Recurrent Artificial Neural Networks (RANN) for forecasting of forward interest rates

Amine Bensaid, Bouchra Bouqata
School of Science and Engineering, Al-Akahawayn University in Ifrane (AUI) 53000, Morocco

Ralph Palliam
School of Business Administration, Al-Akahawayn University in Ifrane (AUI) 53000, Morocco

Received March 2000

There are numerous methods for estimating forward interest rates as well as many studies testing the accuracy of these methods. The approach proposed in this study is similar to the one in previous works in two respects: firstly, a Monte Carlo simulation is used instead of empirical data to circumvent empirical difficulties; and secondly, in this study, accuracy is measured by estimating the forward rates rather than by exploring bond prices. This is more consistent with user objectives. The method presented here departs from the others in that it uses a Recurrent Artificial Neural Network (RANN) as an alternative technique for forecasting forward interest rates. Its performance is compared to that of a recursive method which has produced some of the best results in previous studies for forecasting forward interest rates.

Introduction

In the market where government obligations of various maturities bear coupons at different rates and where ordinary income and capital gains are subject to unknown and varying effective tax rates, the term structure of default-free interest rates is not directly observable. This is consistent with Jordan (1984), Livingstone (1989), McCulloch (1971) and Ronn (1987). At the same time, for financial research and practice, it is essential to know accurately the term structure of spot rates and the underlying term structure of forward rates. Buono, Gregory-Allen & Yaari (1992) contend that this cannot be obtained from the yield curve of treasury strip. This is due to the distinct and separate markets where those obligations are traded. Therefore, the term structure underlying the coupon bond market must be estimated from bonds traded in a given market.

According to Lee (1986) and Kishimito (1989), the term structure estimates are used for the management of fixed-income security portfolios and for pricing interest rate-contingencies claims such as fixed-income securities and options. They are also used in testing theories about the term *structure* itself. This is expounded by Brenan & Schwartz (1986). Furthermore, they are used as inputs to Monte Carlo simulations to value complex claims such as mortgage-based securities (Dattatreya & Fabozzi, 1989).

There are numerous methods for estimating forward rates as well as many studies testing their accuracy (Buono, Gregory-Allen & Yaari, 1992). The accuracy depends on either knowing the true underlying forward rates or the true distribution of errors associated with those rates. However, in the empirical data, the true distribution of the errors is unknown, which probably leads to biased statistical tests. Therefore, in this study, a Monte Carlo simulation was used instead of empirical data to circumvent the empirical difficulties. Buono, Gregory-Allen & Yaari (1992) contend that Monte Carlo simulation allows for a definition of a set of true forward rates with a known distribution of errors for comparison with the accuracy of various methods of estimating those rates. In this study, one was concerned with estimation of forward rates rather than bond prices to meet user objectives. Recurrent Artificial Neural Network is presented as an alternative technique for forecasting forward rates. Its performance is compared with that of a recursive method which has produced some of the best forecasting results in a previous study by Buono, Gregory-Allen & Yaari (1992).

Estimation methods

The notations below are used for the two methods:

- N_T: number of bonds maturing in period T.
- T_i: maturity period for bond i, T = 1,… , 20 (20 six-month periods).
- $P_{i,T}$: price per one dollar face value of bond i maturing T periods hence.
- $C_{i,t}$: cash flow per one dollar face value of bond i to be paid t periods hence (t = 1, …, T).
- $Y_{i,T}$: yield to maturity of bond i.
- $r_{i,t}$: forward interest rate on bond i over period t-1 to t.
- r_t: the arithmetic average of all forward rates over period t–1 to t (for bonds i = 1, …, N).
- $R_{i,t}$: spot interest rate on bond i over period 0 to t.

Where:

$$(1+R_{i,t})^t = (1+r_{i,1})(1+r_{i,2})\dots(1+r_{i,t}) \tag{1}$$

The ex-coupon price of bond i maturing at T is:

$$P_{i,T} = \sum_{t=1}^{T} \frac{C_{i,t}}{(1+Y_{i,T})^t} \tag{2}$$

which can be restated as a function of the spot interest rates:

$$P_{iT} = \sum_{t=1}^{T} \frac{C_{i,t}}{(1+R_{i,T})^t} \tag{3}$$

or function of the forward interest rates:

$$P_{i,T} = \sum_{t=1}^{T} \frac{C_{i,t}}{\prod_{s=1}^{t} (1 + r_{i,s})} \qquad (4)$$

