PREDICTABILITY OF COMMODITY FUTURES RETURNS ON EMERGING MARKETS: A NONLINEAR APPROACH

by

Dmytro Popov

A thesis submitted in partial fulfillment of
the requirements for the degree of

MA in Financial Economics

Kyiv School of Economics

2012

Thesis Supervisor: ______________ Professor Olesia Verchenko

Approved by ______________________
Head of the KSE Defense Committee, Professor Irwin Collier

____

____

____

Date ______________________

Kyiv School of Economics

Abstract

PREDICTABILITY OF COMMODITY
FUTURES RETURNS ON
EMERGING MARKETS: A
NONLINEAR APPROACH

By Dmytro Popov

Thesis Supervisor: Professor OlesiaVerchenko

The ability to predict the returns on financial instruments, either stocks, bonds or derivative contracts, carried away academics and practicing investors for decades. This research aims to assess the validity of EMH for commodity markets of emerging countries. We claim, that given the defining characteristics of emerging economies, additional forecast efficiency can be extracted through the use of non-linear techniques GARCH and Artificial Neural Networks. Our results show, that commodity futures returns are indeed predictable in short-term horizon, and the forecasting accuracy is satisfactory. This implies that the weak form of the EMH does not hold for emerging commodity markets, and additional benefits can be obtained through this inefficiency. We also describe the implications of such a result for hedgers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my girlfriend, who helped a lot with this research. I would never complete this work without her timely support and love.

Chapter 1

INTRODUCTION

The ability to predict the returns on financial instruments, either stocks, bonds or derivative contracts,interested academic and practicing investors for decades. Dozens of various financial and econometric approaches have been utilized to study this topic, but the issue of the returns' predictability is still under consideration.

The challenges connected with making consistent and preciseforecasts of the returns on financial instruments arise from two main groups of sources. The first is a complex nature of the market forces which determine the returns. An analyst who tries to forecast the future movement of the prices of the particular financial instrument (e.g. stock) should be aware that its dynamics is sensitive to changes in global business conditions, random circumstances and market sentiments.

The second source of difficulties in making good forecasts of returns on financial instruments is a changeability of observed market characteristics in response to the availability new forecasting methods.According to the efficient market hypothesis (EMH), any opportunity to consistentlymake profits is instantly discovered by numerous sofisticated market agents, and thus dissapers. In other words, the EMH states that no technique can consistantly beat the market, because after it becomes public, it is used by everyone, thus negating its potential gains.

The purposes of return forecasting vary across utilized instruments – to extract potential gains (for stocks and bonds) or tohedge possible risks (for derivatives). The information about the future returns on financial instrument might be used both in speculative and arbitrage trading strategies. An investor may seek for mispriced instrument and then use herforecast of the return to make potential gains out of this mispricing.

This study concentrates on the predictability of returns on futures contracts. There are several motivations for this research. The most basic motivation for assessing the predictability of futures returns is to assess the validity of EMH on the emerging commodity markets. Samuelson (1965) showed that under the assumption of risk-neutral agents and efficient market, the returns on futures are unpredictable. Therefore, the evaluation of the Samuelson hypothesis would provide arguments in favor (or against) the validity of the EMH for emerging futures markets.

Most of the classical studies of the market efficiency use statistical techniques, which rely on the assumption of the linearity[1] of the underlying data generating process (DGP). Namely, they presume that the information included in the series of past returns linearly affects the future ones. However, recent developments in financial econometrics indicate the non-linear nature of the DGP of financial time series for developed markets. Taking into account the defining characteristics of emerging financial markets, the non-linearity is likely to be even more apparent there. Consequently, statistical approaches, which rely on the linear structure of the underlying DGP's, are likely to produce much less precise results.

The second objective of this research comes from the fact that expected returns can be used by investors in formulating their hedging strategies. To see how it works, consider an investor with a long position on the spot market who wants to hedge her risks. The relative size of the offsetting short position in futures

---

[1] The comprehensive literature review is presented in the works of Fama(1970), (1991), (1998)

market, the hedge ratio, is conventionally determined (see Baillie and Meyers (1991) or Johnson (1960)) so that the total risk associated with the portfolio return is minimized. Although this approach is intuitively simple, it is based on the premises of normality and unpredictability (pure martingale nature) of the futures returns. If these assumptions hold, then this min-variance hedge ratio maximizes the utility of the investor[2].

However, recent studies show that the returns are likely to be distributed not normally, but exhibit fat-tails behavior. Some authors (for example – Cecchetti et al. (1988) or Hsin et al (1994)) show, that conventional min-variance hedge ratio in this case has to be modified to include investors expectation of futures return. The hedge ratio is then the sum of two components –pure hedging term (minimum-variance hedging ratio), and the speculative part, which is a function of the risk-aversion parameter, expected return and its conditional variance. In cases, when the investor is not infinitely risk averse, such a hedge ratio provides a possibility of trade-off between risk-minimization and return. Goyet et al. (2007) show, that in case of predictable futures returns, the inclusion of the expected return into the hedging ratio would have a significant influence on the hedging strategy of the investor.

Recent studies (Gandhi, Saadi, and Ngouhouo(2006), Lim, Brooks, and Hinich(2008)) show a clearly nonlinear nature of the data generating processes of series of stock returns on emerging markets.Although this does not necessary mean that the dynamics of the returns on futures contracts is also non-linear, but it creates sound motives for preliminary testing for linearity.

In this research, two following non-linear approaches are used – ARCH/GARCH and Artificial Neural Networks (ANNs). ARCH/GARCH

[2] The review of the futures hedge-ratios is given by Chen, Lee, and Shrestha(2003)

models are frequently used in the literature to model financial time series (e.g.Bera, Bubnys, and Park(1993)). The returns on financial instruments usually exhibit the pattern of volatility clustering – when sharp shocks are followed by the high variation in the series of returns, and vice versa. ARCH/GARCH models aim to eliminate the inefficiency of the traditional models that is due to violation of the homoscedasticity assumption by simultaneously modeling the dependencies in the series and the pattern of their conditional variance.

ARCH/GARCH models showed well results for developed financial markets (Miffre 2002). However, the nature of emerging financial markets may imply certain limitations for the use of ARCH/GARCH approach in assessing the predictability of futures returns. In particular, since ARCH/GARCH models are parametric specifications, they are likely to fail to capture highly irregular phenomena of futures returnson emerging financial markets (i.e. wild market fluctuations). This issue is likely to be resolved by the ANN approach.

The ANN is a semiparametric, computationally intensivetechnique, which theoretically can approximate any arbitrarily complex functional form of the underlying DGP (in practice much time and computer resources are required to model very intricate relations). Several studies (Dhamija and Bhalla 2010)have shown the superiority of neural networks over the conditional heteroscedasticity models in forecasting financial time series. Moreover, the performance of an ANN is not subject to any distributional or functional form assumptions, as in conventional econometric methods. The main disadvantage of the ANN approach is that it is a black-box type model. There are no clearly observable parameters, and consequently, the hypothesis testing is significantly limited.

The analysis of the predictability includes two basic goals to be reached. The first one, is to check whether series are essentially predictable. The judgment about the predictability is made upon the significance (joint and separate) of GARCH parameters. The second goal is to select a model which provides more robust and accurate forecasts. This is done through the comparison of the forecasts produced by GARCH and NNET methods.Since both approaches have their drawbacks and advantages, there is no prior anticipation which one would perform better.

Our data consist of the daily prices of futures contracts for the commodity markets of Argentina, China, India and the Republic of South Africa for the period from 02.01.2004 to 24.02.2012. These are the largest commodity exchanges among the emerging countries and represent very geographically diversified parts of the globe, thus capturing the differences in commodity-specific characteristics (cost of carry, supply shortages, transition costs etc.).

There are several studies of validity of the EMH for commodity markets of selected countries.Kaur and Rao(2010)examined the Indian commodity market for weak-form efficiency. They used autocorrelation and runs tests for prices of four types of agricultural commodities and found that markets are weakly-efficient. Phukubje and Moholwa(2006) studied South-African futures market with a random-walk model (basically, they assessed the predictability of returns through linear RW regression), and find no evidence of inefficiency. As noted earlier, these linear approaches may be inappropriate for the financial time-series.

The commodities chosen for the analysis are representatives of the agricultural sector: grains – including corn, barley and wheat, and oilseeds – represented by soybeans (further details in Data description). The data are taken from the free daily statistics of the correspondent commodity exchange and the datasets of IMF Primary Commodity Prices unit.

The remainder of the paper is structured in the following way: Chapter 2 gives a review of the literature; Chapter 3 outlines theoretical and empirical framework of the research; Chapter 4 provides the description of the analyzed dataset; Chapter 5 presents the estimation results. Conclusions and inferences are given in the Chapter 6.

C h a p t e r 2

LITERATURE REVIEW

The literature behind this research can be divided onto two broad categories: several works which provide a theory against futures' returns predictability and works which discuss the empirical evidence of it. The latter can be in turn subdivided onto those, which asses the predictability explicitly, and those which report it indirectly, as a consequence of other findings. Although the determinants of the returns vary across different classes of financial instruments, the main argument against return predictability stems from EMH. Therefore, the literature review is not limited to commodity futures only, but discusses also the predictability of returns on stocks and bonds.

The theoretical ground for the research is mainly provided by the works of Samuelson (1965), who showed that given the assumptions that agents are risk neutral, have well behaved common preferences and are rational, the prices of futures are unpredictable. This argument is based on the Law of Iterated Expectations and implies the fact that futures returns are unforeseeable given the market information. The rejection of this fact gives the possibility for hedgers to form their expectations about the returnson futures and therefore face a tradeoff between minimizing portfolio variance and maximizing the expected return.The mean-variance hedge ratio, which incorporates the expected return and the risk-aversion parameter, allows adjusting the hedging strategy for investors preferences.

This argument of non-predictability, though based on strong and elegant economic theory,seemes to have nothing to do with reality. For instance,Hirshleifer(1988) estimated the expected returns on commodity futures allowing them to vary across holdings of hedgers and report significant results. He shows that both systematic and commodity specific risks influence the expected futures prices. The findings of predictability are also supported by study of Ferson and Harvey (1991) who use multi-beta pricing model, with risk factors related to the stock market, unexpected inflation, consumer expenditures and interest rates based on cross-sectional data to confirm that the variation in equity returns is predictable.

Refering to the evidence of return predictability, Fama(1991) argues that it is important to establish the links between expected returns and business conditions. He suggests that linking the expected futures returns to macroeconomic variables may explain a lot. According to the results, the expected returns for bonds and stocks are lower, when economic conditions are strong and vice versa.

Several empirical studies report ambiguous results of the link between the investor sentiment and return on securities, particularly – Wang (2001) studied how trader-position-based sentiment index can be used for forecasting future prices in six major agricultural futures markets, and found, that large speculator sentiment predicts reverse price movements. On the contrary, Solt and Statman(1988) found no statistically significant relation between sentiment of investment newsletter writers and subsequent stock returns. Since the data on investor sentiments is not available for most of the emerging financial markets, this factor would not be used in the research.

In a similar line of research, Bessembinder and Chan (1992) report the predictability of futures' returns for different markets by basically using simple regression methods with various macroeconomic variables as predictors, such as inflation, term structure of interest rates etc. They show that statistically significant predictions can be made for the US commodity market. Their work then is extended by the study of De Roon, Nijman, and Veld(2000), who report that cross-market hedging pressures are important in explaining commodity futures returns. They build a model, which places limits on direct market participation and allow both hedging pressures and non-idiosyncratic risk to influence futures returns.

Miffre(2002) revealed that satisfactory forecasts can be done by the use of GARCH methods for Canadian wheat futures. The market efficiency is tested with a conditional multifactor model, which allows for shifts in diversifiable risk of the futures contracts. She found that on average 86% of the predictable variation of futures returns can be explained by conditional risk, and the latter 12% is relegated to the conditional residuals.

