Neural networks as a semiparametric option pricing tool

Michaela Baruníkova*

Institute of Economic Studies, Charles University, Prague.

e-mail: babenababena@gmail.com

Jozef Baruník

Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Institute of Economic Studies, Charles University, Prague.

Institute of Economic Studies, Charles University, Prague.

e-mail: barunik@utia.cas.cz

Abstract. We study the ability of artificial neural networks to price the European style call and put options on the S&P 500 index covering the daily data for the period from June 2004 to June 2007. We divide the data set into several categories according to moneyness and time to maturity. We then price all options within the categories. The results show that neural networks outperform benchmark ad hoc Black-Scholes model with significantly lower pricing errors across all categories for both call and put options. Moreover, the differences between ad hoc Black-Scholes and neural networks errors widen with deepness of moneyness or longer time to maturity. We show that neural networks, even without the volatility input, can correct for the Black-Scholes maturity and moneyness bias.

Keywords: Keywords: option valuation, neural network, S&P 500 index options

JEL classification: JEL: C13,C14, G13

1. Introduction

Options belong to the wide family of derivatives, price of which is determined the by underlying security price. They can effectively reduce the risk as they allow investors to fix a price for future transaction. Much of the success and growth of the market for options and other derivatives is attributed to Black and Scholes (1973) and Merton (1973), who derived a closed-form option pricing formula through a dynamic hedging argument and no-arbitrage condition. Bernstein (1998) points out that the model was widely in use by practitioners before it was recognized through publication in academic journals. Since then the original formula has been generalized, extended, and applied to a vast array of securities. For review of conventional option pricing techniques, see e.g. Bates (2003).

* *Corresponding Author*

Pricing based on the hedging/no-arbitrage condition approach depends heavily on the price of the underlying asset and its volatility. Thus the misspecification of the stochastic process driving the stock price produces the systematic pricing and hedging errors. Therefore, the success of parametric pricing methods strongly depend on the ability to capture the dynamics of price process of the underlying asset. However this dynamics is stochastic, and the volatility varies over time. Thus any analytic formula is difficult to formalize.

In this paper, we use alternative data-driven method for pricing derivative securities - semiparametric neural networks. The term semiparametric is explained by the fact that the basis functions are parametric, yet the parameters are not the object of interest since we may need an infinity of them to estimate the function in the usual nonparametric sense. Neural networks are emerging computational technology that provide a complex method for exploring the dynamics of various economic and financial applications. Most studies have focused on prediction of financial data as neural networks are effective for input and output relationship modeling of noisy data containing nonlinearities. Among the most recent references, we mention Medeiros et al. (2005); Black and McMillan (2004); Jasic and Wood (2004); Rapach and Wohar (2005); Baruník (2008). McNelis (2005) provides a good literature review of other applications to finance. Based on the universal approximation theorem, neural networks are able to improve the option pricing as they are able to approximate any function (Hornik et al., 1989). The data is allowed to determine both the dynamics of the process of the underlying asset and its relation to the price of derivative with no assumptions on the underlying process. When properly trained, the neural networks then become the derivative pricing formula (Hutchinson et al., 1994).

In the application, we use the set of European style S&P 500 index call and put options covering the period from June 2004 to June 2007. We follow broad range of authors who use S&P 500 data, i.e. Bakshi et al. (1997); Dumas et al. (1996); Garcia and Gencay (2000); Heston and Nandi (2000). S&P 500 index is a broad index of 500 stocks and it serves as a good approximation of the U.S. stock market. Moreover, S&P 500 index options belong to most liquid options traded in the U.S. and world markets.

We follow an unique approach when testing neural networks performance as we divide the data into several detailed categories according to moneyness and time to maturity. Then, we test the ability of neural networks to find the option pricing formula for both calls and puts within all categories without any assumptions on volatility. As a benchmark, we use ad hoc Black-Scholes model, where volatility is not identical across the moneyness and maturities. We model the volatility of the underlying asset as the anualized standard deviation for the period corresponding exactly to the days to expiration. With historical volatility that matches the true days to expiration and is daily updated, Black-Scholes competes with neural networks much better than the original version of pricing formula.

In contrast to Hutchinson et al. (1994); Anders et al. (1998); Bennell and Sutcliffe (2003); Amilon (2003), we particularly train and test the neural networks performance within the narrowly defined moneyness and time to maturity categories in order to show that neural networks can efficiently compete even to well performing ad hoc Black-Scholes model in all categories. The networks correct for the Black-Scholes maturity and moneyness bias. We further use only strike price, close price and time to maturity as inputs for the neural networks to show that there is no need to use a problematic volatility component in option pricing.

