HYBRID WAVELET- SVMs FOR MODELING DERIVATIVS VALUATION

HSING-WEN WANG SHIAN-CHANG HUANG

Department of Business Administration, College of Management, National Changhua University of Education
No. 2, Shi-Da Rd., Changhua 500, Taiwan, R.O.C.
e-mail: shinwen@cc.ncue.edu.tw

Abstract

Due to the rapid grow up of transaction volume of derivatives in the financial market, the Black-Scholes options pricing model (BSM) is played an important role recently and widely applied in various options contract. However, this theoretical model limited by the influences of many unexpected real world phenomena caused due to its six unreasonable assumptions, which often make the miss-pricing result because of the difference of market convention in practical. If we were to soundly take these phenomena into account, the pricing error could be reduced. In this paper, we provide a signal-decomposition oriented framework via wavelet analysis to improve the precision of BSM using integrated wavelet-based feature extraction with support vector machines (WSVMs). We investigate the techniques for transforming the noticeable signal from the mark to market price into estimating the option fair value and hence gain better precision estimation then pure support vector machine, in which has recently been introduced as a new technique for solving a variety of time series forecasting. Compare with the original GARCH method, adaptive neural-based fuzzy inference system (ANFIS) and pure SVMs, the performance of the presented method show the best. Using evidence from the warrants market in Taiwan, it supports our claims. This paper helps to provide an alternative way to refine the options valuation.

Keywords: Black-Scholes, Wavelet analysis, SVMs, GARCH, ANFIS.

1. Introduction

Since the Black-Scholes options pricing model (BSM) was proposed in 1973 [1], it has become the foundation for the development of modern derivative commodity pricing theories, and has been widely adopted by the financial industry [2]. Nevertheless, in terms of its actual application, it is limited by a number of presumptions and hypotheses that are derived from the model itself, and that lead to many unexpected phenomena when the model is established. These bring considerable influence to bear on the applicability, precision and effectiveness of that model [3]. Recently, a hybrid approach that integrates artificial neural networks, fuzzy inference, and other machine learning techniques has been suggested to improve the option valuation accuracy. The results of comparative studies indicate that the artificial intelligence (AI) shows better prediction accuracy [4]. However, AI has a difficulty in explaining the causes of prediction result due to the lack of explanatory power and suffers from difficulties with generalization because of overfitting. In addition, it needs too much times and efforts to construct a best architecture.

Through an empirical study it is discovered that the BSM assumptions are actually different from the practical situation [5], which ignore the volatility skewed and volatility clustering phenomena that influence the real mark to market price of options. As a result, the more serious the bias transmission of the pricing information from volatility behavior is the larger variance that the BSM pricing would generate. In the study of numerous generalized autoregressive conditional heteroscedasticity (GARCH) models are tried to reduce the bias via forecasting the volatility. Unfortunately, the usual GARCH (1, 1) includes only one time horizon, and that is not enough to replicate the multi-horizon complexity of options market trading [6]. On the other hand, wavelet analysis provides an effective way to decompose the trading behavior (e.g. market price of underlying) with feature extraction represented to be different frequency spectrum on every important time scales for the multi-horizon revealed. However, it is lack of making good prediction by itself.

Recent researches still paid little attention to combines the wavelet features into the SVMs [7][8]. In this paper, our main innovative idea is to deal with the collinear problem of input explanatory variables through the orthogonal decomposition process, so as to enhance the capability of features capturing before the nonlinear mapping of input vectors into the high-dimensional feature space (FS). The features were fed into the SVMs, in which corresponds to a linear method in a very high dimensional FS that is nonlinearly related to the input space. Besides, SVMs demonstrates its powerful and superior prediction performance (Cristianini, 2000) under taking few computational resources [9]. In view of this, this study has adopted the WSVMs model to perform the non-parametric options valuation.

The remainder of the paper is organized as follows. Section 2 describes wavelet analysis with SVMs. In Section 3, we discuss the powerful artificial intelligence framework, ANFIS. Section 4 introduces the Black-Scholes formula with GARCH prediction model. Section 5 describes the mark to market observations of covered warrants in our empirical study and their practical findings. The conclusions are given in Section 6.

2. Wavelet-based feature extraction with support vector machines

In order to take the implied trading behavior into account, we decompose the mark to market price of underlying of options using wavelet analysis, and then the extracted time scale features serve as inputs of a SVM model to perform the nonparametric estimation process as valuation model.

