Chapter 6

6.1. Introduction

Finance is one of the fastest developing areas of the modern corporate world. This, together with the sophistication of modern financial products, provides a rapidly growing impetus for the new mathematical models and modern mathematical methods; the area is an expanding source for novel and relevant real world mathematics. The stock exchange, “share market” or a “bourse” is a mutual organization for traders or “stock brokers” who trade in different company securities and stocks. Companies or businesses have to be “listed” in the bourses in order for any trading or exchange in their “shares” or equities to be carried out. Stock markets are also the place for trading in units and bonds issued by the government. Bombay Stock Exchange (BSE) index and National Stock Exchange (NSE) index belong to India. The BSE index is a market capitalization weighted index of 30 stocks of sound Indian and multinational companies.

A time series (or time sequence) is a sequence of real numbers, each number representing a value at a time point. In its simplest form, a time series is a collection of numerical observations made at a discrete and equal time intervals. Usually each observation is associated with a particular instant of time or interval of time, and it is this that provides the ordering. The observations could equally well be associated with points along a line, but whenever they are ordered by a single variable, we refer to it conventionally as time series. We generally assume that time series values are equally spaced. One may observe the various properties of the time series signals like trends, abrupt changes, drifts and self similarities etc. Some of the important situations where time series occurs are, in the field of economics where one is exposed to stock market quotations, weekly inflations rates or foreign exchange rates.

The analysis of the experimental data that is observed at different points in time is called time series analysis. Time series analysis comprises methods that attempt to understand such time or to make predictions. Time series predictions make use of a model to predict future data points based on the past available data points as elements of a given data.
The main objective of work in this chapter is to test a novel architecture called Adaptive Network based Fuzzy Inference System, or simply ANFIS, proposed by Jyh-Shing Roger Jang (1993) to study the prediction of the financial time series of BSE 100 using the concepts of wavelets and Neuro fuzzy. We used the time series data for a period from 06-01-1986 to 12-08-2004.

We have divided this chapter into five sections. In the first section above, we have given a brief introduction to the stock market and time series. In the second section we give some preliminaries required for the subsequent discussion. This section has been divided into three subsections. In the first subsection, we discuss the wavelet tools such as discrete wavelets and multiresolution etc. and in the second and third subsections, the elements of neuro fuzzy and ANFIS (Adaptive Network based Fuzzy Inference System) are described. The third section explains the procedure adopted for the prediction of BSE closings using wavelets and neuro fuzzy. Results obtained with the procedural application on the time series data are given in the fourth section. In the last section of this chapter, we make the concluding remarks. The results of the work presented here are given in Kumar and Manchanda (accepted).

6.2. Preliminaries

In this section, we describe the wavelet methodology and ANFIS technique which are required for the prediction procedure.

6.2.1. Wavelet Tools

The theory of wavelets is a refinement of Fourier analysis which enables to simplify the description of a cumbersome function in terms of a small number of coefficients. The wavelet transform has been found to be particularly useful for analyzing signals which can best be described as non periodic, noisy, intermittent, transient and so on. Its ability to examine the signal simultaneously both in time and frequency is a distinct feature of wavelet analysis. Wavelet analysis has been used for predicting the prices of oil time series (Yousefi et al. (2005)), meteorological pollution (Stanislaw and Garanty (2007)) and wind speed etc. We may recapitulate the following concepts.
Discrete wavelet transform (DWT) is commonly introduced using a matrix or computation form.

In matrix form we can represent the DWT through an orthogonal matrix

$$W = [W_1^T, W_2^T, W_3^T, ..., W_j^T, V_j]^T,$$

where $V_j$ is a scaling function, $j$ is the largest level of transform and $t$ indicates transpose.

A DWT is applied to a vector $X$ of observations as $d = WX$ and decomposes the data into sets of wavelet coefficients.

$$d = [d_1^T, d_2^T, d_3^T, ..., d_j^T, c_j^T]^T$$

with

$$d_j = W_jX \text{ and } C_j = V_jX.$$ 

Wavelet transform leads to an additive decomposition of a signal into a series of different components describing smooth and rough features of the signal. In fact we have

$$X = W^t d = \sum_{j=1}^{J} W_j^t d_j + V_j^t c_j = \sum_{j=1}^{J} D_j + C_j, \quad (6.1)$$

where $D_j$ is the detail of the signal describing changes at the scale $\tau_j$ and $C_j$ is smooth component associated with the variations $\tau_{j+1}$ and higher.

DWT provides a mechanism to represent a time series data or signal $f$ in terms of coefficients that are associated with particular scales and therefore is regarded as a family of effective instrument for signal analysis. The decomposition of a given signal $f$ into different scales of resolution is obtained by the application of the DWT on $f$.