The organization of the paper is as follows. After introduction to option pricing and theoretical framework of Black-Scholes, we provide a brief introduction to neural networks. After the methodology is presented, we apply it to S&P 500 index options pricing.

2. Option pricing and theoretical framework of Black Scholes model

In this section, we briefly introduce the methodology of Black-Scholes model to help readers understand the concept of option pricing. Option is a contingent claim when the option holder (writer) has the right (but no obligation) to buy or sell the underlying instrument (that can be asset, equity, index, swap, etc.) at or before a specified date at a specified price. Thus options allow traders either to speculate on future events and/or to reduce the exposure to the financial risk. Basically, two kinds of options are traded on the option exchanges: American type and European type. The former may be exercised any time before its expiration date while the latter can be exercised only on its expiration date. In this paper, we will restrict ourselves to the European-style options.

The true option value, or the option fair price, is the puzzle under consideration. It should reflect the intrinsic value as the potential profit that would arise from the instantaneous exercise of the option, and the time value of the option, which is the price of the possibility that the price of the underlying asset would change to the investors benefit.

The formula derived by Black and Scholes (1973) and Merton (1973) in early 1970's is the most important formula for pricing options even after the years of successive research, as it helps to understand the option pricing. Black and Scholes transformed the option pricing problem into the task of solving a partial differential equation (PDE) with a boundary condition. The price of the underlying asset is assumed to follow the Geometric Browian motion with constant drift and volatility. Using Ito's lemma, the assumption of no arbitrage, and continuous trading, authors showed that the price of any contingent claim written on the underlying solve the parabolic partial differential equation/footnote For further detail see the original paper Black and Scholes (1973). Authors proved that PDE together with the pay-off of the option as a boundary condition has an analytical solution. The solution is well known as Black-Scholes formula:

$$C = S\Phi(d_1) - Xe^{-r(T-t)}\Phi(d_2), \tag{1}$$

$$d_1 = \frac{\ln(S/X) + (r+\sigma^2)(T-t)}{(\sigma\sqrt{T-t})}, \tag{2}$$

$$d_2 = \frac{\ln(S/X) + (r-\sigma^2)(T-t)}{(\sigma\sqrt{T-t})}, \tag{3}$$

where $\Phi(.)$ represents cumulative normal distribution function, S is a price of an underlying asset, X is a strike price or exercise price, r is a risk-free interest rate, σ is a volatility and $(T-t)$ time to expiration.

Using put-call parity a formula for put options is derived as follows:

$$P(X) = Xe^{-r(T-t)}\Phi(-d_2) - S\Phi(-d_1). \tag{4}$$

The Black-Scholes approach to option pricing led to great boom of derivatives trading in 1970s and 80s respectively. Even thought the formula is still very popular, its original version leads to an errors in pricing of the derivatives. Bates (2000) showed that the distribution implicit in the option prices is negatively skewed in contrast to the lognormal distribution assumed by the Black-Scholes model. Moreover, the instantaneous volatility is not identical across the moneyness and maturities (Macbeth and Merville, 1979; Rubinstein, 1985; Corrado and Su, 1997). Misspecification of the process driving the stock price S is one of the main drawbacks of the framework. The key parameter of the model σ is assumed to be constant, but research in past decades show that we need to allow σ to vary in time.

The choice of Black-Scholes model as a benchmark model has, indeed, its justification. Although the model has its drawbacks, there is a growing body of evidence, that if an assumption of constant volatility is relaxed, the model performs very well (as first shown as Chesney and Scott (1989)).Consequently a term ad hoc Black-Scholes model has estabilished in the literature for a modification of the original version using the daily updating of volatility input. Various authors showed that ad hoc Black-Scholes outperforms the deterministric volatility function models (e.g. Dumas et al. (1996) amongst others). Heston and Nandi (2000) show that it competes well with their closed-form GARCH (1,1) option pricing model. More recently, Christoffersen and Jacobs (2004) find that the ad hoc Black-Scholes model beats Heston (1993) theoretical model if parameters are updated daily. Berkowitz (2010) provides further argumentation on justification of ad hoc Black-Scholes option pricing model with frequent parameters updating. In our work, we use a specific form of ad hoc Black-Scholes model with historical volatility computed for the time interval equal to the option expiration, as the proxy for future volatility.