Recursive method

The set of r_t is derived from Equation (4) in a recursive manner [17], starting with r_1. For each bond i (i = 1,...,N_T) maturing at T=1,

$$P_{i,1} = \frac{C_{i,1}}{(1 + r_{i,1})} = \frac{C_{i,1}}{(1 + R_{i,1})} \qquad (5)$$

which implies

$$r_{i,1} = \frac{C_{i,1}}{P_{i,1}} - 1 \qquad (6)$$

r_1 is define by:

$$r_1 = \frac{\sum_{i=1}^{N_1} r_{i,1}}{N_1} \qquad (7)$$

For each bond i maturing at T=2,

$$P_{i,2} = \frac{C_{i,1}}{(1 + r_{i,1})} + \frac{C_{i,2}}{(1 + r_{i,1})(1 + r_{i,2})} \qquad (8)$$

The average rate r_1 derived in (7) is substituted and it yields for every bond maturing at T=2,

$$P_{i,2} = \frac{C_{i,1}}{(1 + r_1)} + \frac{C_{i,2}}{(1 + r_1)(1 + r_{i,2})} \qquad (9)$$

which implies a set of period 2 forward rates:

$$r_{i,2} = \frac{C_{i,2}}{(1 + r_1)\left[P_{i,2} - \frac{C_{i,1}}{(1 + r_1)}\right]} - 1 \qquad (10)$$

These rates are averaged over the set of bonds i = 1,...,N_2 and then substituted in the price expression for three-period bonds, and so on. In general, the forward rate period is given by:

$$r_t = \frac{\sum_{i=1}^{N_t} r_{i,t}}{N_t} \qquad (11)$$

where

$$r_{i,t} = \frac{C_{i,t}}{\prod_{s=1}^{t-1} (1 + r_s) \left[P_{i,t} - \sum_{s=1}^{t-1} \frac{C_{i,s}}{\prod_{q=1}^{s} (1 + r_q)} \right]} - 1 \qquad (12)$$

Recurrent Artificial Neural Networks (RANN)

Zurada (1992) in presenting artificial neural systems suggests that the back-looping of an intermediate layer or of the output layer, allowed in recurrent networks, makes it possible to take past information into account. The forecast provided by a recurrent network depends on the example which is presented as input, but also on the preceding example and so on and so forth. The forecast provided at time t, thus, depends on the whole history preceding t. This property is interesting since for each series presented as input to the network, all the series' past values are taken into account for the forecast. Zurada (1992) states that this deals with the problem of determining the relevant lags of the series.

A multilayered network with looping back of the output layer is illustrated in Figure 1.

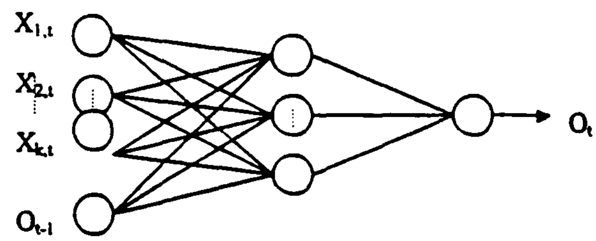

Figure 1 Recurrent network architecture with looping back of the output layer

The network considered here contains k inputs ($X_{1,t}$, $X_{2,t}$, ..., $X_{k,t}$) at time t, h hidden neurons and one output layer O_t at time t. On the input layer a back-looping neuron O_{t-1} corresponds to the output neuron at time t-1. Initially, this neuron's value is taken randomly.

The algorithm used, in this study, is the quick-propagation learning algorithm which is consistent with Fahlman's (1988) empirical study on fast learning variations on back-propagation. Input/output pairs are presented to the network. The input to RANN is three-dimensional; it consists of the price $P_{i,t}$ of bond i that will mature at time t, the corresponding coupon cash flow $C_{i,t}$ and the forward rate $r_{i,t-1}$ at time t-1. The output is the forward rate r_t at time t. The forward rate r_t at time t is the average of the N_t forward rates given by RANN for the N_t input/output patterns at time t as in Equation (11).

