Stock Indices Analysis Based on ARMA-GARCH Model

Weiqiang Wang[1], Ying Guo[2] ,Zhendong Niu[1], Yujuan Cao[1]
[1] School of computer science, Beijing Institute of Technology, Beijing, China
[2] School of Management and Economics, Beijing Institute of Technology, Beijing, China
(weiqiang.wang1983@gmail.com)

*Abstract–***The generalized autoregressive conditional heteroskedasticity (GARCH) model has become the most popular choice in the analysis of time series datas. In this paper, an autoregressive moving average （ARMA）- GARCH model was built, and it also provided parameter estimation，diagnostic checking procedures to model, and predict Dow and S&P 500 indices data from 1988 to 2008,which extracted from yahoo website, and also compared with the GARCH conventional model, experimental results with both two data sets indicated that this model can be an effective way in financial area.**

Keywords – **ARMA-GARCH model, time series, DOW,S&P 500**

I. INTRODUCTION

The Dow Jones Industrial Average is one of several stock market indices, created by nineteenth-century Wall Street Journal editor and Dow Jones & Company co-founder Charles Dow. It is an indices that shows how certain stocks have traded. It is the second-oldest U.S. market indices.

The S&P 500 is a value weighted indices published since 1957 of the prices of 500 large-cap common stocks actively traded in the United States. Almost all of the stocks included in the indices are among the 500 American stocks with the largest market capitalizations.

During the past three decades, Two major classes of models were studied by econometricians for the purpose of forecasting. They are the statistical time series models and structural econometric models. Linear time series models such as the Box-Jenkins autoregressive integrated moving average (ARIMA) models were among the first to be developed and subsequently widely studied. Despite its simplicity and versatility in modelling several types of linear relationship such as pure autoregressive, pure moving average and autoregressive moving average (ARMA) series[1], such type of models was constrained by its linear scope.

The nonlinear serially dependent ARCH/GARCH and EGRACH group of models is widely accepted among econometricians and time series statisticians as the premier model of stock market returns, especially so for the GARCH(1,1) model. This wide acceptance rests on two bodies of empirical evidence. First, a number of statistical tests easily reject the null hypothesis of a linear process; this evidence against a linear process has been accumulating since the mid-1980s. Second, the parameter estimates of a GARCH(1,1) process are statistically significant when a model is estimated on various examples of realized stock market returns and individual stock issues. This statistical significance of the parameter estimates is apparently sufficient evidence for the vast majority of empirical investigators to accept these models as true.

Although ARIMA model takes care of the nonstationary behavior of stock, models such as ARMA, ARIMA have a constant variance, and thus cannot capture the suddenly nature of the stock, which is a very important characteristics of the stock indices.

This article is structured as follows: the next section we provide a brief introduction to the ARCH, GARCH, and ARMA-GARCH model, the next presents the two datasets from yahoo website in Section 3. In section 4 we show how we can fit a new model by ARMA-GARCH model, and show parameter estimation and diagnostic checking of ARMA-GARCH model. In section 5 we do the prediction of the model, Finally, a summary concludes this article.

II. MODEL

A. ARCH Model

In econometrics, an autoregressive conditional heteroscedasticity(ARCH) [2] model considers the variance of the current error term to be a function of the variances of the previous time periods' error terms.

Specifically, let ε_t denote the returns (or return residuals, net of a mean process) and assume that, $\varepsilon_t = \sigma_t z_t$ where $z_t \sim iidN(0,1)$ and where the series σ^2 are modeled by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 \qquad (1)$$

and where $\alpha_0 > 0$ and $\alpha_i \geq 0, i > 0$.

B. GARCH Model

GARCH model has become important in the analysis of time series data, particularly in financial applications when the goal is to analyze and forecast volatility[3].

If an ARMA model is assumed for the error variance, the model is a GARCH model.

The GARCH model is an extension of the ARCH model. A GARCH model with order p $\geqslant$ 0 and q $\geqslant$ 0 is defined as:

$$Z_t = \sigma_t \varepsilon_t \qquad (2)$$

In that case, the GARCH(p, q) model (where p is the order of the GARCH terms σ^2 and q is the order of the ARCH terms ε^2) is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_p\varepsilon_{t-p}^2 + \beta_1\sigma_{t-1}^2 + ... + \beta_q\sigma_{t-q}^2 \quad (3)$$

Generally, when testing for heteroskedasticity in econometric models, the best test is the White test. However, when dealing with time series data, the means to test for ARCH errors and GARCH errors [4].