If return predictability is systematic Ferson et al. (2003) show that it can be explained either by market inefficiency or rational hedging responses to changes in intertemporal business environment. Additionally, these authors provide an empirical argument that most regressions in finance may contain spurious bias, which casts some suspicion on the use of simple regression analysis in financial research.

The article by DuongandKalev(2006) refers explicitly to the Samuelson Hypothesis (Samuelson 1965) and test it on intraday data for five futures exchanges (European LIFFE, Canadian WCE, Japanese TOCOM, US MGEX and Chinese DC) for the period of 1996-2003 and find, that the predictability of returns is limited for agricultural futures. The author also finds a support for an argument provided by Bessembinder et al. (1996) that the volatility of futures prices are more likely to increase as the contract approaches maturity in those markets, which exhibit a negative correlation between a spot price and changes in carry cost.

As it was mentioned above, the predictability of futures returns has direct implications for the investor's ability to hedge their risky positions – by altering a hedge ratio. The hedge ratio, or the relative position in futures markets, is conventionally determined independently of expected returns, and is calculated so that it minimizes risk associated with portfolio returns (min-variance (MV) approach). This approach is inconsistent with mean-variance framework, because it completely ignores preferences of the agents or assumes that they are infinitely risk averse. The reason why it is actually used is that it provides extremely easy formula of the hedge ratio, which does not require any sophisticated estimation techniques.

The strategies which incorporate both the expected return and variance of hedged portfolio have been proposed by several authors (for example – Cecchetti et al. (1988) or Hsin et al (1994)). These strategies are consistent with mean-variance framework, but are contingent on the existence of a precise forecast of futures returns, and consequently – on returns predictability.

Goyet et al. (2007) investigate the predictability of futures using semiparametric approach, with an assumption, that expected returns on futures depend nonparametrically on a combination of predictors. They first collapse the independent variables (lagged values of future returns and lagged spot returns) into a single-index variable with weights determined by average derivative estimator, proposed by Stoker (1986). The results show that the calculated indices of four commodity futures studied in the paper contain statistically significant information concerning the expected returns on agricultural futures. Additionally, the authors prove that the modification of the conventional min-variance hedge ratio with the inclusion of expected futures returns makes significant difference to investors hedging strategies.The main drawback of such an approach lies in the choice of smoothing parameters, both for ADE and for kernel estimator. Essentially, the output depends on stringent assumptions on smoothness of the underlying density for the ADE and on the choice of optimal smoothing factor for the kernel itself. This significantly reduces the accuracy of the produced forecasts.

A recent paper byTimmermann published a paper (2008) discusses the issue of return predictability for stocks. He uses the series of the US stock returns and proposed adaptive forecast combination approach. His conclusion was that although the return predictability is possible for short-term horizons, the relatively weak degree of predictability even during such periods makes predicting returns an extraordinarily challenging task. These findings were supported by the recent study by KonstantinidiaandSkiadopoulos(2011) who do not find any strong evidence of statistically predictable patterns in the evolution of volatility of futures prices (VIX).

Brief description of the overviewed literature is presented in Table 1.

Table 1. Brief description of the main studies on the returns' predictability

Year	Author(s)	Market(s)	Methodology	Result
2011	E. Konstantinidia and G. Skiadopoulos	US	ARIMA, ARMA, VAR, PCA	Unpredictable
2008	Allan Timmermann	US	adaptive forecast combination approach	Unpredictable
2007	Goyet et al.	US	semiparametric (Kernel regression with ADE estimators)	Predictable
2006	Duong and Kalev	Five Developed Exchanges	regression (intraday data)	Limited predictability
2002	Miffre	Canada	GARCH	Predictable
2001	Wang	US	regression (sentiment index)	Predictable
1992	Bessembinder and Chan	US, Canada	regression (macroeconomic factors)	Predictable for US

Year	Author(s)	Market(s)	Methodology	Result
1991	Ferson and Harvey	US	regression (CAPM condt.)	Predictable
1991	Fama	US	regression (business conditions)	Predictable
1988	Hirshleifer	US	regression (CAPM)	Predictable
1988	Solt and Statman	US	regression (newspaper sentiments)	Cannot reject unpredictability
1965	P. Samuelson	-	theory	Unpredictable

Summing up, there is strong theoretical argument, that futures returns are unforeseeable given market efficiency, suggesting that the hedge ratio based on MV framework is optimal. At the same time, numerous empirical studies report that the predictions of futures returns are significant. The evidence of futures returns predictability would lead to considerable corrections of investors hedging strategies.

Most of the literature deal with the predictability of returns on mature financial markets, with relatively stable economic relationships and long data horizons. This research aims to assess the predictability of futures' returns on emerging financial markets, which are subject to wild market shocks, sudden losses of liquidity and other irregular phenomena. This implies that one would expect non-trivial, highly non-linear nature of financial time series there, and consequently justifies the use of sophisticated non-linear estimation techniques. The set of dependent and independent variables is specified equivalently to the definition by Goyet et al. (2007), except series of returns on the commodity spot market. The reason for this is that spot prices of the underlying commodities are not available on emerging financial markets. It is proposed, that the returns on Commodity Research Bureau spot grains index (CRB (Grains)) would be a good proxy for spot returns on commodities. It covers the prices of grains and oils from various commodity exchanges including emerging ones.

Chapter 3

METHODOLOGY

As for methodology, in this studytwo approaches are used – artificial neural networks, usually classified as a semiparametric estimation method, and conditional heteroscedasticity estimation techniques (ARCH/GARCH).

Conditional heteroscedasticity models are frequently used to predict time series of financial returns, since they take into account the excess kurtosis (fat tail behavior) and volatility clustering, which are frequently present in financial time-series.

The ANN approach serves to be an alternative to ARCH/GARCH models, because the latterhave limited power in explanation of highly unanticipated events that can lead to significant structural changes. The dominance of either approaches is not strictly stated in literature, but in some applications neural networks proved to be better (Dhamija and Bhalla 2010). Moreover, the ANN procedure does not rely on any distributional assumption, as most of the nonparametric techniques (such as Kernel regression). Neural network, with its ability to find complex mappings between inputs and outputs, is expected to reveal if there is any predictable variation in series of futures returns.

The tested hypothesis comes from the Samuelson (1965)hypothesis (null of unpredictability), mentioned above, and may be formulated as follows:

$$H_{0,x}: E\ [\Delta f_t | x_{t-1}] = 0, \qquad (1)$$

where, f_t is the natural logarithm of the futures' pricefor the selected commodity and $x_{t-l} = (x_{1,t-1}, \dots, x_{q,t-l})'$ is a vector of explanatory variables (the standard set of factors consisting of lagged futures returns Δf_{t-l} and exogenous factors).The essence of the stated hypothesis comes from the martingale statement of Samuelson (1965), which implies that changes in futures prices of a financial instrument from $t-1$ to t are orthogonal to the information available at $t-1$ (or equivalently – unpredictable).

Since GARCH is a parametric specification, the significance of parameters can serve as a tool for testing of the null. In this set-up, the forecasts of ANN approach serve as an alternative model specification, which is expected to improve forecasting accuracy.

3.1 Preliminary nonlinearitytests

The use of nonlinear techniques is justified by the fact that they can be applied to processes with non-linear nature and may find complex relationships between observed variables. Therefore, the nonlinearity of the studied process is a preliminary condition for usage of the nonlinear modeling techniques. This suggests the use of testing procedures, which would indicate the nonlinear nature of the input series.

The Bispectrum(Brickett, Hinich, and Patterson 1988) test for nonlinearity utilizes the fact that properly normalized bispectrum of the time series is constant over all frequencies (zero under normality).

The bispectrum of a time series is the Fourier transform of its third-order moments. Let x_t bea stationary stochastic time-series of the form

$$x_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \qquad (2)$$

where $\{\epsilon_t\}$ is $iid(0,\sigma_\epsilon^2)$, μ is a constant, ψ_i are real numbers and $\sum_{i=0}^{\infty} [|\psi]_i| < \infty]$. Then the third moment of x_t is defined as

$$c(u,v) = g \sum_{k=-\infty}^{\infty} \psi_k \psi_{k+u} \psi_{k+v}, \qquad (3)$$

where u and v are integers, $g = E(\epsilon_i^3)$, $\psi_0 = 1$ and $\psi_k = 0$ for $k < 0$. The Fourier transforms of (3) are

$$b_3(\omega_1,\omega_2) = \frac{g}{4\pi^2} \Gamma[-(\omega_1+\omega_2)] \Gamma(\omega_1) \Gamma(\omega_2), \qquad (4)$$

where $\Gamma(\omega) = \sum_{u=0}^{\infty} \psi_u \exp(-i\omega u)$ with $i = \sqrt{-1}$, and ω_i are frequencies. Let the spectral density function of x_t is given by

$$p(\omega) = \frac{\sigma_\epsilon^2}{2\pi} |\Gamma(\omega)|^2, \qquad (5)$$

Therefore, thefunction (6) has to be constant

$$b(\omega_1,\omega_2) = \frac{|b_3(\omega_1,\omega_2)|^2}{p(\omega_1)p(\omega_2)p(\omega_1+\omega_2)} = \frac{g^2}{\sigma_\epsilon^6} = \Gamma^2[(\omega_1,\omega]_2), \qquad (6)$$

Hence, if the process $\{x_t\}$ is linear, then the Fischer's skewness function $\Gamma[(\omega_1,\omega]_2)$is constant for all frequency pairs (ω_1,ω_2). The absence of the unit root in the series under study is a necessary requirement for the validity of the test.

The test uses the estimates of (6) over a suitably chosen grid of points and applies Hotelling'sT^2 statistics to check the H_0 of zero bispectrum (for Gaussian series). If the test shows that series are not Gaussian, the linearity test is required. For this purpose, the empirical distribution of (6) is compared to the theoretical one – the linearity of the series under analysis suggests that these distributions are statistically equal. In practice, the theoretical interquartile range of (6) is compared to the empirical one with a conventional z-test, under the null that they are equal.

However, the linearity of time series reported by the test does not necessary constrain the estimation techniques to only linear ones. There are several reasons for this – main of which is that, the test has a weak power in detecting non-linearities in the variance (ARCH effects), thus additional tests for ARCH/GARCH are required.

Another non-linearity test was built by McLeod and Li(1983)and is used to exactly test for ARCH effects in the data, and therefore is a necessary condition for using conditional heteroscedasticity models.

The test is applied to the squared residuals of an ARMA(p,q) model, to check for model adequacy. The test statistic is as following:

$$Q(m) = T(T+2)\sum_{i=1}^{m}\frac{\hat{\rho}_i^2(\varepsilon_t^2)}{T-i}, \qquad (7)$$

where T is a sample size, m is a properly chosen number of autocorrelations used in the test, ε_t – the residual series, and $\hat{\rho}_i^2(\varepsilon_t^2)$ is a lag-i ACF of ϵ_t^2.

For adequate underlying model, $Q(m)$ is asymptotically chi-squared distributed with $m-p-q$ degrees of freedom. The null hypothesis of the statistics is $H_0: \beta_1 = \cdots = \beta_m = 0$, where β_i is the coefficient of ϵ_{t-i}^2 in the linear regression

$$\varepsilon_t^2 = \beta_0 + \beta_1\epsilon_{t-1}^2 + \cdots + \beta_m\varepsilon_{t-m}^2 + v_t, \qquad (8)$$

The Ljung-Box test is particularly sensitive to the choice of the lag order m. Following the approach proposed by Tsay(2005), the lag order of the test equals to the $log(T)$, where T is a sample size. Based on the outcome of the test and graphical representation of the autocorrelation function for squared residuals of an ARMA model the presence of the ARCH effects is ascertained, thus constraining the set of possible techniques to non-linear ones.

3.2 ARCH/GARCH model

Recent studies have shown that most of the financial time series do not fit to the assumptions of the simple regression analysis. Particularly, the homoscedasticity assumption is frequently violated. Moreover, the variance of the financial series usually exhibits a clustering volatility pattern – that the large shock in date is followed by large variance of the process and vice versa. Campbell, Lo, and MacKinlay(1997)argued that it is both logically inconsistent and statistically inefficient to use statistical models that are based on the assumption of constant volatility over some period when the actual series of squared residuals changes through time.

