Expert Systems with Applications 42 (2015) 7245–7251

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model

Werner Kristjanpoller [a,*], Marcel C. Minutolo [b,1]

[a] Departamento de Industrias, Universidad Técnica Federico Santa María, Av. España 1680, Valparaíso, Chile
[b] Department Management, Robert Morris University, 324 Massey 6001 University Blvd Moon Township, PA 15108, United States

ARTICLE INFO

Article history:
Available online 4 May 2015

Keywords:
Gold price volatility
Artificial Neural Network
GARCH models

ABSTRACT

One of the most used methods to forecast price volatility is the generalized autoregressive conditional heteroskedasticity (GARCH) model. Nonetheless, the errors in prediction using this approach are often quite high. Hence, continued research is conducted to improve forecasting models employing a variety of techniques. In this paper, we extend the field of expert systems, forecasting, and model by applying an Artificial Neural Network (ANN) to the GARCH method generating an ANN–GARCH. The hybrid ANN–GARCH model is applied to forecast the gold price volatility (spot and future). The results show an overall improvement in forecasting using the ANN–GARCH as compared to a GARCH method alone. An overall reduction of 25% in the mean average percent error was realized using the ANN–GARCH. The results are realized using the Euro/Dollar and Yen/Dollar exchange rates, the DJI and FTSE stock market indexes, and the oil price return as inputs. We discuss the implications of the study within the context of the discipline as well as practical applications.

1. Introduction

The ability to forecast the volatility of security prices is a major challenge given their economic and financial importance. In this context, the ability to predict gold price (spot and future) volatility with greater precision is important for commodity markets and for the world economy. The common method is to apply generalized autoregressive conditional heteroskedasticity (GARCH) models to forecast volatility. Nonetheless, the errors in prediction using this approach are often quite high. Resultant errors in forecasting have the potential to deliver great economic loss for those using a flawed model. Additionally, shortcomings in modeling approaches contribute to greater inefficiencies in the market. Hence, improved modeling approaches are continuously sought in order to reduce risk and improve market efficiencies.

Traditional research to reduce errors in models has sought to include variables that appear to be more important in explaining gold price volatility for the period studied thereby improving the explanatory power of a particular model; however, this approach typically lacks the capability to forecast outside of the sample period. This study innovates by changing the focus of the models to the capability to forecast future volatility as opposed to explaining. The results of this study are therefore more useful in the prediction of the gold spot price and the future gold price volatility. This should be important for government agents in countries with an economy related to the gold, and for investors wanting to make investment decisions in the commodity market (spot and future) to get a better asset allocation and portfolio diversification.

In an earlier study, Tully and Lucey (2007) modeled the price volatility of gold using an Asymmetric Power (AP) GARCH model, concluding that the most relevant variables influencing gold price volatility were oil prices and the FTSE. Kristjanpoller, Fadic, and Minutolo (2014) demonstrated that an expert system, in particular the ANN–GARCH, increases the accuracy of volatility forecasts predicted by GARCH models. The expert system is sensitive to behavior between variables such that the results are improved forecasts. The ANN–GARCH approach creates the possibility to determine the influences that variables exert the result of which is incremental improvements in the accuracy of forecasts as compared to the classical form of their relationship in the fit of the model. The ANN–GARCH approach incorporates as input to the ANN model the GARCH forecasts but allows for the possibility to incorporate other inputs variable into the ANN. Thus, financial variables that are significant for the price or the volatility of the price of gold in the classical models can be incorporated. This fact is very important, since the focus is not on the behavior *in*-sample, but the influence of the

* Corresponding author. Tel.: +56 (32) 2654571; fax: +56 (32) 2654815.
E-mail addresses: werner.kristjanpoller@usm.cl (W. Kristjanpoller), minutolo@rmu.edu (M.C. Minutolo).
[1] Tel.: +1 412 397 5451.

http://dx.doi.org/10.1016/j.eswa.2015.04.058
0957-4174/© 2015 Elsevier Ltd. All rights reserved.

variables is measured according to its contribution in the forecast *out*-of-sample. Building on these previous works, this study employs an ANN–GARCH model to forecast volatility beyond the sample period to the out-of-sample population thereby improving current knowledge and abilities.

The significance of this current work is based on two findings. First, we determine the improvement of accuracy in forecasting using a hybrid model as opposed to traditional GARCH models. In particular, we demonstrate greater accuracy of the hybrid ANN–GARCH forecasts of gold price volatility for different periods and for spot and future gold prices over the GARCH alone. Second, the ability to include financial variables as inputs of the ANN allows for the determination of the influence that the variables have on the estimation of the gold price volatility, spot and future for different horizons. This work is important since improved accuracy of gold price estimations and of gold price volatility will result in investor decision making, improve the efficiency of the market, and contribute in the future price allocation.

The remainder of this work is divided into four additional sections. In the next section, we provide a brief review of the literature on forecast models, neural networks, and various applications that have analyzed gold spot prices and gold future prices for evidence of which macroeconomic and other economic variables influence them. In the following section the methodology is detailed and the data used are analyzed. The results are then presented and interpreted. Finally, the last section summarizes the main conclusions of this study.

