

The Performance of Hybrid Artificial Neural Network Models for Option Pricing during Financial Crises

David Liu, Siyuan Huang

Department of Mathematical Sciences, Xi'an Jiaotong Liverpool University

Abstract: this paper provides a novel research on the pricing ability of the hybrid ANNs based upon the Hang Seng Index Options spanning a period of from Nov, 2005 to Oct, 2011, during which time the 2007-2008 financial crisis had developed. We study the performances of two hybrid networks integrated with Black-Scholes model and Corrado and Su model respectively. We find that hybrid neural networks trained by using the financial data retained from a booming period of a market cannot have good predicting performance for options for the period that undergoes a financial crisis (tumbling period in the market), therefore, it should be cautious for researchers/practitioners when carry out the predictions of option prices by using hybrid ANNs. Our findings have likely answered the recent puzzles about NN models regarding to their counterintuitive performance for option pricing during financial crises, and suggest that the incompetence of NN models for option pricing is likely due to the fact NN models may have been trained by using data from improper periods of market cycles (regimes), and is not necessarily due to the learning ability and the flexibility of NN models.

Key words: option pricing; hybrid artificial neural network; financial crisis; options

1. Introduction

ASince it was published in 1973, Black-Scholes model has been regarded as the foundation of a large number of conventional pricing models. Widely accepted though it is, researchers also find that Black-Scholes model has significant flaws due to the presumed prerequisites which are not feasible in the real financial market. According to empirical evidence, the distribution of option returns has a higher peak and heavier tails than that of normal distribution, and the implied volatility generated from Black-Scholes model shows a convex curve which is widely known as the “volatility smile” rather than a horizontal line as assumed (Kou, 2002). This to a large extent leads to a systematical bias when carrying out in-the-money and out-of-the-money option pricing, which has been proved in a multitude of research (Black,1975, MacBeth, and Merville, 1979), and biases might even change over time. (Rubinstein,1985). Therefore, later researchers focused much of their attention on varies of modified models in order to get a model fitted better for the real financial transactions when omitting some of the assumed conditions in Black-Scholes model(Kou, 2002, Saurabha, and Tiwari,2007, Bakshi, Cao, and Chen,1997). However, most of

these models are either too complex to apply in reality or have poor out-of-sample performance (Bakshi, Cao, and Chen, 1997).

Apart from the parametric and semi-parametric models, Hutchinson, Lo and Poggio (1994) proposed a nonparametric approach. They used an artificial neural network (ANN) to simulate the relationships between the input parameters and the derivative prices. This approach discards the restriction of Black-Scholes model and thus becomes more adaptive and flexible (Hutchinson, Lo, and Poggio, 1994). The comparison between artificial neural network and Black-Scholes model in derivatives pricing reveals that the neural network model is better in out-of-sample predicting with lower hedging errors (Amilon, 2003). Based on artificial neural network, later researchers (Andreou, Charalambous, and Martzoukos, 2006) also tested the forecasting performance of some hybrid models such as hybrid network with Black-Scholes model or Corrado and Su model (Corrado, and Su, 1996), hybrid network with Autoregressive Conditional Heteroscedasticity (ARCH) or with Generalized ARCH (GARCH) (Hajizadeh, Seifi, Fazel Zarandi, and Turksen, 2011). The results indicate that hybrid artificial neural networks model performs better than conventional parametric models in many cases (Andreou, Charalambous, and Martzoukos, 2006, Hajizadeh, Seifi, Fazel Zarandi, and Turksen, 2011, Lajbcygier, and Connor, 1997).

The purpose of this paper is to examine the effectiveness of the hybrid neural network models for option pricing especially during financial crises. This paper differs from previous research in two aspects. First, it focuses on the causes that would result in the incompetence of NN models (both hybrid neural networks and standard neural networks), while most of the previous research dealt with the comparisons of NN models with parametric models. Second, as it is widely recognized that volatility turns out to increase during financial crisis (Chong, 2011, Ramlall, 2010), and Black-Scholes model mispricing worsens when underlying volatility is high especially for the deep out-of-the-money options (Gencay, and Salih, 2003), it is of practical importance to investigate the effect of financial crisis on hybrid models and traditional neural network model respectively, while most previous researches have mainly been conducted under the condition that market volatility is relatively stable. As indicated by Lento and Gradojevic (2013) recently that little research has been conducted on the effectiveness of neural network option pricing models during the financial crises. Past literature (Lento and Gradojevic, 2013, Gencay and Gibson, 2009, Gradojevic, Gencay and Kukulj, 2009) suggest that the simpler BS model is however superior to the NN models during the periods of high volatility. In this regard, researchers (Lento and Gradojevic, 2013, Gencay and Gibson, 2009, Gradojevic, Gencay and Kukulj, 2009) suggest that more thorough research be needed in order to identify the causes for such counterintuitive performance of NN models.

