# WARRANTS PRICE FORECASTING USING KERNEL MACHINE & EKF-ANN: A COMPARATIVE STUDY


**HSING-WEN WANG[1] JIAN-HONG WANG[2, 3] TSE-PING DONG[4] SHENG-HSUN HSU[5]**

[1] Department of Business Administration, National Changhua University of Education
[2] Department of Information Engineering, Chungchou Institute of Technology
[3] Department of Communication Engineering, National Chung Cheng University
[4] Graduate Institute of performing Arts, National Taiwan Normal University
[5] Department of Information Management, Chunghua University
e-mail: wanghong@dragon.ccut.edu.tw



## Abstract

Due to the six unreasonable assumptions companioned with the Black-Scholes options pricing model (BSM), which often make the miss-pricing result because of the difference of market convention in practical. This study try to combine the BSM and extended Kalman filters-based artificial neural networks (EKF-ANN) to deal with the limitation of consideration of the influences from many unexpected real world phenomena. If we were to soundly take these phenomena into account, the pricing error could be reduced. In this paper, we try to make a comparative study with examined the forecasting accuracy between the BSM-based kernel machines (KM-BSM) and the BSM-based EKF-ANN (EKF-ANN-BSM). From the evidence of Taiwan Warrants market, we found that the performance indicates the KM is superior to the others, and the hybrid EKF-ANN-BSM framework is also better than the pure EKF-ANN. The results show that the KM-BSM and hybrid model could significantly reduce the normalized root-mean-square-errors (NRMSE) of forecasting, it helps to provide an alternative way to refine the options valuation.



**Keywords**: Black-Scholes, kernel machines, extended Kalman filters, neural networks, Warrants.


## 1. Introduction

Warrants are one kind of options that has been traded in Taiwan for nine years, which are one of the most popular derivatives in financial studies in recent years. The well known formula for pricing Warrants is a series of BSMs. Since the 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].

Therefore, it has risen the widely applications of on-line trading system with decision support functions to consider the market phenomena, such as the programming trading systems. Traditionally, many efforts are devoted to employ the neural networks architecture based artificial intelligence (AI) methods to obtain more accurate pricing at real-time forecasting, however, they lacked mostly for several drawbacks of: (1) over-fitting on training data would cause a poor forecasting capability; (2) no standard criteria to determine the huge amount of networks parameters and initial weights; (3) easy to trapped into local optima etc. Numerous types of neural networks encountered the problems; (4) has a difficulty in explaining the causes of prediction result due to the lack of explanatory power; (5) suffers from difficulties with generalization because of over fitting; and (6) In addition, it needs too much times and efforts to construct a best architecture.

At the same time, a hybrid approach that integrates artificial neural networks, fuzzy inference, extended Kalman filters, backpropagations and other machine learning techniques has been suggested to improve the Warrants valuation accuracy. The results of comparative studies indicate that the adaptive neural-based fuzzy inference system (ANFIS) shows better prediction accuracy [4] with its powerful ability to find global optimal solution then different kind of neural networks in the past. Recently, KM has been gaining popularity to depend on the speedy convergence and approximate global optimal solution through its state-of-the-art techniques for regression on prediction [5]. This paper try to reveal the performance variations from the two different style models come from dissimilar learning mechanism and hence the conclusion is given for chosen selections.

The remainder of the paper is organized as follows. Section 2 describes the literatures on pricing the Taiwan Warrants. In Section 3, we discuss the powerful artificial intelligence framework, the hybrid EKF-ANN-BSM framework. Section 4 introduces the Kernel Machines with BSM. 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. Literatures on Pricing the Taiwan Warrants

Lin [6] focused to find whatever could affect the variations of time value or could provide other content of information and result the change of time value. Feng [7] empirically analyzed five stock options in Taiwan to examine whether there were exists mis-pricing between the open price and bid-ask price. Also, he has tested riskless arbitrage opportunity and the theoretical and empirical mis-pricing extent by using boundary conditions, the put-call parity and the BSM.