2. Literature review

In recent years, authors have focused on modeling and forecasting volatility in financial series since it is crucial for the characterization of markets, portfolio optimization and asset valuation; the case of gold is no exception. There are numerous studies whose focus is on gold, whether it is in the analysis of its spot price, future price, or volatility. In this study, the volatility of the spot price and the future price of gold is modeled using an ANN–GARCH model to determine the important macroeconomic variables for forecasting.

Many types of models have been used to forecast volatility, but the most widely used are the ARCH models proposed by Engle (1982), and then generalized by Bollerslev (1986); they have led to significant improvements in the modeling of time series. Later Kroner, Kneafsey, and Claessens (1995) tried to predict the volatility of the daily price of cacao, corn, cotton, gold, silver, sugar, and wheat in the long term by using the combination of the GARCH model and the ISD (Implied Standard Deviation) model. In addition, Tully and Lucey (2007) investigated macroeconomic influences on gold using an AP–GARCH model. Their research examined the spot price and the future price of gold from 1983 to 2003 using macroeconomic variables. They paid special attention to two periods – around 1987 and 2001 – when there were shocks to the stock markets. The results showed that the AP–GARCH model delivered the most appropriate fit to the data, and the most important explanatory variable was the dollar.

Trück and Liang (2012) examined different models that can be used to predict volatility (GARCH, TARCH, TGARCH, ARMA) in order to study their behavior in the gold market and to evaluate the performance of these models. Their results showed that for prediction, both within and outside the sample period, the TARCH models provided the best results.

Finally, Creti, Joëts, and Mignon (2013) contributed to the field by studying the relationship between commodities and stocks. They focused on the dynamics of the correlation between both markets, and analyzed whether these correlations evolved according to the situation of "optimism or pessimism" in the stock market. The methodology they used was the dynamic conditional correlation approach (DCC) GARCH introduced by Engle (2002), which allows the assessment of changes over time in the correlations between returns on commodities and stocks.

Other tools used in the study of price volatility are Artificial Neural Networks. Several researchers have used them for the study of gold. For example, Grudnitski and Osburn (1993) studied the impact of general economic conditions and the expectations of the investors on the S&P 500 index and the prices of gold futures, modeling these prices like a neural network. Later on, Parisi, Parisi, and Díaz (2008) conducted a study in which they analyzed the Recursive Neural Network and the Rolling Neural Network as modifications to the traditional neural networks, and applied these new strategies to predict variations in the price of gold. Another study is that of Yazdani-Chamzini, Yakhchali, Volungevičienė, and Zavadskas (2012), who investigated the ability of the ANFIS (Adaptive Network Fuzzy Inference System) model to capture changes in the gold price, and then evaluated the model's performance compared to those of the ANN and the Autoregressive Integrated Moving Average (ARIMA) model. Their results showed that the ANFIS and ANN methods are powerful tools to model the price of gold and can produce better results than the ARIMA model.

Since we want to predict the future, it is necessary to consider the implied volatility of gold prices in addition to their realized volatility and history. Hamid and Iqbal (2004) predicted the volatility of prices of a group of commodities using a neural network, and then they compared their results to the implied volatility obtained by using the Barone-Adesi and Whaley options pricing model (1987) to contrast both predictions with the actual volatility. The conclusions of the article indicated that the predictions made by the neural network substantially improved the predictions obtained through the implied volatility.

Szakmary, Ors, Kyoung Kim, and Davidson (2003) also analyzed the implied volatility. Their work represents an important contribution because they used data from futures markets, where options about futures and underlying future contracts are traded on the same basis. The results obtained in the study indicated that even though the implied volatility performed well, it was not an unbiased forecaster of future volatility. Neely (2003) used high-frequency data (every 30 min) of future spot prices of gold as well as new econometric techniques, including long memory models, to examine why implied volatility is an inefficient and biased forecaster of actual volatility. According to his study, none of the suggested explanations previously mentioned in the literature (inaccurate estimation of volatility, overlapping samples, selection of the sample, etc.), can plausibly explain this bias and inefficiency.

Finally, it is necessary to know what other variables affect the gold market, since one of the requirements for using neural networks is the correct selection of explanatory variables. There are numerous studies that have analyzed the impact of macroeconomic variables on the price of gold; for example, Batten, Ciner, and Lucey (2010) examined the relationship between the volatility of four metals actively traded in the markets (gold, silver, platinum and palladium) with key macroeconomic factors in the global economy, such as the price of oil and inflation. These studies provide clear evidence that the same macroeconomic factors have an influence on the volatility of the series of precious metal prices, but there is limited evidence of the feedback from the price volatility of other commodities.

Shafiee and Topal (2010) analyzed the behavior of gold prices from all existing records to see what factors have helped increase its value, given that the gold market has attracted a lot of attention and its price is nearly the highest in history. The most important variables that explain gold price behavior are the price of oil and inflation. There has been a high correlation between the prices of gold and oil, close to 85%, in the last four decades. However, the study showed that the relationship between the gold price and the accumulated inflation was only approximately 9% over the last four decades, and that there never has been a statistically significant relationship between gold and inflation.