The notion of fat tails or leptokurtic feature of financial series as well as clustering volatility is partially cured by the use of the ARCH/GARCH processes. The setup of the model starts from the following regression equation

$$\Delta f_t = \beta_0 + \beta_1 \Delta f_{t-1} + \cdots + \beta_p \Delta f_{t-p} + \beta_{p+1} \Delta s_{t-1} + \cdots + \beta_{p+n} \Delta s_{t-n} + \varepsilon_t, \quad (9)$$

This regression equation is then tested for ARCH effects with the test of (7), and if the test reveals positive results (the null of adequate model is rejected) then the autoregressive pattern for conditional variance of the residuals of (9) is assumed:

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 h_{t-1}^2 + \cdots + \beta_p h_{t-p}^2, \quad (10)$$

where h_t^2 is a conditional variance of the process (9). Simultaneous estimation of (9) and (10) constitutes a GARCH(p,q) process proposed by Bollerslev (1986). The estimation of (9) can be done with the use of simple OLS, but in this caseestimates would be inefficient. Therefore, the maximum-likelihood method is used, assuming conditionally normal distribution. Since the additional constraints significantly increase the computational time, it is impossible to ensure stability of (10). Consequently, it is a common approach not to take (p,q) higher than (2,2).

In practice, the variance model described by (10) has several serious drawbacks. First of all, it equivalently treats the positive and negative shocks to the variance of the residuals, thus neglecting so-called leverage effect (when the variance of the series responds differently to positive and negative shocks). This is captured by the Threshold ARCH (TARCH) model (Zakoian 1991), where (10) is changed to:

$$h_t^2 = \alpha_0 + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 d_{t-1} + \beta_1 h_{t-1}^2, \qquad (11)$$

where, $d_t = \begin{cases} 1, & if\ \varepsilon_t < 0 \\ 0, & otherwise \end{cases}$, is a leverage parameter.

The other drawback of the conventional ARCH model is that it allows the variance to take negative values, which does not have any statistical meaning. In order to fix this issue, the EGARCH model was developed (Nelson 1991), which through logarithmic specification takes into account the leverage effect and constrains variance to be positive.

$$\ln(h_t^2) = \alpha_0 + \alpha\left|\frac{\varepsilon_{t-1}^2}{h_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}^2}{h_{t-1}} + \beta[\ln(h]_{t-1}^2), \qquad (12)$$

In this research, all three models are considered. The choice between models is based upon the information criteria (Akaike and Shwartz).

3.3 Artificial neural network approach (ANN)

A neural network is an adaptive system that changes its structure as it learns from endogenous and exogenous inputs. ANNs are very powerful tool in contemporary financial analysis, since they are able to detect the underlying functional form of dependencies within a dataset without any prior knowledge about this relationships.

The neural network estimation involves the following steps:

1. Network creation – building network topology, defining number of hidden layers etc.

The structure of the ANN is usually expressed as a multi-layer perceptron. The network consists of several layers of processing units (called neurons or nodes). The input values are first processed within a neurons of the input layer, and the output values of this neurons are then fed to the neurons of the hidden layer. Each connection between layers has corresponding parameter, which indicates the strength of this connection – the so-called weight. Changing the weights in a specific manner constitutes the learning of the network, or basically mapping patterns in the input layer to target values of the output layer.

In this study, the closed-loop, diffusion (out-sample predictions can be made based on previous lagged values) NARX (Nonlinear Autoregressive with Exogenous input) algorithm proposed by Hopfield (1982) is used:

$$\Delta f(t) = g\big(\Delta f(t-1), \dots, \Delta f(t-n), \dots, \Delta s(t-1), \dots, \Delta s(t-n)\big), \qquad (13)$$

where, Δs represents the change in log spot prices of the underlying commodity. Since for many markets, the spot price data are not available, the Commodity Spot Index may be used as a proxy. The orders of the lags are chosen through the process of testing alternative model specifications so that the autocorrelation in residuals are eliminated..

The topology of the neural network used in the research (Figure 1) consists of one basic input – lagged futures' returns (the order of the lag n is chosen so that the autocorrelations in residuals is removed.

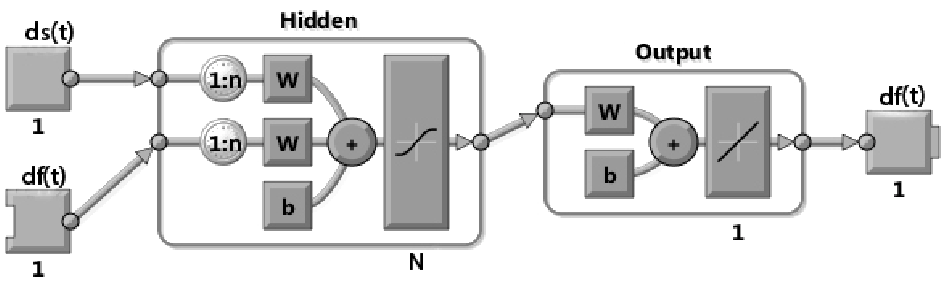

Figure 1. The topology of the artificial neural network

The network includes one hidden layer, which consists of hidden nodes (neurons) each defined as a function of inputs:

$$h_j = f_j\left(b_{0j} + \sum_{i\to j} w_{ij}x_i\right), \qquad (14)$$

where x_i is the value of the i^{th} input node, w_{ij} – weight of the i^{th} input, b_j – bias, $f_i(\cdot)$ – activation function.

The number of hidden neurons N is often a debatable question - setting too few hidden nodes results in high training errors and high generalization errors (the latter arise from under-fitting). From the other hand, setting too many hidden neurons leads to low training errors but still too high generalization error (due to over-fitting). The rule of thumb used in this study, developed by Wanas et al.

(1998), is to take number of hidden neurons equal to $log(T) + 2$, where T is a size of the training sample. According to these authors, this number produces the lowest level of entropy in the resulted series.

2. Configuring network – arranging the network so that it is compatible with the proposed problem.

In practice, the functioning of the neural network consists of three stages: on the first stage, the input x_i is multiplied by the scalar weight w_{ij} and forms a product (again a scalar) $x_i w_{ij}$, then the scalar bias b_j is added to form net input (the bias is simply shifting the transition function to the left – it is just like a weight which has constant input of 1), and afterwards net input passes through the transfer function f_i and forms an output to the next layer. For the purposes of this research, the weight function is just a summation of weighed inputs and their biases. The parameters w_{ij} and b_j, are adjustable, so that the network may exhibit some desired properties (in this study – the minimum of MSE – mean squared error).

The transfer function appears at two stages of the built neural network – in hidden and output layers. The activation function which is commonly used in the hidden layer is a sigmoid function (12).

$$f_i = \frac{1}{1 + e^{-z_j}}, \qquad (15)$$

The main reasons for choosing this functional form are that it is continuously differentiabile (the differentiability property is desirable for network learning), and that its first derivative is bell-shaped (corresponds to the most common type of non-linearity observed in financial data). The output function is usually linear.

3. Training the network – tuning the adjustable network parameters (weights and biases), to optimize network performance.

The original dataset is subdivided onto three patches: the Training set (the largest one – usually 70% of the data) – used for computing gradient and adjusting network weights and biases; the Validation set – the set used to halt the training of the network and the Testing set – used mostly for comparison of different models.

The training of the network involves backpropagation algorithm: first, reproduction of a trained network, in order to get network output; then obtain the resulting MSE and adjust the parameters of the network in order to reach a global minima of the network error. In this study we will use the Levenberg-Marquadt which is a generalization of Gauss-Newton algorithm and gradient descent, and produces the fastest output (the iterative computations of the network may take long time until the desired network performance (minimum error) is reached).

4. Validating the results –measuring network generalization and halting the training when the generalization stops improving.

The purpose of the validation set is to minimize overfitting of the network. In other words, the training of the network is performed only while the increase of network accuracy on the training set increases the accuracy of prediction on the validation set. If the additional training iteration still improves the network performance on the training set but does not increase the accuracy of prediction on validation set, then the training process should be stopped, to avoid overfitting.

The main shortcoming of the ANN approach, similarly to any other semiparametric technique, is that it is difficult to interpret obtained results (Basically the parameters of the network – weights and biases have no meaning in terms of interpretation). The argument against such kind of critique comes from the fact that assessing the predictability of the futures' returns does not require any rigorous interpretation or disclosure of dependencies. Moreover, the functional form of the relationship, even if revealed, may be too complex to interpret.

3.4Test of predictability and assessment of model accuracy

The test of the null of (1) can be reformulated in terms of used models:

$$\Delta f_t = \beta_0 + \beta * \mu(\Delta f_{t-1}, \Delta s_{t-1}), \quad (16)$$

where μ is a function of input variables available at $t-1$ and described by either model (GARCH or ANN). The test of the null is then equivalent to test of the joint significance of coefficients β_0, β for the correspondent model.

For GARCH models this can be done by Wald test of joint significance, described by the equation:

$$W = \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})}, \quad (17)$$

where, $\hat{\theta}$ are the parameters of unrestricted model, and θ_0 – of restricted. The test statistic is distributed Chi-squared, with degrees of freedom equal to the difference in number of parameters in $\hat{\theta}$ and θ_0. Unfortunately, due to the nature of the model, Wald test is not available for ANN approach. Therefore, the predictability of series of futures returns is judged primarily on the significance of GARHC coefficients.

The estimation procedure is divided on three steps – the preliminary analysis of the series, specification of the models based on in-sample fit and analysis of GARCH coefficients and the assessment of forecasting power based upon one-step-ahead rolling-window forecasts. The latter is done in the following way:

1. The estimation window is defined as a first $Win = T - 100$ observations, where T is a sample size;
2. The model is evaluated for the window, and the on-step ahead forecast is produced;
3. The window is rolled one step ahead;
4. Steps 2 and 3 are repeated iteratively, so that the 100 points of forecasted are obtained.

The accuracy of forecasts is evaluated based on mean absolute deviation, mean squared error and root mean squared error coefficients. In addition, the Theil coefficient is estimated for both of the models.

$$U(h) = \sqrt{\frac{\sum_{j=0}^{m-h}[fe_{t+j}(h)]^2}{\sum_{j=0}^{m-h}[[\Delta f_{t+h+j} -]\Delta f_{t+j}]^2}}, \quad (16)$$

where, fe_t is a forecast error of the model. Theil coefficientis used to compare the fit of the given model to the one of the RW specification. The closer the value to zero, the better the fit of the model compared to naïve prediction.

The comparison of the models, in terms of forecasting power, is done through Diebold-Mariano test. The test statistic is

$$DM = \frac{\bar{d}}{\left[\frac{2\pi \hat{f}_d(0)}{T}\right]^{\frac{1}{2}}}, \qquad (17)$$

where, $\bar{d} = \frac{1}{T}\sum_{t=1}^{T}(fe_{1t} - fe_{2t})$ is a sample mean loss differential between forecasts 1 and 2 (given by MSE or MAD) and $\hat{f}_d(0)$ is a spectral density of loss differential $d_t = fe_{1t} - fe_{2t}$. The spectral density is defined as a Bartlett Kernel, which guarantees the costiveness of the long-term variance:

$$\hat{f}_d(x) = \begin{cases} 1 - |x|, for\ |x| \leq 1 \\ 0 \qquad , otherwise \end{cases}, \qquad (18)$$

The null hypothesis that $E(d_t) = 0$ for all t is rejected if the value of the test exceeds the critical value of a standard unit Gaussian distribution.

The main advantages of the test is that it is model-dree, applicable to multiperiod forecast with non-Gaussian, non-zero mean, serially correlated and contemporaneously correlated forecast errors.