In this research, we have not only examined the effectiveness of traditional neural networks on option pricing during financial crises, but also the hybrid-neural networks (i.e. neural networks coupled with parametric option pricing models). The research of this kind is novel and of great importance for the application of hybrid neural networks for option pricing.

In the next part of the paper, a brief review of the feed forward back-propagation (BP) networks and the hybrid ANN models will be provided, as well as BS and CS model. Then,

training and testing of the BP neural network model, the hybrid ANN with BS model, and the hybrid ANN with CS model will be conducted using the collected data of Hang Seng Index Option. In the Results section, in-sample forecasting results from each model will be provided. Finally, we will draw a brief conclusion of this research.

2. Methodology

Artificial Neural Network

Among various of neural networks, back-propagation networks have gained a wide application especially in modeling and forecasting [16]. Therefore, we adopt BP networks for this research.

Network Structure

There are four input nodes and only one output node which denotes the difference between the theoretical price and the market price of the option over their respective strike prices.

In determining the number of nodes in hidden layer, several criteria have been used such as the mean squared error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). According to the BIC expression developed by Kashyap [17] and Schwarz [18] independently:

$$BIC = M \ln(MSE) + P \ln(M) \quad (1)$$

The AIC formula proposed by Akaike [19] and then adjusted by Box and Jenkins [20] is

$$AIC = M \ln(MSE) + 2P \quad (2)$$

where $P = \sum_{i=0}^{L-1} N_{i+1}(N_i + 1)$ (3)

$$MSE = SE / M \quad (4)$$

P is the number of weights and biases in network;

L is the output layer ($L = 2$ in this research);

N_i is the number of nodes in the i th layer

M is the number of data pair used in training process

Transfer Function

In BP multilayer networks, three kinds of transfer function are usually used. They are log-sigmoid transfer function, tan-sigmoid transfer function, and linear transfer function, and it is essential for these transfer functions to have derivatives in BP networks.

The log-sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

The tan-sigmoid function:

$$g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (6)$$

The linear function:

$$h(x) = x \quad (7)$$

In this research, the linear transfer function is used in the output layer because it is expected that the option prices should not be restricted into a specific range as (0,1) in log-sigmoid function or (-1,1) in tan-sigmoid function. Tan-sigmoid function is used in the hidden layer.

The outputs of layer two (option prices) can be expressed as

$$O = h \left[b_2 + \sum_{j=1}^m w_j f_j \left(b_j + \sum_{i=1}^n w_i^j I_i \right) \right] \quad (8)$$

where n is the total number of nodes in layer 0 (input layer);

m is the total number of nodes in layer 1 (hidden layer);

I_i is the i th input in layer 0;

w_i^j is the weight assigned to input I_i by node j in layer 1;

b_j is the bias of j th node in layer 1;

f_j is the tan-sigmoid transfer function of j th node in layer 1;

w_j is the weight assigned to the output of j th node in layer 1 by the single node in layer 2 (output layer);

b_2 is the bias of the single node in layer 2;

h is the linear transfer function of the single node in layer 2.

O is the final output from layer 2 which is the forecasted option price.

Hybrid Neural Net with Black-Scholes Model (BS-ANN)

The Black-Scholes formula for option pricing is

$$C^{BS} = S_0 N(d_1) - Ke^{-r} N(d_2) \quad (9)$$

$$\text{where } d_1 = \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)T}{\sigma\sqrt{T}} \quad (10)$$

$$d_2 = \frac{\ln(S_0 / K) + (r - \sigma^2 / 2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \quad (11)$$

C^{BS} is the European call option price on a non-dividend-paying asset at time 0;

S_0 is the price of the underlying asset at time 0;

K is the strike price;

r is the continuously compounded risk-free rate;

σ is the volatility of the underlying asset in one year;

T is the time to maturity, and

$N(x)$ is the cumulative probability distribution function.