Elder, Miao, and Ramchander (2012) studied the impact of macroeconomic news on returns, actual volatility, and volume in the futures of gold, silver, and copper; concluding that economics news positively influenced the realized volatility and volume for all three metals. They also analyzed the time it takes for news to be fully absorbed by the market, and whether the metals market responds asymmetrically to unexpected macroeconomic news, finding evidence that several news announcements exerted an asymmetric impact on the market activity variables.

Some recent studies applied different model to forecast volatility. For instance, Haugom, Langeland, Molnár, and Westgaard (2014) applied the Heterogeneous Autoregressive model of the Realized Volatility (HAR-RV) to the volatility of the US oil market. Their results demonstrate that including market variables improved forecasting precision. Fernandes, Medeiros, and Scharth (2014) used parametric and semiparametric heterogeneous autoregressive to predict the volatility index (VIX) of the Chicago Board Options Exchange (CBOE) and the results indicated that it is difficult to beat the pure HAR process, given the persistent nature of the VIX index. Monfared and Enke (2014) applied a hybrid model GJR–GARCH Neural Network to forecast the volatility of the Nasdaq, concluding that the hybrid model predicts well in extreme events given that the structure process of volatility is complex and that the approach is a good method to use along with the CVaR.

In additional research, Muzzioli, Ruggieri, and De Baets (2014) compare different fuzzy regression methods to estimate the implied volatility smile function, determining that the fuzzy regression is better predictor method than the classical approach based on cubic splines. Vortelinos (2015) examines nonlinear models to forecast the volatility of seven financial markets, being the Heterogeneous Autoregressive (HAR) model has the best performance. Efimova and Serletis (2014) modeled energy markets volatility using the GARCH approach concluding that univariate and multivariate models yield similar estimates but univariate models produce more accurate forecasts. Azadeh, Moghaddam, Khakzad, and Ebrahimipour (2012) applied a flexible algorithm based on Artificial Neural Network and fuzzy regression to predict the oil price, concluding that the ANN model has the best predictions measure in terms of mean absolute percentage error (MAPE). Boyacioglu and Avci (2010) applied adaptive neural fuzzy inference system (ANFIS) in Istanbul stock market to predict the earnings per share, concluding successful about the monthly forecasting. Svalina, Galzina, Lujic, and Šimunovic (2013) using ANFIS to predict the close price in Zagreb Stock Market index, obtaining information that is useful for predicting within its limits.

While all of the above mentioned studies have continued to produce increasingly more accurate results there is still room for improvement. In the next section, we present the ANN–GARCH model and the data used to test it. Then we present the findings and discuss the implications.

3. Methodology and data

Since classical models, such as OLS, are not suitable for situations in which the variance is heteroskedastic, which is the case for financial time-series, Engle (1982) introduced ARCH (autoregressive conditional heteroskedasticity) models, in which the current error variance is a function of the terms of error of previous periods. Subsequently Bollerslev (1986) generalized these models, and proposed the GARCH (generalized autoregressive conditional heteroskedasticity) process, that consists in a symmetric model in which conditional variance depends on the conditional variances from previous periods, as well as depending upon the square of perturbations. One of the reasons these models have been widely used in the financial literature is because they capture volatility clustering.

To illustrate, suppose P_t is an index of prices from a financial series and r_t its return or percentage price variation, where the index t denotes an observation to the daily closing.

$$r_t = \log P_t - \log P_{t-1}$$

For the series of returns, the GARCH model may be expressed as follows:

$$r_t = \mu + \varepsilon_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

where

$p \geq 0,\ q > 0,\ \alpha_0 > 0,\ \alpha_i \geq 0,\ i = 1, \ldots, q$ and $\beta_i \geq 0,\ i = 1, \ldots, p$.

Conditions on the parameters are taken to ensure that the conditional variance of the GARCH (p,q) is always positive. In addition, to ensure a finite expected value of the variance it is assumed that:

$$\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i < 1$$

Realized volatility is the forecasted parameter through a GARCH (1,1) model, an AR(1) model and the Artificial Neural Network model and is computed as the sample variance log returns in a 21 d window to the future (approximately one month of transactions), as shown by the equation.

$$RV_t = \frac{1}{21} \sum_{i=t+1}^{t+21} (r_i - \bar{r}_t)^2$$

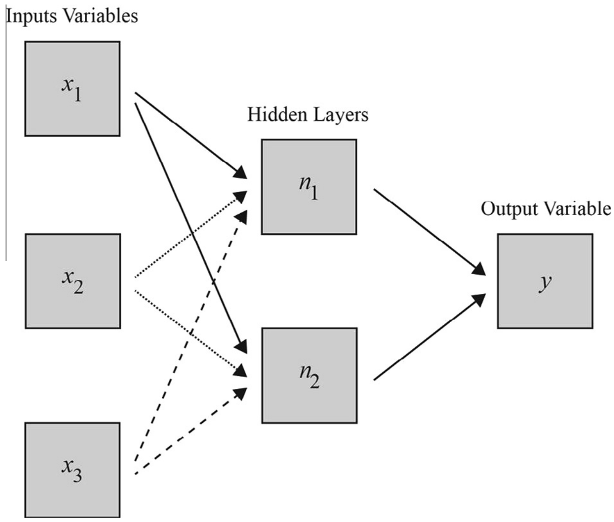
In this study, a GARCH model (1,1) is used with a moving window length of 252 d back (one year of transactions); in addition, an autoregressive model of order 1 is used for the mean equation. With these considerations, equations are determined in the following way:

$$r_t = c + \theta_1 r_{t-1} + \varepsilon_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Artificial Neural Networks are a powerful non-parametric tool used for signal filtering, recognition of patterns and interpolation, among many other applications. Within their characteristics, they are also able to tolerate data with errors and find nonlinear associations between the parameters of the model. In addition, one of their major advantages over other econometric methods is that it is not necessary to take the functionality of the model, which means you do not need to make an assumption about the functional relationship among the variables; however, it is necessary to incorporate the appropriate variables to be able to make a good estimate.