Wavelet analysis has been applied in various fields for many years [10][11]. In this section we give a short introduction to the wavelet analysis. We use the multiresolution decomposition (MRD) [11][13] through this paper. MRD is implemented based on the technical skills from Bruce & Gao [14]. The details are described as follows.

2.1 Multi-resolution Decomposition (MRD)

Any function $f(t)$ in $L^2(R)$ can be decomposed by a sequence of projections onto the wavelet basis. The wavelet representation of the signal or function $f(t)$ in L^2(R) can be written as:

$$f(t)=\sum s_{J,k}\phi_{J,k}(t)+\sum d_{J,k}\psi_{J,k}(t)+\sum d_{J-1,k}\psi_{J-1,k}(t)+...+\sum d_{1,k}\psi_{1,k}(t), \quad (1)$$

where ϕ is the father wavelet and ψ the mother wavelet. $\phi_{j,k}$ and $\Psi_{j,k}$ are scaling and translation of ϕ and ψ, defined as:

$$\phi_{j,k}(t)=2^{-j/2}\phi(2^{-j}t-k)=2^{-j/2}\phi((t-2^{j}k)\times 2^{-j}) \quad (2)$$

$$\psi_{j,k}(t)=2^{-j/2}\psi(2^{-j}t-k)=2^{-j/2}\psi((t-2^{j}k)\times 2^{-j}) \quad (3)$$

In the representation J is the number of multiresolution components, and $s_{J,k}$ are call the smooth coefficients, and $d_{j,k}$ are called the detailed coefficients. If we define:

$$S_J(t)=\sum_k s_{J,k}\phi_{J,k}(t)\ ,\ D_j(t)=\sum_k d_{j,k}\varphi_{j,k}(t), \quad for \quad j=1,2,...,J \quad (4)$$

The equation (4) are called the smooth signal and the detail signals, respectively, which constitute a decomposition of s signal into orthogonal components at different scales. A signal $f(t)$ can thus be expressed in terms of these signals:

$$f(t)=S_J(t)+D_J(t)+D_{J-1}(t)+...+D_1(t) \quad (5)$$

2.2 Support Vector Machines (SVMs)

SVMs are a radically difference type of classifier which have attracted a great deal of attention lately due to the novelty of the ideals that they bring to pattern recognition and their significant results in practical problems. Given a training set $D=\{x_i,t_i\}_{i=1}^N$ with input vectors, $x_i=(x_i^{(1)},\ldots,x_i^{(n)})^T\in\mathbf{R}^n$ and target labels: $t_i\in\{-1,+1\}$, the SVMs classifier satisfies the conditions:

$$\begin{cases}\omega^T\varphi(x_i)+b\geq +1, & \text{if } t_i=+1\\ \omega^T\varphi(x_i)+b\leq -1, & \text{if } t_i=-1\end{cases} \quad (6)$$

which is equivalent to: $t_i\left[\omega^T\varphi(x_i)+b\right]\geq 1,\ \ i=1,...,N$, where ω represents the weight vector and b the bias. The nonlinear function $\varphi(\cdot):\mathrm{R}^n\to\mathrm{R}^{n_k}$ maps the input or measurement space to a high-dimensional, and possibly infinite-dimensional, FS. Comes down to the construction of two parallel bounding hyperplanes at opposite sides of a separating hyperplane $\omega^T\varphi(x)+b=0$ in the FS, with the margin width between both hyperplanes equal to $2\times\|\omega\|^{-2}$. In primal weight space, the classifier then takes the form: $y^t(x)=\mathrm{sgn}(\omega^T\varphi(x)+b)$. But, on the other hand, it is never evaluated in this form. One defines the optimization problem as:

$$\underset{\omega,b,\xi}{\mathrm{Min}}\ \tau(\omega,\xi)=\frac{1}{2}\omega^T\omega+C\sum_{i=1}^{N}\xi_i \quad (7)$$

$$s.t.\ \begin{cases}t_i(\omega^T\varphi(x_i)+b)\geq 1-\xi_i, & i=1,...,N\\ \xi_i\geq 0, & i=1,...,N\end{cases} \quad (8)$$

where, ξ_i : soft margin needed to allow misclassifications in the set of inequalities; $C\in\mathrm{R}^+$: tuning hyperparameter, weighting the importance of classification errors vis-à-vis the margin width. The solution of the optimization problem is obtained after constructing the Lagrangian. From the conditions of optimality, one obtains a quadratic programming problem in the Lagrange multipliers α_i. A multiplier α_i exists for each training data instance. Data instances corresponding to non-zero α_i are called support vectors. As is typical for SVMs, we never calculate ω or $\varphi(x)$. This is made possible due to Mercer's condition, which relates the mapping function $\varphi(x)$ to a kernel function $K(\cdot,\cdot)$ as follows. For the kernel function $K(\cdot,\cdot)$, Then construct the SVM classifier as: $K(x_i,x_j)=\varphi(x_i)^T\varphi(x_j)$.(Cristianini and Shawe-Taylor, 2000).

3. Artificial Intelligence Modeling using ANFIS

The neural-based fuzzy inference system under consideration is ANFIS [4], which is a first-order Sugeno model. The i^{th} If-Then rule of Sugeno model is:

$$R_i : If\ x_1\ \ is\ \ \tilde{A}_{i1}\ \ and \cdots and\ \ x_n\ \ is\ \ \tilde{A}_{in}$$
$$then\ \ \underbrace{f_i = c_{i0} + c_{i1}x_1 + \cdots + c_{in}x_n}_{first-order\ \ consequent\ \ equation} \quad (9)$$

In equation (9), $\tilde{A}_{il}$ is a fuzzy set (the MFs set as Gaussian function); f_i is the i^{th} first-order consequent equation. For more detail descriptions of ANFIS could be found in [4]. In addition, the basic functions of each layer in ANFIS are summarized as follows: Layer1: Fuzzification; Layer2: Firing strengths; Layer3: Normalization; Layer4: TSK outputs; Layer5: Summation.

The fine-tune procedures of ANFIS include applying recursive least-squares estimator and steepest descent algorithms for calibrating both premise and consequent parameters iteratively. The two-phase learning starts from the consequent parameters. The updating formula for estimating consequent parameters is

$$P^{(k+1)} = P^{(k)} - \frac{P^{(k)}a^{(k+1)}(a^{(k+1)})^T P^{(k)}}{1+(a^{(k+1)})^T P^{(k)} a^{(k+1)}},\ c^{(k+1)} = c^{(K)} + P^{(k+1)}a^{(k+1)}\left\{t^{(k+1)} - (a^{(k+1)})^T c^{(k)}\right\}. \quad (10)$$

In equation (10), vector c contains the estimated consequent parameters, elements of vector a are the normalized firing strength of each rule multiplies its corresponding inputs, and $t^{(k+1)}$ is the target value for the $(k+1)^{th}$ training pattern. The initial conditions for this iterative process are $c(0)=0$ and $P(0)=\Gamma I$, where I is an identity matrix and γ is a large positive value. A detail description of equation (10) can be also found in [4].

The second stage of learning involves the renewing premise parameters. Define the sum of squared errors for the k^{th} training pattern as $E^{(k)} = (t^{(k)} - O_5^{(k)})^2$ and $O_5^{(k)}$ is the actual output produced by the presentation of the k^{th} pattern.

4. Black-Scholes Formula with GARCH Estimation Model

A Black-Scholes Formula evaluation model is indicated in equation (11):

$$C = \quad S \times N(d1) - ke^{-r \times T} N(d2)\ ,$$
$$d_1 = \frac{\ln(S/K) + r \times T}{\sigma\sqrt{T}} + \frac{1}{2}\sigma\sqrt{T}\ ,\ d_2 = d_1 - \sigma\sqrt{T}\ . \quad (11)$$

where, C: the fair value of options, S: the spot price of the underlying asset, K: the strike price, r: the instantaneously risk-free rate, T: maturity, σ : the underlying return on the instantaneous standard deviation, $ln(.)$: the natural-log, and N (.): the accumulated properties of the standardized normal distribution.

In equation (11), a GARCH [15] model is employed here to estimate the parameter, σ , for theoretical fair value prediction of options. In the conditional mean part, r_t, is represented as $r_t = \alpha + \beta r_{t-1} + r x_{t-1} + \varepsilon_t$, which describes the causal relationship between current returns and lagged returns of underlying. r_t, is the daily return at time t , x_{t-1} is the lagged underlying return, thus, the innovation at time t could be written as $\varepsilon_t \sim N(0, \sigma_t)$. We model the volatility, σ , in equation (11) to time-variant volatility, σ_t , in order to follow the univariate GJR-GARCH(1,1) process:

$$\sigma_t = a_0 + a_1\varepsilon^2_{t-1} + c_1 S_{t-1}\varepsilon^2_{t-1} + b_1\sigma_{t-1}, \quad (12)$$

where

$$S_{t-1} = \begin{cases} 1 & if \quad \varepsilon_{t-1} < 0 \\ 0 & if \quad \varepsilon_{t-1} > 0 \end{cases} \quad (13)$$

Depending on whether ε_{t-1} is over or under the threshold value of zero, ε^2_{t-1} would has different effects on the conditional variance σ^2_t . The asymmetric impacts of ε_{t-1} on σ^2_t has known as the leverage effects.