From most of the empirical results on pricing options and Warrants show that: (1). The five stock options would violate their lower bounds when trading costs are added in, and in-the-money options are less efficient than out-of-money and at-the-money options. (2).The stock options satisfy the put-call parity and do not exist riskless arbitrage opportunity. Nevertheless, for the bid-ask price composed of options has large gap, it would be explained that the stock options market in Taiwan is not liquid enough to be arbitraged. (3).Due to the less and less trading volumes of stock options, the increasing liquidity risk has made the theoretical price overvalued. And the option prices calculated in terms of implied volatility could be more close to the market price than those in terms of historical volatility. The results revealed that the BSM needs to be refined with certain framework to match the real derivatives market.

Through an empirical study it is discovered that the BSM assumptions are actually different from the practical situation [8], 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 many studies, a 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 [9].

## 3. The Hybrid EKF-ANN-BSM Framework

Extended Kalman filters-based artificial neural networks contain various frameworks; the typical one is named ANFIS, which is therefore employed in our study. The basic functions of each layer in ANFIS are summarized as follows: Layer1: Fuzzification; Layer2: Firing strengths; Layer3: Normalization; Layer4: TSK outputs; Layer5: Summation. It is a first-order Sugeno model. The $i^{th}$ If-Then rule of Sugeno model in the premise part and consequent part is as follows:

$$
\begin{aligned}
&R_i : If\ x_1\ \ is\ \ \tilde{A}_{il}\ \ and \cdots and\ \ x_n\ \ is\ \ \tilde{A}_{in} \\
&\quad then\ \ \underbrace{f_i = c_{io} + c_{i1}x_1 + \cdots + c_{in}x_n.}_{first-order\ \ consequent\ \ equation}
\end{aligned} \tag{1}
$$

In equation (1), $x_i$ is the input pricing factors; $\tilde{A}_{il}$ is a fuzzy set (the MFs set as Gaussian function); $f_i$ is the $i^{th}$ first-order consequent equation; $c_{in}$ is the coefficients of input variables, which is estimated by extended kalman filters; $n$ is the number of input factors. Obviously, the original linear polynomial equation in the consequent part is quite not enough to identify the highly nonlinear options (Warrants) price behavior and so as its returns or residuals. We perform the following BSM in the consequent part instead of $f_i = c_{io} + c_{i1}x_1 + \cdots + c_{in}x_n$ is as:

$$
\begin{aligned}
&R_i : If\ x_1\ \ is\ \ \tilde{A}_{il}\ \ and \cdots and\ \ x_n\ \ is\ \ \tilde{A}_{in} \\
&\quad then \\
&\quad \underbrace{\begin{aligned} f_i = S \times N\Big([\ln(S/K) + (r \times T)] \times \sigma\sqrt{T} + 0.5 \times \sigma\sqrt{T}\Big) \\ - ke^{-r \times T} N(d_1 - \sigma\sqrt{T}). \end{aligned}}_{BSM\ \ consequent\ equation}
\end{aligned} \tag{2}
$$

where, $f_i$: fair value of options; $S$: spot price of underlying; $K$: strike price; $r$ : instantaneously risk free rate; $T$: maturity; $\sigma$ : underlying return of instantaneously standard deviation; $ln$(.): natural-log; $N$ (.): accumulated properties of standardize normal distribution.

The reduced form of $f_i$ would be:

$$
\begin{aligned}
f_i = {} & c_{i0} + c_{i1}x_1 \times N\Big([\ln(c_{i1}x_1 / c_{i2}x_2) + (c_{i3}x_3 \times c_{i4}x_4)] \times c_{i5}x_5 \times \sqrt{c_{i4}x_4} + 0.5 \times c_{i5}x_5\sqrt{c_{i4}x_4}\Big) \\
& - c_{i2}x_2 e^{-c_{i3}x_3 \times c_{i4}x_4} N(\{[\ln(c_{i1}x_1 / c_{i2}x_2) + (c_{i3}x_3 \times c_{i4}x_4)] \times c_{i5}x_5 \times \sqrt{c_{i4}x_4} \\
& + 0.5 \times c_{i5}x_5\sqrt{c_{i4}x_4}\} - c_{i5}x_5\sqrt{c_{i4}x_4}).
\end{aligned} \tag{3}
$$