Each neural network connects a group of input variables $\{x_i\},\ i = 1, \ldots, k$ with a group of one or more output variables $\{y_j\},\ j = 1, \ldots, k$ and zero, one, or more so-called hidden layers. Neurons are connected between the layers for connections that are activated by reaching a threshold, because the evaluation of the function of transfer is based on the input parameters. Each layer can have a different number of neurons. The input and output can be continuous, discrete, binary variables, or a combination of all of them.



This study uses the back propagation algorithm which is an algorithm for supervised learning which seeks to minimize the quadratic error by descent maximum gradient. It is based on "back propagation" of errors.

To estimate the neural network model, it is necessary to define the input variables, the characteristic parameters of the network, and the length of window available. The independent variables used as input for the Artificial Neural Network are the daily variations of Euro/Dollar and Dollar/Yen exchange rates, the stock market index returns of the Dow Jones Industrial (DJI) and the Financial Time Stock Exchange (FTSE) and the daily price variation of oil. The initial parameters of the ANN are three layers and five neurons per layer. The realized volatility to forecast is $t + 22$, which implies the standard deviation, calculated 22 d from today, taking the last 21 data (unknown today) to calculate it. The rolling window length is 252 d. In all the models tested, the GARCH forecast and the square of gold price variation are included.

The first models are built from two input series, the GARCH forecast and the square of the gold price return. These two inputs variables are basic because the GARCH forecast is the variable which is improved through the ANN–GARCH and the square gold price return is a good predictor for the realized volatility. Then the other variables are added stepwise. To choose the order, the variables are ranked from the correlation matrix of the explanatory variables with the realized volatility. After this analysis, to measure the robustness of the results, the rolling window length, the period of the realized volatility to forecast and the combination of input variables are modified. In particular, the volatility forecast is varied to 21 d, 14 d and 28 d. These models are applied to forecast the gold spot price realized volatility and gold future price realized volatility.

One of the advantages that an ANN–GARCH model has over other techniques is that ability that the ANN has to learn from the GARCH forecasting errors. Further, the ANN allows the user to feed the network with additional variables to improve the forecasting results. We recognize, however, that the ANN–GARCH is less flexible than an NFIS model. One other disadvantage of the current method is that the backpropagation algorithm of the ANN lacks the ability to learn of its own forecasting error as in other dynamic networks. However, the ANN–GARCH is rather simple to implement which gives the approach an added advantage.

To test the results, the forecasted values are compared against the realized volatility and the following measures are used for determining the error: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Deviation (MAD or MAE) and Mean Absolute Percentage Error (MAPE). These four measures are relevant in order to analyze the errors; however, for this particular case the MAPE is the most relevant for two reasons. The first reason is that MPAE provides a comparison between the realized volatility and the forecast by the percentage of error (and not for the error value), thus making it is easy to measure the precision. The second reason is that since the objective is to forecast the volatility which is not constant but heteroskedasticity, the relative analysis is appropriate.

The data sets analyzed in this paper are the Gold Spot Price and the Gold Future Price (Generic 1st 'GC' Future) from Bloomberg. The sample period for the Gold Spot Price data is from September 6, 1999 to March 20, 2014 for a total of 3,836 observations. The Gold Future Price data are from September 6, 1999 to April 15, 2014 for a total of 3,665 observations.

Table 1 shows some descriptive statistics for the Gold Spot Price returns and the Gold Future Price returns. For both cases the mean is close to 0 (0.0424% and 0.0444%) and the standard deviation is around one. These values are typical for stationary series since they possess a mean close to 0 with small variations. The Augmented Dickey Fuller test (ADF) is used to analyze the stationarity, which turns out to be significant at 1% for the series; therefore, it is possible to conclude that the series are stationary and thus one can make projections from them.

The Kurtosis is much higher than 3 (value for which the series is usually Mesokurtic), indicating that the series presents a high degree of concentration around the central values of the variable. This shows that fat-tailed distributions are necessary to correctly describe the conditional distribution of the returns.

The skewness is small and negative, showing that the lower tails of the empirical distributions of returns are longer that the upper tails, meaning negative returns are likely to be well below the average. LM(12) is the Lagrange Multiplier test for ARCH (ARCH-LM) effects in the OLS residuals from the regression of the returns on a constant. The null hypothesis of non-existence of ARCH effects is rejected for both series.

