasting (RBF) is an expert system that integrates judgment and statistical procedures to combine forecasts. It consists of condition-action statements (rules) where conditions are dependent upon features of the historical time series and upon domain knowledge. These rules yield weights appropriate to the forecasting situation as defined by the conditions. In effect, RBF uses structured judgment and statistical analysis to tailor extrapolation methods to the situation. Empirical results on multiple sets of time series show that RBF produces forecasts that are more accurate than those produced by traditional extrapolation methods or by an equal-weights combination of extrapolations. RBF is most useful when one has good domain knowledge, the domain knowledge has a strong impact (as often happens for longer-range forecasts), the series is well-behaved (such that patterns can be identified), and there is a strong trend in the data. When these conditions do not occur, RBF neither improves nor harms forecast accuracy.

Earlier versions of RBF were developed and presented in Collopy and Armstrong (1992), hereafter referred to as C&A. For the M3-IJF Competition, several modifications were made to the original version of RBF. In this paper, we describe this refinement and application of RBF to the competition. The second section provides background on RBF as it was developed for annual data. In the sections that follow, we describe the changes to RBF that were made from its inception in 1992 until the M3-Competition.

[*] Corresponding author. Tel.: +1-312-362-8495. *E-mail address:* adya@umbc.edu (M. Adya).

2. Rule-based forecasting

The rule-base in C&A consisted of 99 rules that used 28 features of time series to combine forecasts from four simple and widely understood extrapolation methods (random walk, linear regression, Holt's exponential smoothing, and Brown's exponential smoothing). Table 1 provides the details of these features as reported in C&A. Some of these features can be determined analytically by rules contained in RBF. For instance, the direction of the recent trend can be determined by fitting Holt's exponential smoothing to the historical data. Similarly, fitting a regression line on past data will indicate the direction of the basic trend. Features such as causal forces and start-up series rely on the experts' domain knowledge. These features represent knowledge of past events and future expectations that influence the series. Finally, several features of RBF rely on an analyst's examination and characterization of the plots of a series. Changing basic trend, unusual last observations, and level discontinuities are examples of such features. Table 1 summarizes the features used by RBF.

Table 1
Rulebased forecasting relies on 28 time-series features

Domain knowledge	Historical data	
Causal forces	*Types of data*	*Uncertainty*
Growth	Only positive values possible	Coeff. Of variation about trend >0.2
Decay	Bounded (e.g., percentages,	Basic and recent trends differ
Supporting	asymptotes)	
Opposing	Missing observations	
Regressing		*Instability*
Unknown	*Level*	Irrelevant early data
	Biased	Suspicious pattern
Functional form		Unstable recent trend
Multiplicative	*Trend*	Outliers present
Addictive	Direction of basic trend	Recent run not long
	Direction of recent trend	Near a previous extreme
Cycles expected	Significant basic trend(t>2)	Changing basic trend
Forecast horizon		Level discontinuities
Subject to events	*Length of series*	Last observation unusual
Start-up series	Number of observations	
Related to other series	Time interval (e.g., annual)	
	Seasonality	
	Seasonality	

C&A tested the initial version of RBF using 126 annual time series from the M-competition. These series were well-behaved in that they had strong trends, modest uncertainty in the trend, and few instabilities such as unstable trends, level discontinuities, and unusual observations. Moreover, since the data were annual, the causal forces acting on the series could be expected to play a stronger role. RBF proved more accurate than alternative methods, including the random walk, the typical methods used in the M-competition, and equal weights combining.

Despite the conclusions in C&A, there were concerns about the generalizability of RBF to other forecasting situations. First, RBF had been developed and calibrated on only 126 time series. Second, RBF relied on the identification of 28 features of time series. Since several of these features required judgmental coding, the process of feature identification was costly and restricted the applicability of RBF to newer, larger collections of data. Finally, RBF had been developed, tested, and validated on annual time series. Its applicability to shorter periods such as monthly and quarterly time series had not been examined. The M3-Competition provided a situation where RBF could be tested on larger data sets that included shorter-period data.