C. ARMA-GARCH model

The mixture of ARMA-GARCH model is similar to the mixture of AR-GARCH model proposed in [5]. We can see each component of the mixture model can be denoted as a normal ARMA series:

$$y_{t,j} = \sum_{r=1}^{R} b_{rj} y_{t-r,j} + \sum_{s=1}^{S} a_{sj}\varepsilon_{t-s,j} + \varepsilon_{t,j} \quad (4)$$

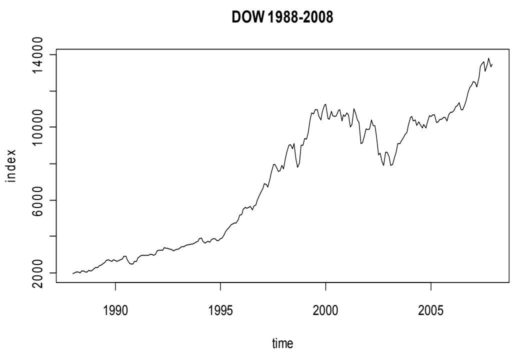
Furthermore, each residual term $\varepsilon_{t,j}$ is assumed Gaussian white noise with variance denoted by the GARCH model.

$$\sigma_{t,j}^2 = \delta_{0j} + \sum_{q=1}^{Q} \delta_{qj}\varepsilon_{t-q,j}^2 + \sum_{p=1}^{P} \beta_{pj}\sigma_{t-p,j}^2 \quad (5)$$

where δ_{0j} ψ>ψ0 for qψ=‑1, ‹Qψand β_{pj} >ψ0 for pψ=‑1‹ ‹ψP.

III. Data

We consider daily prices on two stock market indices, namely the Standard & Poor's 500(S&P 500) indices and the Dow Jones Industrial Average indices. We have data from 01/01/1988 to 01/01/2008, which results in 5044 daily observations. Both daily indices are plotted in Fig 1.



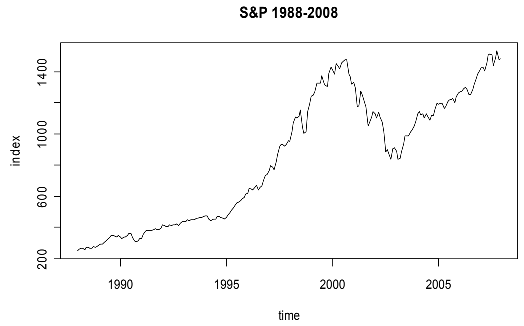


Fig. 1. Dow and S&P daily price

From Fig 1,for the Dow Jones indices we can see there is obviously a strong upward trend except the peak period around 2000 which because The uncertainty of the 2000s brought a significant bear market, characterized first by extreme fear on the part of newer investors. In this display, observation suggests that the variation series is not constant over time and that there is a trend as well as a fluctuation and seasonal pattern in the data. This DOW time series have to be determined as non stationary.

And for the S&P 500 Indices, we also can see the upward trend before 2000, The indices reached an all-time intraday high of 1,552.87 in trading on March 24, 2000, and then lost approximately 50% of its value in a two-year bear market, spiking below 800 points in July 2002 and reaching a low of 768.63 intraday on October 10, 2002.

IV. Model fitting

From Fig 1, evidence shows that there is substantial other correlation that needs to be modeled. Clearly we need at least one order of differentiation.

A. Parameter Estimation

Applying the GARCH model and ARMA-GARCH model to our DOW and S&P data, the following table shows:

TABLE I
PARAMETER ESTIMATION

	AIC	BIC	SIC	HQIC
S&P Garch(1,1)	14.13	14.19	14.13	14.16
S&ARMA(1,1)+ Garch(1,1)	9.08	9.16	9.08	9.11
DOW Garch(1,1)	18.28	18.33	18.28	18.30
DOWARMA(1,1)- Garch(1,1)	13.38	13.46	13.38	13.41

For the model identification, we used statistics software R, it can be chosen using Akaike Information Criteria[6] (AIC) and Bayesian Information Criteria (BIC),and SIC etc. One point we should emphasize is that metrics like AIC and BIC not only evaluate the fit between values predicted by the model and actual measurements, but also penalize models with larger number of parameters.

From the results of the AIC and BIC, we can see ARMA-GARCH model is obviously better and suitable than GARCH model.