The Jarque–Bera Normality Test indicates that model errors are not normally distributed, which implies that the empirical distribution of the daily returns of the gold prices exhibit significantly heavier tails than in a normal distribution.

The descriptive statistics of the independent variables are shown in Table 2. All the variables have a positive mean except the Japanese Yen. The DJI was more profitable on average than the FTSE in the period analyzed. The oil price shows a high daily return and also a high standard deviation. All the series are stationary.

The correlations of the independent variables with the realized volatility are very low and in some cases they are negative. Table 3 shows the correlation matrix between the independent variables and the realized volatility for the gold spot price and gold future price for three time horizons, 14 d, 21 d and 28 d. Also, the GARCH forecast is included as an independent variable because it is one of the inputs for the Artificial Neural Network.

4. Results analysis

To obtain better results in the forecasting of the gold price's realized volatility, the maximum daily change in the ANN–GARCH predictions is adjusted. This threshold is fixed at 10%; if the ANN–GARCH volatility forecast in day t is higher than 1.1 times than the previous volatility forecast (t-1), then the forecast in day t is limited to 1.1 times the previous forecast (t-1).

Following Kristjanpoller et al. (2014), the first ANN–GARCH model keeps the GARCH forecast and the squared gold price return (spot or future depending on case) fixed and adds the other

Table 1
Descriptive statistics of the logarithmic returns of gold prices.

Descriptive statistics of the logarithmic returns of gold prices									
	Mean (%)	Standard deviation (%)	Min (%)	Max (%)	Skewness	Kurtosis	Normality test	ADF	LM(12)
Spot	0.0424	1.1682	−7.80	9.77	−0.10	8.44	4743.55	−62.80	276.42
Future	0.0444	1.2073	−9.82	8.89	−0.16	8.88	5308.74	−60.26	179.70

Note: The Normality Test is the Jarque–Bera test which has a $\chi^2(q)$ distribution with 2 degrees of freedom under the null hypothesis of normally distributed error. The 5% critical value is therefore 5.99. The stationarity test used is the ADF, and its critical value at 5% is −2,86. The LM(12) statistic is the ARCH LM test up to the twelfth lag and under the null hypothesis of no ARCH effects it has a $\chi^2(q)$ distribution where q is the number of lags. The 5% critical value is 21.03.

Table 2
Descriptive Statistics of the Independent Variables.

Variable	Mean (%)	Standard deviation (%)	Min (%)	Max (%)	Skewness	Kurtosis	Normality test	ADF
DJIA	0.01050	1.1911	−8.201	10.508	−0.0627	11.16	10595.47	−47.745
FTSE	0.00157	1.2319	−9.263	9.384	−0.1528	9.32	6376.07	−29.861
EUR	0.00691	0.4505	−2.207	2.532	−0.0591	4.97	6198.33	−24.583
JPY	−0.00183	0.7163	−6.168	3.662	−0.9282	9.30	6860.10	−12.342
Oil	0.04001	2.3870	−16.551	21.277	−0.0956	8.37	4602.85	−62.789
Sq. Gold Spot	0.0137	0.0373	0.000	0.9550	10.1951	170.93	4554529	−19.369
Sq. Gold Future	0.0146	0.0409	0.000	0.9644	11.1442	186.837	5236799	−11.769

Sq. Gold Spot is the square of gold spot price return and Sq. Gold Future is the square of gold spot future return. *Note:* The Normality Test is the Jarque–Bera test which has a $\chi^2(q)$ distribution with 2 degrees of freedom under the null hypothesis of normally distributed error. The 5% critical value is therefore 5.99. The stationarity test used is the ADF, and its critical value at 5% is −2,86.

Table 3
Correlation matrix between the independent variables and the volatility for different horizons.

	Independent variables						
Volatility	GARCH	DJIA	FTSE	EUR	JPY	Oil	Sq. gold price
14 d							
Spot	0.2561	−0.0695	−0.0734	−0.0535	−0.0404	−0.0434	0.2336
Future	0.1857	−0.0582	−0.0618	−0.0461	−0.0360	0.0166	0.2035
21 d							
Spot	0.2436	−0.0709	−0.0723	−0.0432	−0.0461	−0.0415	0.2398
Future	0.1377	−0.0617	−0.0658	−0.0349	−0.0385	0.0119	0.2142
28 d							
Spot	0.2194	−0.0618	−0.0664	−0.0485	−0.0447	−0.0430	0.2425
Future	0.1135	−0.0512	−0.0556	−0.0438	−0.0351	0.0079	0.2139

GARCH means the GARCH model forecast and it is included because is one of the inputs for the ANN. The squared gold price depends on the price analyzed, spot or future, to calculate the correlations.

variables stepwise according to the correlation criteria, using the initial parameters in the ANN. The GARCH model applied to forecast the gold spot price volatility has a MAPE equal to 0.8664, while for the case of prediction of gold future price volatility the MAPE is 1.0700. The initial results are such that there is a great deal of opportunity for improvement. The results demonstrate that the GARCH model applied to predict the gold volatility has an error in the forecasting which could potentially be reduced using the hybrid ANN–GARCH. Table 4 illustrates the improvement in prediction accuracy. The results in Table 4 demonstrate that all combinations for ANN–GARCH for the two volatilities have reduced MAPE. In particular, the best result for the gold spot price volatility is the model ANN-GARH that includes all the variables. In this case, the MAPE is reduced by 25% as compared to the GARCH forecast, an important improvement for forecasting because of the greater precision. For the gold future price volatility model, the ANN–GARCH with all variables is the best model, reducing the MAPE by 38%. The improvements demonstrated by the ANN–GARCH method means a greater precision in forecasting than the spot price volatility. The results are presented in the Table 4. In most of the cases the MAD and MSD increase, indicating a better GARCH forecast in the periods with high volatility.