3. Enhancement and refinement of rulebased forecasting

The replicability and refinement of RBF had been restricted due to the costs involved in coding and reconciling the time-series features. We estimate that it took about 5 min to code and reconcile 10 features of a time series when two experts were involved. Consequently, for the M3-Competition, we estimated took over 200 h of

feature coding and reconciliation. To address this, Adya, Armstrong, Collopy and Kennedy (2000) developed heuristics to automatically identify six of the 10 features. Coding the remaining four domain-based features required under a minute.

Several other changes were made to RBF in the course of the .competition. These included corrections to the rule-base, the elimination of Brown's exponential smoothing as one of the component methods, modifications to the rules on causal forces, and the inclusion of seasonality and additional rules to handle short-period data. These changes are discussed in the next sections.

3.1 Corrections to RBF

RBF was originally implemented in Pascal running under Apple's Hypercard, which restricted its usability across other platforms. To counter this limitation, we converted the code to C. During this process, we identified several errors in the reporting and implementation of RBF in C&A. These errors are reported in Adya (2000) and the corrected set of rules is available on the web site http://www.research.umbc.edu/adya/rbf.html/. In all, there were 10 rules with inconsistencies between the code and the reporting. Six rules were correctly implemented in RBF's code but incorrectly reported in C&A. Four were correctly reported but were incorrectly implemented in the working version of RBF. We tested the corrected version of RBF against the results reported in C&A. There were no significant changes in the results. Details of the comparison are available in Adya (2000).

3.2 Heuristic identification of time series features

Rule-based forecasting is based on the premise that the features of time series can be reliably identified. Judgmental coding and manual reconciliation of these features has been an expensive undertaking. In Adya et al. (2000), we addressed this issue by identifying and automating heuristics for the detection of six of these 10 features. These were outlier, unusual last observation, changing basic trends, level discontinuity, unstable recent trend, and functional form.

We developed heuristics that rely on simple statistical measures such as first differences and regression estimates. For instance, the identification of a change in historical trend uses a comparison of slopes in various parts of the historical data. If there is a large difference in slopes, a change in the basic trend is assumed to have occurred.

Although there were differences between expert and automatic coding of individual features across 122 series, automating the identification of features caused little or no decline in accuracy across multiple error measures. The system resulting from this enhancement of RBF is now referred to as RBF(A).

3.3 Elimination of Brown's parameter estimation rules

Combining forecasts yields substantial benefits (Clemen, 1989). In empirical studies, the combined forecast error is almost always substantially lower than the average error of the component forecasts and it is sometimes better than the best component (Armstrong, 2001). RBF combined forecasts from random walk, linear regression, Holt's exponential smoothing, and Brown's exponential smoothing. Thirty-four rules in RBF(A) were used to determine smoothing coefficients for Brown's linear exponential smoothing method. This represents a sizable share of the rules. Following the principle of Occam's Razor we examined the effects of removing Brown's to reduce the complexity of the rule-base. Eliminating such a large portion of the rule-base could, however, be damaging to RBF(A)'s performance. Consequently, RBF(A) without Brown's was not expected to perform as well as RBF(A) with Brown's.

RBF(A) rules were modified so that Browns' weight allocations were assigned to Holt's exponential smoothing. Errors for the 36 time series from validation sample V3 in C&A were compared. Results are presented in Table 2. As hypothesized, RBF(A) without Brown's performed only slightly worse than RBF(A) with Brown's on all the measures except Median APEs. This evidence supported the elimination of Brown's. RBF(A) rules were trimmed from 99 to 64 and the combination used random walk, linear regression, and Holt's exponential smoothing.