All the variables, but volatility, are easily obtainable from the market. However the forecasting accuracy is based on the volatility estimation. Therefore we use the ad hoc Black-Scholes pricing model with daily updating volatility as an input. Historical volatility is computed for every day, for each option separately. It is defined as an annualized standard deviation of the log-returns of the underlying asset prices over the n days, where n equals to remaining time to maturity of given option. We believe that volatility updated daily improves the Black-Scholes pricing so that it becomes competitive to neural networks.

3. Neural Networks

In this section, we introduce data driven method of derivative pricing where the data will determine the dynamics of the price of the underlying asset and its relation to the derivative security. Assumptions of constant volatility and lognormal distribution of the underlying process are relaxed thanks to this approach. On the basis of the universal approximation theorem, we assume that network is capable to learn the true option pricing formula (Hutchinson et al., 1994). The neural network can also be trained on the real data and optimal model with optimal weights becomes the derivative pricing model. We expect that the neural network can better approximate the price of derivative through learning process than Black-Scholes formula, and can be used to minimize error of hedging or pricing of the derivatives.

Greatest advantage of the neural network approach is that networks do not rely on the restrictive parametric assumptions described above, they are robust to the specification errors that plague parametric models, and more importantly, they are also adaptive and respond to structural changes in the data generating process. Finally, they are flexible enough to encompass a wide range of the price dynamics. On the other hand, the advantages come to cost at large amounts of data needed to best optimalization of weights. Therefore, the approach is not appropriate for newly issued instruments. There is another cost - if the underlying assets prices are well understood and can be analytically expressed, networks will probably not outperform the Black-Scholes. The first drawback turns out to diminish if we consider that there are always amounts of derivatives available to the same asset on the market, thus the newly issued derivative can often be replicated using this data as the underlying process is identical. Another drawback we need to mention is that the computational burden of neural network approach is significantly higher when compared to simple parametric pricing models as Black-Scholes.

3.1. What is a Neural Network?

A neural network relates a set of input variables, say, $\{x_i\}_{i=1}^{k}$ to a set of one or more output variables, say, $\{y_j\}_{j=1}^{k^*}$. The difference between network and other approximation methods is that the approximating function uses one or more so-called hidden layers, in which the input variables are squashed or transformed by a special function. In this paper, we use logsigmoid transformation. While this approach may seem *esoteric* or maybe even *mystical* at first glance, it may be used as a very efficient way to model nonlinear processes. The reason we turn to neural networks is straightforward. It is the goal of the pricing problem to find an approach or method that best prices the options data generated by stochastic underlying processes.

3.2. Feedforward Neural Network

The most widely used neural network in financial applications with one hidden layer (Hornik et al., 1989) is the feedforward neural network and contains two neurons, three input variables, and one output. The general feed-forward or multilayered perception (MLP) network can be described by the following equations:

$$n_{k,t} = \omega_{k,0} + \sum_{i=1}^{i^*} \omega_{k,i} x_{i,t} \tag{5}$$

$$N_{k,t} = \Lambda(n_{k,t}) = \frac{1}{1 + e^{-n_{k,t}}} \tag{6}$$

$$y_t = \gamma_0 + \sum_{k=1}^{k^*} \gamma_k N_{k,t} \tag{7}$$

where $\Lambda(n_{k,t})$ is the logsigmoid activation function. There are i^* input variables $\{x\}$ and k^* neurons. $\omega_{k,i}$ represents a coefficient vector or *input weights* vector. Variable $n_{k,t}$ is squashed by the logsigmoid function and becomes a neuron $N_{k,t}$ at time t. Then the set of k^* neurons are combined linearly with the vector of coefficients $\{\gamma_k\}_{k=1}^{k^*}$ to form the final output, which is the forecast y_t. This model is the workhorse of the neural network modeling approach in finance as almost all researchers start with this network as the first alternative to linear models.

In contrast to classical linear models, there are two additional neurons which process the inputs to improve the predictions in the model. Connections between the input variables and the neurons, also called *input neurons*, and the connections between the neurons and the output, the *output neurons*, are called *synapses*. For the purpose of this study, the hidden layer always uses the logsigmoid transfer function. The reader might note that the simple linear regression model is just a special case of the feedforward neural network. Namely a network with one neuron which contains a linear approximation function.