When the forecast horizon is changed, the GARCH model MAPE is 0.9834 for 14 d and 0.8195 for 28 d in the case of gold spot price volatility and 1.0507 and 1.0747 respectively in the gold future price volatility. These results confirm that the gold future price volatility is harder to predict than the gold spot price volatility. When the hybrid model is applied, the results show that for the gold spot price volatility the best model incorporates four variables (leaving out the DJI) for both 14-day horizons. The model with only two variables (Euro/Dollar and Yen/dollar) is best for the 28-day

Table 4
Performance results for forecast models.

Models	Spot price volatility (21 d)			Future price volatility (21 d)		
	MAPE	MAD	MSD	MAPE	MAD	MSD
GARCH	0.8664	8.2E−05	1.8E−08	1.0700	9.2E−05	2.7E−08
ANN–GARCH 1	0.7165	9.6E−05	2.6E−08	0.7155	9.9E−05	2.4E−08
ANN–GARCH 2	0.7133	8.9E−05	2.0E−08	0.7009	9.8E−05	2.4E−08
ANN–GARCH 3	0.7084	9.5E−05	2.5E−08	0.7133	9.4E−05	2.3E−08
ANN–GARCH 4	0.6938	9.2E−05	2.3E−08	0.7072	1.0E−04	2.7E−08
ANN–GARCH 5	0.6493	8.8E−05	2.1E−08	0.6621	9.6E−05	2.4E−08

All the ANN–GARCH models have as input the square gold price volatility and the GARCH forecast. ANN–GARCH 1 uses as extra input Euro/Dollar, ANN–GARCH 2 adds the Yen/Dollar over the inputs of ANN–GARCH 1. ANN–GARCH 3 adds FTSE, ANN–GARCH incorporates Oil price and ANN–GARCH adds the DJI completing all variables as input. For each model the number of forecasts is 3,383 for the spot case and 3,188 for the future analysis.

horizon, but its performance is very similar to the model with the same four variables. For the gold future price volatility case the best models continue being the ANN–GARCH with all variables as input. The improvement in the MAPE for gold spot price volatility is 26% and 18%, for the horizon of 14 d and 28 d, respectively. In the case of the future price volatility forecast, the MAPE is reduced by 31% and 42% for the horizons of 14 d and of 28 d. The results are presented in the Table 5.

To test the robustness of the results, the forecasts are also done with a longer rolling windows of 504 d for the three horizons, that is approximately two years, keeping all the others parameters at their initial values. In the cases of the 21-day and 28-day spot price volatility forecasts, the best results are reached with the five variables. For the 14-day forecast the optimal model includes Euro/Dollar, Yen/Dollar and FTSE for the spot price volatility and future price volatility. In the case of 21-day gold future price volatility, the results for the models with three, four and five variables are very similar and for the 28-day forecast horizon, it is clear that the best performance is for the model with all variables. The results can see in Table 6.

Finally, to check whether the sort chosen in the first model, which was related to the correlation, influenced the results obtained and as an innovation in the methodology, the models are re-calculated following a new algorithm. The algorithm starts calculating the models with the two fixed variables and one of the five variables (Euro/Dollar, Yen/Dollar, FTSE, Oil price and DJI); thus in the first round five models are calculated. The variable associated with the best of these five models is then incorporated as fixed for the next round. In the second round the models have three fixed variables and new ones are calculated adding one of the four remaining variables. From the best of these, the next variable is selected to add as fixed and the third round is started. With this algorithm the variables are incorporated into the models as they improve the model.

For series, spot and future, the models for the first round had two fixed variables, the GARCH forecast and the square price return. For the spot price volatility in the first round the best forecast results are reached with the model including the DJI; then for the second round the DJI is incorporated as fixed and the best model in the second round includes the oil price return. In the third round keeping as fixed the GARCH forecasts, square price return, DJI and the oil price returns, the best result is for the model which adds the Yen/Dollar exchange rate, but the MAPE is not improved. However, the fourth round is calculated because the third round has not yet reached the performance obtained by the model with all variables (Table 4). In the fourth round including the FTSE, the MAPE improves compared with the third round MAPE, but it does not reach the results with all variables, confirming the finding that the model with the all the variables is the best model to forecast through ANN–GARCH the gold spot price volatility. The model with the lowest MAPE for future price volatility in the first round is the one that includes the FTSE. Given the second round results, the variable to add as fixed is the Euro/Dollar exchange rate and in the third round the variable selected is oil price returns. As the spot models in the third round do not improve the MAPE, the fourth round is calculated. In the fourth round the MAPEs also do not improve when the remaining variable is added, being the same phenomenon than the spot price volatility. The results for the different round for spot and future price are presented in the Table 7.

