Tridib Bose

Writing Sample

Volatility Methodologies: A Review

Abstract

Many important time-series economic variables require special care in modeling if they suffer from volatility. Among the advancements made in time-series and financial econometrics over the past five decades, estimation techniques dedicated to measuring and forecasting volatile variables have received remarkable attention. This article evaluates the evolution of such methodologies by studying the research done in this field. The theoretical developments considered here are certainly motivated by the technological advancements in computation and information explosion of the past decades. The future pathway seems to be in the same direction.

Keys: *volatility*, *ARCH*, *GARCH*, *Stochastic volatility*, *Realized volatility*, *RiskMetrics*, *Artificial Neural Network*.

Introduction

"A sign . . . can be seen in the development of standard econometric textbooks. The early versions, available in the 1970s and well into the 1980s and sometimes beyond, would make virtually no mention of time series methods, apart from a possible brief mention of an AR(1) model or a linear trend. Today many textbooks cover almost nothing but time series methods with considerable attention being paid to ARCH, cointegration, fraction integration, nonlinear models including neural networks, White robust standard errors, regime switching models, and causality . . ."

- Clive W. J. Granger (2010, p. 9)[6]

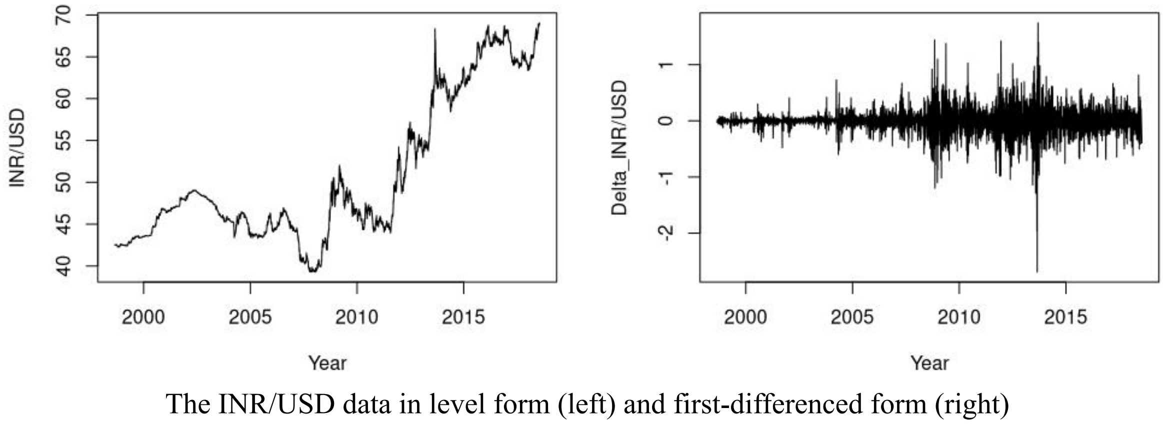
The estimation methods of volatile variables have undergone a fascinating development over the past five decades. Although the field of time-series econometrics is enriched because of advancements made in computation and data-collection of decreasing interval lengths, the credit goes to the econometricians for factoring in volatility into the theory. Especially in the case of financial economics where most of the data suffer from time-varying variance, modeling volatility is very helpful to estimate the associated risk. The discussion here focuses on the development of these methodologies.

While the developments in the field of time-series econometrics began around a hundred years ago, it wasn't until 1970, when Box-Jenkins published their book[3], and the door to several forecasting applications was open. They provided an iterative method to estimate the data-generating process by including identification before and diagnostic after the estimation. However, addressing volatility in the modern sense began only after 1982, when Engel introduced the ARCH model.[1] It established a new platform for analyzing volatile data with a novel outlook in viewing volatility estimation as a separate modeling process.

The body of this article is divided into 6-parts, preceded by an introduction and succeeded by a conclusion. The purpose of this article is to identify the factors behind the evolution of volatility estimation methodologies, and each section explores the idea behind the reasoning to achieve it. The first section addresses ARCH and GARCH models, beginning with introducing volatility and problems associated with it. The second section addresses stochastic volatility models. The third and fourth sections briefly introduce implied volatility and address realized volatility, respectively. The fifth section addresses RiskMetrics exponential weighted moving average. The final section discusses some of the methods applied in artificial neural network applications.