In order to be able to approximate the target function, the neural network has to be able to "learn". The process of learning is defined as the adjustment of weights using a learning algorithm. The most common way to train a neural network is by learning an algorithm called *backpropagation* or *error-backpropagation*. The main goal of the learning process is to minimize the sum of the prediction errors for all training observables. The training phase is thus an unconstrained nonlinear optimization problem where the goal is to find the optimal set of weights of the parameters by solving the minimization problem:

$$\min\{\Psi(\omega) : \omega \in \mathbb{R}^n\}, \tag{8}$$

where $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable error function. There are several ways of minimizing $\Psi(\omega)$, but basically we are searching for the gradient $G = \nabla\Psi(\omega)$ of function Ψ which is the vector of the first partial derivatives of the error function $\Psi(\omega)$ with respect to the weight vector ω. Furthermore, the gradient specifies direction that produces the steepest increase in Ψ. Negative of this vector thus gives us the direction of steepest decrease. Stochastic gradient descent backpropagation learning algorithm, as well as other methods, will not be discussed in any further detail in order to keep the length of the paper under control.

Besides the popular steepest descent algorithm, the conjugate gradient algorithm is another search method that can be used to minimize the network error function $\Psi(\omega)$ in conjugate directions. This method puts into use the orthogonal and linearly independent non-zero vectors and in some cases brings better convergence results than the previous method.

4. Application to S&P 500 Index Options

One of the usage of neural networks in financial markets modeling is an improvement of forecasts of the stock returns. However, much stronger implications may be made when considering usage of neural networks in another area - pricing and hedging of the derivatives. As discussed above, Black-Scholes model is based on assumptions that are unrealistic. One solution to the problem is to re-estimate the model every day with new updated volatility which will be set to constant. The other approach may be usage of semiparametric neural networks. On the basis of universal approximation theorem, neural networks should be able to price the options with significantly lower error even compared to ad hoc Black-Scholes.

4.1. Experimental Setup

One of the major issues in option pricing is modeling volatility. While neural networks have the ability to learn complex nonlinear patterns from the historical data, we can relax any assumptions on the volatility. Neural networks should be able to model any nonlinear relationship between the price of an option and the variables that affect its price. If all market participants would use unique pricing formula to price the options, neural network would be able to learn it from the data.

We would like to show that neural networks are able to compete ad hoc Black-Scholes model. We use ad hoc Black-Scholes with unique historical volatilities matching the days to expiration, as this model performs well (see the section 2 for argumentation). Another aim is to find if neural networks can be used for pricing of all option categories within the moneyness and time to expiration. As we expect neural network to learn also volatility from the data, our inputs to neural network will only be S/X ratio, $(T-t)$, r and we will model output option price. If ad hoc Black-Sholes is able to describe the data set well, then neural network should be able to recover the Black-Scholes formula but it will not price the options with lower error. In the situation when Black-Scholes price diverge from real option prices, like for example during high volatility periods, neural network should be able to outperform the well established model.

We use the basic generalized feed-forward network architecture, with one hidden layer, logsigmoid transformation functions and conjugate gradient optimization algorithm. The general rule for partitioning the data for training (in-sample), validation and testing (out-of-sample) is 60%, 15% and 25%, respectively.

4.2. Evaluation of Tested Models

For evaluation of individual model, we use various loss functions including mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The loss functions are expressed as follows:

$$MAE = 1/N \sum_{n=1}^{N} |y_n - \hat{y}_n| \tag{9}$$

$$RMSE = \sqrt{1/N \sum_{n=1}^{T} (y_n - \hat{y}_n)^2} \tag{10}$$

$$MAPE = 1/N \sum_{n=1}^{N} |\frac{y_n - \hat{y}_n}{y_n}|, \tag{11}$$

where y_n is real option price, and $\hat{y}_n$ is modeled option price, N is number of observations.

Using these evaluation criteria, we select the model with the lowest error statistic. However, for comparison of the performance of two models, i.e. $\hat{y}_n^i$ and $\hat{y}_n^j$ we use Diebold and Mariano (1995) approach. Diebold-Mariano (DM) statistics tests the null hypothesis of equal predictive accuracy. The statistics is based on the difference of loss functions of two compared models. For further details see Diebold and Mariano (1995).