Table 2
Results of eliminating Brown's from A-RBF (validation sample V3 from RBF)

	GMRAE			MdRAE			MAPE			MdAPE		
	1-yr	6-yr	Cum	1-yr	6-yr	Cum	1-yr	6-yr	Cum	1-yr	6-yr	Cum
Random walk	1.00	1.00	1.00	1.00	1.00	1.00	8.92	28.36	19.39	5.61	25.39	19.06
Equal weights	0.88	0.74	0.75	0.82	0.63	0.63	8.72	32.45	22.46	4.90	19.29	13.22
RBF(A) with Brown's	0.49	0.46	0.60	0.56	0.51	0.61	6.17	18.56	15.43	2.49	12.91	12.21
BRF(A) without Brown's	0.66	0.56	0.63	0.67	0.62	0.66	6.67	19.86	16.38	2.83	13.89	11.81

4. Causal force assumptions

Rule-based forecasting benefits from domain knowledge, in particular from the identification of causal forces that are acting on the series. Causal forces assess the net directional effect of the various factors expected to affect the trend over the forecast horizon. For instance, in forecasting the sales of computers, several factors such as rising incomes, increasing population, improvements in product capabilities, and reductions in prices can be expected to address the net directional effect. In practical situations, the forecaster has sufficient expertise in the domain to identify the causal forces.

Armstrong and Collopy (1993) examined the impact of causal forces on 104 annual series from the M-competition and found that the use of causal forces improved accuracy. For 1-yearahead forecasts, the use of causal forces reduced the Median APE by 4% and the Geometric Mean of the Relative Absolute Error by more than 2%. For the 6-year-ahead forecast, the MdAPEs improved by 12% and for the GMRAEs by 15%.

For series in the M3-Competition, causal force information was sparsely available. For instance, several series in the M3 data set were labeled simply as `Sales'. Consequently, we assumed that causal forces were unknown for all of the series. We expected that the accuracy of forecasts produced from RBF(A) would decline because the rules were developed, refined, and calibrated under conditions of identifiable causal forces. Consequently, we examined the rules that related to unknown causal forces and performed calibrations to find improved parameters for these rules. For instance, Rule 40 suggests that if the causal forces are unknown, then the weight on random walk should be increased by 5% and reduced from regression (making the trend extrapolation more conservative).

To accommodate the lack of domain information in the competition, we followed the `wind tunnel' approach where we selected a set of series to examine the impact of changes in rules (Armstrong, Adya & Collopy, 2001). We examined the impact of rule calibrations on a set of test series from the M3-Competition. These changes were validated on an ex-ante basis. 'In this calibration, we did not consider rules that related to mechanical adjustments. This narrowed our calibrations to three rules -40, 76, and 89. Rules 40 and 76 related to trend adjustments for the short and long models, respectively. Rule 89 pertained to trend damping. Calibrations of Rule 89 produced the most significant improvements in forecast accuracy on the validation data and produced the following modified rule:

Rule 89: If causal force is unknown, then damp the trend by 40%.

(The original implementation of this rule in C&A damped the trend by 5%.)

5. Adjustments for the shorter-period data

RBF and subsequently RBF(A) had been developed and tested on annual data. For the M3-Competition, we had an opportunity to work with shorter-period data. However, redefining the entire rule-base for shorter-period data would be a costly and time consuming process. Therefore, we made three major modifications to RBF(A) to

accommodate shorter-period data: seasonal adjustments of shorter-period data, recalibration of Rule 89, and the introduction of a new rule. No other changes were made to RBF(A) in its application to shorter-period data.

5.1. Seasonal adjustments

We implemented a simple version of seasonal adjustment. As a first step, the series are log transformed, detrended, and deseasonalized using the simple moving average approach. All feature detections and model fits were performed on the deseasonalized data. Once RBF(A) rules have produced a forecast, the forecasts were reseasonalized.

Michele Hibon provided selective seasonal adjustments for those series that indicated such a requirement based on a statistical criterion. Her seasonal factors were made available to all competitors. We generated a set of forecasts using these seasonal factors. We found that selectively adjusting short-period data did not improve our aggregate results and, consequently, reverted back to our original approach.

5.2. Causal force recalibrations for shorterperiod data

Calibration of the causal force adjustments was done for shorter-period data as well. This included quarterly and monthly series and the category `other', which included weekly and daily data. The damping factor was adjusted to accommodate the change in forecast horizon from 6 for annual data to 8 for quarterly and 18 for monthly and other data. Consequently, we modified rule 89 as follows:

For quarterly series:

Rule 89: If causal force is unknown, then damp the trend by 10%.

(The original implementation of this rule in C&A damped the trend by 5%.)