Thus, comparing with the conventional GARCH model, our model should be ARMA(1,1)-GARCH(1,1),and the coefficient are as follows:

TABLE II
COEFFIENCIENT OF ARMA-GARCH MODEL

DOW	mu	ar1	ma1	omega	alpha1	beta1
	31.1	1.0	0.2	461.3	0.2	0.8

S&P	mu	ar1	ma1	omega	alpha1	beta1
	3.9	1.0	0.22	4.6	0.23	0.81

B. Diagnostic Checking

Diagnostic checking is necessary to ensure the best forecasting model has been built. The coefficient estimates are highly significant and the estimated value is small. However, in order to check the estimated ARMA-GARCH model[7], there are some other things to check. Let's look at the time series plot of ACF first.

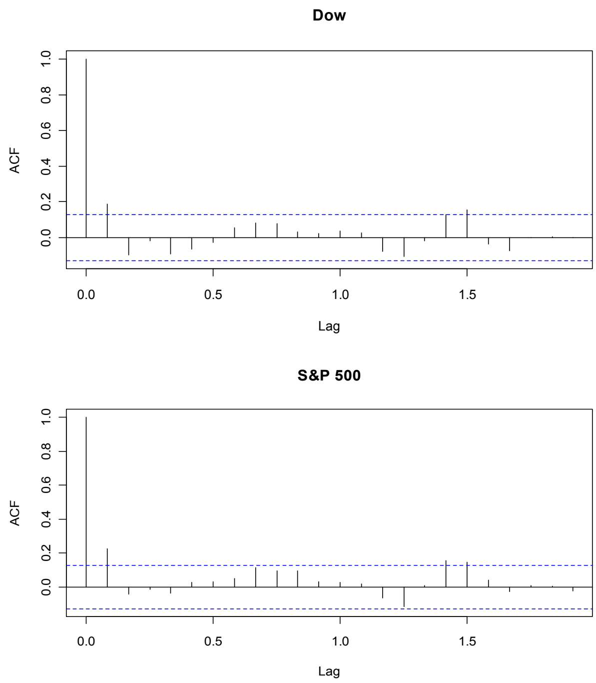


Fig. 2.ACF of Dow and S&P 500 indices

To take a further look, we can see the residuals as follows:

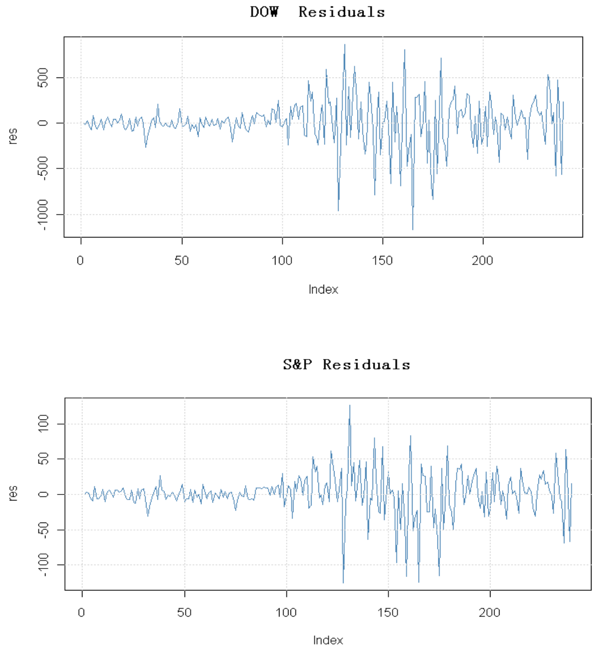


Fig.3. Residuals of Dow and S&P 500 indices

From the Fig 3, we can see the residuals are fluctuation around zero, results are all in the reasonable intervals.

To evaluate the goodness of fit for the GARCH model, we plotted the Histogram and Q-Q (t-distributed) plot of the GARCH residuals.

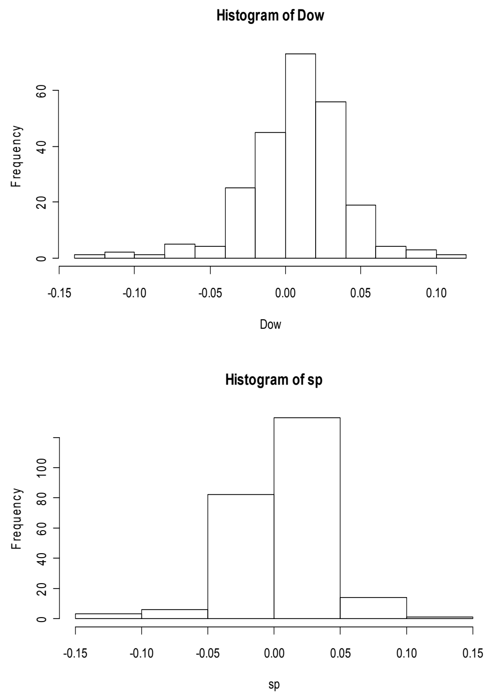


Fig.4. Histogram of Dow and S&P 500 indices

By plotting the histogram of the residual we can see the center totally nearby zero, and the shape of histogram appears "good shaped".

