Expert Systems with Applications 39 (2012) 840–848

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A fuzzy extended DELPHI method for adjustment of statistical time series prediction: An empirical study on dry bulk freight market case ☆

Okan Duru [a,*], Emrah Bulut [b], Shigeru Yoshida [b]

[a] Department of Maritime Transportation and Management Engineering, Istanbul Technical University, Tuzla 34940, Istanbul, Turkey
[b] Department of Maritime Logistics, Kobe University, Higashinada 658-0022, Kobe, Japan

ARTICLE INFO

Keywords:
Fuzzy-DELPHI
Decision support systems
Forecasting support systems
Dry bulk shipping
Consensus forecasts

ABSTRACT

This paper investigates the forecasting accuracy of fuzzy extended group decisions in the adjustment of statistical benchmark results. DELPHI is a frequently used method for implementing accurate group consensus decisions. The concept of consensus is subject to expert characteristics and it is sometimes ensured by a facilitator's judgment. Fuzzy set theory deals with uncertain environments and has been adapted for DELPHI, called fuzzy-DELPHI (FD). The present paper extends the recent literature via an implementation of FD for the adjustment of statistical predictions. We propose a fuzzy-DELPHI adjustment process for improvement of accuracy and introduced an empirical study to illustrate its performance in the validation of adjustments of statistical forecasts in the dry bulk shipping index.

1. Introduction

Over the past 50 years, scholars have conducted substantial research on accurate group decisions and the improvement of forecasting tasks. The structure and design of group decision-making is critical for ensuring efficiency and participation stability, as it is a recurring challenge in traditional group meetings. Dalkey and Helmer (1963) were the first scholars to draw attention to the efficiency of group decisions by suggesting the DELPHI consensus building methodology (Linstone & Turoff, 1975; Rowe & Wright, 1999). The DELPHI method has become widely accepted and adopted for many different fields in which group consensus decisions are needed (Fischer, 1981; Parenté, Anderson, Myers, & O'Brien, 1984; Rowe & Wright, 1996). While these studies have focused on conventional group meetings, DELPHI has also been implemented for forecasting purposes, such as forecasting economic and political issues (Parenté et al., 1984), project selection problems (Yang & Hsieh, 2009), and the assessment of gas storage services (Bonacina, Cretì, & Sileo, 2009). The main attributes of DELPHI are its anonymity, iterative process, feedback process and consensus of group members via equal participation in outcomes (Rowe & Wright, 1999).

DELPHI is employed for forecasting purposes by utilizing expert judgments as one of the judgmental forecasting tools. Expert prediction has been suggested by many researchers for forecasting tasks (Goodwin & Wright, 1993; Lawrence, Edmundson, & O'Connor, 1985; Lawrence & O'Connor, 1992). Ex ante studies of expert prediction have attempted to average the individual interpretations. However, DELPHI consensus offers a unique improvement to expert groups via its iterative and anonymous process, which provides revisions of single forecasts according to overall intentions and prevents individual domination of the group.

The term 'consensus' is one of the critical concepts of DELPHI-like procedures. A consensus defines the degree of agreement on the intended decision task. In most cases, uncertainty is considered to be the opposite of consensus (Zarnowitz & Lambros, 1987). Uncertainty is a term that describes the dispersion of the outcomes. A DELPHI procedure strives to produce a collective group decision, but there may be significant differences in the participants' predictions. There may be no meaningful consensus if the standard deviation of the outcomes reflects a highly diffuse set of results. Despite the iterative structure of the DELPHI procedure, recurring sessions may not assure a consensus of the group. The manner by which to maintain a group consensus is one of the critical questions in studies of DELPHI-like systems.

In most cases, the variable that the group is attempting to predict may take a wide range of values, and it is a crucial task to predict data from a volatile background. Zarnowitz and Lambros (1987) presented an analysis of consensus and uncertainty, and one of their important results was that volatility and uncertainty are highly correlated. Consensus building can be a particularly difficult task for the facilitator when the data regarding the objective variable indicate a non-stationary pattern. Zadeh (1965) introduced fuzzy set theory, whose primary purpose is to reduce system uncertainty in several applications. Fuzzy logic techniques have

☆ An earlier version of this paper appeared as part of the ICCAE 2010 Conference Proceedings, published under IEEE copyright.

* Corresponding author. Tel.: +81 80 3830 0190; fax: +81 78 431 6259.
E-mail address: duruokan@yahoo.com (O. Duru).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.07.082

provided considerable improvements in the fields of automation and computer intelligence, which use many uncertain variables. Fuzzy set theory is an alternative for increasing predictive performance. It ensures consensus incentives and compiles expert expectations.

The implementation of fuzzy DELPHI methods (FDM) has been investigated by many researchers. Kaufmann and Gupta (1988) introduced an FDM procedure designed to define optimistic, moderate and pessimistic valuations of experts via triangular fuzzy numbers (TFN). In their technology forecasting study, Ishikawa et al. (1993) collected responses on a scale of expert judgments. TFNs use the selected scales as their grades of membership. The final crisp result is calculated by max–min FDM and fuzzy integration. An important conclusion of this study is that the FDM ensured consensus in only one round.

Chang, Huang, and Lin (2000) proposed a new FDM approach to assess managerial talent. This study used fuzzy statistics and ensured continuous, mathematically explicit membership functions. Chang et al. (2000) reviewed the previous FDM works and concluded that the FDM provides an opportunity to treat fuzziness arising from the forecasting objective and experts themselves. Fuzzy numbers provide a way to process groups of variables rather than the crisp estimates of the conventional DELPHI method.