In the hybrid model of artificial neural network with Black-Scholes model, values of parameters of σ , S_0 , K , r and T are collected and put into Black-Scholes model in order to get C^{BS} . The differences between those forecasted values C^{BS} and the real market values C^M will be put into artificial neural network afterwards as training targets. With other inputs σ , S_0 / K , r and T , ANN will generate the predicted difference $\frac{C^{BS} - C^M}{K}$. The final forecasted result is the sum of the multiplication of this predicted difference from ANN with its corresponding strike price K and the theoretical value of BS model, which can be expressed as $K(\Delta(C^{ANN} / K) + C^{BS} / K)$.

Hybrid Neural Net with Corrado and Su Model (CS-ANN)

Corrado and Su [10] adjusted the Black-Scholes model with a Gram-Charlier series expansion in order to take into consideration the skewness and kurtosis deviations from normality in stock returns. While it is similar to the semi-parametric option pricing model proposed by Jarrow and Rudd [21], but Corrado and Su's model are easier to report and interpret [10]. Evidence shows that this semi-parametric model performs significantly better than Black-Scholes model [5]. Hence, the latter model will be used in this paper.

Corrado and Su [10] modified the standard normal density function by a Gram-Charlier series expansion. The modified density function is

$$g(z) = n(z) \left[1 + \frac{\mu_3}{3!} (z^3 - 3z) + \frac{\mu_4 - 3}{4!} (z^4 - 6z^2 + 3) \right] \quad (12)$$

where $z = \frac{\ln(S_t / S_0) - (r - \sigma^2 / 2)T}{\sigma\sqrt{T}}$ (13)

$n(z)$ is the standard normal probability density function;

S_t is a random stock price at time t .

μ_3 is the nonnormal skewness;

μ_4 is the nonnormal kurtosis.

The option pricing formula after modification is

$$C^{CS} = C^{BS} + \mu_3 Q_3 + (\mu_4 - 3) Q_4 \quad (14)$$

where $C^{BS} = S_0 N(d_1) - Ke^{-rt} N(d_2)$ (15)

$$Q_3 = \frac{1}{3!} S_0 \sigma \sqrt{T} \left[(2\sigma\sqrt{T} - d_1) n(d_1) + \sigma^2 T N(d_1) \right] \quad (16)$$

$$Q_4 = \frac{1}{4!} S_0 \sigma \sqrt{T} \left[(d_1^2 - 1 - 3\sigma\sqrt{T} d_2) n(d_1) + \sigma^3 T^{3/2} N(d_1) \right] \quad (17)$$

Similarly, the differences between forecasted values from Corrado and Su model with market values over their corresponding strike price will also be put into the network as targets in the hybrid model of artificial neural network with Corrado and Su model. The final forecasted result from this model is $K(\Delta(C^{ANN} / K) + C^{CS} / K)$.

Training Process for option pricing

Among all kinds of training algorithms, Levenberg-Marquardt (LM) modified algorithms is the fastest and is also the default training function in MATLAB [21]. Due to the large amount of memory spaces it requires, LM algorithm is usually used in small or medium sized network. Considering that the size of the network here is not large, gradient descent LM training algorithm will be used in this research.

To train the network, the weights and biases were first initialized and the initial values are generated automatically by the network. With these weights and biases, the network will calculate the option price and the error of the network E which is defined as

$$E = f(w) = \sum_{i=1}^M e_i^2 = \sum_{i=1}^M (O_i - T_i)^2 \quad (18)$$

where M is the total number of input and output pairs as defined previously. Then, the weights and biases will be adjusted after each iteration following

$$x_{t+1} = x_t - \lambda \nabla f \quad (19)$$

where λ is the learning rate. As explained by Hagan and Menhaj [23] in their paper, the derivatives of $f(w)$ is expressed by Jacobian matrix in LM algorithm.

$$J(x) = \begin{bmatrix} \frac{\partial e_1}{\partial x_1} & \dots & \frac{\partial e_1}{\partial x_r} \\ \dots & \dots & \dots \\ \frac{\partial e_M}{\partial x_1} & \dots & \frac{\partial e_M}{\partial x_r} \end{bmatrix}$$

where x_i represents the element in parameter column vector X containing the weights and biases.

$$\text{Hence, } \nabla f(w) = J^T e \quad (20)$$

$$\text{and } \nabla^2 f = J^T J + \sum_{i=1}^M e_i \nabla^2 e_i \quad (21)$$

So, the parameter matrix will be adjusted through equation

$$X_{t+1} = X_t - [J^T J + \mu I]^{-1} J^T e \quad (22)$$

in each iteration, where μ is a learning parameter.