Normally, the most common test of normality is Q-Q plot. we would like to determine if outliers exist. We can take a look at the normal Q-Q plot formed by residuals, from the Q-Q plot below, we can see there are not many outliers, and almost all the points are laid on the line.

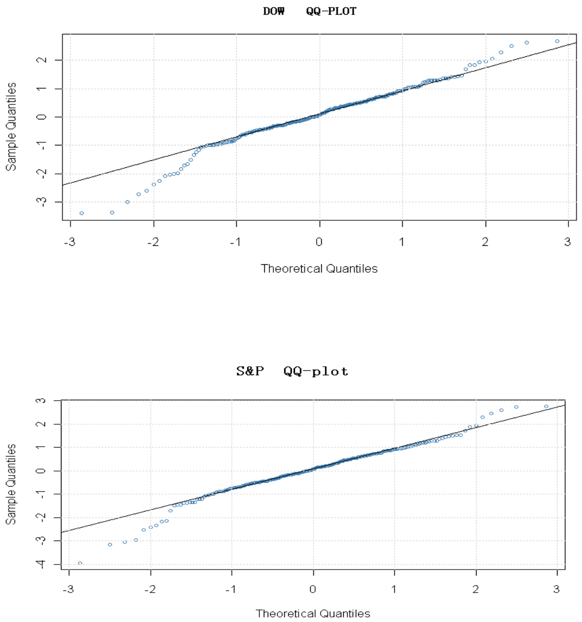


Fig. 5. QQ-PLOT of Dow and S&P 500 indices

TABLE III
Standardised Residuals Tests

DOW			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	32.73165	7.80573e-08
Shapiro-Wilk Test	R	W	0.9699429	5.80263e-05
Ljung-Box Test	R	Q(10)	12.66138	0.2432256
Ljung-Box Test	R	Q(15)	15.99925	0.3821015
Ljung-Box Test	R	Q(20)	26.12278	0.1617951
Ljung-Box Test	R^2	Q(10)	5.925618	0.8214733
Ljung-Box Test	R^2	Q(15)	12.30569	0.6557585
Ljung-Box Test	R^2	Q(20)	16.77428	0.6675783
LM Arch Test	R	TR^2	8.500406	0.7449056

S&P 500			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	45.12342	1.590645e-10
Shapiro-Wilk Test	R	W	0.9687331	3.974364e-05
Ljung-Box Test	R	Q(10)	10.14911	0.4275105
Ljung-Box Test	R	Q(15)	13.35924	0.5745714
Ljung-Box Test	R	Q(20)	21.20197	0.3853318
Ljung-Box Test	R^2	Q(10)	9.062306	0.5262009
Ljung-Box Test	R^2	Q(15)	15.88776	0.3895499
Ljung-Box Test	R^2	Q(20)	18.19338	0.5746709
LM Arch Test	R	TR^2	9.176007	0.6878282

From Table III, we can see the Jarque-Bera[7] and the Shapiro-Wilk test proved normal distributed residuals performed normal, the Ljung-Box test[11] proved the residuals and squared residuals have significant autocorrelations, and the Lagrange-Multiplier ARCH test[8] test the residuals have conditional heteroskedasticity.

Above at all, through the diagnostic checking including the residual, ACF , Histogram ,QQ-plot of the residuals, the experiment results indicate that model expresses very well. We could conclude that stock indices can be represent very well by ARMA(1,1)-GARCH(1,1) model.

V. Prediction of ARMA-GARCH model

One of the major aspects in the investigation of heteroskedastic time series is to produce forecasts.

For a GARCH(p,q) process, the h-step-ahead forecast of the conditional variance $\hat{\omega}^2_{t+h|t}$ is computed recursively from:

$$\hat{\sigma}^2_{t+h|t} = \hat{\omega} + \sum_{i=1}^{q} \hat{\alpha}_i \varepsilon^2_{t+h-i|t} + \sum_{j=1}^{p} \hat{\beta}_j \sigma^2_{t+h-j|t} \qquad (6)$$

Where $\varepsilon^2_{t+i|t} = \sigma^2_{t+i|t}$ for i>0 while $\varepsilon^2_{t+i|t} = \varepsilon^2_{t+i}$ and $\sigma^2_{t+i|t} = \sigma^2_{t+i}$ for i<0[9].

