Modeling Stock Market Industrial Sectors as Dynamic Systems and Forecasting

Salim Lahmiri
ESCA School of Management, Morocco & University of Quebec at Montreal, Canada

INTRODUCTION

A great deal of research has focused upon the interdependence among international financial markets; particularly; in terms of how shocks or variations are transmitted across global stock markets (Narayan & Smyth, 2004; Fraser & Oyefeso, 2005; Cappiello, Engle, & Sheppard, 2006; Chen & Shen, 2007; Yu & Hassan, 2008; Ahlgren & Antell, 2010; Qiao, Li, & Wong, 2011) based on the assumption that the simultaneous dynamics of asset returns are well captured by a linear vector-autoregression (VAR) or a vector error-correction model (VECM) of Johansen (1988, 1991). The obtained empirical results indicated that the world stock markets have become more closely linked in recent years due to the development of global cooperation, financial market liberalization and the advances in information processing technology (Qiao, Li, & Wong, 2011). The VAR and VECM (Johansen, 1988, 1991) are popular statistical methods since they can systematically describe the dynamics among several variables, and were widely employed for multivariable systems modeling and forecasting in science (Chandra & Al-Deek, 2009; García-Ascanio & Maté, 2010; Wong & Ng, 2010; Xu & Moon, 2013) and economics (Jore, Mitchell, & Vahey, 2010; Clark, 2011; Das, Gupta & Kabundi, 2011; Korobilis, 2013).

The advantage of using VAR and VECM as statistical models is to account for the linear dynamics of the temporal causal relationship between stock markets that compose a system (Masih & Masih, 1999; Laopodis, 2011). In particular, they allow for examining both the short and long term dynamic causal linkages among variables of the system (Masih & Masih, 1999). Moreover, they are useful to capture long-run information often ignored in systems (Masih & Masih, 1999; Laopodis, 2011). For instance, VECM searches evidence of cointegration between variables in the system to find if they share common stochastic trend at long-run, or share deviations at short-run, or share both (Laopodis, 2011). The importance of a statistically significant long-run relationship among variables in the system lies in the existence of common stochastic trend, and; therefore; they tend to revert to such relationship after some short-run fluctuations (Laopodis, 2011). As a result, VECM provide long-run system parameters useful to ascertain the fundamental information content of the variables in the system (Tswei, 2013). In addition, if the short-run co-movements exist among stock markets that form the system under study; then, equity managers may design their portfolios to try to win the market (Chu, 2011). All these types of information regarding the relationships between stock markets that form the system could be used to predict each stock market future value.

BACKGROUND

Although linear statistical VAR and VECM have been widely used to model several stock markets as one dynamic system, they have several drawbacks. First, they are based on linearity and stationarity assumptions of the underlying data. Second, they are not robust to noisy data and are not adaptive. Therefore, as statistical linear systems, the VAR and VECM are not appropriate to describe the linkages between such noisy and nonstationary stock markets. Third, forecasts based on these linear systems could be inevitably imprecise. As a dynamic system, the linkages between different stock markets could be better modeled by artificial neural networks (Rumelhart, Hinton & Williams, 1986; Haykin, 2008) for forecasting purpose. Indeed, they are intelligent systems capable to model noisy and nonstationary variables; in particular, to approximate any continuous function up to certain accuracy (Cybenko, 1989; Funahashi, 1989).

DOI: 10.4018/978-1-4666-5888-2.ch376

Contrary to the prediction of financial time series with traditional models such as VAR and VECM that fit linear prediction models to track patterns in historical data, intelligent systems; also called machine learning; use a set of data samples to track patterns in order to approximate the underlying function that generates the data (de Oliveira et al., 2013). In the literature, different intelligent methods have been applied to make financial market predictions; including artificial neural networks (Kara et al., 2011; de Oliveira et al., 2013; Lahmiri et al., 2014a,b), support vector machines (Wen et al., 2010; Kara et al., 2011; de Oliveira et al., 2013; Lahmiri et al., 2014a,b), decision trees (Tsai & Chiou, 2009), evolutionary computation (Allen & Karjalainen, 1999; Korczak & Roger, 2002; Chiam et al., 2009), fuzzy logic (Chang & Liu, 2008; Chu et al., 2009; Zhang et al., 2009), and case-based reasoning (Goswami, Bhensdadia, & Ganatra, 2009), to name few.

In this study, the radial basis function neural network (RBFNN) (Broomhead & Lowe, 1988; Powell, 1987; Ghodsi & Schuurmans, 2003) is suggested as an alternative system to the conventional statistical linear VECM to simultaneously model linkages among stock markets and to forecast their future values.

There exit many artificial neural network architectures to design a system for multi-stock market predictions such as the well-known multi-layer perceptron or backpropagation neural network (BPNN) for short, recurrent neural networks, and Bayesian neural networks. The reader may refer to Haykin (2008) for an overview of various artificial neural network architectures. The RBFNN is chosen because of its ability to form a unifying link between function approximation, regularization, noisy interpolation, classification and density estimation (Ghodsi & Schuurmans, 2003). In addition, training radial basis function networks is usually faster than training the well known backpropagation neural network (BPNN) (Ghodsi & Schuurmans, 2003). The forecasting performances of the RBFNN and VECM are evaluated based on the root mean of squared errors (RMSE).

Previous works have focused on investigation of the empirical interdependence between stock markets both in developed and emergent financial markets (Narayan & Smyth, 2004; Fraser & Oyefeso, 2005; Cappiello, Engle, & Sheppard, 2006; Chen & Shen, 2007; Yu & Hassan, 2008; Ahlgren & Antell, 2010; Qiao, Li, & Wong, 2011). However, the design of complex systems to simultaneously model and predict industrial sectors within a given financial market remains an important issue for industrials and managers. Therefore, this article deals with the problem of industrial sectors modeling and forecasting within a single stock market. For comparison purpose, data from developed and developing countries are used to implement RBFNN and VECM and to conduct out-of-sample forecasting task. In particular, daily data from NASDAQ and Casablanca Stock Exchange are considered to perform our simulations. Since the RBFNN is capable to perform an exact interpolation of a set of data points in a high-dimensional space (Powell, 1987), it could be more appropriate than the conventional linear VECM to model such complex and noisy systems composed of numerous and different financial time series.

The remaining of this article is organized as follows. Next section presents the methods used to conduct our research; namely the statistical approach based on the VECM as a linear system, and the intelligent approach based on the RBFNN as a nonlinear system. Then, the data and results are presented and discussed; and, followed by a section that presents the future research directions. Finally, we conclude.

METHODS

Statistical Approach

Unit Root Test

Prior to estimating the VECM, unit root test is employed to ensure that prices of all industrial sectors are integrated to the same degree. In other words, it is needed to determine whether each industrial sector price time series are stationary or not; that is, if it contains a unit root. We evaluated the unit root hypothesis by employing the conventional Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979) which is a popular used methodology to examine the presence of stationarity in a given time series. This technique tests the null hypothesis of a unit root against the alternative of stationarity. The test will be performed with intercept (see equation 1) or intercept and trend (see equation 2) based on the time trend of the series. The ADF test equations are given as follows: