EXTENDING THE FEATURE SET OF A DATA-DRIVEN ARTIFICIAL NEURAL NETWORK MODEL OF PRICING FINANCIAL OPTIONS

Luis Montesdeoca and Mahesan Niranjan
School of Electronics and Computer Science
University of Southampton, University Road SO17 1BJ
Southampton, UK
Email: {*ljm1e14,m.niranjan*}*@soton.ac.uk*

***Abstract*—Prices of derivative contracts, such as options, traded in the financial markets are expected to have complex relationships to fluctuations in the values of the underlying assets, the time to maturity and type of exercise of the contracts as well as other macroeconomic variables. Hutchinson, Lo and Poggio showed in 1994 that a non-parametric artificial neural network may be trained to approximate this complex functional relationship. Here, we consider this model with additional inputs relevant to the pricing of options and show that the accuracy of approximation may indeed be improved. We consider volume traded, historic volatility, observed interest rates and combinations of these as additional features. In addition to giving empirical results on how the inclusion of these variables helps predicting option prices, we also analyse prediction errors of the different models with volatility and volume traded as inputs, and report an interesting correlation between their contributions.**

1. Introduction

Pricing derivative contracts is a challenging problem in financial engineering because contracts mature at some point in the future and there are multiple sources of uncertainty between the current time at which a fair price for the contract needs to be determined and the point at which it may be exercised. The celebrated Black-Scholes model of options pricing makes specific assumptions about a stochastic process model of the underlying asset price and other factors relating to it [1]. These assumptions lead to a solvable differential equation and result in a closed form pricing formula for certain simple derivative instruments. For more complex contracts, where analytical solutions are not possible, Monte Carlo simulation and numerical analysis based methods have been developed. There is significant interest in research literature and wide practical applications of these topics [2].

In this context, the work of Hutchinson, Lo and Poggio in 1994 [3], could be seen as an elegant development from a machine learning or data-driven modelling point of view. Their non-parametric approach discards the restriction of the Black-Scholes model and hence it becomes more adaptive and flexible. The authors showed that a non-parametric artificial neural network, specifically a Radial Basis Functions (RBF) model, can be trained to approximate the complex relationship between the prices of an options contract and the underlying asset. In particular they used only the normalized asset price and the time to maturity of the contract as inputs to the network and further showed that the derivatives of the mapping learned by the network faithfully reproduced the hedge ratio (Δ, Delta), a widely used parameter in balancing portfolio risk. Neural networks are powerful non-linear approximators, and the RBF architecture itself has found a wide range of applications including speech classification [4], time series prediction [5] and financial engineering [6] among others. RBF is easily deployed in problems that require sequential learning and adaptive model complexity as demonstrated in the resource allocating network [5], [7], [8], and their generalization properties have been analysed in [9], [10].

In addition to the demonstration that options prices may be well-approximated, Hutchinson *et al.*'s work, which forms the basis of the present study, is notable for a second reason that is of interest in financial engineering. Their work showed that the derivatives of the learned model, which is easily computed for the RBF model analytically, turned out to be good approximations to the hedge ratio, commonly known to practitioners by the Greek letter *Delta* (Δ). This ratio determines the construction of a risk neutral portfolio in which the uncertainty induced by a stochastic process model of asset price changes may be cancelled out. In a later development of Hutchinson *et al.*'s work, one of us showed that the RBF model, as used in this context, and the Black-Scholes model itself, may be cast as dynamical systems, and the unknowns in the model inferred in a sequential setting using the Extended Kalman Filter (EKF) algorithm [11]. A broader review of the uses of neural networks, with currency options as the application is given in [12].

A further topic in empirical finance is the relationship between the traded volume of an asset and its volatility. Do assets that are traded more show greater price fluctuations, and hence greater uncertainty? While it is tempting to expect such a relationship, there may be no theoretical grounds to reach such a conclusion. Various empirical studies have attempted to test if volume traded and volatility are correlated [13], [14]. For instance, in [15] for futures on the London

International Financial Futures Exchange (LIFFE) a positive correlation between these two variables was found. Besides, in [16] found that trading volumes are associated with equity market volatility, in all the markets studied of the BRIC (Brasil, Russian, India and China) countries.