3. Hong Kong (HK) Hang Seng Index Options

The data adopted in this research ranges from 1st Nov, 2005 to 31st Oct, 2011 and covers the subprime mortgage crisis in 2008 during which period of time, a large number of stocks' prices experienced a slump [27]. The financial option we considered in this research is the Hang Seng Index Option traded on Hong Kong Stock Exchange based upon the Hang Seng stock market index.

Parameters: T and S/K

The performances of the aforementioned three models are tested separately in several groups by considering that the predicted bias of BS model differs across maturity and moneyness [6]. Based upon their time to maturity, the data are separated into long term options ($T > 180$ days), medium term options ($60 < T \leq 180$ days), and short term options ($T \leq 60$ days). Further, based upon the moneyness of the option, S_0/K , each of the previous three groups are further separated into three subgroups as did in Bakshi, Cao and Chen's research [6]: Out of-the-money (OTM)

($S_0/K \leq 0.97$), At-the-money (ATM) ($0.97 < S_0/K < 1.03$), and In-the-money (ITM) ($S_0/K \geq 1.03$). Consequently, in total 9 groups of data consist of 116,467 observations are summarized in Table 1.

Parameters: risk free rate, historical volatility, skewness and kurtosis

One year deposit rate covering the time period of the collected data ranges from 2.25% to 4.14%. To simplify the model, a risk free rate of 3% is used in this paper. Yearly historical volatility of the returns in the past 60 days (including the present day) is calculated as the volatility of the present day. Skewness and kurtosis of the returns are also calculated using historical data of the past 60 days. It is assumed that there are 252 trading days per year.

Table 1. Hong kong index options

Moneyiness	Time to maturity			Subtotal
S/K	(0,60]	(60,180]	>180	
OTM	HK\$113.52	HK\$295.95	HK\$521.58	[74094]
No. of obvs	[41288]	[27304]	[5502]	
ATM	HK\$568.10	HK\$950.76	HK\$1686.55	[27883]
No. of obvs	[16286]	[9911]	[1686]	
ITM	HK\$1763.36	HK\$2433.84	HK\$4628.34	[14490]
No. of obvs	[10401]	[3172]	[917]	
Subtotal	[67975]	[40387]	[8105]	[116467]

Pricing options under financial crisis

One of the main purposes of this research is to test the forecast performances of neural network models for option prices during the 2008 financial crisis. Therefore, the data from each of the nine groups above are used for training the networks but also based upon different periods of time that are consistent with the developing stages of the subprime mortgage crisis in 2008 as shown in Figure 1.

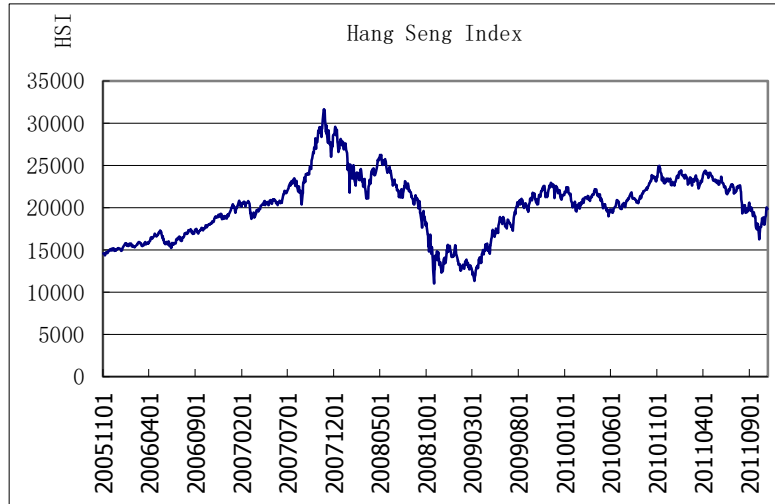


Fig. 1 Trend of Hang Seng Index from Nov, 2005 to Oct, 2011

We now roughly split the time in consideration into four stages reflecting different periods of development of the 2008 financial crisis in Hong Kong. These can be seen in Table 2.