Based on the parameters estimated from the time series using the ARMA-GARCH model and also on the information obtained from the last time instant of the stock data, we proceed to forecast the stock inedx for the next time instant[10].

The following result shows the 10 step ahead forecast for the DOW and S&P 500 indices modeled by ARMA(1,1)- GARCH(1,1) model.

TABLE IV
Forecast of Dow and S&P 500 indices

DOW	MeanForecast	meanError	SD
1	11845.88	2581.861	1691.019
2	10303.48	3425.319	1688.436
3	10303.48	3425.319	1685.867
4	10303.48	3425.319	1683.313
5	10303.48	3425.319	1680.774
6	10303.48	3425.319	1678.250
7	10303.48	3425.319	1675.740
8	10303.48	3425.319	1673.245
9	10303.48	3425.319	1670.764
10	10303.48	3425.319	1668.297

S&P	MeanForecast	meanError	SD

1	1292.661	258.0161	207.1376
2	1102.370	347.5169	208.0580
3	1102.370	347.5169	208.9755
4	1102.370	347.5169	209.8902
5	1102.370	347.5169	210.8022
6	1102.370	347.5169	211.7115
7	1102.370	347.5169	212.6181
8	1102.370	347.5169	213.5221
9	1102.370	347.5169	214.4234
10	1102.370	347.5169	215.3222

Results show that ARMA(1,1)-GARCH(1,1) model provides excellent prediction, and this model is good fitting in our stock indices datasets.

VI. CONCLUSION

In this paper, we presented a ARMA(1,1)-GARCH(1,1) model which has been fitted and provide parameter estimation, diagnostic checking procedures to this model, and predict Dow and S&P 500 indices data extracted from yahoo website, and also compare with conventional GARCH model, the experimental results reaffirm that the ARMA-GARCH model is better fitted than other models for our data. We observed that ARMA(1,1)-GARCH(1,1) model fitted the DOW and S&P 500 indices data very well and this is confirmed by the goodness of fit test which is given by the Histogram, Q-Q plot with the good squared correlation R2.A forecasting of our datasets illustrate that the model will be helpful to predict the Dow and S&P 500 composite price indices. So we can see ARMA-GARCH model has board applicability in the financial field.

REFERENCES

[1] . H. Y. Kwok, C. M. Chen, and L. Xu, “Comparison between mixture of ARMA and mixture of AR model with application to time series forecasting,” in Proceedings of Fifth International Conference on Neural Information Processing, 1998, pp. 1049–1052.

[2].Robert F. Engle. "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation", Econometrica 50:987-1008, 1982. (the paper which sparked the general interest in ARCH models)

[3]. Ling S., McAleer M., (2002); Necessary and Sufficient Moment Conditions for theGARCH(r,s) and Asymmetric Power GARCH(r,s) Models, Econometric Theory 18, 722–729.

[4]. R. A. Davis P. J. Brockwell, Introduction to Time Series and Forecasting, Springer, New York, 2002.

[5]. Lanne, M., and P. Saikkonen (2007): “A multivariate generalized orthogonal factor GARCH model,” Journal of Business and Economic Statistics, 25, 61–75.

[6]. He, C., and T. Ter¨asvirta (2004): “An extended constant conditional correlation GARCH model and its fourth-moment structure,” Econometric Theory, 20, 904–926.

[7]. Engle R.F., (2001); GARCH 101: An Introduction to the Use of ARCH/GARCH models in Applied Econometrics, forthcoming Journal of Economic Perspectives.

[8]. Brooks C., Burke S.P., Persand G., (2001); Benchmarks and the Accuracy of GARCH Model Estimation, International Journal of Forecasting 17, 45-56.

[9]. Bollerslev T., Engle R.F., NelsonD.B., (1994); ARCH Model, Handbook of Econometrics Volume IV, Chapter 49, 2961–3031 edited by Engle and McFadden, Elsevier Science.

[10]. McLeod, A. I. and Li, W. K. (1983) “Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations” Journal of Time Series Analysis 4, 269-73.

[11]. Bollerslev T., Chou R.Y., Kroner K.F., (1992); ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence; Journal of Econometrics 52(5), 5–59.