Simulation of the data

In order to compare the recursive method and the RANN method, a Monte Carlo simulation with 100 trials for 100 bonds that will mature at each semi-year t, t = 1,..., 20 was used. The choice of 100 trials for 100 bonds, which is consistent with previous studies, was made on the assumption that this would provide statistically significant and reliable findings. Each trial has several steps as in Buono, Gregory-Allen & Yaari (1992):

1. Four profiles of forward interest rates were generated (increasing, decreasing, flat and humped) matching the four kinds of yield curves observed in practice as follows:

   (a) For the increasing term structure:

   $$r_t = r_1 + \frac{\log(t)}{100}, \text{ for } t = 2, ..., 20, \qquad (13)$$

   $r_1 = 2.5\%$

   (b) For the decreasing term structure:

$$r_t = r_1 - \frac{\log(t)}{100}, \text{ for } t = 2, \ldots, 20, \quad (14)$$

$r_1 = 7.5\%$

(c) For the flat term structure:

$$r_t = r_1, \text{ for } t = 2, \ldots, 20, \quad (15)$$

$r_1 = 2.5\%$

(d) For the humped (random walk) term structure:

$$r_t = \phi_t, \text{ for } t = 2, \ldots, 20, \quad (16)$$

where ϕ □→ U(0, 2% to 5%).

The coefficient 1/100 in (a) and (b) is a scaling factor to assure feasible rates. Similarly, the range of the uniform distribution in (d) is chosen experimentally to maximize realism. So far all but the random walk are smooth functions. Smoothness is appropriate for future spot rates, which average forward rates, but would be a restrictive assumption for forward rates. Therefore, the forward rates was perturbed in (a) through (c) by:

$$r^*_t = r_t + \tau, \quad \text{for } t = 1, \ldots, 20 \quad (17)$$

where $\tau \rightarrow N(0.001, 1)$.

The mean of the normal distribution is chosen for realism.

2. Arbitrarily chose a range of 2% and 6% for the semi-annual coupon rates, but randomly assigned a coupon rate to each issue within that range. Specifically, the coupon rate of bond i is determined as:

$$C_i \rightarrow U(0.025, 0 \text{ to } 0.05),\ i = 1, \ldots, N \quad (18)$$

where U is a uniform distribution with mean 0.025 and range 0 to 0.05.

3. Actual forward rates and cash flows were used to calculate the price of the bonds as reflected in equation (19).

$$P_{i,t} = \sum_{t=1}^{T} \frac{C_{i,t}}{\prod_{s=1}^{t} (1 + r_s^*)} + \varepsilon_i, \quad (19)$$

To mimic reality, two components were added to the price: homoscedastic and heteroscedastic.

$\varepsilon \rightarrow N(0, \sigma)$

with:

σ = constant = 1, if ε_i is homoscedastic. In addition
$\sigma \rightarrow U(0.0015, 0.001 \text{ to } 0.002)$, if ε_i is heteroscedastic, corresponding to a randomly distributed noise with a mean of 0.0015 and a range in the interval [0.001,0.002].

Results

RANN was trained with different topologies (4-3-1, 4-4-1, 4-5-1, ..., 4-8-1), momentum (0.1, ..., 0.9) and learning rate (0.1, ..., 0.9). The threshold taken was equal to 0.0001 and the maximum number of epochs was equal to 10000. The best performance was given by three hidden neurons, momentum = 0.9 and learning rate = 0.5. The sigmoid function was used for the hidden neurons as activation function:

$$f(x) = \frac{1}{(1 + \exp(-x))}, \quad (20)$$

and for the output neuron the linear function:

$$f(x) = \frac{x}{10}, \quad (21)$$

The whole set of data of (100*20 = 2000) patterns (20 semi-years) was divided into two subsets: the first 2/3 for training RANN and last 1/3 for the testing.

The Square Average Mean Absolute Error (SAMAE) was used in estimating the performance of RANN and the recursive method since SAMAE penalises large errors to a greater extent. It is more sensitive to unusually large errors. The statistical properties of SAMAE are widely accepted in the studies cited.

The summary results of the finding is displayed in Table 1.

The comparison of errors in forward rates estimates on homoscedastic and heteroscedastic simulated data reveals that:

- The results for the recursive method and RANN are comparable for homoscedastic noise.
- RANN is more accurate than the recursive method with heteroscedastic noise.