Table 2. Stages of development of the financial crisis

Stage	Period	Market Features
1	1 st Nov 2005 – 31 st Oct 2007	Before financial crisis (booming)
2	1 st Nov 2007 – 15 th Mar 2009	Financial crisis (tumbling)
3	16 th Mar 2009 – 9 th Feb 2010	Fast recovering (booming)
4	10 th Feb 2010 – 31 st Oct 2011	Slow recovering (stabilising)

We have designed in total 6 simulation sets (as shown in Table 3) for the nine groups of the index options mentioned earlier. In each of the sets 1, 2 and 3, data from Stage 1 are used for training. For example, when using the BS-ANN model, the input data for the model are obtained from Stage 1, and the targeted outputs are also from the same stage. After training, the network is then used as an option pricing tool for Stage 2, 3, and 4 separately. For Simulation sets 4, 5, and 6, we combine two or three stages in time to train neural network models, and the trained models are then used for prediction of the option prices for other stage(s). For instance, in Simulation Set 6, data from Stage 1, 2, and 3 are altogether used to train the network, which is later used to simulate the option prices in Stage 4.

Table 3. The simulation sets

Simulation set	Training set	Prediction
1	Stage 1	Stage 2
2	Stage 1	Stage 3
3	Stage 1	Stage 4
4	Stage 1&2	Stage 3
5	Stage 1&2	Stage 4
6	Stage 1, 2,&3	Stage 4

4. Results and discussions

To provide a thorough comparison among the three models, four different measures are used:

$$ME = \frac{1}{N} \sum_{i=1}^N (F_i - A_i) \quad (23)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - A_i)^2} \quad (24)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |F_i - A_i| \quad (25)$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{F_i - A_i}{A_i} \right| \quad (26)$$

Where

N is the number of data pair used in simulate;

F_i is the forecasted option price;

A_i is the actual price from the market.

Preliminary results

The first purpose of this research is to test the prediction performance of the models in each stage of the financial crisis. Firstly, we want to know whether learning of prices from a certain stage (s) would cause any side effect to the prediction of prices. Secondly, we want to find out whether hybrid networks possess any advantages over the traditional network.

Table 4: RMSEs of each simulation set in artificial neural network

RMSE ANN		L			M			S		
Training	Prediction	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
S1*	S2	1355.3636	1556.9918	2161.878	1310.0963	639.66857	764.30342	2016.1897	1263.0071	992.20655
S1	S3	1151.3077	669.38867	1132.3559	721.71177	142.59979	512.38419	275.47231	154.62877	341.35742
S1	S4	616.53639	294.20669	393.17118	221.07265	151.68939	230.3813	88.452636	103.34483	133.69532
S1&2	S3	127.35899	244.95334	262.33144	76.446726	187.9516	179.77449	42.461521	100.16175	135.73236
S1&2	S4	133.91734	230.56396	271.07116	94.902816	156.27682	185.11979	48.611721	102.7587	107.72082
S1,2&3	S4	120.36331	214.16583	292.1646	90.784591	149.08585	175.12264	46.484094	93.824601	102.90797

*S1 stands for Stage 1

Table 5. RMSEs of each simulation set for the Hybrid Model CS-ANN

RMSE CS-ANN		L			M			S		
Training	Prediction	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
S1	S2	2141.7422	957.8311	1094.4773	631.63212	427.12016	499.41196	298.01252	941.10333	327.1581
S1	S3	653.59865	528.46051	814.59158	195.98336	299.05957	426.30136	110.20774	594.77594	245.86609
S1	S4	108.99772	277.93967	511.64853	111.16594	180.21061	240.48015	59.75144	302.4292	160.21959
S1&2	S3	159.97553	276.70231	761.67637	99.257598	156.35721	235.8928	245.15136	123.54767	144.29526
S1&2	S4	95.637359	247.086	295.20516	91.217341	182.34225	187.23117	72.516084	104.2104	110.78911
S1,2&3	S4	91.624076	238.74704	2364.6613	92.565105	158.98428	172.54063	48.00654	101.89797	103.78162