The fine-tune procedures of EKF-ANN-BSM 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 using Extended Kalman Filters is:

$$P^{(k+1)} = P^{(k)} - \frac{P^{(k)} a^{(k+1)} (a^{(k+1)})^T P^{(k)}}{1 + (a^{(k+1)})^T P^{(k)} a^{(k+1)}}, c^{(k+1)} = c^{(K)} + P^{(k+1)} a^{(k+1)} \{t^{(k+1)} - (a^{(k+1)})^T c^{(k)}\}. \quad (4)$$

In equation (4), vector $c$ contains the estimated consequent parameters, ($c_{i0}$~$c_{i5}$), 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 $\gamma$ is a large positive value.

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. Kernel Machines with BSM

This research chooses the support vector machine regression function (SVMRF) [10] from various KMs. In order to take the implied trading behavior into account, we extracted time scale features with SVMRF model to perform the nonparametric estimation process as valuation model. SVMRF is 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 SVMRF classifier satisfies the conditions:

$$\begin{cases} \omega^T \varphi(x_i) + b \geq +1, & \text{if } t_i = +1 \\ \omega^T \varphi(x_i) + b \leq -1, & \text{if } t_i = -1 \end{cases} \quad (5)$$

which is equivalent to: $t_i \left[ \omega^T \varphi(x_i) + b \right] \geq 1,\ i = 1, \ldots, N$, where, $\omega$ represents the weight vector and $b$ is 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:

$$\underset{\omega, b, \xi}{\mathrm{Min}}\ \tau(\omega, \xi) = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^N \xi_i \quad (6)$$

$$s.t. \begin{cases} t_i (\omega^T \varphi(x_i) + b) \geq 1 - \xi_i, & i = 1, \ldots, N \\ \xi_i \geq 0, & i = 1, \ldots, N \end{cases} \quad (7)$$

where, $\xi_i$ : 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, $\alpha_i$. A multiplier, $\alpha_i$, exists for each training data instance. Data instances corresponding to non-zero, $\alpha_i$, are called support vectors. As is typical for SVMRF, we never calculate, $\omega$, 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 SVMRF classifier as: $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ [5].

SVMRF demonstrates its powerful and superior prediction performance [5] under taking few computational resources [11]. In view of this, this study has adopted the SVMRF model to perform the non-parametric options valuation.

# 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 two comparative models, namely, EKF-ANN-BSM and SVMRF-BSM, 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. All factors are normalized and than fed into the models. 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 positions.

The EKF-ANN-BSM model is converged after training 10,000 epochs. Variables of premise include stock price ($S$), exercise price ($K$), volatility estimated with GARCH (1, 1) ($\sigma$) [12], time to maturity ($T$), interest rate ($r$), etc. The consequence is implied volatility. There are six MFs for each variable. The SVMRF-BSM is also trained in a batch manner and input-output factors treated the same as EKF-ANN-BSM model [13]. The valuation results for SVMRF-BSM is displayed in Table 2, in which, we adopt the NRMSE as the performance index of the comparative valuation of five valuation model for six warrants case study. It is showed that although all of the combination models have better outcomes than the original, pure or individual method, the proposed SVMRF-BSM significantly outperforms the other models. These results are also corresponding to the discoveries from recently studies [14] [15]. For the comparisons with the literatures, the Box-Jenkins statistical model [16] is took into account.