1. ARCH and GARCH models

Several time-series economic variables like inflation, exchange rate, stock prices, and several other financial data follow a random-walk pattern in their level form. Their differenced form seems stationary - showing a constant mean, but the variance is unstable. The swings in the differenced forms follow a pattern that often can not be modeled with the usual ARIMA (autoregressive integrated moving average) modeling. This poses threats to the forecasting of these variables since the fluctuating variance would lead to greater uncertainty in predictions.



The INR/USD data in level form (left) and first-differenced form (right)

In such cases, the conventional heteroscedasticity tools can't be used. Heteroscedasticity is the violation of the OLS assumption of constant variance of the residual/error term. It is usually associated with the cross-sectional data and is modeled assuming the variance of the residual term being associated with a function of the independent variable(s). In doing so, one would regress the residual or residual-squared on the linear or nonlinear function of the independent variables and then perform GLS (Generalized Least Square) to obtain unbiased estimates. In the field of time-series econometrics, however, heteroscedasticity is involved quite differently.

Unlike in cross-sectional data, heteroscedasticity can be modeled in time-series data even in the case of a univariate dataset. With the ARCH model, the conditional variance of the current period is associated with the variable's values in the past period(s). The ARCH model was introduced as a "new class of stochastic processes . . . [which] are mean zero, serially uncorrelated processes with nonconstant variances conditional on the past, but constant unconditional variances."[1]

Building the ARCH model began with noticing that if the mean of an AR(1) process depends on the previous period, then the variance would also depend on the previous period. But, the conditional variance in usual econometric models doesn't exhibit such a relationship. For the AR(1) model be

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t,$$

we have

$$E(y_t) = \alpha_0 + \alpha_1 E(y_{t-1}),\ var(y_t) = \frac{\sigma_u^2}{1-\alpha_1^2}\text{, and } var(y_t|y_{t-1}) = \sigma_u^2,$$

where u_t is the $IIDN(0, \sigma_u^2)$ (independent and identically distributed random normal), or white noise, and $var(u_t) = \sigma_u^2$ is the residual variance. As can be seen, the conditional variance is constant and doesn't depend on the past values. This means that no previous period information is incorporated in the forecasting of the current period.

Afterward, a simple approach to model heteroscedasticity was presented. For the model be $y_t = u_t x_{t-1}$, where x_t represents an independent variable, we would have the mean zero and variance depending on the past period of x_t, ie. $var(y_t) = \sigma_u^2 x_{t-1}^2$. In this case, however, yet the variance changes with past period x_t, the mean is still zero. To introduce a model in which the mean and variance both are not independent of each other, a simple case of a bilinear model by Granger and Andersen[7] was presented as $y_t = u_t y_{t-1}$, where the conditional variance would be $var(y_t|y_{t-1}) = \sigma_u^2 y_{t-1}^2$. The problem with such modeling, however, is that the unconditional variance in this model is unstable: it would be either zero or infinity.

The final building block for the ARCH model is presenting the required model as

$$y_t = u_t \sqrt{\beta_0 + \beta_1 y_{t-1}^2},$$

where $var(u_t) = 1$. This is an ARCH(1) process. It would have zero mean, while the conditional variance would be as

$$\sigma_t^2 = var(y_t|y_{t-1}) = \beta_0 + \beta_1 y_{t-1}^2.$$

A striking feature of such a model is that the conditional variance of the current period varies with the value of the past period(s), due to which the forecast variance would now account for past information more appropriately. For estimation purposes, the regression model considered is

$$\sigma_t^2 = \beta_0 + \beta_1 y_{t-1}^2 + e_t,$$

where e_t is the IIDN error term. With this, the estimated coefficients could be used for hypothesis testing and verifying whether the ARCH process exists. If β_1 is found not significant, it would mean that the series is not an ARCH process. Similarly, the regression model for an ARCH(p) process would be

$$\sigma_t^2 = \beta_0 + \beta_1 y_{t-1}^2 + \beta_2 y_{t-2}^2 + \beta_3 y_{t-3}^2 + ... + \beta_p y_{t-p}^2 + e_t.$$

In practice, however, $\sigma_t^2 = var(y_t | y_{t-1}, y_{t-2}, ...)$ is not observed and squared estimated y_t $\left(\hat{y}_t^2\right)$ is used as a proxy. The estimated y_t would be the realization of a volatile random variable, and the hypothesis testing of the slope coefficients would reveal whether an ARCH process is present or not. In forecasting the UK inflation, Engle (1982) used the differenced logarithm of price (ie. the inflation indicator) to estimate the ARCH process. One may also include the independent variables in the estimation so that, along with ARCH, heteroscedasticity due to the independent variables would be treated as well.