C h a p t e r 4

DATA DESCRIPTION

The dataset used in the analysis is composed of the daily closing prices of futures contracts of four emerging economies – Argentina, China, India and Republic of South Africa. The data comes from include free sources (national commodity exchanges[3]), and the datasets of IMF Primary Commodity Prices[4](for those exchanges which provide data partially, or which are only in local language).The periods covered mainly depend on the starting date of operation of selected exchange. The longest period, covered by futures prices is for South-African futures contracts and is from 02.01.2004 to 24.02.2012.

The main reason for selection of exactly these four countries is the data availability. Most of other organized commodity markets (including the Russian one) have rather low liquidity. The four selected four economies represent geographically distant parts of the world, thus capturing the differences in commodity-specific characteristics (such as cost of carry, supply shortages, transition costs etc.). Moreover, none of the selected commodity exchanges is subject to any strict regulations or any other irregular impacts.

The commodities chosen for the analysis belong to the agricultural sector: grains – such as corn, barley and wheat and oilseeds – such assoybeans. The main reason, why these futures contracts were chosen, is that theyall have the same maturity – 2months, and that agricultural commodities are frequently subject to exogenous impacts (e.g. weather conditions) which may impact the futures price.

[3]http://english.czce.com.cn; http://www.dce.com.cn; http://www.bseindia.com; http://www.ncdex.com; http://www.safex.co.za; http://www.rofex.com.ar.

[4]http://www.imf.org/external/np/res/commod/index.aspx

The returns are used in logarithmic form - $f_t = \ln\left(\frac{p_t}{p_{t-1}}\right)$, because this transformation allows to avoid asymmetric treatment of upward and downward price movements. Following the conventional practice, the contracts are chosen in such a way that, when the previous contract reaches maturity in month m (number of the month), it is replaced by the next expiring contract of the last day of month $m-1$. This helps to avoid thin trading and expiration effects. To avoid the effects of differences in the exchange rates, all contracts are priced in USD per metric ton. The descriptive statistics of the series of daily futures returns is presented in the Table 2.

Table 2. Descriptive statistics for the daily futures' returns on selected commodities

Symbol	Name	# obs	Mean	Std. Dev.	Skewness	Kurtosis	P(S+K)*
CRB_G	CRB Futures Index (Grains)	2125	0,03%	1,68%	-0,12	4,74	0.00
GE	Corn (Argentina)	592	0,04%	1,40%	0,24	6,07	0.00
GK	Soybeans (Argentina)	1095	0,02%	1,87%	0,57	28,96	0.00
WR	Wheat -Winter (China)	1774	0,04%	0,80%	3,95	44,45	0.00
XV	Corn (China)	1937	0,05%	0,83%	3,40	36,30	0.00
OD	Corn - Feed (India)	938	-0,01%	2,84%	0,09	33,35	0.00
Z1	Barley (India)	1359	0,03%	1,83%	-3,45	71,85	0.00
OF	Barley - Feed (India)	933	0,02%	2,11%	-0,65	13,87	0.00
HX	Wheat (South Africa)	2125	0,02%	1,64%	-0,70	8,15	0.00
HV	Corn (South Africa)	2125	0,03%	2,03%	-0,06	4,62	0.00
HZ	Soya (South Africa)	2125	0,01%	1,86%	-0,58	19,02	0.00

*the p-value for Kurtosis and Skewness jointly equal to the values of Normal distribution

The number of observations ranges from 933 for Indian Barley to 2125 of South African wheat, corn and soybeans. Relatively low number of observations for some of the contracts places significant doubts on the validity of the output of neural network, since it requires big data samples. Theoretically, low data sample should expand the confidence interval of the predicted series, thus leading to insignificant predictions.

The commodity index of the Commodity Research Bureau (Thomson Reuters) serves as a proxy for commodity spot prices. This index represents the price of the basket of grains including oilseeds. The logarithmic returns of this index are used as an exogenous input in the chosen estimation models.

The mean values of the returnsare positive, meaning that selected futures contracts were profitable on average. The values of Skewness and Kurtosis are far different from the Normal distribution (zero and three) and this is also supported by the p-value of their joint normality test – the null hypothesis, that series are normal is rejected. This supports the leptokurtic behavior of the financial series, which is frequently observed in the data, and brings another argument in favor of nonlinear estimation techniques.

The prediction horizon considered in this paper is one day – the transformation of futures prices into weekly data (the literature usually considers weekly and daily data) would leave too little datasets. Assuming that a hedger rebalances her hedging position every day (which is frequently observed in practice), the one-day horizon seems to be appropriate.

Chapter 5

EMPIRICAL RESULTS

This chapter outlines the estimation of non-linear forecasts of the commodity futures returns. The estimation procedure involves several steps to be undertaken. First, data has to be pre-processed – stationarity and nonlinearity tests have to be carried-out, preliminary structures of the models have to be determined, the data has to be properly normalized (for ANN estimation). Then the estimation itself is carried out – this includes the set up and pruning of the models, discussion of the adequacy of obtained results and the null (1) of predictability is evaluated.Finally, the forecasting power of the models is examined – several accuracy tests are undertaken, the forecasts are compared to themselves and to naïve models.

5.1 Preliminary tests

As noted earlier, GARCH and ANN estimation routines presume the nonlinearity of the series under study. This justifies the importance of the preliminary linearity tests.

The non-linearity tests applied in this research require several conditions to be met. The preliminaryrequirement of the Hinich test is the absence of the unit root in the characteristic equations of the processes under consideration. Additionally, unit root test (here – Augmented Dickey-Fuller test) is required for GARCH estimation, since the order of integration has to be correctly specified. Brief output of the ADF test is presented in the Table 3 (more rigorous test results are in Table A.1). The graphical analysis of the studied series does not give any reasons for the inclusion of drift or trend into ADF regression, so conventional ADF test is used. The results show that no unit root is present (all p-values are less than 0.05) hence the original series of returns may be utilized in further analysis.

First, we apply the Hinich non-linearity test. The test should reveal whether the data is generated by non-linear DGP, or equivalently, whether the application of non-linear estimation techniques is justified. The output of the Hinich test is presented in the following table.

Table 3. The output of the Hinichbispectrum test for Normality and Linearity

Symbol	Name	ADF p-value	Gaussianity		Linearity	
			Chi2	P-value (Chi2)	Z	P-value (Z)
GE*	Corn (Argentina)	0,0000	139,13	0,0000	-1,16	0,8777
GK	Soybeans (Argentina)	0,0000	984,08	0,0000	23,62	0,0000
WR	Wheat -Winter (China)	0,0000	4542,29	0,0000	8,43	0,0000
XV	Corn (China)	0,0000	3973,82	0,0000	18,77	0,0000
OD	Corn - Feed (India)	0,0000	522,66	0,0000	3,60	0,0002
Z1	Barley (India)	0,0000	2207,85	0,0000	19,96	0,0000
OF	Barley - Feed (India)	0,0000	350,84	0,0000	7,52	0,0000
HX*	Wheat (South Africa)	0,0000	982,55	0,0000	-1,40	0,9195
HV*	Corn (South Africa)	0,0000	440,57	0,0000	-4,77	1,0000
HZ	Soya (South Africa)	0,0000	1049,58	0,0000	6,71	0,0000

* the null of linearity cannot be rejected

Under the assumption of Normality (Gaussianity) the test statistics is χ^2 distributed with 2 degrees of freedom. According to the output of the test (Table 3) none of the studied series is normally distributed, whichconfirms revious results for the values of Kurtosis and Skewness. Unexpectedly, the test for linearity cannot reject the null for some of the series – namely Corn (Argentina), Corn and Wheat (South Africa), which may imply that the series are generated by the linear underlying DGP. However, as noted earlier, the Hinich test has limited power in detecting ARCH effects in the data. Hence, the Ljung-Box test can reveal additional information about the linearity of the studied returns.

In addition, the correct structure of the model has to be specified. While the lag structure of the ANN models are chosen in the process of estimation (through minimization of information criteria holding the absence of autocorrelations in residuals), part of the GARCH model specification can be done before the GARCH estimation.

According to the methodology of Ljung-Box test, the linear model, which removes autocorrelation in residuals, has to be specified. Table 4 contains information on correspondent number of lags of exogenous input (spot price index CRB) and ARMA(r,m) lags for every equation. The lag order was chosen in such a way, that it maintain high (above 0.05) p-value for Q-test for residuals while meeting minimum values for information criteria (Akaike (AIC) and Shwartz (BIC) criteria are reported in Table 4). The null hypothesis for the test is no autocorrelation in squared residuals of these properly specified equations.

Table 4. Ljung-Box test for ARCH effects

Symbol	**Lag(s) of CRB index**	**ARMA (r,m)**		**AIC**	**BIC**	**p(Q) for e**	**p(Q) for e^2**
		AR(r)	**MA(m)**				
GE	1	1	1	-3362,762	-3340,853	0,7766	0,0000
GK	1, 2	1	1	-5605,151	-5580,163	0,5285	0,0000
WR*	1	0	1	-12088,8	-12066,88	0,5415	0,9682
XV*	1, 2	2	0	-13118,58	-13085,17	0,421	0,9705
OD	1	4	1	-4046,445	-4007,704	0,158	0,0000
Z1*	1	2	2	-7028,705	-6992,208	0,2169	1,0000
OF	1	1	0	-4565,75	-4546,401	0,3037	0,0005
HX	1	1	1	-11539,96	-11511,65	0,1159	0,0000
HV	1	2	0	-10667,46	-10639,16	0,2023	0,0000
HZ	1	2	2	-10995,68	-10956,06	0,7295	0,0000

* the null of no autocorrelation in squared residuals cannot be rejected (no ARCH effects)

The output of the test shows, that for seven out of ten series of returns ARCH effects are present in the data. Namely, the p-values of the test for Winter Wheat (China), Corn (China) and Barley (India) are close to one, meaning that there are no significant ARCH effects.

What is interesting is that the linear results of the two tests do not overlap. This means that the series, which do not contain ARCH effects according to Ljung-Box test, are still non-linear in nature according to the Hinich test. This in turn implies benefits of using either non-linear methodology (ANN or GARCH).

5.2 GARCH approach

Estimation of the GARCH-type models includes the specification of the equation, which describes the evolution of the variance. It may be a linear autoregressive process (conventional GARCH) or it can be modified to fit the purposes of the research. We consider three types of the generalized ARCH models – GARCH, EGARCH and TGARCH. The latter two allow to account for so-called leverage effect – when series of returns respond differently to positive and negative shocks.

Although these ARCH specifications are extremely popular in modeling returns, there are no formal tests to decide which one is better.Therefore, we use Akaike (AIC) and Shwartz (BIC) information criteria to measure their relative goodness of fit (Table 5).

Table 5. Choice between (G)ARCH, EGARCH and TARCH models based on AIC/BIC information criteria

Tag	Information criteria	GARCH Lags (p,q)	(G)ARCH	EGARCH	TARCH
GE	AIC	1,1	**-3394,9**	-3393,6	-3360,1
	BIC		**-3364,2**	-3362,9	-3329,4
GK	AIC	1,1	**-5747,2**	-5702,2	-
	BIC		**-5712,2**	-5667,3	-
OD	AIC	2*,2*	-4079,5	**-4114,2**	-
	BIC		-4045,6	**-4065,7**	-
OF	AIC	1,1	**-4665,8**	-4632,2	-4569,6
	BIC		**-4636,8**	-4603,1	-4540,6
HX	AIC	1,1	**-11785,4**	-11747,9	-11670,4
	BIC		**-11745,7**	-11708,3	-11630,8
HV	AIC	1,1	**-10838,5**	-10821,7	-10738,2
	BIC		**-10798,9**	-10782,0	-10698,5
HZ	AIC	2,0	**-11132,36**	-	-
	BIC		**-11081,41**	-	-

* only second lag is included

The ARMA specification for models presented in Table 5 are equivalent to the ones used for the Ljung-Box test (Table 4) with the addition of GARCH(1,1) equation for variance (notice that series which do not contain ARCH effects are not modeled using GARCH). According to the results of the test, conventional GARCH specification yielded the lowest values for chosen information criteria for most of the series, suggesting that they do not contain pronounced leverage effects. Only series of returns on OD (Feed corn (China)) futures contract seem to have different response to positive and negative shocks (implied by EGARCH model). Moreover, adequate GARCH model for OD series has different in terms of included lags specification for the variance equation – EGARCH(2*,2*) (*only second lag is included). The lowest value of information criteria for HZ series was obtained with a simple ARCH(2) model.