Table 5
Performance results for forecast models for different forecast horizons.

Models	Spot price volatility		Future price volatility	
	14 d	28 d	14 d	28 d
GARCH	0.9834	0.8195	1.0507	1.0747
ANN–GARCH 1	0.7519	0.7275	0.7557	0.7238
ANN–GARCH 2	0.7596	0.6693	0.7890	0.7513
ANN–GARCH 3	0.7753	0.6830	0.7230	0.7260
ANN–GARCH 4	0.7148	0.6717	0.7504	0.6397
ANN–GARCH 5	0.7307	0.6748	0.7222	0.6353

All the ANN–GARCH models have as input the square gold price volatility and the GARCH forecast. ANN–GARCH 1 uses as extra input Euro/Dollar, ANN–GARCH 2 adds the Yen/Dollar over the inputs of ANN–GARCH 1. ANN–GARCH 3 adds FTSE, ANN–GARCH incorporates Oil price and ANN–GARCH adds the DJI completing all variables as input. For each model the number of forecasts is 3,391 (14 d) and 3,377 (28 d) for the spot case and 3,186 (14 d) and 3,172 (28 d) for the future analysis.

Table 6
Performance results for forecast models for different forecast horizons with two years of data in the input.

Models	Spot price volatility			Future price volatility		
	14 d	21 d	28 d	14 d	21 d	28 d
GARCH	0.8780	0.7702	0.7195	0.8466	0.7482	0.7026
ANN–GARCH 1	0.6699	0.7280	0.6397	0.7425	0.6913	0.7305
ANN–GARCH 2	0.7196	0.6367	0.6871	0.6282	0.6375	0.6168
ANN–GARCH 3	0.6394	0.6444	0.6184	0.5826	0.5813	0.6000
ANN–GARCH 4	0.6685	0.6371	0.6177	0.6348	0.5840	0.5932
ANN–GARCH 5	0.6571	0.6180	0.6012	0.6548	0.5875	0.5600

All the ANN–GARCH models have as input the square gold price volatility and the GARCH forecast. ANN–GARCH 1 uses as extra input Euro/Dollar. ANN–GARCH 2 adds the Yen/Dollar beyond the inputs of ANN–GARCH 1. ANN–GARCH 3 adds FTSE, ANN–GARCH 4 incorporates Oil price and ANN–GARCH 5 adds the DJI, completing all variables as input. For each model the number of forecasts is 3,124 (14 d), 3,115 (21 d) and 3,110 (28 d) for the spot case and 2,986 (14 d), 2,980 (21 d) and 2,973 (28 d) for the future analysis.

Table 7
Performance results for forecast models by rounds.

Gold price		Round			
Volatility	Variable	First	Second	Third	Forth
Spot	EUR/USD	0.7165	0.7352	0.7262	0.7581
	USD/JPY	0.7985	0.7119	0.7180	–
	FTSE	0.6973	0.7367	0.7197	0.7086
	Oil	0.6937	0.6908	–	–
	DJI	0.6923	–	–	–
Future	EUR/USD	0.7155	0.6376	–	–
	USD/JPY	0.6967	0.7468	0.6950	0.6975
	FTSE	0.6638	–	–	–
	Oil	0.8713	0.7134	0.6607	–
	DJI	0.7284	0.6886	0.7606	0.6959

5. Conclusions

This paper has two main contributions. The first contribution is the demonstration o the hybrid model ANN–GARCH capability to forecast volatility is confirmed. Through the examination of the method's predictions of gold price volatility we were able to demonstrate improvements over classical means of forecasting. Second, we demonstrated an innovative method to determine which financial variables are the most important in affecting the volatility of gold spot prices and future prices.

The results show that the ANN–GARCH model improves the forecast results as compared with that of the GARCH model by 25% for the gold spot price volatility and by 38% for the gold future price volatility. The best results were found in the 21-day volatility forecasts when including the Euro/Dollar, Yen/dollar, FTSE variation, DJI variation and oil price returns as input variables to the ANN. Additionally, for forecasts of 14-day and 28-day future price volatility, the results show the best performance is when all the variables are included in the model.

When sensitivity analysis is applied over the parameters, the results tend to be the same. One analysis was to change the order of entry of the variables to model, changing the correlation criteria by a multi-round analysis. The conclusion was the same – the model ANN–GARCH with all variables is the best predictor. The stability of the results suggests that the model will hold well out-of-sample. In the analysis with 504 d of data as input, the results for 21-day and 28-day spot and future price volatility confirms that the model with all variables has the best performance. Only for the 14-day volatility forecast do the results indicate that the model with Euro/Dollar, Yen/Dollar and FTSE (and without the DJI and oil prices) is the best model to forecast the spot price volatility and future price volatility.

Given the results, it can be concluded that it is possible to improve the GARCH forecasting method with an ANN–GARCH model for all horizons and for the spot and future gold price volatility. Also, it is possible to determine the main financial variables that influence the gold price volatility. These findings open the opportunity to test if all variable that influence in the classical model fit (in-sample) are those that influence in the forecasting (out-of-sample).