The GARCH model was proposed by Bollerslev (1986) as a natural extension of the ARCH model. It was mentioned that the extension can be considered to resemble the extension of AR to the ARMA process. One thing to note is that including huge numbers of parameters might lead to a negative degree of freedom, which further leads to negative estimated parameter variances. The ARCH model's empirical testing often requires long lags, leading to such problems. While a fixed lag structure is imposed to combat this problem, the GARCH model extends the ARCH model allowing for a long memory as well as a more flexible lag structure. The GARCH(p,q) model is defined as

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i}^2 + \sum_{j=1}^{q} \gamma_i \sigma_{t-i}^2.$$

Afterward, the GARCH(1,1) model is presented as

$$y_t = u_t \sqrt{\beta_0 + \beta_1 y_{t-1}^2 + \gamma_1 \sigma_{t-1}^2},$$

where the conditional variance of this data-generating process is now

$$\sigma_t^2 = \beta_0 + \beta_1 y_{t-1}^2 + \gamma_1 \sigma_{t-1}^2,$$

in which the lagged conditional variance itself is included. This process is ARCH if $\gamma_1 = 0$ and is a white-noise process if $\beta_1 = \gamma_1 = 0$. Further, with some arithmetics and assumptions, the GARCH model is then reduced to a form of ARMA in y_t^2, which can then be used in estimation too (Davidson and MacKinnon (2004)). In the case of the GARCH(1,1) model, for example, taking $var(y_t) = \sigma_t^2 = y_t^2 - \varepsilon_t$, such that ε_t is the difference between y_t^2 and its conditional expectation, we have the

$$y_t^2 = \beta_0 + (\beta_1 + \gamma_1)\, y_{t-1}^2 + \varepsilon_t - \gamma_1 \varepsilon_{t-1}.$$

Since ε_t has mean zero and is serially uncorrelated[2], the conditional variance can be now considered as an ARMA process in y_t^2. After this, the model can be tested for GARCH, and even estimated with GLS. However, instead of least square estimation techniques, the maximum likelihood technique is popular and is the proposed method in such estimation.

Among many extensions of (G)ARCH model (see [6]: Chapter 8 - Glossary to ARCH (GARCH)), the Integrated GARCH (IGARCH) model[5] was introduced by Engle and Bollerslev (1986) as a particular type of GARCH model in which the sum of the coefficients of the dependent variables are restricted to be 1, ie.

$$\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \gamma_j = 1.$$

A notable feature of this model is that the multi-period forecast of the variance depends only on the one-period forecast of the variance. Another extension is the GJR-GARCH model[8], also known as GJR and sign-GARCH model, which was proposed by Glosten, Jagannathan, and Runkle (1993). It is specified as

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i}^2 + \sum_{i=1}^{p} \alpha_i I_{t-i} y_{t-i}^2 + \sum_{j=1}^{q} \gamma_i \sigma_{t-i}^2,$$

where I_t is a dummy variable such that it is 0 for positive returns (the y's) and 1 for negative returns. This attribute is added due to the finding that a negative shock affects volatility much more than a positive shock, and including I_t as an interaction variable would hence capture volatility more efficiently.

The development of time-series models concerning volatility has a quite intrinsic connection with the decades of the late 70 to the early 00s. It is interesting how beginning with efforts for a simple gathering of econometricians spawned - not only research marvels in volatility and even econometrics as a whole - but also several duplications of the effort worldwide with similar results.[6]

2. Stochastic Volatility Models

The stochastic volatility (SV) model[9] takes the volatility modeled as a logarithm of the conditional variance as a stochastic process. The process considered is

$$y_t = u_t \sigma_t,$$

such that

$$log(\sigma_t^2) = \beta * log(\sigma_{t-1}^2) + \eta_t.$$

In this case, the logarithm of the volatility parameter σ_t is an AR(1) process, meaning that the logarithm of the conditional variance depends on the logarithm of the past variance, and the

η_t is a white noise error term. However, the approach towards this kind of modeling is much more about the nature of the financial variables. Note that this is just one type of SV model.