An underlying theoretical premise, known as the Mixture of Distribution Hypothesis (MDH), introduced by Clark [17], suggests that daily trading volume and price changes are driven by the same flow of information. Starting from this there are empirical studies that have attempted to explore this relationship. For example, Wen-Cheng *et al.* [18] suggest a bidirectional relationship between the two using a bivariate vector autoregressive methodology, and Jain [19] suggests a strong contemporaneous relation between trading volume and volatility with hourly values. However, Karpoff [20] reports an asymetric relationship, which is also supported by the work in [21]. Similar explorations have been carried out on the Korean and New York Stock Exchange data [22], [23]. Nevertheless, in various commodity future contracts analyzed in the Chinese market, the correlation between return and trading volume are close to zero [24] and on intra-day trading data on S&P 500 index a negative correlation is reported [25].

In this work, we extend Hutchinson *et al.*'s work by asking the question if expanding the feature set to include additional variables of interest help in improving their data driven model of options pricing. Specifically, we include combinations of historic volatility, volume traded and interest rates as inputs in addition to the normalized asset price and time to maturity, as illustrated in Figure 1. Empirical work we carried out, on a range of data with much wider scope than the original work, shows this to be the case. We follow this up with an analysis of how much the volume traded and volatility of asset help in predicting options prices and demonstrate an intriguing correlation between their relative contributions.

2. Model & Data

In this section, we describe the architecture of the Radial Basis Functions (RBF) model as introduced in [3], our extensions to the feature sets and the data used in training the networks. We use substantially more data than [3], who restricted their analysis to options drawn on the $S\&P500$ Index only.

Let the input data vector to the RBF model given by the vector

$$\boldsymbol{x} = \left[\frac{S}{X} \quad T - t \right]^T, \tag{1}$$

where S denotes the underlying asset price, X, the strike price of the contract and $T - t$ is the time to maturity (the difference between the time of maturity, T, of the contract and the present time t). With this vector of features as input, each of the basis functions in the RBF model for predicting options prices is written as

$$\boldsymbol{\phi}_k = \sqrt{\left[\left(\boldsymbol{x} - \boldsymbol{m}_k\right)^T \quad \textstyle\sum_k^{-1} \quad \left(\boldsymbol{x} - \boldsymbol{m}_k\right) + b_k \right]}, \tag{2}$$

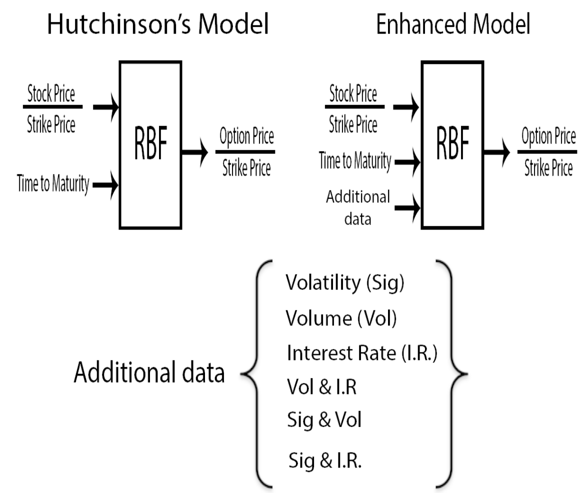


Figure 1. The data-driven neural network scheme of approximating options prices. Left: Model used by Hutchinson *et al.* using two inputs [3]; Right: our models with the combination of additional variables considered.

where, ϕ_k denotes the response of a nonlinear basis function which is parameterized by local mean $\boldsymbol{m}_k$ and covariance matrix Σ_k and b_k, a local bias term.