Table 6. Estimated coefficients (p-values in parenthesis) of the ARMA/(G)ARCH/EGARCH models

Tag	Lags of CRB Index		AR(r) Lags				MA(m) Lags		ARCH		GARCH	
	1	2	1	2	3	4	1	2	1	2	1	2
GE	0,004	-	0,993	-	-	-	-0,985	-	0,106	-	0,735	-
	(0,927)	-	(0,000)	-	-	-	(0,000)	-	(0,002)	-	(0,000)	-
GK	0,042	-	0,942	-	-	-	-0,960	-	0,109	-	0,686	-
	(0,252)	-	(0,000)	-	-	-	(0,000)	-	(0,014)	-	(0,000)	-
OD	0,272	-	-0,680	-0,022	0,033	-0,067	0,610	-	-	0,646	-	77,514
	(0,000)	-	(0,000)	(0,721)	(0,478)	(0,189)	(0,000)	-	-	(0,000)	-	(0,086)
OF	0,091	-	0,058	-	-	-	-	-	0,050	-	0,943	-
	(0,015)	-	(0,151)	-	-	-	-	-	(0,000)	-	(0,000)	-
HX	0,224	-	0,470	-	-	-	-0,595	-	0,068	-	0,911	-
	(0,000)	-	(0,008)	-	-	-	(0,000)	-	(0,021)	-	(0,000)	-
HV	0,294	-	-0,101	-	-	-	-0,056	-	0,078	-	0,886	-
	(0,000)	-	(0,000)	-	-	-	(0,000)	-	(0,000)	-	(0,000)	-
HZ	0,209	-	0,909	-3,772	-	-	-0,958	0,444	0,086	0,111	-	-
	(0,000)	-	(0,000)	(0,123)	-	-	(0,000)	(0,058)	(0,000)	(0,000)	-	-
WR*	0,056	-	-	-	-	-	-0,376	-	-	-	-	-
	(0,000)	-	-	-	-	-	(0,086)	-	-	-	-	-
XV*	0,079	-0,424	-0,435	-0,053	-	-	-	-	-	-	-	-
	(0,000)	(0,000)	(0,008)	(0,005)	-	-	-	-	-	-	-	-
Z1*	0,050	-	0,745	-0,998	-	-	-0,739	0,991	-	-	-	-
	(0,114)	-	(0,000)	(0,000)	-	-	(0,000)	(0,000)	-	-	-	-

* estimated by ARMA(r,m) model

Most of the coefficients reported in the Table are highly significant – almost all of the GARCH(p,q) parameters are significantly different from zero under ordinary significance levels (0.05 and 0.1). Surprisingly, parameters of lagged CRB index returns in GE, GK (Corn and Soybeans of Argentina) and Z1 (Indian Barley) equations turn out to be statistically not different from zero. It may be caused either by peculiarities of Argentina's commodity market (for GE and GK) or by underlying nonlinear relationship between lagged CRB index returns and returns to futures contracts on these commodities.The p-values of the joint significance of the parameters (not reported here) for each of the equations are equal to zero, implying that all the equations contain significant information about the futures returns.

Since WR, XV and Z1 series are not subject to ARCH effects, correspondent parameters in Table 6 come from linear ARMA(r,m) models. Although, given the results of the bispectrum test, this linear estimation makes little sense, it may further serve as a comparable in model accuracy evaluation.

All the equations reported in Table 6 lead to elimination of autocorrelation in squared standardized residuals, thus to removal of ARCH effects. Moreover, for almost all of the series the value of kurtosis decreased (in comparison to kurtosis implied by ARMA estimation) towards the one of normal distribution (kurtosis comparison and model adequacy tests are in Table A.2). The reason why kurtosis remain high mainly comes from the way how it is specified by GARCH models (it depends on the value of parameter α_1 of variance equation). Moreover, GARCH estimation depends on the assumption about the errors distribution. All the GARCH parameters provided in Table 6 are derived assuming normally distributed residuals. Although this assumption is not fully credible, it still yields better results than in case of t-distribution or generalized error distribution (evaluation of alternatives is done using information criteria and resulting decrease in kurtosis).

All the upper analysis implies that ARMA/GARCH models are adequate and provide significant results. This means, that the null (1) of unpredictability is rejected. Next part of forecast estimation is devoted to Artificial Neural Network approach. Since it does not depend on any distributional assumptions, it may better account for underlying nonlinear dependencies and consequently provide more accurate forecasts.

5.3 Artificial Neural Networks approach

Similarly to GARCH modeling, the ANN estimation procedure starts with model specification, which is also done through the information criteria. The lag is chosen as the one that minimizes AIC/BIC criteria while maintaining model adequacy (no autocorrelations in residuals). As a result, the highest lag order of the autoregressive part (Table 7) is 15, meaning that returns on futures contracts now have an influence on returns in three weeks ahead.

Table 7. ANN models specification based on AIC/BIC information criteria

Tag	Lagorderoffuturesreturns	Lagorderof CRB index	AIC	BIC
GE	15	2	-4793,59	-4122,91
GK	10	6	-8616,00	-7801,24
WR	8	2	-17054,63	-16457,20
XV	10	3	-18447,34	-17606,44
OD	10	2	-6583,75	-5968,60
Z1	8	2	-10635,69	-10067,31
OF	10	3	-7067,17	-6409,15
HX	8	6	-17512,65	-16601,14
HV	6	2	-16594,29	-16022,47
HZ	5	2	-16972,93	-16457,73

Although the lag structure of ANN models does not seem to be meaningful, the overall performance of the models seems to be satisfactory. All the specifications left no autocorrelation in residuals, and almost all of them (except the model for HV) managed to successfully remove ARCH effects. Moreover, the value of Kurtosis of the residuals in general is much closer to the one implied by the normal distribution then in case of GARCH estimation (kurtosis comparison and model adequacy tests are in Table A.3).

For HV model, the ANN approach failed to eliminate the autocorrelation in squared residuals, meaning that they are not pure white noise. This implies possible inefficiency and a loss of forecast accuracy compared to GARCH model.

5.4 Comparison of forecasting accuracy

The evaluation of the forecasting performance of the two models is a necessary step for assessing the economic significance of obtained forecasts. For this purpose, the one-step-ahead rolling-window forecasts for 100 data points were estimated. In other words, the returns forecast which is not different from the naïve static prediction ($\Delta f_1(t+1) = \Delta f_1 t)$ does not bring any benefit for investors hedging strategies and tells nothing about the efficiency of the correspondent commodity market. Table 8 contains the set of accuracy coefficients calculated for both built models.

Table 8. Coefficients of out-of-sample forecasting accuracy of the GARCH and ANN models

	(E)GARCH				ANN			
Tag	**MAD**	**MSE**	**RMSE**	**U**	**MAD**	**MSE**	**RMSE**	**U**
GE	0,0102	0,00023	0,0151	0,7394	0,0105	0,00022	0,0149	0,6937
GK	0,0091	0,00020	0,0141	0,7060	0,0093	0,00020	0,0141	0,6918
OD	0,0176	0,00057	0,0238	0,6021	0,0175	0,00053	0,0230	0,5820
OF	0,0133	0,00029	0,0170	0,6424	0,0142	0,00033	0,0183	0,6888
HX	0,0110	0,00022	0,0147	0,5854	0,0158	0,00039	0,0198	0,6313
HV	0,0161	0,00038	0,0196	0,6456	0,0110	0,00020	0,0140	0,5573
HZ	0,0125	0,00025	0,0159	0,6240	0,0127	0,00026	0,0162	0,6336
WR*	0,0034	0,00003	0,0051	0,6612	0,0035	0,00003	0,0050	0,6528
XV*	0,0034	0,00002	0,0045	0,6832	0,0035	0,00002	0,0045	0,6892
Z1*	0,0097	0,00022	0,0148	0,6433	0,0107	0,00025	0,0158	0,6861

The analysis of the forecasting power also serves as a model robustness check and the indication of possible data-mining. Namely, if the model simply fits the sample data, it would show poor results for out-of-sample forecasts (the summary table of forecasting accuracy for in-sample forecasts is in Table B.1).

In general, both models showed similar results in terms of forecasting power (both in-sample and out-of-sample). The values of Theil inequality measure are almost equivalent for ANN results and GARCH results, except the HZ and Z1 series (which are estimated by the linear ARMA model). Additionally, there is no clear dominance of the either model in terms of the mean absolute deviation. The important result is that both models beat the naïve prediction in terms of forecasting power (all Theil coefficients are less than unity).

The graphical analysis of the forecasting power of the models is provided on Figure 1.

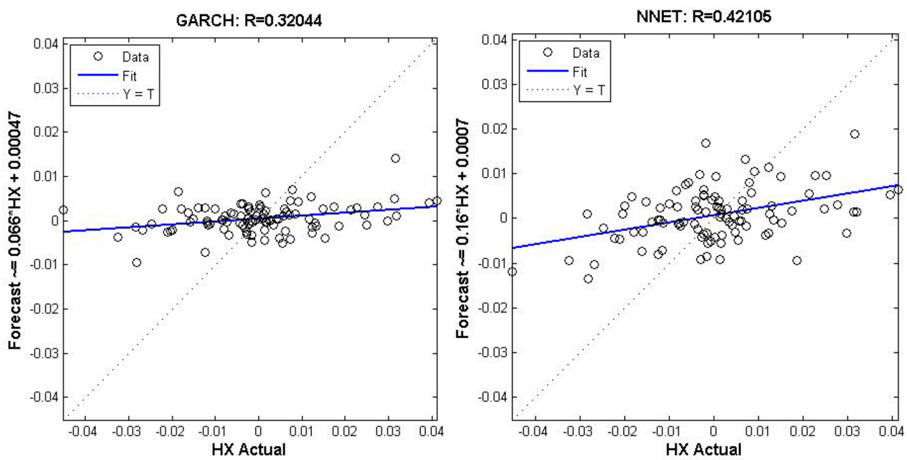

Figure 2.Target-output plots for Wheat (South Africa)

The Y-axe represents the forecasted value of HX series (out-of-sample) for one of the models (GARCH on the left and ANN on the right) and the X-axe stands for initial HX returns. R-value, reported above each diagram, is a measure of linear relationship between fitted and initial value. Essentially, the closer the regression line to the main diagonal of the Y,X plane, the better the fit of the model.

According to Figure 1, the ANN model provided slightly better fit then GARCH one. In general, GARCH forecasts exhibit similar behavior – the R-value is lower than for ANN forecast and fitted values are stretched along the regression line (for more details consider Table B.1). This indicates that GARCH provides a bit less accurate but more robust forecast as compared to ANN forecasts.

The values provided in the Table 8, and R-values reported on the Figure 1 are essentially not very informative in terms of comparison, since the underlying distributions are unknown. Diebold-Mariano test provides inference on whether the forecasts provided by the models are equal in terms of forecasting power. The output of the test is provided in Table 9.

Table 9. The p-values of Diebold-Mariano test for equal forecasting power

Tag	GE	GK	OD	OF	HX	HV	HZ	WR	XV	Z1*
MSE: p(DM)	0,721	0,9948	0,4299	0,1057	0,1007	0,7242	0,2428	0,6783	0,7431	0,0013
MAD: p(DM)	0,7106	0,4809	0,9218	0,2095	0,9006	0,6093	0,357	0,5115	0,393	0,0005

The p-values of the test indicate that the null of equal forecasting power cannot be rejected for any of the series except Barley India (Z1). This means that both methods can be applied for forecasting futures returns on emerging markets.