Introducing the SV models in 1986, Taylor noted several interesting features of the financial data, on which he developed his theories. The fluctuation of the yearly and half-yearly standard deviations in different financial variables was presented. In the case of prices, a smaller standard deviation would mean that the change in prices is lower, and hence the corresponding standard deviation can be compared to check for, or *measure*, risk in two alternative investments. Also, one may take the proportion of two period's standard deviations serially, and whether their geometric mean is closer to 1 or not would imply that the standard deviations don't change very much. The skewness and kurtosis estimates of several financial variables were presented, after which it was asserted that the proposed model *should* have the probability distribution fitting such parameters (symmetric distribution, long tail, and high kurtosis). It was also asserted that the series has a non-linear data-generating process.

Before presenting such models, Taylor (1986) argued about the specific background of the financial variables. For example, the price would vary along with the activities within a market - which can be attributed to the volatility in the data. A variable in each period has a probability distribution before it is realized/observed, and the conditional standard deviation of that probability distribution is also determined by the market. For the variable in question, consider it is the realized from Y_t and its standard deviation is realized from V_t, a model proposed by several researchers is of the form

$$Y_t = \mu + V_t U_t,$$

where U_t is standard normal. This is presented as a general model for variance. Next, models capturing "variance jumps" were presented.

Furthermore, it was argued that if only extraordinary (political) events affect the variance changes, then changes in the variance would not occur often, if not repeatedly. Therefore it seems that it would be more helpful to regard V_t as non-stationary since the standard deviation in two consecutive periods is not necessarily the same. Also, a higher conditional variance would imply a higher change in the mean of Y_t. So, one must consider models for v_t (which is realized V_t) such that it is expected to change in each period, although $v_t - v_{t-1}$ is small most of the time.

The first set of models presented were assumed to have the standard deviation not dependent on the series Y_t, in which the lognormal autoregressive model stated above was presented. The lognormal distribution is rightly-skewed, which was fit for the findings in the standard deviations of several financial variables mentioned before in the book (see [9]: chapter 2). Considering $ln(V_t) \sim N(a, b^2)$, the model presented before is the simplest AR(1) form.

The second set of the models for v_t includes the term y_t and began with the ARCH model since the ARCH model includes past period squared y_t. Afterward, models in which the absolute mean deviation of y_t depends on the lagged absolute mean deviation of y_t were presented, along with the ARMACH (autoregressive moving average conditional heteroscedastic) model. The interesting thing about the SV models is that it not only encompasses the previous models in volatility but also open scope to model volatility concerning the behavior of the (conditional) standard deviation of the real-world data.

The estimation of the given SV models includes MLE, MM (method of moments), and even estimating a linearized version of the original process. Estimating parameters using a generalized method of moments produce asymptotically normal as well as consistent estimates. The ML estimation is said to have progressed very well due to developments in Monte Carlo Markov Chain procedures and numerical methods based on *importance sampling*.[10]

3. Implied Volatility

Implied volatility is a method in which a given value of a financial asset is used to determine the underlying volatility of the Black-Scholes (BS) model. The BS model begins with assuming that an asset follows a geometric (stochastic) Brownian motion, and builds an equation that gives a mathematical relationship between European option-call or option-put (C) and the spot price (S), the strike price (K), risk-free interest rate (r), and the volatility (σ) of that asset. The model has gained a significant position in financial economics since the early 70s. The BS equation to calculate the call/put option price is given as

$$C_t = S_t N(d_1) - Ke^{-r(T-t)} N(d_2),$$

where T is expiration time, and the N denotes the cumulative distribution function of a normal distribution, such that

$$d_1 = \frac{1}{\sigma_t \sqrt{T-t}} \left[ln(\frac{S_t}{K}) + (r + 0.5\sigma_t^2)(T - t) \right],$$

$$\text{and } d_2 = d_1 - \sigma_t \sqrt{T - t}.$$

While the BS equation assumes constant volatility, in later works, several researchers found this assumption to be false.[11] The volatility in this model is the conditional standard deviation of the stock's return. However, for a given option call/put price, one may reverse the BS equation to obtain estimates of the volatility parameter, which is how the implied volatility is measured.[12] Yet being a parametric method, it is unique in terms of model specification. Being entirely dependent on the BS model can also be considered as a limitation since there is no scope to expand this theoretically as in models stated before: the reliability of this method is hence dependent on the applicability of the BS model.