The basis function locations and local covariance matrices are estimated by fitting a Gaussian Mixture Model (GMM) to the distribution of input data, making the model sensitive to its local density. For simplicity, in our implementations, we set the bias terms b_k to zero. In addition to these nonlinear terms, the model includes a linear part as well. Thus least squares problem to solve is shown by the simultaneous equations:

$$\begin{pmatrix} \phi_{11} & \dots & \phi_{14} & x_{11} & x_{12} & 1 \\ \phi_{21} & \dots & \phi_{24} & x_{21} & x_{22} & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \phi_{n1} & \dots & \phi_{n4} & x_{n1} & x_{n2} & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_4 \\ w_1 \\ w_2 \\ w_0 \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \tag{3}$$

where n is the number of observations and $c_j, j = 1, ...n$ are the output call option prices normalized by their strike prices. The vector of unknowns $\left[\lambda_1 \quad ... \quad \lambda_4 \quad w_1 \quad w_2 \quad w_0 \right]^T$ is estimated by linear least squares, which is the main computational advantage of using a fixed non-linear, linear in unknowns RBF model. A pseudo-inverse solution to the problem is often used

$$\boldsymbol{w} = \left(\boldsymbol{P}^T \boldsymbol{P}\right)^{-1} \boldsymbol{P}^T \boldsymbol{\Phi}, \tag{4}$$

where $\boldsymbol{P}$ is the so called design matrix and $\boldsymbol{\Phi}$, the vector of outputs. Often, for reasons of numerical ill-conditioning and to avoid over-fitting by the model, a regularization term in the form of a diagonal matrix is added before the inversion of $\boldsymbol{P}^T \boldsymbol{P}$:

$$\boldsymbol{w} = \left(\boldsymbol{P}^T \boldsymbol{P} + \gamma \, \boldsymbol{I}\right)^{-1} \boldsymbol{P}^T \boldsymbol{\Phi}, \tag{5}$$

where γ controls how much regularization is applied.

Once the model is trained, the resulting output is given by,

$$\widehat{c} = \boldsymbol{\Phi\lambda} + \boldsymbol{w^T x} + w_0 \tag{6}$$

Following the work of Hutchinson *et al.* we have defined the above model with four basis functions. In our own work, we tested the effect of the choice of the number of basis functions, hence the model complexity, and found four to be a reasonable number. We also confirmed this by running the RBF model with the Akaike information criterion (AIC) [26] as the method of model selection. Beyond that the model showed clear signs of overfitting as shown in Figure 2.

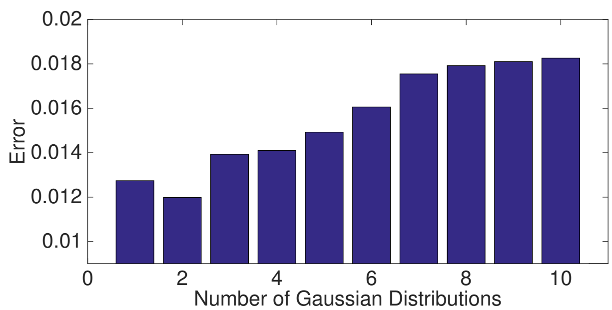


Figure 2. Modelling errors as a function of the number of Gaussian basis functions used on a small subset of the data used. Consistent with [3], a small number of basis functions was considered sufficient.

For empirical evaluation, we used the daily prices of 21 call and 11 put options written on the Financial Times Stock Exchange (FTSE) index and nine call and seven put options of the popular software companies Apple and Microsoft. Additionally, we considered minute-by-minute intraday call and put options prices on a contract on Apple with a strike price of 95 and 100, maturing in September 2015 and June 2016 respectively. Table 1 lists the range of options considered and includes their strike prices and dates of maturity.

For selection of training data, for any of the options we took a temporal window of 40% of the data as the training set and evaluated the model performance on the next point in time. We moved this window one sample at a time and repeated the training and testing. Thus all results we quote are based on single sample unseen data. The reason for this choice of window, rather than a randomized training/test partition as often used in machine learning is that the financial data is expected to have temporal structure and in any practical application, one is likely to apply a trained model in the next point in time. We computed volatility over a window of 25% of the data immediately preceding a point of analysis.