C h a p t e r 6

CONCLUSIONS

In thisresearchwe evaluate the hypothesis of unpredictability of commodity futures returns on emerging markets. The series of daily returns for four commodity markets (Argentina, China, India and Africa) were analyzed. It was found, that the series of futures returns do exhibit nonlinear nature and thus the application of nonlinear models is justified. Moreover, as expected, most of the series do contain ARCH effects in the variances, thus GARCH models should be applied.

The in-sample estimation of GARCH model revealed that most of the coefficients are individually statistically significant, and all of them are jointly significant for each equation. Comprehensive diagnostic testing was conducted. The residuals of the GARCH models contained no ARCH effects and autocorrelation, meaning that models are adequate.

This implies that the validity of information efficiency should be rejected given the data analyzed, or that the statistically significant forecasts can be made based on implicit market information.

The forecasting power of both models was analyzied, and no evidence of the superiority of either model was found. The accuracy coefficients are almost the same, and the Diebold-Mariano test failed to reject the hypothesis of equal forecasts. The graphical analysis of the results (based on input-output diagrams) also suggests no clear evidence which model is better. In general, GARCH yielded more robust forecasts (based on graphical analysis the forcasted values were more grouped), while the R-values of measure of fit are on average higher for neural network, indicating better fit.

The violation of EMH hypothesis means that possible abnormal returns can be extracted by forecasting futures returns. Although this is unlikely to happen consistently for a long period of time, the application of either GARCH or ANN approach for forecasting futures returns for given markets would bring extra benefit to the portfolio strategy of the investors.

The predictability of futures returns on emerging markets also implies, that the hedging strategies of investors may be significantly optimized, allowing them to face a tradeoff between risk and expected return. Moreover, since the results are common for all the studied emerging markets, it is reasonable to expect similar behavior of futures returns in Ukraine, when commodity market would be sufficiently mature.

WORKS CITED

Baillie, Richard T. and Robert J. Meyers. 1991. Bivariate estimation of the optimal commodity futures. *Journal of Applied Econometrics* 6:100-124.

Bera, Anil K., Edward Bubnys, and Hun Park. 1993. ARCH Effects and Efficient Estimation of Hedge. *Advances in Futures and Options Research* 6:313-328.

Bessembinder, Hendrik and Kalok Chan. 1992. Market Efficiency and Returns to Technical Anlysis. *Financial Management* 27:5-17.

Bessembinder, Hendrik, Jay F. Coughenour, Paul J. Seguin, and Margaret M. Smeller. 1996. Is here a Term Structure of Futures Volatilities? Reevaluating the Samuelson Hypothesis. *Journal of Derivatives* 4:45-58.

Bollerslev, Tim. 1986. Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics* 31:307-327.

Bonilla, Claudio A., Rafael Romero-Meza, and Elizabeth Guitierrez. 2007. Nonlinear Beahavior of the Chilean Capital Markets. *Applied Economics Letters* 14:987-991.

Brickett, Patrick L., Melvin J. Hinich, and Douglas Patterson. 1988. Bispectral-Based Tests for the Detection of Gaussianity and Linearity in Time Series. *Journal of the American Statistical Association* 83(403):657-664.

Campbel, John Y., Andrew W. Lo, and Craig A. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Cecchetti, Stephen G., Robert E. Cumby, and Stephen Figlewski. 1988. Estimation of the Optimal Futures Hedge. *Review of Economics and Statistics* (70):623-630.

Chen, Sheng-Syan, Cheng-few Lee, and Keshab Shrestha. 2003. Futures hedge ratios: a review. *The Quarterly Review of Economics and Finance* 43:433-465.

Chongfu, H. and M. Claudio. 2004. A diffusion-neural-network for learning from small samples. *International Journal of Approximate Reasoning* (35):137–161.

De Roon, Frans A., Theo E. Nijman, and Chris Veld. 2000. Hedging Pressure Effects in Futures Markets. *Journal of Finance* 55:1437-1456.

Dhamija , Ajay K. and Vinod K. Bhalla. 2010. Financial Time Series Forecasting: Comparison of Neural Networks and ARCH Models. *International Research Journal of Finance and Economics* (49):185-202.

Duong, Huu N. and Petko S. Kalev. 2006. An Intraday Analysis of the Samuelson Hypothesis for Commodity Futures Contracts. *Monash University*.

Efron, Bradley and Robert Tibshirani. 1993. *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Fama, Eugene F. 1970. Efficient capital markets: A review of theory and empirical work. *Journal* 25:383-417.

Fama, Eugene F. 1991. Efficient Markets: II. *Journal of Finance* 46:1575-1617.

Fama, Eugene F. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49(3):283-306.

Fernandez, Viviana. 2010. Commodity futures: How Efficient are financial markets? *Fourteenth Annual Conference of the Central Bank of Chilie* 1-23.

Ferson, Wayne E. and Campbell R. Harvey. 1991. The Variation of Economic Risk Premiums. *Journal of Political Economy* 99:385-415.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin. 2003. Spurious Regressions in Financial Economics? *Journal of Finance* 58:1393-1413.

Gandhi, Devinder, Samir Saadi, and Ibrahim Ngouhouo. 2006. Testing for Nonlinearity & Modeling Volatility in Emerging Capital Markets: The Case of Tunisia. *International Journal of Theoretical and Applied Finance* 9(7):1021-1050.

Heskes, Tom. 1997. Practical Confidence and Prediction Intervals. *Advances in Neural Information Processing Systems* 9:176-182.

Hirshleifer, David. 1988. Residual risk, Trading Costs, and Commodity Futures Risk Premia. *The Review of Financial Studies* 1:173-193.

Hopfield, J.J. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of National Academy of Sciences of USA V.79, No.8* 2554-2558.

Hsln, Chin-Wen, Jerry Kuo, and Cheng-Few Lee. 1994. A New Measure to Compare the Hedging Effectiveness of Foreign Currency Futures Versus Options. *Journal of Futures Markets* 14:685-707.

Johnson, Leland L. 1960. The theory of hedging and speculation in commodity futures. *Review of Economic Studies* (27):139-151.

Kaur, Gurbandini and Narasimha D. Rao. 2010. Efficiency of Indian Commodities Market: A Study of Agricultural Commodity Derivatives Traded on NCDEX. *Working Paper Series* 1-21.

Konstantinidi, Eirini and George S. Skiadopoulos. 2011. Are VIX Futures Prices Predictable? An Empirical Investigation. *International Journal of Forecasting* 27(2):543-560.

LeRoy, Stephen F. 1989. Efficient Capital Markets and Martingales. *Journal of Economic Literature* 27:1583-1621.

Lim, Kian-Ping, Robert D. Brooks, and Melvin J. Hinich. 2008. Nonlinear Serial Dependence and the Weak-form Efficiency of Asian Emerging Stock Markets. *Journal of International Financial Markets, Institutions and Money* 18(5):527-554.

Lim, Kian-Ping, Muzafar S. Habibullah, and Melvin J. Hinich. 2009. The Weak-Form Efficiency of Chinese Stock Markets: Thin Trading, Nonlinearity and Episodic Serial Dependencies. *Journal of Emerging Market Finance* 8(2):133-163.

McLeod, Ian A. and Wai K. Li. 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis* 4:269–273.

Miffre, Joelle. 2002. Economic Significance of the Predictable Movements in Futures Returns. *Economic Notes* 31:125–142.

Nelson, Benjamin D. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* (59):347-370.

Phukubje, Pierce M. and Motlatjo B. Moholwa. 2006. Testing for weak-form efficiency in South African futures market for wheat and sunflower seeds. *Agricultural Economics Association of South Africa (AEASA)* 45(2).

Samuelson, Paul A. 1965. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6:41-49.

Solt, Michael E. and Meir Statman. 1988. How useful is the sentiment index? *Financial Analysts Journal* 44:45-55.

Stoker, Thomas M. 1986. Consistent Estimation of Scaled Coefficients. *Econometrica* 54:1461-1481.

Timmermann, Allan. 2008. Elusive return predictability. *International Journal of Forecasting* 24(1):1-18.

Tsay, Ruey S. 2005. *Analysis of Financial Time Series*. Hoboken, NJ: John Wiley & Sons, Inc.

Ville de Goyet, Cédric d., Geert Dhaene, and Piet Sercu. 2007. Testing Futures Returns Predictability: Implications for Hedgers. Department of Accountancy, Finance and Insurance, Katholieke Universiteit, Leuven.

Wanas, Nayer, Gasser Auda, Mohamed S. Kamel, and Fakhreddine Karray. 1998. On the Optimal Number of Hidden Nodes in a Neural Network. *Electrical and Computer Engineering* 2:918-921.

Wang, Changyun. 2001. Investor Sentiment and Return Predictability in Agricultural Futures Markets. *The Journal of Futures Markets* 21 (10):929-952.

Worthington, Andrew C. and Helen Higgs. 2003. Tests of Random Walks and Market Efficiency in Latin American Stock Markets: An Empirical Note. *School of Economics and Finance Discussion Papers and Working Papers Series* (157):1-14.

Zakoian, Jean-Michel. 1991. Threshold heteroskedastic models. *Technical report, NSEE.*

APPENDIX A

Table A.1 The detailed results of Dickey-Fuller test for unit root

Tag		Lagorder 0	1	2	3
GE	ADF p-val	0,0000	-	-	-
	DW p-val	0,5306	-	-	-
GK	ADF p-val	0,0000	-	-	-
	DW p-val	0,1579	-	-	-
WR	ADF p-val	0,0000	-	-	-
	DW p-val	0,1742	-	-	-
XV	ADF p-val	0,0000	-	-	-
	DW p-val	0,1160	-	-	-
OD	ADF p-val	0,0000	-	-	-
	DW p-val	0,6127	-	-	-
Z1	ADF p-val	0,0000	0,0000	-	-
	DW p-val	0,0785	0,6074	-	-
OF	ADF p-val	0,0000	-	-	-
	DW p-val	0,4813	-	-	-
HX	ADF p-val	0,0000	-	-	-
	DW p-val	0,9755	-	-	-
HV	ADF p-val	0,0000	0,0000	0,0000	0,0000
	DW p-val	0,0008	0,0273	0,0108	0,8106
HZ	ADF p-val	0,0000	-	-	-
	DW p-val	0,3528	-	-	-

Table A.2 P-values of Ljung-Box test for GARCH residuals, and the comparison of pre-GARCH and aafetr-GARCH kurtosis

Tag	**p(Q) for e**	**p(Q) for e^2**	**ARMA Kurtosis(e)**	**GARCH Kurtosis(e)**
GE	0,76	0,11	6,02	6,03
GK	0,80	0,96	26,75	24,82
OD	0,66	0,29	33,28	27,40
OF	0,68	0,93	13,87	10,74
HX	0,75	0,89	8,58	7,40

HV	0,60	0,92	4,51	4,08
HZ	0,45	0,95	17,42	17,86

Table A.3 P-values of Ljung-Box test for GARCH residuals, and the comparison of pre-GARCH and aafetr-GARCH kurtosis

Tag	**p(Q) for e**	**p(Q) for e^2**	**ARMA Kurtosis(e)**	**ANN Kurtosis(e)**
GE	0,33	0,11	6,02	4,91
GK	0,80	0,96	26,75	16,43
OD	0,45	0,32	33,28	23,41
OF	0,21	0,96	13,87	13,05
HX	0,40	0,21	8,58	8,35
HV**	0,15	0,00	4,51	4,22
HZ	0,11	0,34	17,42	14,04
WR*	0,27	0,97	45,09	33,47
XV*	0,39	0,83	39,77	35,12
Z1*	0,22	1,00	71,59	69,02

APPENDIX B

Table B.1 Coefficients of in-sample forecasting accuracy of the GARCH and ANN models