4. Realized Volatility

All the above-mentioned models are attempting to model volatility with parametric methods. The realized volatility (RV), also known as historical volatility, is a nonparametric approach: it doesn't assume that the time series follows any particular probability distribution for the estimating volatility. For a series be S_t, the realized volatility up to time T would be

$$\sigma = \sqrt{\frac{1}{T-1}\sum_{1}^{T}(R_t - \overline{R})^2},$$

where

$$R_t = ln(S_t/S_{t-1}),$$

and $\overline{R}$ is the sample mean of R_t. A central feature of financial returns data that reflects volatility is the second moment[14], which is the time-varying conditional variance. The RV approach exploits this fact in modeling - the low-frequency volatility estimates computed from high-frequency data.

In establishing RV theoretically, Andersen, Bollerslev, and Diebold (2003) noted that in financial literature, volatility within a particular period in high-frequency data calculated from given or realized sample variances is apparent. Several researchers used intraday data to compute the daily variance. However, the theoretical justifications for the changing volatility were absent for those measures. To make RV a theoretically sound measure of changing volatility, the connection between it as a financial tool and the theory of quadratic variation must be established.

Afterward, RV is defined as the second moment (uncentered) of a return process over a particular period, and it was shown that the estimate is unbiased and consistent with the help of the theory of quadratic variation. According to the theory of quadratic variation, the sum of the differences of subsequent observations of a stochastic process following Brownian motion would be finite even when the number of observations would increase indefinitely. The return process is indeed the difference between the natural logarithm of the prices. Along with proving the normality of the RV error term, the RV's nature of normal mixture distribution and moments of the RV measures was also discussed.

The authors concluded by remarking on the advantages of the non-parametric methods to estimate volatility and the gradual shift of researchers towards these methods from the parametric models. The reason for this shift is the availability of high-frequency data and greater efficiency in computation. Even in the high dimensional multivariate environment, non-parametric methods like RV can handle information more effectively for modeling and forecasting.

5. RiskMetrics - Exponentially Weighted Moving Average

Availability of data in smaller time intervals (weekly, daily, and even hourly) has allowed performance improvements in modeling volatility not only in short periods but even in longer monthly-quarterly periods. But it became apparent that the standard volatility models are not very well suited for this high-frequency data: the information in the intraday data can't be handled by volatility models designed for daily data, while the models for intraday data don't perform well in forecasting volatility in daily data. That is why applications shifted to simpler models like exponential smoothing methods and realized volatility models.[14]

J.P. Morgan's RiskMetrics (1996) introduced the EWMA methodology, in which weights are exponentially assigned to each squared return. While in the RV model, each squared return is equally weighted, in EMWA, they are exponentially weighted as

$$\sigma = \sqrt{(1 - \lambda) \sum_{t=1}^{T} \lambda^{t-1} R_t^2},$$

where $\lambda \in (0, 1)$ is termed as the decay factor. Note that each subsequent exponent of λ would be smaller, which is why such a model would place higher weightage on recent returns. As the market experiences a shock due to which the returns would fluctuate more than before, EWMA would be able to capture the change more efficiently than the equally weighted RV models. The term $(1 - \lambda)$ is an approx of $1/\sum_{t=1}^{T} \lambda^{t-1}$, and the given model can be expressed as

$$\sigma^2 = \frac{\sum_{t=1}^{T} \lambda^{t-1} R_t^2}{\sum_{t=1}^{T} \lambda^{t-1}},$$

in which case it is easy to see how the volatility is calculated as the weighted sum of the squared returns. Another interesting aspect of EMWA is that it can also be transformed in a recursive form as

$$\sigma_{t+1}^2 = \lambda \sigma_t^2 + (1 - \lambda) R_t^2,$$

which makes it closer to the IGARCH(1,1) model. Pafka and Kondor (2018) argued this as the reason for the success of the model in estimating volatility.