TABLE 1. CALL AND PUT OPTIONS USED IN THIS STUDY WITH DIFFERENT TIME TO MATURITY AND NUMBERS OF TRADING DAYS. THE STRIKE PRICES AND INDICATION OF WHETHER THE OPTION IS A CALL OR PUT IS IN THE NAME OF THE CONTRACT.

Opt.	Exp.	Days	Opt.	Exp.	Days
C6600	09-13	171	C6600	12-14	397
C6700	09-13	290	C6700	12-14	359
C6800	09-13	171	C6800	12-14	397
C6700	06-14	352	C6900	12-14	299
C6750	06-14	85	C7000	12-14	387
C6800	06-14	525	P6000	09-13	462
C6900	06-14	291	P6300	09-13	525
C7000	06-14	288	P6400	09-13	525
C7100	06-14	171	P6500	09-13	250
C6600	09-14	462	P6600	09-13	513
C6400	09-14	462	P6700	09-13	171
C6800	09-14	462	P6800	09-13	525
C6900	09-14	299	P6600	06-14	513
C7000	09-14	387	P6700	06-14	352
C7200	09-14	462	P6750	06-14	85
C7400	09-14	342	P6800	06-14	525
CAPL50	07-15	239	CMSF39	07-15	236
CAPL85	07-15	239	PAPL120	07-15	239
CAPL90	07-15	239	PAPL130	07-15	238
CAPL95	07-15	239	PAPL135	07-15	238
CAPL105	07-15	236	PAPL140	07-15	238
CAPL110	07-15	236	PAPL150	07-15	236
CAPL115	07-15	236	PMSF45	07-15	236
CAPL120	07-15	236	PMSF47	07-15	236

3. Empirical Results

3.1. Prediction performance with additional features

Evaluating the options listed in the Table 1 demonstrate that Hutchinson *et al.*'s model with additional inputs relevant to the pricing of options may have enhanced the accuracy of approximation. This values are shown in Table 2 for call and put options. An example of this improvement is illustrated in Figure 3.

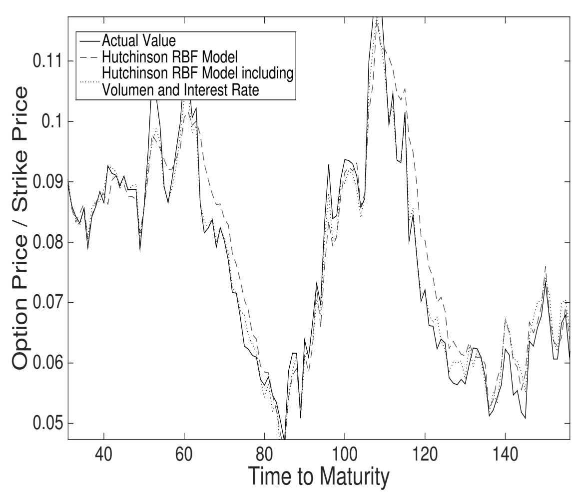


Figure 3. Price of a FTSE index option (strike price 6800, expiring Sept. 2013) and its approximations from Hutchinson *et al.*'s model and a model that includes interest rate and trading volume as additional inputs.

TABLE 2. AVERAGES OF MEAN SQUARED MODELLING ERRORS OF THE VARIOUS MODELS ON THE DIFFERENT OPTION CONTRACTS. THE DISTRIBUTION OF MEAN SQUARED ERRORS IS SHOWN IN FIGURE 4. EACH VALUE IN THE TABLE SHOULD BE SCALED BY 10^{-6}.