The error measures (RMSE) of each simulation set are extracted and summarized in Table 4, 5 and 6. It is clear that the worst performing excises lie with Simulation Set 1 (shown in Table 3). In this set, the booming stage before the financial crisis was used for training the networks, which were then used for predicting the prices for the tumbling period of the market. This phenomenon can be seen in all three networks regardless the network is hybrid or not. Largest errors always occurred in Set 1. The pricing performance become generally better when the networks (trained by using Stage 1) are used to predict option prices for Stage 3 and 4 which represent for a recovering economics after the financial crisis, though there are some large errors when predicting OTM options. The performance becomes better because the later stages in the market bear similar features to those in Stage 1, which was that market prices were climbing up in one way or the other. When we used more stages as training sets, for instance, using Stages 1, 2 and 3 as a whole, the prediction of prices of options gets better. Take the RMSE of BS-ANN

in long term, out-of-the-money group as an example. The RMSE of Stage 4 with a training set of Stage 1 is 829.50198, and that of the same stage with a training data from Stage 1 and Stage 2 together is 123.92778. The difference is as large as over 700. As seen in the errors of RMSEs with a training set of Stage 1 and Stage 2 or a training set of Stage 1, 2, and 3 altogether are much smaller. This holds true almost for all the groups.

On the other hand, the network trained only by Stage 1 can generate better results (smaller errors) when used to prediction option prices for Stage 3 or Stage 4, though the results are a little worse than those obtained by the networks trained by using more than 2 stages. All the phenomena indicate that the choice of a training set for a network is crucial for its prediction performance. Researchers need to pay attention to the financial data from which period of the economic cycles that is retained for training, and for which period that is to be used for prediction. Of course, large quantity of training data could certainly generate smaller errors, however, the improvement could be limited as seen in our simulation that a network trained by using a combination of Stages 1, 2 and 3 didn't outperform much over a network trained by using Stages 1 and 2. In our view, it was the diluteness effect of combination of booming period (S1) and crashing period (S2) in the market that resulted in better prediction for a stable period (S4), not necessarily because of larger data amount. Of course, the later certainly helps.

Table 6: RMSEs of each simulation set in Hybrid Model 2 (CS-ANN)

RMSE CS-ANN		L			M			S		
Training	Prediction	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
S1	S2	2141.7422	957.8311	1094.4773	631.63212	427.12016	499.41196	298.01252	941.10333	327.1581
S1	S3	653.59865	528.46051	814.59158	195.98336	299.05957	426.30136	110.20774	594.77594	245.86609
S1	S4	108.99772	277.93967	511.64853	111.16594	180.21061	240.48015	59.75144	302.4292	160.21959
S1&2	S3	159.97553	276.70231	761.67637	99.257598	156.35721	235.8928	245.15136	123.54767	144.29526
S1&2	S4	95.637359	247.086	295.20516	91.217341	182.34225	187.23117	72.516084	104.2104	110.78911
S1,2&3	S4	91.624076	238.74704	2364.6613	92.565105	158.98428	172.54063	48.00654	101.89797	103.78162

Table 7: Error measures of each simulation set 1

Simulation set 1	L			M			S			
	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM	
ANN	1003.7967	1089.6223	1560.8882	1035.444	-344.9254	347.35144	1143.9516	-480.5624	461.71309	
ME	BS-ANN	-4712.07	472.79527	1168.2261	331.85378	712.51359	339.06811	1135.3272	397.59058	203.52689
	CS-ANN	1309.4909	670.06779	853.75597	412.33942	254.25163	348.13177	149.94286	498.09419	210.15869
	ANN	1355.3636	1556.9918	2161.878	1310.0963	639.66857	764.30342	2016.1897	1263.0071	992.20655
RMSE	BS-ANN	14294.774	698.25948	1463.6269	511.14249	1047.4885	514.23247	2132.325	518.03605	311.04159
	CS-ANN	2141.7422	957.8311	1094.4773	631.63212	427.12016	499.41196	298.01252	941.10333	327.1581
	ANN	1061.8612	1140.5856	1596.7049	1056.535	466.31536	563.07748	1198.6332	675.27138	604.8783
MAE	BS-ANN	5517.7467	542.73745	1222.0666	380.11521	754.01412	386.51086	1149.2039	414.52023	252.86513
	CS-ANN	1399.9303	721.57394	871.5729	468.02413	334.88126	388.9443	206.6553	738.78484	264.56355
	ANN	79.742693	0.4461355	0.3956542	70.544843	0.299637	0.2036564	180.28338	0.9648541	0.3444915
MAPE	BS-ANN	664.61821	0.2115112	0.2739962	13.689559	0.4926544	0.1478211	175.275	0.7243695	0.1545263
	CS-ANN	39.720518	0.2828634	0.2054983	15.724776	0.2287239	0.1468097	33.54445	1.0964372	0.1648266