The original contribution of the present paper is that it improves the FDM for financial time series forecasting and implements an algorithm for defining the final crisp adjustments of statistical benchmark methods. Fuzzy numbers are applied to incorporate the scale of adjustments. An expert group is required to adjust the statistical average forecast provided by linguistic terms that are linked to percentage changes. These adjustments are transferred to intended fuzzy sets. The crisp result is calculated as the centroid of the final average fuzzy set.

This paper is organized as follows: Section 2 reviews the data on the shipping freight market and detects the unit roots of seasonality in the BDI data. Section 3 presents the forecasting models and procedures. Section 4 examines the empirical results and compares the performance of several results. Section 5 concludes the present paper and provides recommendations for future studies.

2. The Baltic Dry Index and motivation for fuzzy-DELPHI forecasting

Shipping is an important method of world transportation, and most merchandise and industrial products are carried by commercial ships. The price of shipping service (freight rate) is an equilibrium price of the ship's owner–charterer negotiations and is listed in USD per metric ton or USD per day. Freight rates also have a considerable effect on final product prices. In the industry-retail market linkage, shipping margins play a critical role, and in some cases, they are the cause of large price fluctuations (Metaxas, 1971). The Baltic Dry Index (BDI) is a leading global economic indicator that traces prices for shipping bulk cargoes such as iron ore, coal and grains from producers such as Brazil, Australia and South Africa to markets in the US, Europe, Japan and China. The BDI is useful in that it offers a prognostic for determining where prices for raw materials shipped in dry bulk are rising and, therefore, provides a future look into inflationary or deflationary trends in the prices of goods ('The economists' magazine, October 16th, 2008).

The BDI index is a cumulative price of several standard contracts, including those of various dry bulk routes, ship sizes and types of contract (voyage domain or time domain charters). A panelist committee of the Baltic Exchange defines BDI levels for trading days based on the reported fixtures, or when there is no reported contract, the committee estimates the market.

The prediction of freight rates has many fields of interest, including transport planning, fixing the prices of finished goods, and financing ships; furthermore, it affects industries linked to shipping. Despite the value of freight forecasting, the task of prediction has some limitations and difficulties in practice in the shipping industry. Previous forecasting studies have suggested some univariate methods and causal models, but the recent behavior of shipping freights has shown highly volatile and sporadic fluctuations. It is subject to changes based on political and behavioral aspects whose prediction is purely a matter of judgment (Duru & Yoshida, 2008a; Duru & Yoshida, 2008b). Due to the lack of accurate statistical results, judgmental forecasting and its improved methods are considered suitable for revealing subjective factors. Duru and Yoshida (2009) first applied classical DELPHI for short term freight market forecasting (they also used BDI series) and DELPHI suggested particularly in analysis of sporadic conditions. Duru, Bulut, and Yoshida (2010) also applied fuzzy-DELPHI method for forecasting dry bulk freight market. In this method, an automatic forecasting model, X12 ARIMA, is used for statistical extrapolation. However, X12 ARIMA has limitations on modeling and in this method some processes are applied unnecessarily such as seasonal differencing.

Rather than conventional DELPHI, the proposed fuzzy-DELPHI method improves uncertainty on decision space and also combines advantages of statistical methods. Since experts are required to adjust statistical benchmark results, quality of statistical extrapolation is also somewhat important. Therefore, the present paper extends the work with the state-of-art methods of univariate time series modeling. Forecasting of levels of data is performed by Box–Jenkins type autoregressive integrated moving average (ARIMA) method and volatility of series is also investigated and predicted by generalized autoregressive conditional heteroscedasticity (GARCH) model. These methods will be discussed in dept on Section 3.

The proposed method contributes to the literature in a number of ways. First, this paper investigates accuracy of fuzzy-DELPHI consensus method for financial forecasting problem. Second, the proposed method does not ignore capabilities of statistical methods on extrapolation by recognizing historical patterns and combines them by expert consultation process. Finally, more complex methods of statistical extrapolation are used rather than previous studies.

In the present study, the fuzzy-DELPHI (FD) method was implemented for the BDI data, and predictions were calculated for 2009:11–2010:04 term (monthly average) by statistical methods and expert judgments. The empirical study was conducted in the middle of September 2009 with 11 experts, mostly from the shipbroking and managerial professions. The subjects were asked to define an adjustment to the statistical benchmark forecast.

The BDI raw data consists of monthly series from 1999:1 to 2009:10. Fig. 1 shows the BDI data history in the sample. The yellow part between 2009:01 and 2009:10 indicates testing period for statistical extrapolation and the green part between 2009:11 and 2010:04 indicates prediction period. Descriptive statistics of the raw BDI data and the first difference series of BDI data for the period from January 1999 to October 2009 are reported in Table 1. Normally, scholars expect that differencing operation increases normality of data, but results clearly indicate rejection of assumption. Jarque–Bera statistics show that both series distribute irregularly. Coefficient of variation for differenced series is excessively high which can not be illustrated by tangible variables. Fig. 1 and Table 1 expose that freight market has very high volatility and statistical prediction should contain processes for modeling volatility.

Table 2 presents results of unit root tests for levels and differenced BDI series. Tests are based on Augmented Dickey–Fuller (ADF) and Phillips–Perron testing procedures (Dickey & Fuller,