In choosing an appropriate (default) value for λ, RiskMetrics found its value to be 0.94 for the daily dataset and 0.97 for the monthly dataset, by an optimization technique concerning RMSE (root mean square error) of 480 time-series variables. For the forecast error be $R_{t+1}^2 - \sigma_{t+1}^2$, the RMSE would be the square root of the variance of the series of forecast errors. Considering RMSE as the criteria for measuring forecast error, since different values of λ would yield different RMSE: the λ's with the minimum RMSE in each of the variables is taken. The default values of λ in the RiskMetrics are then obtained by weighted average (the weights depending on the minimum RMSE) of the λ's.

There are, however, drawbacks associated with the RiskMetrics methodology. It assumes that the returns are normally distributed, which seems to be unrealistic since several financial returns show a fat-tailed distribution.[16] Also, while the model works very well in theory for one period forecast, the "square root of time rule" exists in the case of multiple period forecasts. It is found that $\sigma^2_{t+T} = T\sigma^2_{t+1}$, which implies that $\sigma_{t+T} = \sqrt{T}\sigma_{t+1}$ - meaning that "multiple day forecasts are simple multiples of the one-day forecasts".[15] Due to this, the quality of the multi-period forecast becomes questionable.

6. Artificial Neural Network

Since its inception in the early '40s, artificial neural network (ANN) has found its way in several applications we use currently, and is especially popular in pattern recognition software. This algorithm is inspired by the human brain's neural system[17], in which not only inputs are processed to yield output, but also feedbacks are used in training the system. The algorithm may have several layers between the input and output, in which mathematical functions are used to assign weights to the passing information during the process. These are called the transfer function, which may be linear or nonlinear, and the layers are termed as hidden layers. ANN is known to be *data-hungry*, in the sense that they don't perform as well with limited data - which is why applying it to the abundant financial datasets would not be a problem.

In machine learning, ANN is used for *training* an algorithm to recognize patterns. However, ANN can also be trained for detecting statistical patterns. For training, the inputs (or features) are passed on in the hidden layers with random weights assigned, and then it is checked whether the desired output is the true one or not, and with feedback, the weights are altered. The weights can determine how much and which of the inputs are significant in determining the output. Measures like mean squared error can be used to determine the success of the training. The hidden layers can be set with any transfer function, which can be determined by training the algorithm with different transfer functions.

ANN can even be used along with parametric methods. For example, Donaldson and Kamstra (1996) created a nonlinear GARCH model based on ANN which was found to be capturing volatility effects overlooked by the GARCH model and its extensions. A popular transfer function such as

$$Y = a + \frac{1}{1+e^{-(c+bX)}}$$

is included in a GJR model. For a particular set of parameters (a, b, c), Y would be termed as an *information node*. A node might be active or inactive depending on whether Y is positive, negative, or depending on any other attribute as desired. Such a node is then included in the GJR model with X being a uniform distribution. It was then found that the ANN model didn't overfit the data, which is why superior results were obtained in one-step-ahead out-of-sample forecasts.

Another example is forecasting with hybrid neural network models provided by Kristjanpoller, Fadic, and Minutolo (2013) that demonstrated the improvement of the forecasting performance of GARCH models by using ANN models. For different parameters of the nodes, ANN models were found to be robust and consistent. The input variables used to estimate the ANN model were forecasted RV and GARCH(1,1) estimates, squared log return, and index return. For each index, the input variables are calculated for 21 days as

$$RV_t = \frac{1}{21}\sum_{i=t+1}^{t+21}(r_i - r_t)^2,$$

where r_t is log returns average in each window of 21 days, and the log returns are modelled as

$$r_t = c + \theta_1 r_{t-1} + \varepsilon_t \sigma_t \text{ and } \sigma_t^2 = d + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

All this is to illustrate how the ANN technique can be used in modeling volatility apart from the usual modeling methodologies. However, it must also be noted that these machine learning techniques are not entirely independent of the previous volatility models. The methodology of these popular pattern recognition algorithms can certainly be used, but currently, it is done so to amplify the result obtained by traditional models.