Due to the high nonlinear dynamics of options price, the SVMRF-BSM shows slightly better performance than EFK-ANN-BSM but better than BSM-GARCH. The impressive findings support the thoughts on the key features extraction deeply improve the mechanism of SVMRF-BSM. In terms of classification accuracies, our results indicated that AI is superior to that of time series forecaster. This might be due to following reasons: (1).SVMRF-BSM implements the structural risk minimization principle which minimizes an upper bound for the generalization error rather than minimizing the training error. However, EFK-ANN-BSM implements the empirical risk minimization principle, which might lead to worse generalization than SVMRF-BSM; (2).An EFK-ANN-BSM may not converge to global solutions due to its inherent algorithm design. In contrast, finding solutions in SVMRF-BSM is equivalent to solving a linearly constrained quadratic programming problem, which leads to a global optimal solution; (3).In choosing parameters, SVMRF-BSM are less complex than EFK-ANN-BSM. The parameters that must be determined in SVMRF-BSM are the kernel bandwidth $\delta^2$ and the margin *C*. However, in EFK-ANN-BSM, the number of hidden layers, number of hidden nodes, transfer functions and so on must be determined. Improper parameter selection might cause the over-fitting problem.

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 |

Table 2. Comparative valuation performance of two valuation model for six warrants case study by NRMSE (unit: $\times$ 0.01)

| | # 0550 | # 0575 | # 0651 | # 0678 | # 0645 | # 0658 |
|---|---|---|---|---|---|---|
| BSM-GARCH | 0.071240763 | 0.042853237 | 0.047463902 | 0.066378802 | 0.05694051 | 0.059329069 |
| Box-Jenkins statistical model | 0.092843 | 0.0610374 | 0.06153745 | 0.0985563 | 0.0712645 | 0.075428 |
| EKF-ANN | 0.02999548 | 0.042654192 | 0.040469062 | 0.051354704 | 0.022171996 | 0.030477871 |
| EKF-ANN-BSM | 0.029614328 | 0.030817650 | 0.029176479 | 0.041837095 | 0.021138242 | 0.027197363 |
| SVMRF-BSM | 0.02930662 | 0.032232799 | 0.013567201 | 0.031931216 | 0.020185745 | 0.015600854 |

# 6. Conclusions

The success of the proposed SVMRF-BSM model could be attributed to the following three reasons: first, the structure of warrant price is changing over periodical time scale, SVMRF-BSM follow the changing periodical structure, and, second, SVMRF-BSM are capable to capture all the structure-break (or the changing-point) to be the important features. The third reason is the important features enhance the SVMRF-BSM's 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.


# 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 relevant researchers' supports and 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, 2006).

[4] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System (IEEE Transaction on System, Man, and Cybernetics, Vol.23, No3, pp.665-685, 1993).

[5] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines (Cambridge University Press, 2000).

[6] H. C. Lin, Time value of Covered Warrants in Taiwan Stock Market (Unpublished master's thesis, National Chengchi University, 2004).

[7] H. H. Feng, Empirical Studies on Evaluation of Equity Options in Taiwan (Unpublished master's thesis, National Cheng Kung University, 2004).

[8] Emanuel Derman, Kani Iraj and Neil Chriss, Implied Trinomial Trees of the Volatility Smile, (The Journal of Derivatives, 1996).

[9] C. C. Chang, Futures and Options, (YehYeh Book Gallery, 2004).

[10] V. N. Vapnik, The Nature of Statistical Learning Theory (New York: Spring-Verlag, 1995).

[11]Michael Gavridis, Modeling with High Frequency Data: A Growing Interest for Financial Economists and Fund Managers, in Nonlinear Modeling of High Frequency Financial Time Series (John Wiley & Sons, Ltd., 1998).

[12] T. Bollerslev, Generalized Autoregressive Conditional Heteroskedasticity (Journal of Econometrics, Vol. 31, pp. 307-327, 1986).

[13] H. W. Wang and Y. D. Zeng, A Comparative Study of Adjusting Systematic Error via Hybrid Intelligence Valuation Inference System—Evidence from Basket Warrants (working paper, 2006).

[14] N. Terui and H. K. Dijk, Combined Forecasts from Linear and Nonlinear Time Series Models (International Journal of Forecasting, Vol. 18, pp. 421-438, 2002).

[15] G. P. Zang, Times Series Forecasting using a Hybrid ARIMA and Neural Network Model (Neurocomputing, Vol. 50, pp. 159-175, 2003).

[16] G. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control (2nd edition, Holden-Day, 1976).