	(E)GARCH				ANN			
Tag	**MAD**	**MSE**	**RMSE**	**U**	**MAD**	**MSE**	**RMSE**	**U**
GE	0,0093	0,00020	0,0140	0,7204	0,0097	0,00018	0,0135	0,6880
GK	0,0116	0,00035	0,0186	0,6870	0,0112	0,00028	0,0169	0,6234
OD	0,0166	0,00077	0,0277	0,6757	0,0168	0,00068	0,0261	0,6359
OF	0,0140	0,00043	0,0208	0,7381	0,0137	0,00038	0,0196	0,7002
HX	0,1153	0,00026	0,0160	0,6736	0,0112	0,00023	0,0151	0,6369
HV	0,1474	0,00038	0,0196	0,6710	0,1478	0,00038	0,0195	0,6690
HZ	0,0126	0,00033	0,0181	0,6768	0,0126	0,00031	0,0177	0,6590
WR*	0,0044	0,00006	0,0080	0,6919	0,0046	0,00006	0,0077	0,6650
XV*	0,0045	0,00007	0,0081	0,6792	0,0045	0,00006	0,0079	0,6619
Z1*	0,0112	0,00033	0,0181	0,7133	0,0119	0,00034	0,0184	0,7283

APPENDIX C

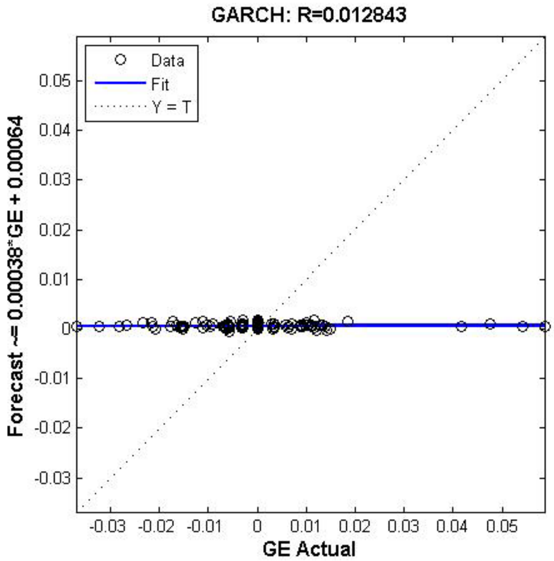

Figure C.1. GARCH target-output plot for Corn (Argentina)

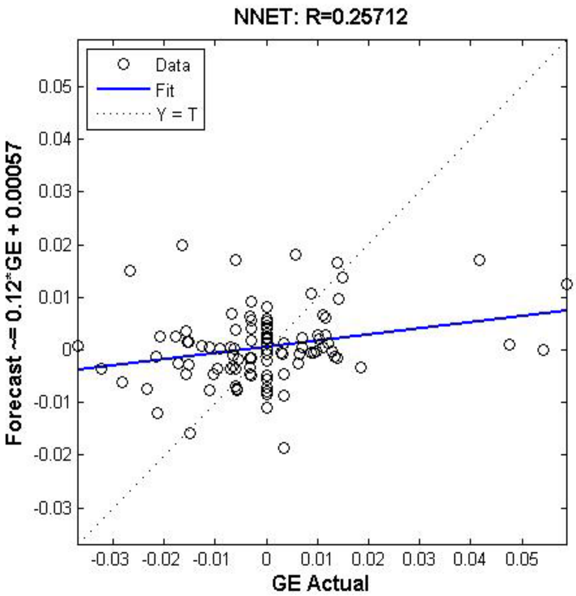

Figure C.2. ANN target-output plot for Corn (Argentina)

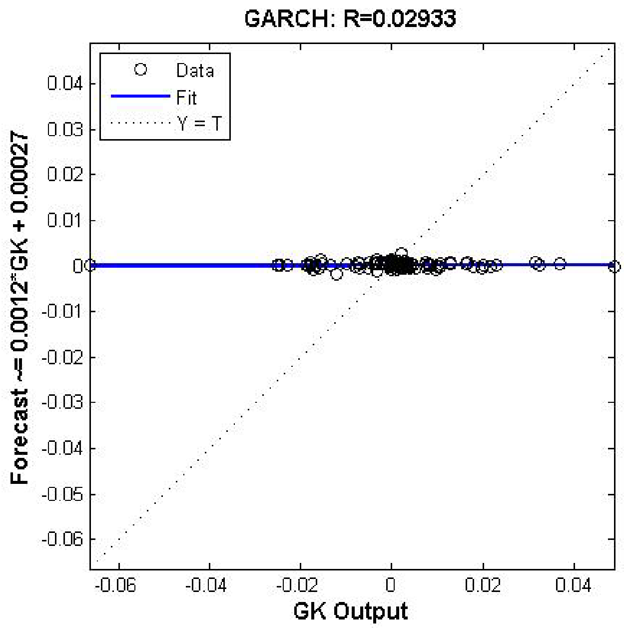

Figure C.3. GARCH target-output plot for Soybeans (Argentina)

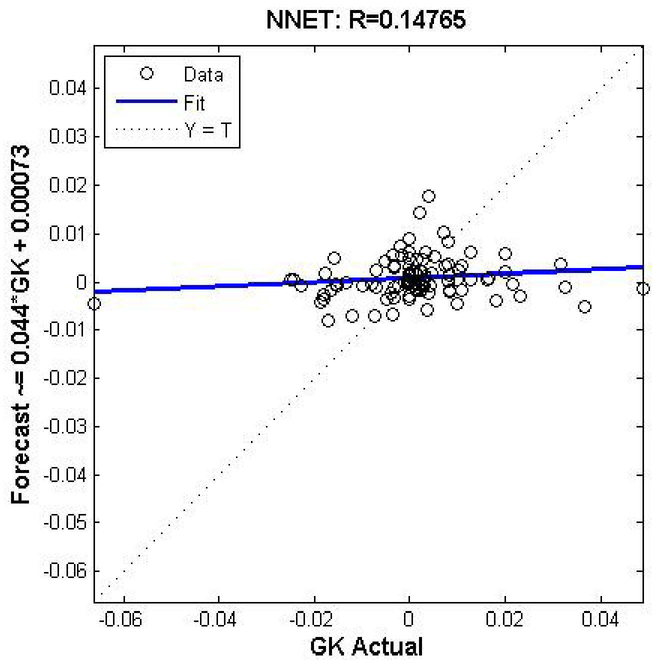

Figure C.4. ANN target-output plot for Soybeans (Argentina)

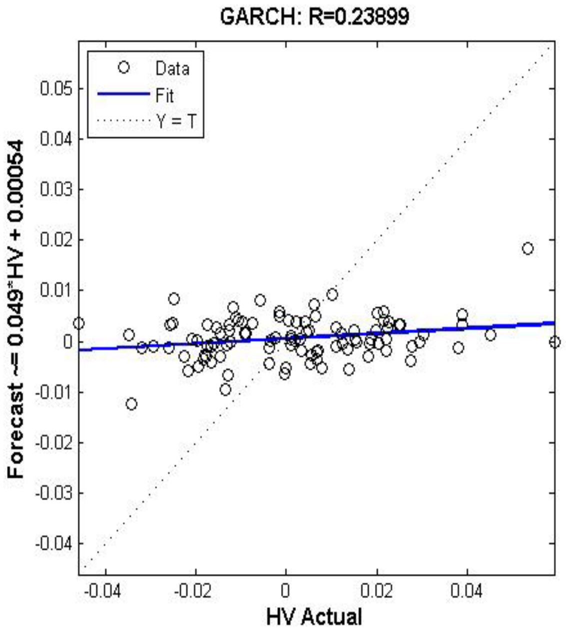

Figure C.5. GARCH target-output plot for Winter wheat (China)

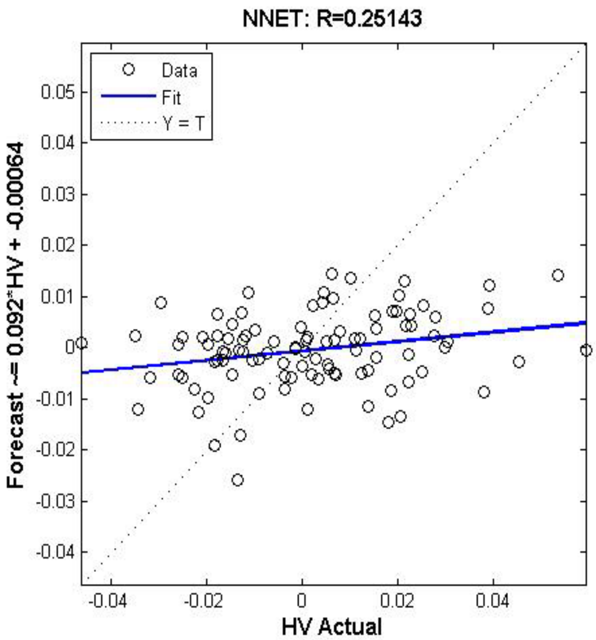

Figure C.6. ANN target-output plot for Winter wheat (China)

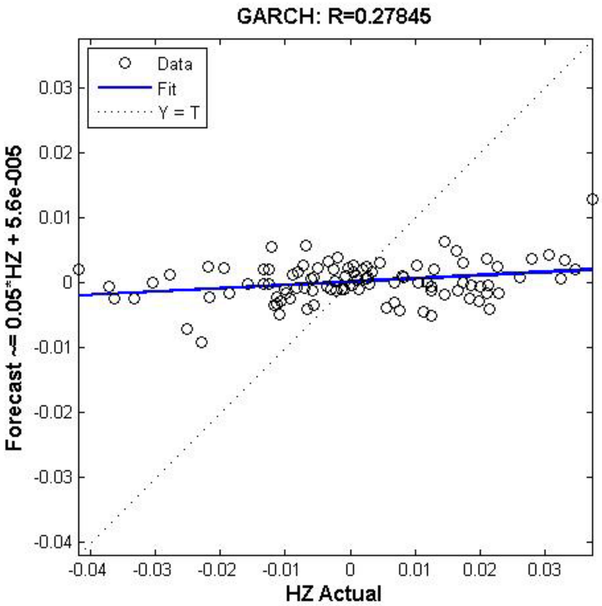

Figure C.7. GARCH target-output plot for Soya (South Africa)

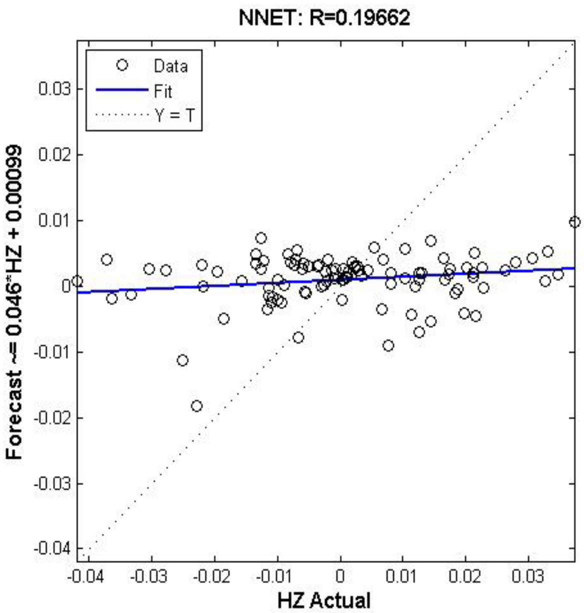

Figure C.8. ANN target-output plot for Soya (South Africa)

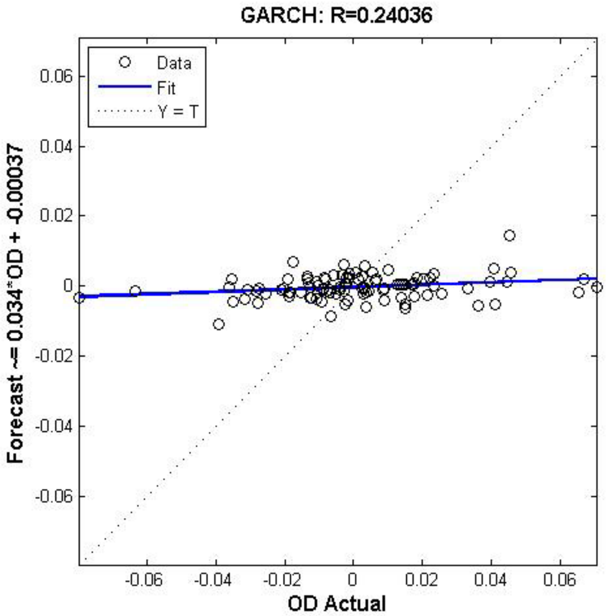

Figure C.9. GARCH target-output plot for Feed Corn (South Africa)