-	Hutch.	Vol.	Sig	Sig+Vol	I.R.	IR+Vol
C	5.19	4.27	5.51	4.79	4.70	4.21
P	9.13	8.78	9.31	9.74	8.38	8.93

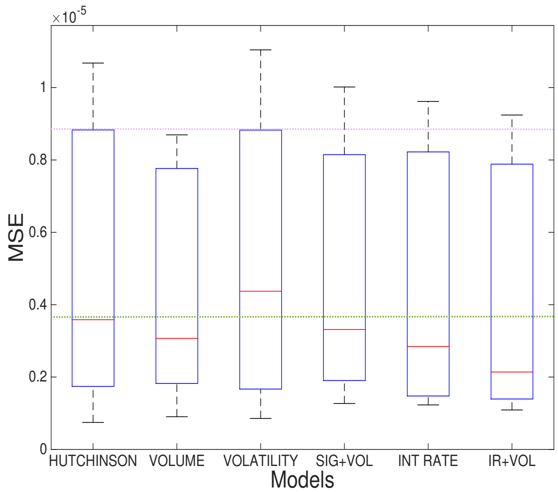


Figure 4. Prediction performance on FTSE index call options for the various models considered. Overall, the largest improvement in prediction accuracies is obtained when volume traded and interest rates are included as additional covariates.

Figure 5 shows performance of RBF predictors on the two equity options. Here, with volatility as additional input we see only a marginal improvement in performance whereas all other models give significant reduction in errors when compared to Hutchinson *et al.*'s model.

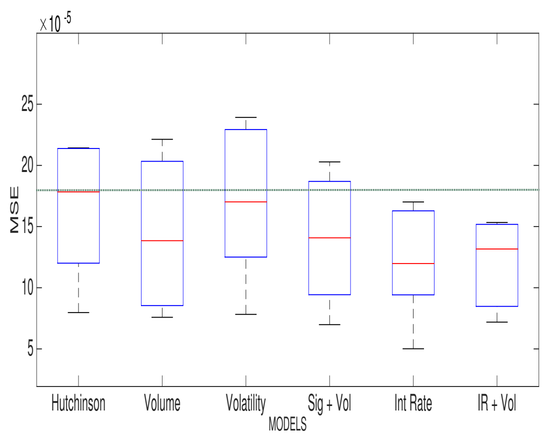


Figure 5. Prediction performance on single equity options. The use of volatility as additional input yields a marginal improvement, whereas all the other combinations of additional variables give significant reduction in error.

Figure 6 shows performance of RBF predictors considering minute-by-minute intraday call and put equity options. Here, we plot the squared error at every point in time (minute) as boxplots of the intraday call option. The corresponding mean squared errors are shown in Table 3.

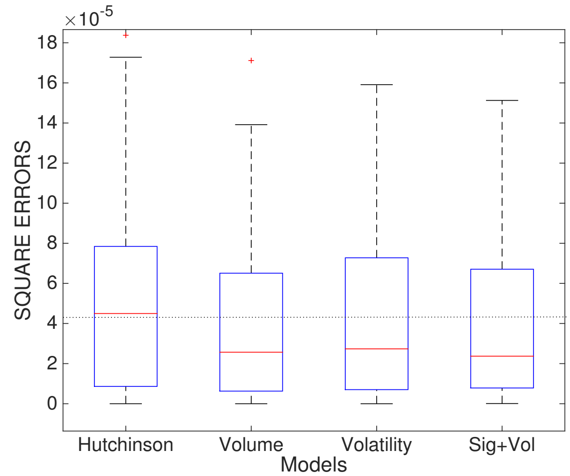


Figure 6. Squared errors of different RBF models on equity call option with intra-day data. The corresponding averages (MSE) are shown in Table 3.

TABLE 3. AVERAGE PREDICTION PERFORMANCES OF VARIOUS MODELS ON INTRA-DAY CALL AND PUT OPTIONS CONTRACT. EACH VALUE IN THE TABLE SHOULD BE SCALED BY 10^{-5}.

-	Hutchinson	Volume	Sigma	Sigma+Volume
C	2.97	2.19	2.64	2.76
P	2.72	2.57	2.83	2.29

3.2. Volume Traded and Volatility

Further we used these models as a tool to explore the relationships between volatility and trading volume. The measurement of this correlation is made by the mean squared errors of the predictive models with volume traded and volatility as additional inputs. Figure 7 (left boxplot) shows the distribution of correlations in mean squared errors of model fitting with volume traded and volatility as additional inputs on athe 48 options which is 0.675. In the right bloxplot in Figure 7 shows the correlation between the corresponding volume trading values and volatility values.