With the other objective in our mind, we now pick Simulation Set 1 (shown in Table 7) as an example (the comparisons in relation to other Simulation Sets can be found in the Appendix). Understandingly, this is the worst case that we have had with all the networks having performed badly for the periods during which a regime switch in the financial market had occurred. As shown before, in this simulation, the booming stage before the 2008 financial crisis was used for training the networks, which were then used for predicting the prices for the tumbling period of the market. What we have observed here is that the hybrid CS-ANN and BS-ANN models performed a little better than the NN models in overall, while the BS-ANN performed worst for OTM options for options with long term maturity. This is somewhat related to the fact that Black-Scholes model mispricing worsens when underlying volatility is high especially for the deep out-of-the-money options [15].

Although network models can make a significant contribution in improving forecast precision, there is a common and significant deficiency in all the three models attributed to this. They all might generate negative prices for out-of-the-money options. A small moneyness ratio S_0/K usually leads to a relatively low option price in the real market as shown in Table 1. In standard neural network model, it is hard for the network to control the extent to which the option price is close to zero. Similar to this, it is also hard for the network to control the difference

between the predicted price from BS (or CS) model and the real price from the market ($\frac{C^{BS} - C^M}{K}$). This deficiency in simulation would probability lead to a negative sum $\Delta(C^{ANN} / K) + C^{BS} / K$, and further a negative predicted option price $K(\Delta(C^{ANN} / K) + C^{BS} / K)$.

Discussions and conclusions

Recent researches on non-parametric models for option pricing have shown NN models to be less effective than parametric models (such as BS model) especially during the periods of financial crises. Lento and Gradojevic [24] ever concluded after pricing options for 1987 and 2008 by using NN models, that “the very advantages of non-parametric models over their parametric counterparts such as learning ability and the flexibility of functional forms largely contribute to the poor performance of non-parametric models when markets are highly volatile and experience a regime shift.”. However, our research suggests that the incompetence of NN models for option pricing is likely due to the fact NN models may have been trained by using data from improper periods of market cycles (regimes), and is not necessarily due to the learning ability and the flexibility of NN models. Further, one should be aware that this research is not only about traditional neural networks, it is mainly about the hybrid ANN models for option pricing, i.e. the ANNs coupled with parametric option pricing models. The research is therefore novel on this aspect and the related results should be mainly referred to hybrid ANN models for options.

In conclusion, using data of Hang Seng Index Options dating from Nov, 2005 to Oct, 2011, this paper tests the performances of a standard neural network and two hybrid neural networks in option pricing. In order to test their performances in different stages of financial crisis, the data is separated into several groups and are forecasted separately. Four error measures including ME, RMSE, MAE, and MAPE are used to compare the performances. Two conclusions can be drawn based up our research: 1) a neural network trained by using the financial training data retained from a booming period of a market cannot have good predicting performance for options for the period which undergoes the financial crisis (tumbling period in the market), therefore, it should be cautious for researchers/practitioners when carry out the predictions of option prices by using either a standard ANN or a hybrid ANN. 2) We observed that Hybrid ANNs performed slightly better than traditional ANNs for option prices in the course of financial crisis except for predicting OTM option prices, while BS-ANN performed worst for OTM options for options with long term maturity.

Acknowledgment

The authors acknowledge that some results of this paper have been published in the third international conference of statistics and operational research, Singapore, 21-22 April, 2013.