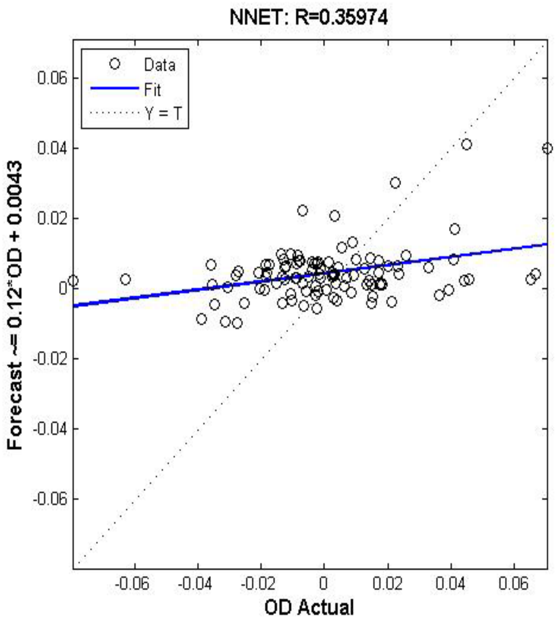

Figure C.10. ANN target-output plot for Feed Corn (South Africa)

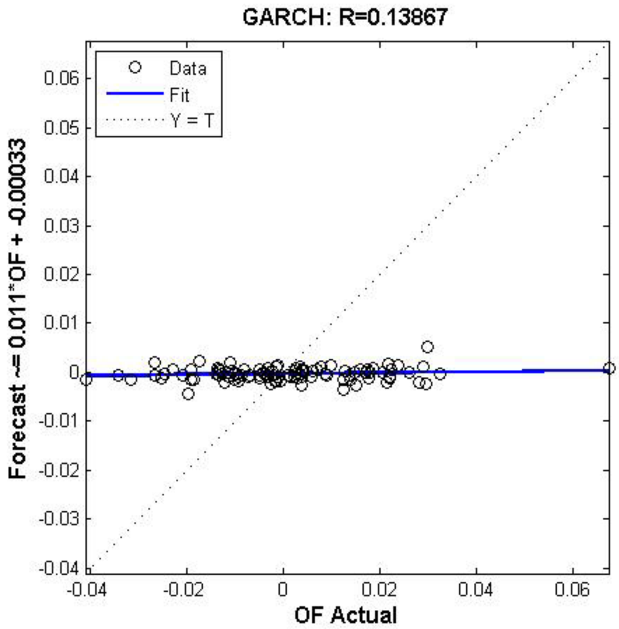

Figure C.11. GARCH target-output plot for Feed Barley (India)

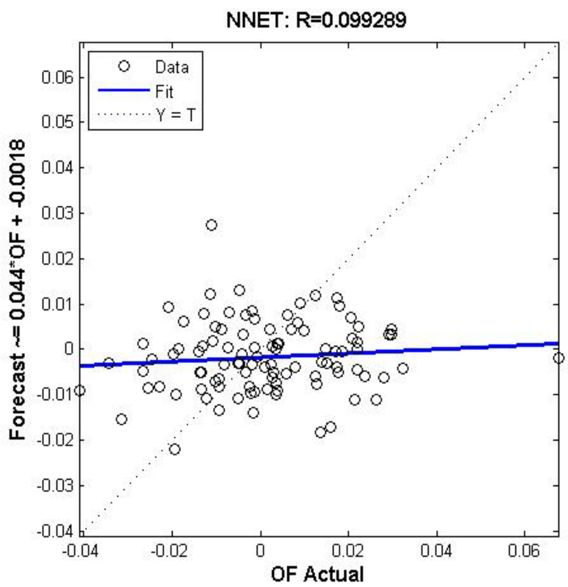

Figure C.12. ANN target-output plot for Feed Barley (India)

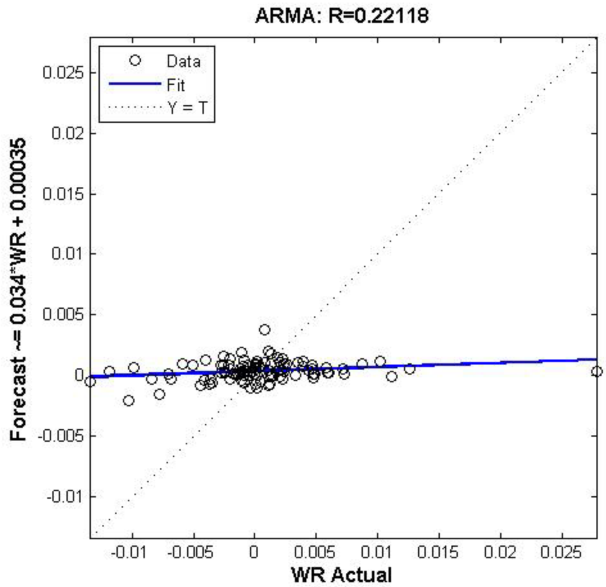

Figure C.13. ARMA target-output plot for Winter Wheat (India)

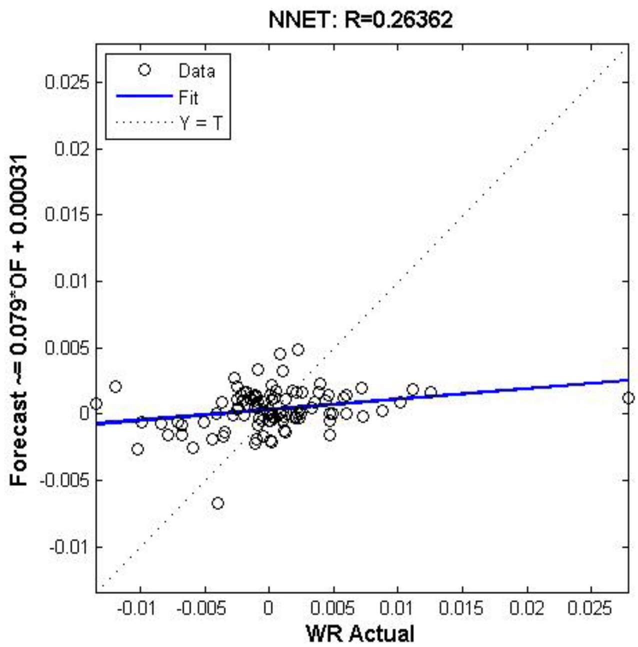

Figure C.14. ANN target-output plot for Winter Wheat (India)

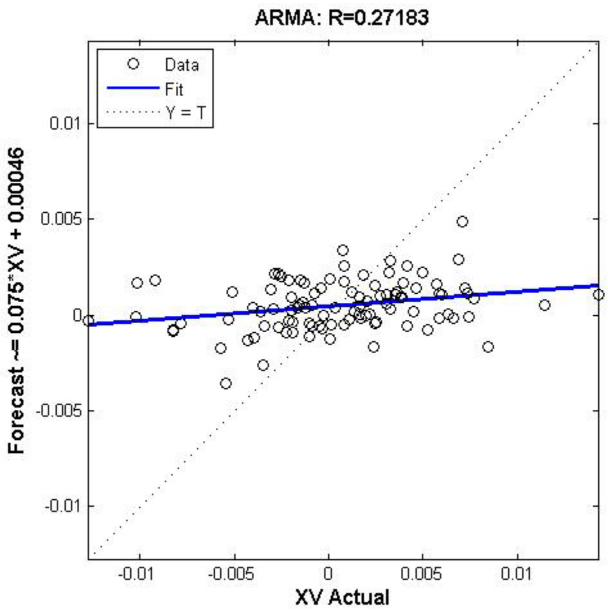

Figure C.13. ARMA target-output plot for Corn (South Africa)

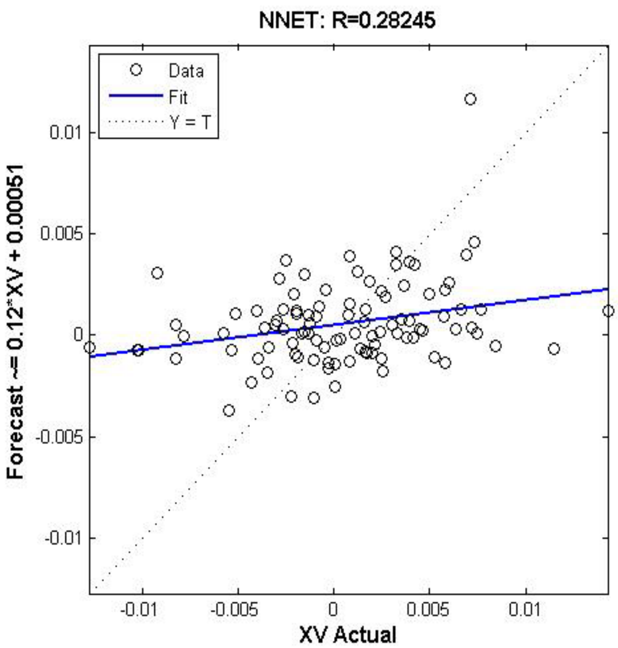

Figure C.14. ANN target-output plot for Corn (South Africa)

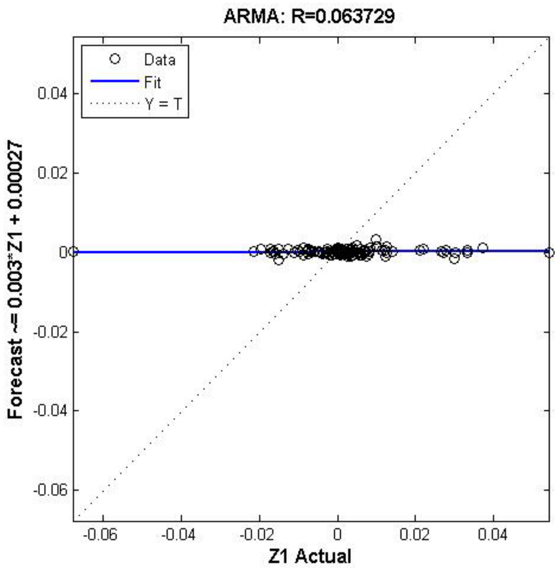

Figure C.15. ARMA target-output plot for Barley (India)

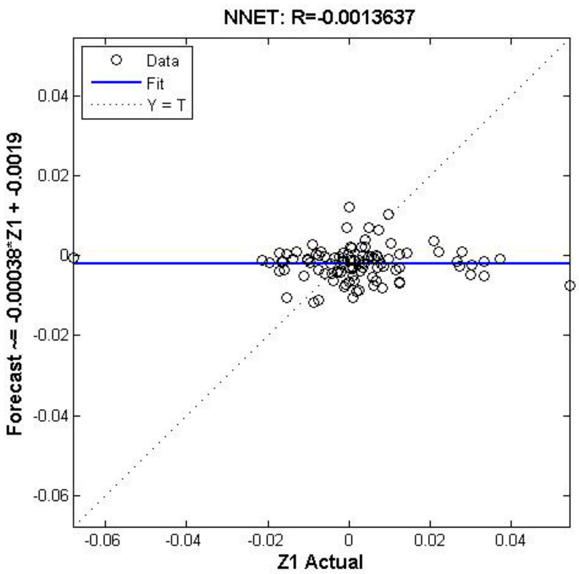

Figure C.16. ANN target-output plot for Barley (India)