The observations analysed for this empirical study are separated into equity options and FTSE 100 index options. Figure 8 shows a substantial correlation between volume traded and volatility in predicting options prices in the FTSE 100 index call options. Evaluating the call options found a Pearson coefficient value of ***0.663*** in the Mean Square Errors on the RBF with volume and volatility as additional inputs . A similar distribution was observed for put options with a correlation of ***0.678*** pricing put derivative options.

Analysing the equity call and put options, Figure 9 shows the Pearson correlation coefficient between Mean Square Errors on the RBF with volume and volatility as additional inputs. The mean on the correlation of each call

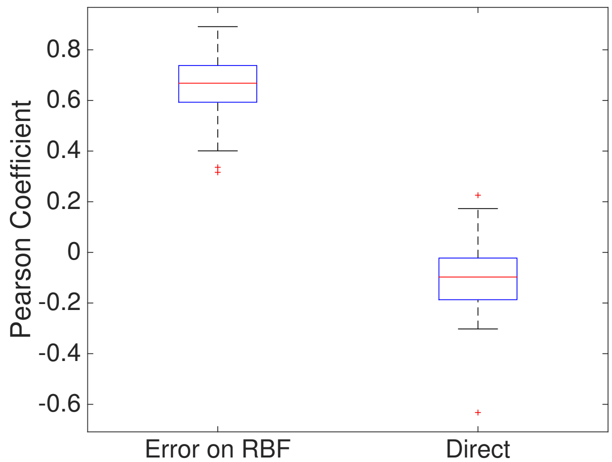


Figure 7. The Pearson correlation values from the mean square errors of the RBF models with volatility and trading volume on all options tested. It demonstrates that the mean on the correlation of each option analysed is over ***0.67*** giving a substantial correlation. The relationship between the corresponding values of trading volume and volatility does not show a correlation

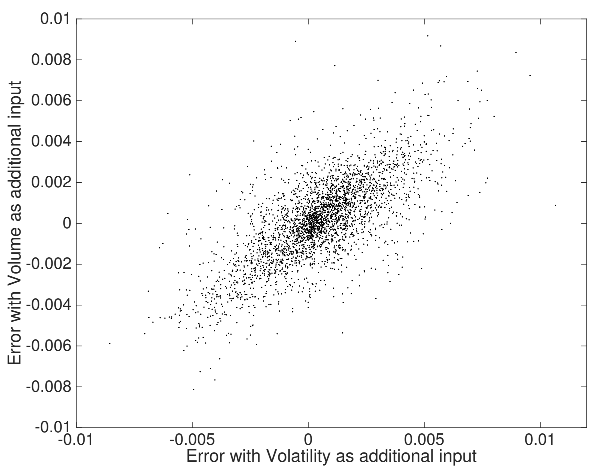


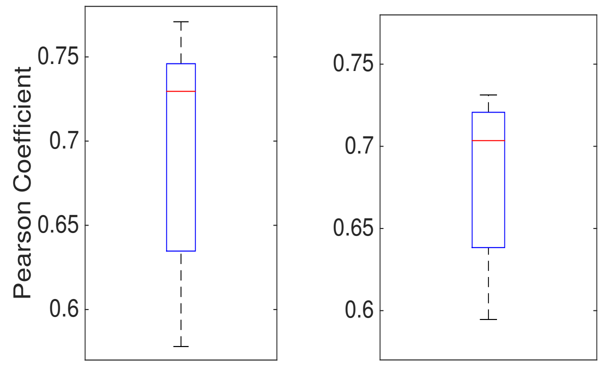
Figure 8. Correlation on the contribution of RBF errors on predicting each call FTSE options when volume and volatility are included as additional inputs. This correlation value is **0.663**.

option analysed is over ***0.6818*** and ***0.6985*** for each put option, which demonstrates a substantial correlation.