5. Empirical Study and Analysis

Taking the Taiwan warrants market as empirical study, generally speaking, the issued covered warrants are mostly based on European-style BSM but in fact contracts are American-style. The targets selected for four comparative models, namely, BS-GARCH, ANFIS-G, BS-SVM, and WSVM, in our evaluation experiment include Concord Securities Group (SG), Yuanta SG, Yuanta SG, Taiwan SG, Masterlink SG and Yuanta SG, respectively, with underlying assets: Mega Holdings and Teco Corp. There are 121 pairs observations for each targets employed here. The period of our experiment extends from 2003 to 2004 with daily data as reported in Table 1. Considering the five-step-ahead estimating with rolling windows to verify the predictive stability instead of one-step-ahead would be helpful for tactical portfolio management and decrease in transaction cost while rebalancing the holding positions.

The ANFIS model is converged after training 1,000 epochs. Variables of premise include stock price, moneyness (S/K), volatility estimated with GARCH (1, 1), time to maturity (T), etc. The consequence is implied volatility. There are six MFs for each variable (Wang & Zeng, 2006). The BS-SVMs is also trained in a batch manner and input-output factors treated the same as ANFIS model. Except for the BS-SVMs, the wavelet analysis applies the Daubechies least asymmetric filters with length eight to decompose the explanatory variables, the daily returns and lagged returns of the underlying market price of warrants (r_{t-1}).

These returns are decomposed into four mutually orthogonal different frequency series, ranging from the lowest- frequency series to the highest- frequency series (selected example see Fig. 1), which served as the inputs of WSVMs to take the implied returns changing-signal into account. The parameters of SVMs-based models used in this study are set as follows: C=100, ε =0.01, and σ =0.8 for the Gaussian kernel. The valuation results for WSVMs are displayed from Fig. 2 to Fig. 7.

Table 1. Descriptions of warrant contracts in Taiwan covered warrant market

Warrant Code #	Warrants Name	Under-lying	Listing Day	Maturity	Exercise Price	Strike Ratio
0550	Concord01	MegaHoldings	4/8/03	3/2/04	21.93	1.04
0575	YuantaA4	MegaHoldings	21/8/03	20/2/04	19.37	1
0651	YuantaB9	MegaHoldings	22/9/03	22/3/04	17.2	1
0678	Taiwan14	MegaHoldings	14/10/03	13/4/04	20.7	1
0645	Masterlink23	Teco	19/9/03	18/3/04	13.09	1
0658	YuantaC4	Teco	25/9/03	24/3/04	14.75	1

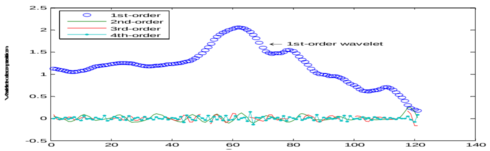

Fig. 1 The wavelet series of warrant code #: 0550 using MRD

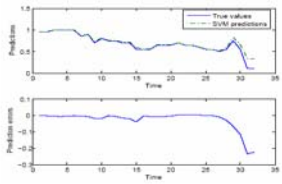
Fig. 2 WSVMs Estimates fair value of #0550

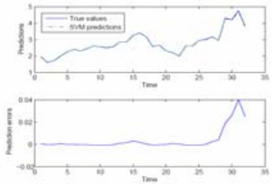
Fig. 3 WSVMs Estimates fair value of #0575

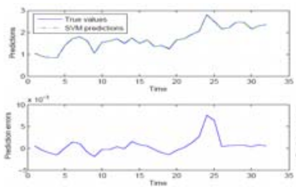
Fig. 4 WSVMs Estimates fair value of #0645

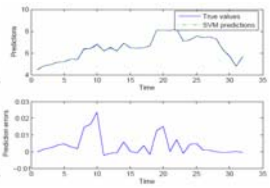
Fig. 5 WSVMs Estimates fair value of #0651