Conclusion

Modeling the changing conditional variance has surely become a thriving research area in time-series econometrics. After beginning with applications on macroeconomic variables, it further gained momentum with modeling the financial variables. The focus on the data-distribution aspects like skewness and kurtosis helped shape the models with simulations. In the data-rich era, the methodology changed from complex parametric measures to more simple non-parametric measures. It even extended to incorporate machine-learning algorithms to improve forecasting, which suggests how the future models would be going along with data-driven technological advancements. With such indications and because heavy computations are getting cheaper and more efficient, one may certainly hope that the future methodologies for volatility estimation would be enriched by the inclusion of algorithmic machine-learning approaches in modeling.

References

[1] Engle, R. F. (1982). "*Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*", Econometrica, 50 (4), 987-1008.

[2] Bollerslev, T. (1986). "*Generalized Autoregressive Conditional Heteroskedasticity*", Journal of Econometrics, 31, 307–327.

[3] Box, G. E. P., Jenkins G. M. (1970). "*Time Series Analysis: Forecasting and Control*", Holden-Day, Inc.

[4] Davidson R., MacKinnon J. G. (2004). "*Econometric Theory and Methods*", Oxford University Press.

[5] Engle R. F., Bollerslev T. (1986). "*Modelling The Persistence of Conditional Variance*", Econometric Reviews, 5(1), 1-50.

[6] Bollerslev T., Russell, J. R., Watson, M. W. (2010). "*Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*", Oxford University Press.

[7] Granger C. W. J., Andersen A. (1978). "*An Introduction to Bilinear Time-Series Models*", Gottingen, Vandenhoeck and Ruprecht.

[8] Glosten, L. R., Jagannathan R., Runkle D. E. (1993). "*On The Relation between The Expected Value and The Volatility of Nominal Excess Return on Stocks*", Journal of Finance, 48(5), 1779-1801.

[9] Taylor S. J. (1986). "*Modelling Financial Time Series*" Wiley.

[10] Broto C., Ruiz E. (2004). "*Estimation Methods for Stochastic Volatility Models: A Survey*", Journal of Economic Surveys, 18(5), 613-649.

[11] Dumas B., Fleming J., Whaley R. E. (1996). "*Implied Volatility Functions: Empirical Tests*", The Journal of Finance, 53(6), 2059-2106.

[12] Canina L., Figlewski S. (1993). "*The Information Content of Implied Volatility*". The Review of Financial Studies, 6(3), 659-681.

[13] Andersen T., Bollerslev T., Diebold F. (2003), "*Parametric and Nonparametric Volatility Measurement*", [in:] L.P. Hansen, Y. Ait-Sahalia (eds.), Handbook of Financial Econometrics, Amsterdam: North-Holland.

[14] Bollerslev T., Andersen T. G., Diebold F. X., Labys P. (2003) "Modeling and Forecasting Realized Volatility", Econometrica, 71(2), 579-625

[15] RiskMetrics Group. (1996). "*RiskMetrics - Technical Document*", J.P. Morgan/Reuters.

[16] Pafka S., Kondor I. (2001). "Evaluating the RiskMetrics Methodology in Measuring Volatility and Value-at-Risk in Financial Markets", Physica A: Statistical Mechanics and its Applications, 299(1-2). 305-310.

[17] Dongare A. D., Kharde R. R., Kachare A. D. (2012). "*Introduction to Artificial Neural Network*", International Journal of Engineering and Innovative Technology, 2(1), 189-194.

[18] Donaldson R. G., Kamstra M. (1997). "*An artificial neural network-GARCH model for international stock return volatility*", Journal of Empirical Finance, 4(1), 17-46.

[19] Kristjanpoller W., Fadic A., Minutolo M. C. (2014). "Volatility forecast using hybrid Neural Network models", Expert Systems with Applications, 41(5), 2437-2442.

Bibliography

- Greene W. H. (2008). "*Econometric Analysis*", Prentice Hall.
- Enders W. (2014). "Applied Econometric Time Series", Wiley.
- Gujarati, D. N. (2004). "*Basic Econometrics*", The McGraw-Hill Companies.
- McAleer M. J., Medeiros M. C. (2008). "Realized Volatility: A Review", Econometric Reviews, 27(1-3), 10-45.
- Kliber, A., Bedowska-Sojka B. (2010). "Realized Volatility Versus GARCH and Stochastic Volatility Models. The Evidence from the WIG20 index and the EUR/PLN Foreign Exchange Market", Statistical Review, 57(4), 105-127.