The ANN is an excellent complement used to improve the effectiveness of the GARCH model to forecast the volatility, defining a hybrid model ANN–GARCH. Also, applying this model it is possible to detect the influence of financial variables important to predict the volatility, focusing in the contribution to explain the behavior out-of-sample and not in-sample as in the classical model fit. Increasing the number of variable does not necessarily increase the performance of the forecasts. Therefore, the ANN–GARCH provides not only improved forecast but appears more parsimonious than other approaches.

The limitations of the proposed model are related to performance of the forecasting which are based mainly in the quality of the data used as input. The results and relationships depend on the historic data as all forecasting models. While there is no perfect forecasting model, the ANN–GARCH incorporates the heteroskedastic concept and the learning from past prediction errors. Another limitation relates to the amount of data necessary to predict the volatility for short term (14, 21 and 28 d) given that 252 d were used. Finally, while this study employed backpropagation as characteristic of the ANN, the results may be improved using other characteristics of learning and feedback which may serve as basis for a new research.

Future research may want to analyze new hybrid models to predict the volatility, incorporating fuzzy logic. Additionally, it may be beneficial to add a step to detect in each iteration the best GARCH model, analyzing EGARCH, IGARCH, T-GARCH, GJR.-GARCH, and other approaches. One may also want to incorporate a Markov switching states to feed the ANN. Finally, given the benefits of ANFIS, one may want to consider the hybrid ANFIS–GARCH model. The ANFIS would have the capability to integrate and simulate knowledge from quantitative and qualitative source to model behavior changes in the volatility which may result in a better model to predict volatility.

References

Azadeh, A., Moghaddam, M., Khakzad, M., & Ebrahimipour, V. (2012). A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting. *Computers and Industrial Engineering, 62*(2), 421–430.

Batten, J. A., Ciner, C., & Lucey, B. M. (2010). The macroeconomic determinants of volatility in precious metals markets. *Resources Policy, 35*(2), 65–71.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics, 31*(3), 307–327.

Boyacioglu, M. A., & Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange. *Expert Systems with Applications, 37*(12), 7908–7912.

Creti, A., Joëts, M., & Mignon, V. (2013). On the links between stock and commodity markets' volatility. *Energy Economics, 37*, 16–28.

Efimova, O., & Serletis, A. (2014). Energy markets volatility modelling using GARCH. *Energy Economics, 43*, 264–273.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business Economic Statistics, 20*(3), 339–350.

Elder, J., Miao, H., & Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking and Finance, 36*(1), 51–65.

Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and predicting the CBOE market volatility index. *Journal of Banking and Finance, 40*, 1–10.

Grudnitski, G., & Osburn, L. (1993). Forecasting S&P and gold futures prices: an application of neural networks. *Journal of Futures Markets, 13*(6), 631–643.

Hamid, S. A., & Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index futures prices. *Journal of Business Research, 57*(10), 1116–1125.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014). Forecasting volatility of the US oil market. *Journal of Banking and Finance, 47*, 1–14.

Kroner, K. F., Kneafsey, K. P., & Claessens, S. (1995). Forecasting volatility in commodity markets. *Journal of Forecasting, 14*(2), 77–95.

Kristjanpoller, W., Fadic, A., & Minutolo, M. (2014). Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications, 41*(5), 2437–2442.

Monfared, S. A., & Enke, D. (2014). Volatility forecasting using a hybrid GJR–GARCH neural network model. *Procedia Computer Science, 36*, 246–253.

Muzzioli, S., Ruggieri, A., & De Baets, B. (2014). A comparison of fuzzy regression methods for the estimation of the implied volatility smile function. *Fuzzy Sets and Systems, 2*(6), 131–143.

Neely, C. J. (2003). Implied Volatility from Options on Gold Futures: Do Econometric Forecasts Add Value or Simply Paint the Lilly?. *Federal Reserve Bank of St. Louis Working Paper Series*, (2003-018).

Parisi, A., Parisi, F., & Díaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. *Journal of Multinational financial management, 18*(5), 477–487.

Svalina, I., Galzina, V., Lujić, R., & Šimunović, G. (2013). An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert systems with applications, 40*(15), 6055–6063.

Shafiee, S., & Topal, E. (2010). An overview of global gold market and gold price forecasting. *Resources Policy, 35*(3), 178–189.

Szakmary, A., Ors, E., Kyoung Kim, J., & Davidson, W. N. III, (2003). The predictive power of implied volatility: Evidence from 35 futures markets. *Journal of Banking and Finance, 27*(11), 2151–2175.

Trück, S., & Liang, K. (2012). Modelling and forecasting volatility in the gold market. *International Journal of Banking and Finance, 9*(1), 3.

Tully, E., & Lucey, B. M. (2007). A power GARCH examination of the gold market. *Research in International Business and Finance, 21*(2), 316–325.

Vortelinos, D. I. (2015). Forecasting realized volatility: HAR against Principal Components Combining, neural networks and GARCH. *Research in International Business and Finance*.

Yazdani-Chamzini, A., Yakhchali, S. H., Volungevičienė, D., & Zavadskas, E. K. (2012). Forecasting gold price changes by using adaptive network fuzzy inference system. *Journal of Business Economics and Management, 13*(5), 994–1010.