An example of the correlation on minute-by-minute intraday call equity option is shown in Figure 10 and its Pearson correlation value is ***0.7199***.

4. Conclusion

In this work, we study a non-parametric neural network model that quantifies the complex relationship between a class of financial instruments known as options and that



(a) Put single options (b) Call single options

Figure 9. Correlation between Mean Square Error values of the RBF models that include trading volume and volatility on put(a) and call(b) equity options.

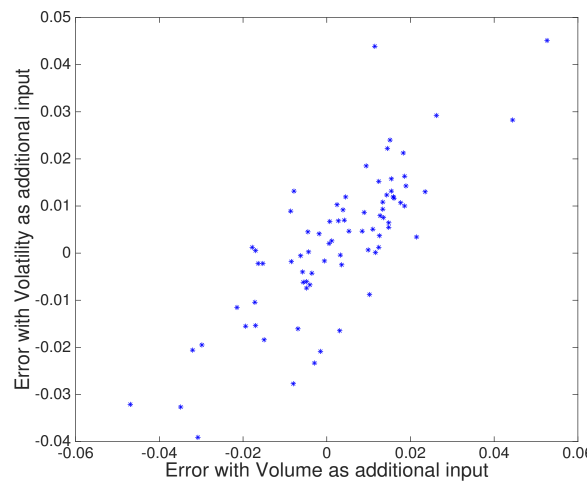


Figure 10. Contribution to prediction error from volatility and volume traded on minute-by-minute intra-day data on equity call option, showing a high correlation of 0.72.

of the underlying asset on which the contract is drawn. Whereas previous work introducing this model uses the asset price and the time to maturity of the contract as its only inputs, we have demonstrated that the inclusion of additional features relating to the contract, namely the volatility, volume of the underlying asset traded and the risk-free interest rate help in improving the accuracy with which the market value of the contract may be predicted. Our empirical results, carried out on index options, two equity options and an intra-day contract covers a substantially wider range than previous authors have considered. Hence the results quoted in this study are more robust.

The results lead to the exploration of an intriguing relationship between the volatility of an asset and the amount of volume traded, a topic that has attracted healthy discussion in empirical finance. Our model makes an interesting contribution to this discussion in that we demonstrate that while the volume and volatility of an asset do not correlate significantly, their contributions to modelling the price of an options contract drawn on the asset do show significant correlation. This has also been demonstrated on the FTSE index, individual equities and intra-day datasets.

The empirical work reported in this paper is on static data, mapping instantaneous values of covariates to the response variable. Apart from the selection of training and test data, we have not explicitly modelled the temporal dynamics. However, with financial time series, this dynamics is important and for this particular model a state space formulation and inference using extended Kalman filter has been shown to be possible [11]. Hence our current work focuses on extending the models with the extended features to a dynamical setting.

References

[1] J. C. Hull, *Options, Futures, and Other Derivatives*, 8th ed. New York: Prentice Hall, 2011.

[2] P. Brandimarte, *Numerical Methods in Finance and Economics*. Wiley, 2006.

[3] J. Hutchinson, A. Lo, and T. Poggio, "A nonparametric approach to pricing and hedging derivative securities via learning networks," *The Journal of Finance*, vol. 49, no. 3, pp. 851–889, 1994.

[4] M. Niranjan and F. Fallside, "Neural networks and radial basis functions in classifying static speech patterns," *Computer Speech and Language*, vol. 4, pp. 275–289, 1990.

[5] V. Kadirkamanathan and M. Niranjan, "A function estimation approach to sequential learning with neural networks," *Neural Computation*, vol. 5, pp. 954–975, 1993.

[6] A. Golbabai, D. Ahmadian, and M. Milev, "Radial basis functions with application to finance: American put option under jump diffusion," *Mathematical and Computer Modelling*, vol. 55, no. 34, pp. 1354 – 1362, 2012.