References

- [1] Kou, S (2002), 'A jump-diffusion model for option pricing', *Management Science*, vol. 48, No. 8, pp. 1086-1101
- [2] Black, F (1975), 'Fact and fantasy in the use of options', *Financial Analysts Journal*, vol. 31, pp. 36-72
- [3] MacBeth, J & Merville, L (1979), 'An empirical examination of the Black-Scholes call option pricing model', *Journal of Finance*, vol. 34, No. 5, pp. 1173-1186
- [4] Rubinstein, M (1985), 'Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978', *Journal of Finance*, vol. 40, No. 2, pp. 455-480
- [5] Saurabha, R & Tiwari, M (2007), 'Empirical study of the effect of including skewness and kurtosis in Black Scholes option pricing formula on S&P CNX Nifty index options', *Munich Personal RePEc Archive*, No. 6329
- [6] Bakshi, G, Cao, C & Chen, Z (1997), 'Empirical performance of alternative option pricing models', *Journal of Finance*, vol. 52, No. 5, pp. 2003-2049
- [7] Hutchinson, J, Lo, A & Poggio, T (1994), 'A nonparametric approach to pricing and hedging derivative securities via learning networks', *Journal of Finance*, vol. 49, No. 3, pp. 851-889
- [8] Amilon, H (2003), 'A neural network versus Black-Scholes: a comparison of pricing and hedging performances', *Journal of Forecasting*, vol. 22, No. 4, pp. 317-335
- [9] Andreou, P, Charalambous, C & Martzoukos, S (2006), 'Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters', *European Journal of Operational Research*, vol. 185, No. 3, pp. 1415-1433
- [10] Corrado, C & Su, T (1996), 'Skewness and kurtosis in S&P 500 index returns implied by option prices', *Journal of Financial Research*, vol. XIX, No. 2, pp. 175-192
- [11] Hajizadeh, E, Seifi, A, Fazel Zarandi, M & Turksen, I (2011), 'A hybrid modeling approach for forecasting the volatility of S&P 500 index return', *Expert Systems with Applications*, vol. 39, No. 1, pp. 431-436

- [12]Lajbcygier, P & Connor, J (1997), 'Improved option pricing using artificial neural networks and bootstrap methods', *International Journal of Neural Systems*, vol. 8, No. 4, pp. 457-471
- [13]Chong, C (2011), 'Effect of subprime crisis on U.S. stock market return and volatility', *Global Economy and Finance Journal*, vol. 4, No. 1, pp. 102-111
- [14]Ramlall, I (2010), 'Has the US subprime crisis accentuated volatility clustering and leverage effects in major international stock markets?', *International Research Journal of Finance and Economics*, No. 39, pp. 157-185
- [15]Gencay, R & Salih, A (2003), 'Degree of mispricing with the Black-Scholes model and nonparametric cures', *Annals of Economics and Finance*, vol. 4, No. 1, pp. 73-101
- [16]Chu, C & Widjaja, D (1994), 'Neural network system for forecasting method selection', *Decision Support systems*, vol. 12, No. 1, pp. 13-24
- [17]Kashyap, R (1977), 'A Bayesian comparison of different classes of dynamic models using empirical data', *IEEE Trans. Automatic Control*, vol. 22, pp. 715-727
- [18]Schwarz, G (1978), 'Estimating the dimension of a model', *The Annals of Statistics*, vol. 6, No. 2, pp. 461-464
- [19]Akaike, H (1974), 'A new look at the statistical model identification', *IEEE Transaction on Automatic Control*, vol. 19, pp. 715-727
- [20]Box, G, & Jenkins, G (1976), *Time Series Analysis Forecasting and Control Revised Edition*, Holden-Day, San Francisco
- [21]Jarrow, R, & Rudd, A (1982), 'Approximate option valuation for arbitrary stochastic processes', *Journal of Financial Economics*, vol. 10, No. 3, pp. 347-369
- [22]Beale, M, Hagan, M & Demuth, H (1992), 'Neural network toolbox user's guide', Math Works, Inc. p. 2-18, 2-32
- [23]Hagan, M & Menhaj, M (1994), 'Training feedforward networks with the Marquardt algorithm', *IEEE Transactions on Neural Networks*, vol. 5, No. 6, pp. 989-993
- [24] Lento C. and N. Gradojevic, (2013) "The effectiveness of Option Pricing Models During Financial Crises", *Rethinking Valuation and Pricing Models*. [Http://dx.doi.org/10.1016/B978-0-12-415875-7.00001-4](http://dx.doi.org/10.1016/B978-0-12-415875-7.00001-4)
- [25] Gencay R. and R. Gibson, (2009) "Model risk for European-Style Stock Index Options", *IEEE Transactions on Neural Networks*, 18, 193-202
- [26] Gradojevic, N. Gencay R. and D. Kukolj, (2009) "Option Pricing with modular neural networks", *IEEE Transactions on Neural Networks*, 20, 626-637

- [27] Kim T. Y. , Hwang C. and J. Lee, (2004) “Korean Economic Condition Indicator Using a Neural Network Trained on the 1997 Crisis”, *Journal of Data Science* 2, 371-381

Received March 15, 2013; accepted November 10, 2013.

David Liu
Department of Mathematical Sciences
Xi'an Jiaotong Liverpool University

Siyuan Huang
Department of Mathematical Sciences
Xi'an Jiaotong Liverpool University