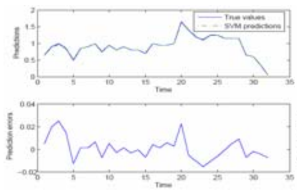
Fig. 6 WSVMs Estimates fair value of #0658

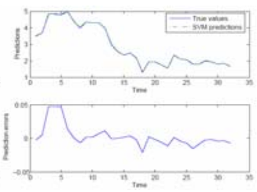
Fig. 7 WSVMs Estimates fair value of #0678

Table 2. Comparative valuation performance of four valuation model for six warrants case study by RMSE

(unit: × 0.01)

	# 0550	# 0575	# 0651	# 0678	# 0645	# 0658
BS-GARCH	0.286328	0.172234	0.190765	0.266787	0.228853	0.238453
ANFIS-G	0.755745	1.074685	1.01963	1.293897	0.55863	0.767899
BS-SVMs	0.7191	0.7909	0.3329	0.7835	0.4953	0.3828
WSVMs	0.0103	0.1042	0.0958	0.0368	0.0393	0.0368

We adopt the root mean square error (RMSE) as the performance index of the comparative valuation of four valuation model for six warrants case study, and illustrated in Table 2. It is showed that the proposed WSVM significantly outperform the other models. The BS-SVMs shows better performance than ANFIS-G but slightly poor than BS-GARCH. The impressive findings support the thoughts on the key features extraction deeply improve the mechanism of SVMs.

6. Conclusions

The success of the proposed WSVMs model could be attributed to the following three reasons: first, the structure of warrant price is changing over periodical time scale, WSVMs employ the MRD to follow the changing periodical structure, and, second, wavelet analysis is capable to capture all the structure-break (or the changing-point) to be the important features. The third reason is the important features from wavelet analysis enhance the SVMs' capability of mapping input data into high dimensional reproducing kernel Hilbert space which has robust topological structures to capture the nonlinear relationship and estimation ability. The interesting results also indicate that smooth signal is the major wave with the longest horizon which represents the main institution traders still considers the long-term allocation instead of intra-day trading or very short-term investment strategy.

Acknowledgement

This paper was supported by the National Science Council under contract number NSC- 95-2516-S-018 -016 -. In addition, the author is very grateful to the anonymous reviewers for their suggestions and comments.

7. References

[1] F. Black and M. Scholes, The Pricing of Options and Corporate Liabilities, Journal of Political Economy. Vol.81, pp. 637-59 (1973).
[2] Linda Canina & Stephen Figlewski, The Informational Content of Implied Volatility, The Review of Financial Studies, Vol.6, No. 3 (1993).
[3] C. John Hull, Options Futures and Other Derivatives, (Prentice-Hall International, 1997).
[4] J. S. R. Jang, 1993, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," IEEE Transaction on System, Man, and Cybernetics, vol.23, no3, pp.665-685.
[5] Emanuel Derman, Kani Iraj and Neil Chriss, Implied Trinomial Trees of the Volatility Smile, The Journal of Derivatives (1996).
[6] Hahn Shik Lee, International Transmission of Stock Market Movements: a Wavelet Analysis, Applied Economics Letters (2004).
[7] V. N. Vapnik, The Nature of Statistical Learning Theory (New York: Spring-Verlag, 1995).
[8] J. Daniel Strauss and Gabriele Steidl, Hybrid Wavelet-Support Vector Classification of Waveforms, Journal of Computational and Applied Mathematics, Vol. 148, pp. 375-400 (2002).
[9] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines (Cambridge University Press, 2000).
[10] G. H. Lee, Wavelets and Wave Estimation: a Review Journal of Economic Theory and Econometrics, Vol. 4, pp. 123-158, 1998.
[11] R. Gencay, F. Selcuk, and B. Whitcher, An Introduction to Wavelet and other Filtering Methods in Finance and Economics, (Academic Press, London, 2002).
[12] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA (1992).
[13] D. Percival and A. Walden, Wavelet Methods for Time Series Analysis (Cambridge University Press, Cambridge, 2000).
[14] A. Bruce and H. Y. Gao, Applied Wavelet Analysis with Splus, (Springer-Verlag, New York, 1996).
[15] T. Bollerslev, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, Vol. 31, pp. 307-327 (1986).
[16] Hsing-Wen Wang and Yao-De Zeng, Modifying Systematic Error for Valuation Models via nonparametric Financial Algorithms, working paper (2006).