For practical applications, the first step of DWT involves the mapping $f$ to its wavelet coefficients and from these coefficients two components viz. approximations (smooth version) and details (noise version) are received. A decomposition of signal $f$ into a low frequency part $a_1$, and a high frequency part $d_1$, is represented by $f = a_1 + d_1$. The same procedure is performed on slower part of the signal $a_1$ in order to obtain a decomposition in finer scales as $a_1 = a_2 + d_2$. A recursive procedure for the decomposition of the low frequency parts (approximations) is illustrated as
This diagram illustrates a wavelet decomposition into $N$ levels and corresponds to

$$f = d_1 + d_2 + d_3 + \ldots + d_{N-1} + d_N + a_N. \quad (6.2)$$

In practical applications, such decomposition is obtained by using a specific wavelet. Several families of wavelets have proven to be especially useful in various applications. They differ with respect to orthogonality, smoothness and other related properties such as vanishing moments or size of the support.

6.2.2. Neuro Fuzzy

An adaptive network, as its name implies, is a network structure consisting of nodes and directional links through which the nodes are connected. Moreover, part or all nodes are adaptive, which means each output of these nodes depends on the parameter(s) pertaining to this node, and the learning rule specifies how these parameters should be changed to minimize a prescribed error measure. An adaptive network is a multilayer feed forward network in which each node performs a particular function (node function) on incoming signals as well as a set of parameters pertaining to this node.

The ANFIS architecture (Jang and Gulley (1995), Jang and Sun (1995), Jang et al. (1997)) and the learning rule of adaptive network in the following section is used and is referred as ANFIS standing for Adaptive Network based Fuzzy Inference System. We have used Matlab for training the system using ANFIS edit to train the wavelet coefficients.

6.2.3. ANFIS Architecture

According to the Takagi and Sugeno (1985) type, the fuzzy inference system has two inputs $x$ & $y$ and one output $z$.

Rule 1: If $x$ is $A_1$ and $y$ is $B_1$, then $f_1 = p_1x + q_1y + r_1$.

Rule 2: If $x$ is $A_2$ and $y$ is $B_2$, then $f_2 = p_2x + q_2y + r_2$. 
Layer 1 Every node $i$ in this layer is a square node with a node function

$$Q_i^1 = \mu_{A_i}(x),$$

where $x$ is the input to node $i$, and $A_i$ is the linguistic label (small, large, etc.) associated with this node function. In other words $Q_i^1$ is the membership function of $A_i$ and it specifies the degree to which the given $x$ satisfies the quantifier $A_i$. Usually we choose $\mu_{A_i}(x)$ to be bell-shaped with maximum equal to 1 & minimum equal to 0, such as the generalization bell function

$$\mu_{A_i}(x) = \frac{1}{1 + [(\frac{x-c_i}{a_i})^2]^b_i},$$

or the Gaussian function

$$\mu_{A_i}(x) = \exp \left[ - \left( \frac{x-c_i}{a_i} \right)^2 \right],$$

where $\{a_i, b_i, c_i\}$ (or $\{a_i, c_i\}$ in the latter case) is the parameter set. As the values of these parameters change, the bell shaped functions vary accordingly, thus exhibiting various forms of membership functions on linguistic label $A_i$. In fact, any continuous and piecewise differentiable functions, such as commonly used trapezoidal or
triangular-shaped membership functions, are also qualified candidates for node functions in this layer. Parameters in this layer are referred to as *premise parameters*.

**Layer 2** Every node in this layer is a circle node labeled $\Pi$ which multiplies the incoming signals and send the product out. For instance

$$W_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i = 1, 2.$$  

Each node output represents the *firing strength* of a rule.

**Layer 3** Every node in this layer is a circle node labeled $\mathbb{N}$. The $i$-th node calculates the ratio of the $i$-th rule’s firing strength to sum of all rules’ firing strengths:

$$\overline{W_i} = \frac{W_i}{W_1 + W_2}, \quad i = 1, 2.$$  

For convenience, outputs of this layer will be called *normalized firing strength*.

**Layer 4** Every node $i$ in this layer is a square node with a node function

$$Q_i^4 = \overline{W_i} f_i = \overline{W_i}(p_i x + q_i y + r_i),$$

where $\overline{W_i}$ is the output of layer 3, and $\{p_i, q_i, r_i\}$ is the parameter set. Parameters in this layer will be referred to as *consequent parameters*.