For monthly and 'other' series:

Rule 89: If causal force is unknown, then damp the trend by 3%.

(The original implementation of this rule in C&A damped the trend by 5%.)

5.3. Introduction of new rules for quarterly, monthly, and other shorter period data

In examining the impact of the rule calibrations discussed above, we found that, for short-period data, Holt's consistently outperformed RBF(A) on the early horizons, particularly for 1- to 5-ahead forecasts. This was possibly a function of the large damping factor that was used to accommodate unknown causal forces. Furthermore, two interesting patterns emerged with the short-period data:

- ?? Under conditions of stability and low uncertainty, RBF(A) performed about as well as Holt's exponential smoothing.
- ?? RBF(A) performed better than random walk, Holt's exponential smoothing, and equalweights combining when basic and recent trends are inconsistent.

To analyze this further, we regrouped the data set on the basis of features present and examined forecast accuracies on these sub-samples. In particular, we identified groups of series that had two or more discontinuities, contrary basic and recent trends, and series with high variability about the trend. Results indicated that Holt's outperformed RBF(A) on series with low variability and no discontinuities. Further analysis revealed that, under the conditions identified above, the starting weights for short model trend should shift entirely to Holt's with 0% weights on the other component methods. Consequently, we added an additional rule for short-period data. This rule suggests that: *If there are no discontinuities and the variation about the trend is less than 0.2, AND the short and long trends are consistent, then put all the initial weight for trend on Holt's else use the initial weighting scheme as proposed in the original version of RBF.*

6. Future enhancements of RBF(A)

As future efforts towards the improvement of RBF(A), we need to understand and refine the impact of seasonal adjustments on RBF(A)'s performance. In particular, modification of seasonal indices according to features of time series should produce improved forecasts.

Another area for research relates to improving performance on short-period series. Possible changes include reintroducing Brown's or adding Gardner's Damped Trend as a component method. Holt's did well on short-period data, particularly on the short-term forecasts. RBF(A), on the other hand, was weaker on these. The inclusion of either or both of these methods should produce improvements on the short model (Armstrong, 2001).

7. Conclusions

Several modifications have been made to enhance the accessibility and simplicity of RBF. In particular, we developed heuristics for automatically identifying features and reduced the number of rules by eliminating Brown's exponential smoothing. The latter change streamlined our rule-base without having a significant negative impact on RBF's accuracy against traditional benchmarks. We also identified several corrections to RBF.

For the M3-Competition, we had to make the assumption of unknown causal force for all series due to lack of sufficient domain information in the series descriptions. The parameters of several rules were modified. For the short-period data, we introduced new rules that increased the component weight of Holt's exponential smoothing under conditions of stability and low uncertainty. Finally, we factored in seasonal adjustments for short-period data. Our findings after the changes were consistent with those reported in C&A.

As expected, RBF(A) does better for annual data than for monthly and quarterly. This finding is particularly so for forecasts relating to the long horizons. Work remains to be done in the area of short-period data for RBF(A). These results are encouraging considering the absence of domain knowledge about the series in the competition.

References

Adya, M. (2000). Corrections to rule-based forecasting: findings from a replication. *International Journal of Forecasting,* 16, 125-128.

Adya, M., Collopy, F., Armstrong, J. S., & Kennedy, M. (2001). Automatic identification of time series features for rule-based forecasting, *International Journal of Forecasting*, 17, 143-157.

Armstrong, J. S. (2001). Combining forecasts, in: Armstrong, J. S., *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Norwell, MA: Kluwer Academic, pp. 417-439.

Armstrong, J. S., Adya, M., & Collopy, F. (2001). Rule based forecasting: Integrating judgment in time series extrapolations: Armstrong, J. S., *Principles of Forecasting: A Handbook for Researchers and Practitioners.* Norwell, MA: Kluwer Academic, pp. 259-282.

Armstrong, J. S. & Collopy, F. (1993). Causal forces: structuring knowledge for time series extrapolation. *Journal of Forecasting* 12, 103-115.

Clemen, R. (1989). Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting* 5, 559-583.

Collopy, F. & Armstrong, J. S. (1992). Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations. *Management Science* 38, 1392-1414.