[7] J. Platt, "A resource-allocating network for function interpolation," *Neural Computation*, vol. 3, no. 2, pp. 213–225, 1991.

[8] C. Molina and M. Niranjan, "Pruning with replacement on limited resource allocating networks by F-projections," *Neural Computation*, vol. 8, no. 4, pp. 855–868, 1996.

[9] S. Holden and M. Niranjan, "On the practical application of VC dimension bounds," *Neural Computation*, vol. 7, no. 6, pp. 1265–1288, 1995.

[10] S. B. Holden and M. Niranjan, "Average-case learning curves for radial basis function networks," *Neural Computation*, vol. 9, no. 2, pp. 441–460, 1997.

[11] M. Niranjan, "Sequential tracking in pricing financial options using model based and neural network approaches," in *Advances in Neural Information Processing Systems 9*, M. Mozer, M. Jordan, and T. Petsche, Eds. MIT Press, 1997, pp. 960–966. [Online]. Available: http://papers.nips.cc/paper/1245-sequential-tracking-in-pricing-financial-options-using-model-based-and-neural-network-approaches.pdf

[12] F. Chen and C. Sutcliffe, "Pricing and hedging short sterling options using neural networks," *Intelligent Systems in Accounting, Finance and Management*, vol. 19, no. 2, pp. 128–149, 2012.

[13] C. Floros and D. Vougas, "Trading volume and returns relationship in Greek stock index futures market: GARCH vs. GMM," *International Research Journal of Finance and Economics*, vol. 12, pp. 98–115, 2007.

[14] T. Wang and Z. Huang, "The relationship between volatility and trading volume in the Chinese stock market: A volatility decomposition perspective," *Annals of economics and finance*, vol. 13, no. 1, pp. 211–236, 2012.

[15] J. Board and C. Sutcliffe, "Information, volatility, volume and maturity: An investigation of stock index futures," *Review of Futures Markets*, vol. 9, no. 3, pp. 532–549, 1990.

[16] K. Pramod and P. Puja, "Stock market volatility and equity trading volume: Empirical examination from Brazil, Russia, India and China (BRIC)," *Global Business Review*, vol. 16, no. 5, pp. 28–45, 2015.

[17] P. Clark, "A subordinated stochastic process model with finite variance for speculative prices," *Econometrica*, vol. 41, no. 1, pp. 135–155, 1973. [Online]. Available: http://www.jstor.org/stable/1913889

[18] L. Wen-Cheng, C. Ming, and F.-J. L., "An empirical study of volatility and trading volume dynamics using high-frequency data," *The International Journal of Business and Finance Research*, vol. 4, no. 3, 2010.

[19] P. Jain and G. Joh, "The dependence between hourly prices and trading volume," *Journal of Financial and Quantitative Analysis*, vol. 23, pp. 269–284, 1988.

[20] J. Karpoff, "The relation between price changes and trading volume: A survey," *Journal of Financial and Quantitative Analysis*, vol. 22, pp. 109–126, 1987.

[21] T. Andersen, "Return volatility and trading volume: An information flow interpretation of stochastic volatility," *Journal of Finance*, vol. 51, pp. 169–204, 1996.

[22] K. Choi, Jiang, Z.H., S. Kang, and S. Yoon, "Relationship between trading volume and asymmetric volatility in the Korean stock market," *Modern Economy*, vol. 3, p. 584589, 2012.

[23] A. Darrat, M. Zhong, and L. Cheng, "Intraday volume and volatility relations with and without public news," *Journal of Banking & Finance*, vol. 31, no. 9, pp. 2711–2729, 2007.

[24] G. Chen, M. Firth, and Y. Xin, "The price-volume relationship in Chinas commodity futures markets," *The Chinese Economy*, vol. 37, no. 3, pp. 87–122, 2004.

[25] T. Wang and Z. Huang, "The relationship between volatility and trading volume in the Chinese stock market: A volatility decomposition perspective," *Annals of Economics and Finance*, vol. 13, no. 1, pp. 211–236, 2012.

[26] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.