This can be explained, on the one hand, by the fact that in the homoscedastic case the perturbance of the prices is not big and the data were generated by the same formula from where the recursive method was drawn. Therefore, the recursive method has a previous knowledge of the relationship existing between the input variables and the output one. Thus, it is more willing to come up with slightly better results than RANN, which does not have any previous information about the correct structure of the data. On the other hand, when introducing the heteroscedastic noise, the result is that the data now admit a noise component that the recursive method does not have any prior knowledge about. There is a new noise component in the data that is not accessible by the recursive method (noisy to the RANN method) whereas in the

Table 1 Forecast error in forward rates averaged over twenty periods

	Homoscedastic noise				Heteroscedastic noise			
Shape	AVG* (Rec)	SD* (Rec)	AVG (RANN)	SD (RANN)	AVG (Rec)	SD (Rec)	AVG (RANN)	SD (RANN)
Increasing	0.002	1.414	0.0018	0.00015	0.0088	0.00015	0.0039	0.00025
Decreasing	0.0021	0.0002	0.0026	0.0003	0.012	0.002	0.0036	0.0001
Random	0.00185	0.00005	0.0085	0.0015	0.0013	0.0001	0.009	0.0001
Flat	0.00043	0.00002	0.00002	0.00001	0.0004	0.00006	0.00002	0.00001

*AVG is the average of the minimum and maximum values of SAMAE for the 100 data sets and SD is its Standard Deviation homoscedastic experiments the recursive method exploits knowledge above the type of noise in the data. The recursive method does not perform as well as RANN, which tries to find the relationship existing in the data and hence is not dependent on certain axioms or criteria. Moreover, artificial neural networks were proven by Zurada (1992) to be good at modelling non-linear relationships and the forward interest rate data involves non-linearity in its structure.

Conclusion

The empirical study conducted in this article for forecasting forward interest rates has lead to the following observations:

- The results for the recursive method and RANN are comparable for homoscedastic noise.
- RANN is more accurate than the recursive method with heteroscedastic noise.

The superiority of RANN is generally present when the term structure of forward rates has a complex shape.

As further work one can site the application of RANN in forecasting future spot rates, which are more complex than forward rates and tend to involve a lot of irregularities in the future. To add to the performance of RANN one can further use fuzzy membership functions for the input data.

References

Buono, M., Gregory-Allen, R.B. & Yaari, U. 1992. The efficacy of term structure estimation techniques: a Monte Carlo study, *The Journal of Fixed Income*, March: 52–63.

Brenan, M.J. & Schwartz, E.S. 1986. A continuous time approach to the pricing of bonds, *Journal of Banking Finance*, 3: 133–156.

Brown, S.J. & Dybvig, P.H. 1986. The empirical application of the Cox, Ingresoll, Ross theory of the term structure of interest rates, *Journal of Finance*, 41: 617–632.

Carleton, W.T. & Cooper, I.A. 1976. Estimation and uses of the term structure of interest rates, *Journal of Finance*, 31: 1067–1083.

Chambers, D., Carleton, W.T. & Waldman, D.W. 1984. A new approach to the estimation of the term structure of interest rates, *Journal of Financial and Quantitative Analysis*, 19: 133–252.

Cox, J.C., Ingersoll, J.E. & Ross, S.A. 1981. A theory of the term structure of interest rates, *Working Paper #19*, Stanford University Graduate School of Business.

Dattatreya, R. & Fabozzi, F. 1989. A simplified model for the valuation of debt option. In *Handbook of fixed-income options*. Chicago: Probus Publishing Company.

Fahlman, S. 1988. Fast-learning variations on back-propagation: an empirical study, *Proc. Connectionist Model Summer School.*

Haugen, R. 1986. *Modern investment theory.* Englewood Cliffs, NJ: Prenctice-Hall, pp. 305–307.

Ho, T.S.Y. & Lee, S. 1986. Term structure movements and interest rate contingent claims pricing, *Journal of Finance*, 41: 1011–1029.

Jordan, J.V. 1984. Tax effects in term structure estimations, *Journal of Finance*, 39: 393–4064.

Kishimito, N. 1989. Pricing contingent claims under interest rate and asset price risk, *Journal of Finance*, 44: 571–589.

Langitieg, T.C. 1980. A multivariate model of the term structure, *Journal of Finance*, 35: 71–97.

Livingston, M. 1979. Bond taxation and the shape of the yield-to-maturity curve, *Journal of Finance*, 34: 189–196.

McCulloch, J.H. 1971. Measuring the term structure of interest rates, *Journal of Finance*, 34: 19–31.

Ronn, E.I. 1987. A new approach to bond portfolio management, *Journal of Financial and Quantitative Analysis*, December: 439–466.

Shaefer, S.M. 1982. Tax-induced clientele effects in the market for British government securities, *Journal of Financial Economics*, 10: 121–159.

Zurada, J.M. 1992. *Artificial neural systems*. West Publishing Company.