**Layer 5** The single node in this layer is a circle node labeled $\Sigma$ that computes the overall output as the summation of all incoming signals, i.e.,

$$Q_i^5 = \text{overall output} = \sum_i \overline{W_i} f_i = \frac{\sum_i W_i f_i}{\sum_i W_i}.$$  

Thus, we have constructed an adaptive network which is functionally equivalent to a type 3 fuzzy inference system. The type-3 fuzzy reasoning is illustrated in Fig. 4.1(a) and the corresponding equivalent ANFIS architecture (type-3 ANFIS) is shown in Fig. 4.1(b). For type-1 fuzzy inference systems, the extension is quite straightforward where the output of each rule is induced jointly by the output membership function and the firing strength.
6.3. Predictability Method

We have used Daubechies db8 wavelet for decomposing the signal (data) under consideration. All signals (wavelet coefficients from $d_1$ to $d_3$ and the core approximation $a_3$ on the third level) are illustrated in the original resolution. We observe the substantial difference of variability of the signals at different levels. The higher is the wavelet level, the lower variation of the coefficients and easier prediction of them. Our main idea is to substrate the prediction task of the original time series of high variability by the prediction of its wavelet coefficients on different levels of lower variability, and then using Matlab code with ANFIS for final prediction of the stock market closing at any time instant $n$. Since most of the wavelet coefficients are of lower variability, we expect the increase of the total prediction accuracy.

The wavelet tool in Matlab wavelet toolbox is used for the process of wavelet decomposition of the time series representing average of the financial time series under consideration. This step involves several different families of wavelets and a detailed comparison of their performance. In our case, we have used the db8 wavelets. A three level wavelet decomposition of the given time series signal $f$ is performed as

$$f = a_3 + d_3 + d_2 + d_1.$$ 

The smooth part of $f$ is stored in $a_3$, and details at different levels are captured by $d_1$, $d_2$, $d_3$. Consequently a decomposition of the time series in three different scales is obtained. Fig.6.1 illustrates the decomposition of the original signal.

The next value is predicted by using the forecasting procedure methodology explained in section 6.2.3. The basic idea is to use the wavelet transform and predict the data by neuro fuzzy for individual coefficients of wavelet transform represented by $a_3$, $d_1$, $d_2$, and $d_3$. The input to the ANFIS architecture to predict the wavelet coefficients is explained in Fig. 6.2.
Figure 6.1: Wavelet decomposition of BSE 100 time series data.

The total predicted closing at an instant \( i \) is given by

\[
F(i) = f_1(i) + f_2(i) + f_3(i) + f_4(i).
\]

Figure 6.2: ANFIS mechanism for prediction procedure.
6.4. Results of Numerical Experiment

The numerical experiment for predicting the stock market closings has been performed for BSE 100. The main experiment has been performed using neuro fuzzy predictor. We decomposed the time series data into wavelet coefficients and then applied the neuro fuzzy for predicting the wavelet coefficients at each level. Out of 4354 points, we train 4300 data points and tested the remaining 54 points using the ANFIS approach through the Daubechies wavelet db8 at level 3.

The trained predicted output is obtained from the decomposed wavelet coefficients by simple summation as

\[ S(N) = D1 + D2 + D3 + ... + DN + AN. \]

Figure 6.3 shows the trend of the stock market closings for the tested data.
Figure 6.4. Plot showing predicted output (obtained using wavelet and ANFIS) and actual output (54 points used for prediction testing).

Figure 6.5. Predicted approximation (A3) versus original approximation.
Figure. 6.6. Predicted versus original detail (D1) output.

Figure. 6.7. Predicted versus original detail (D2) output.
Figure 6.8. Predicted versus original detail (D3) output.

Figure 6.9. The actual data and training data.
Figure 6.10. Error at each point between training data and actual data for 4300 data points (actual error, not in percentage).

Figure 6.11. The testing error for 54 points.
Figure 6.12. The prediction error.

Figure 6.13. The training error in coefficients for db8 wavelet at level 3.
Table 6.1:

<table>
<thead>
<tr>
<th>Wavelet Coefficients</th>
<th>Average Training Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3</td>
<td>7.2494</td>
</tr>
<tr>
<td>D1</td>
<td>2.9937</td>
</tr>
<tr>
<td>D2</td>
<td>3.5688</td>
</tr>
<tr>
<td>D3</td>
<td>1.5247</td>
</tr>
</tbody>
</table>

Table 6.1 shows the average training error for approximation and detail coefficients.

Figure 6.14. The testing error in coefficients for db8 wavelet at level 3.
Table 6.2

<table>
<thead>
<tr>
<th>Wavelet Coefficients</th>
<th>Average Testing Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3</td>
<td>6.3975</td>
</tr>
<tr>
<td>D1</td>
<td>5.9851</td>
</tr>
<tr>
<td>D2</td>
<td>4.6070</td>
</tr>
<tr>
<td>D3</td>
<td>2.7283</td>
</tr>
</tbody>
</table>

Table 6.2 shows the average testing error for wavelet coefficients.

6.5. Conclusion

In this work we have made efforts to predict the closings of the BSE 100 stock market for the 54 points. Satisfactory results were obtained on employing wavelet decomposition with neuro fuzzy. The wavelet transform has explored a new idea of generalization of neural network and specialization to Tagaki Sugeno inference fuzzy logic for training the non stationary data and predicting the output. We are getting considerable low average training and testing errors of the coefficients. The time series data was found unfit to be trained by using only neuro fuzzy because of its variability. We combined the concept of neuro-fuzzy with the wavelet decomposition of data and got better results.

***************