Chapter 1 INTRODUCTION

Time series analysis is an important method for data forecasting to provide important information during decision-making processes, especially in economic developments, population policies, management planning, and financial controls. Nevertheless, the forecast model constructed only by real numbers may not demonstrate the entire trend of a daily or monthly process. Many practical cases are affected by a lot of vague and incomplete information. Since there are so many unpredictable and continuous fluctuations on the process to be predicted, the observed values are discrete instantaneous values which are insufficient to represent the true process. Therefore, the observed data usually are present as not only a single numeric type but also a range of numbers, such as the variation of daily temperature, the fluctuation of the exchange rate, and the level price of petroleum, and so on. Due to the uncertainty of the continuous fluctuations, using intervals to illustrate the uncertain factors in the forecasting model is practical and objective. This dissertation proposes interval time series approaches which utilize the characteristics of interval time series to analyze and forecast interval data.

Since the interval time series is applied to forecast, some definitions of interval operations relevant to interval time series will be given first. In order to deal with a set of dynamic data, the conception of fuzzy is led into interval and a new expression of interval will be introduced to represent the dynamic range of data. The new interval expression integrates the mean value of interval with its distances to the boundaries instead of the conventional notation of interval which is simply represented by the lower boundary and the upper boundary. Based on the expressions of interval and its operations, we are able to make an appropriate interval forecasting for an interval time series by means of the location shift and the length variation of interval data. Furthermore, we also can investigate the performance evaluation of interval

forecasting and the dynamic trend of interval time series. Accordingly, the composition of this dissertation includes three topics regarding interval data, which are the analysis and forecasting of interval time series, the evaluation of forecasting performance for interval data, and the calculation of the fuzzy correlation coefficient.

In Chapter 2, we discuss the interval time series and the forecast technique. Since it is an innovative subject, there is no forecasting analysis about interval time series. Thus, three forecasting approaches are proposed, which are the interval moving average of order k, the weighted interval moving average of order k, and the *ARIMA* forecasting. In order to analyze the interval forecasting, stationary interval time series and non-stationary interval time series are generated by simulation, which are the $AR(1)$ model, the threshold models combining three models of $AR(1)$ with different sets of coefficients, and $ARCH(1)$ model designed by three different variations.

The validity of the forecasting methods is the most concerned issue for policymakers. We concern not only the forecasting of interval lengths but also the location disparity between the forecast intervals and the actual intervals. Hence, the mean squared error of interval and the mean relative interval error are established to illustrate the efficiency of the forecast result. The *ARIMA* forecasting is significant forecasting technique which conveys the best efficiency of the forecast result than the other forecasting methods. If we can establish a superior model construction according to the interval operations and the *ARIMA* method, we can make a superior interval forecasting. For investors, it provides a new forecasting method and also offers more flexible forecast results. Thus, investors can make more objective judgment under correct information.

In Chapter 3, we investigate how to estimate the accuracy of forecast interval results. Since the mean squared error of interval and the mean relative interval error are reasonable ways for explaining the efficiency of forecast results. But they are not sufficient to describe the

forecasting situation when the forecast interval is too wide. By the overlap and the non-overlap parts of the actual intervals and the forecast intervals, we propose the mean ratio of exclusive-*or* to demonstrate the forecasting performance for interval data. Based on the forecast interval, there are four forecast situations: (1) the forecast interval is too wide, (2) the forecast interval is too narrow, (3) the forecast interval inclines to the right, and (4) the forecast interval inclines to the left. By means of the four forecast situations, an example is used to compare the mean squared error of interval, the mean relative interval error, and the mean ratio of exclusive-*or*. It can be found that the mean ratio of exclusive-*or* provides an important efficiency analysis for interval forecasting. By the mean ratio of exclusive-*or* in different forecast situations, it may modify the forecasting method of the center and radius respectively.

Since interval data can be used to provide more objective statistical inferences, we must consider the correlation between the observed data which contains interval data. Although the traditional correlation coefficient is a convenient and reasonable statistic which explains the interrelation of two variables, but is the correlation coefficient appropriate enough to represent the characteristics of the population? Is it too subjective and constrained when the correlation coefficient is evaluated by real-valued data? Consequently, in Chapter 4, we propose an evaluation technique of correlation coefficient for the case that one of the two variables is in the form of interval. The dynamic range of correlation coefficient is expressed by an interval and evaluated by interval operations. Because the operations of interval are complicated, evaluating the correlation coefficient interval is difficult. Since an interval can be viewed as a fuzzy number, we construct the approximation procedure of evaluating the correlation coefficient interval by the aspect of fuzzy. Such correlation coefficient interval is thus called the fuzzy correlation coefficient. At last, an empirical study is given to compare the traditional correlation coefficient and the fuzzy correlation coefficient. The fuzzy correlation coefficient will be proved to be an objective statistic to represent the interrelation of two variables.