4.3. Data

In the paper, we use the set of European-style S&P 500 index call and put options as they belong to most liquid options traded in U.S. an in the world markets as well. S&P 500 index is a broad index of 500 stocks. The data consists of the daily close S&P 500 index price adjusted for dividends, strike price, the date, call/put flag, option expiration dates, the daily best bid and the best offer. The sample contains 491819 unique option prices and 761 index prices in the period from June 2004 until June 2007. Figures 1 and 2 show prices and returns of S&P 500 index, respectively.
Following the empirical practice, we use the midpoint of the bid-offer spread as the

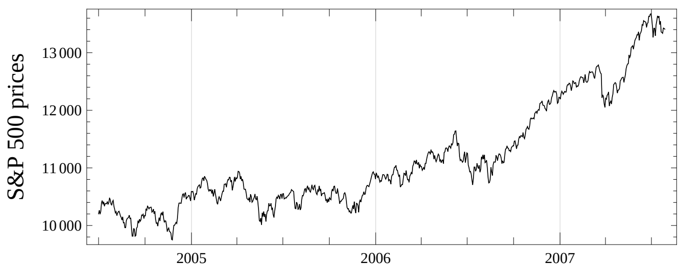

Figure 1: S&P 500 index prices for the period.

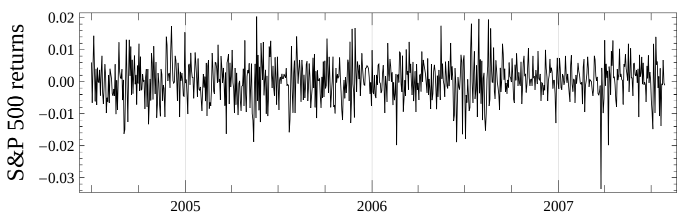

Figure 2: S&P 500 index returns for the period.

option price. Further on, we use the continuously compounded interest rate that is calculated from the continuously compounded zero-coupon[1] interest rates at various maturities.

[1]The zero-coupon curve is derived from BBA LIBOR rates and settlement prices of CME Eurodollar futures

Following Bakshi et al. (1997), we apply the exclusion filters on the data set. First, options with *less than six days to expiration* are excluded in order to prevent the liquidity related bias. Second, price quotes *lower than 0.375*\$ are excluded in order to mitigate the impact of price discreteness on option valuation. Third, the quotes that do not satisfy the no-arbitrage condition $C \geq \max(0, S_t - X_t)$ *for calls and* $P \geq \max(0, X_t - S_t)$ *for puts* are taken away. Almost 30% of the data were excluded.

We further divide the data set into the categories according to the *moneyness and the time to maturity*. We again follow Bakshi et al. (1997) and define the moneyness and time to maturity as follows. *The call (put)* is said to be *in-the-money (out-of-the-money), or ITM(OTM),* if the spot price to strike price ratio $S/X \geq 1.03$. *At-the-money (ATM)* are defined by the $S/X \in \langle 0.97, 1.03)$ and *out-of-the-money (in-the-money), or OTM(ITM)*, for $S/X < 0.97$. A finer partition wit boudnaries of 0.94 and 1.06 respectively includes deep OTM (deep ITM), or DOTM (DITM) categories. The *short-term* maturity option expires in less than 60 days, *long-term* in more than 180 days and *mid-term* has more than or equal 60 and less than 180 days to expiration. Table 1 describes the properties of the prices of proposed 18 categories for which the results will be reported.

The summary statistics is obtained for the daily average bid-offer mid-point option price. Note that the price of the option is increasing with the deepness of the option being in the money, as there is higher chance for the spot to move in desirable direction, and increasing days to the expiration, as its time value increases. The price of call option is in range of \$0.96 for short-term deep OTM to \$ 333.07 for long-term ITM. Put has somewhat narrower range, it's price fall between \$1.86 - \$140. Earlier in the text, we have mentioned that Black-Scholes performs extremely bad when pricing in-the-money or out-of-the-money options. To better understand this moneyness and time to maturity bias, we compute the Black Scholes implied volatility for each category. The theoretical Black-Scholes price is set equal to the averaged best bid-offer mid-point option price and the formula is inverted using the numerical search technique. Table 2 shows the implied volatilities for each category of the S&P 500 index options compared to the historical volatilities computed using different historical window length which corresponds to the time to maturity.

Figures 3 and 4 confirms the well-known Black-Scholes bias. Regardless the time to expiration, the implied volatility exhibits U-shaped pattern across the moneyness as the option goes from deep OTM to deep ITM. Calls exhibit rather "sneer-like" pattern, while puts show traditional "smile". This indicates the most severe mispricing of Black-Scholes for the deep ITM options.

5. Results

The data are randomized for the training and out-of-sample period. Figures 5 and 6 show the RMSE statistic comparison for call options and put options, respectively. For all categories, neural network with strike, close price and time to maturity as inputs outperforms ad hoc Black-Scholes. When we look at the RMSE for both call and put options, the neural network RMSE is flat and low (only very slightly increasing with days to maturity or as the option goes deep in/out-of-the-money).

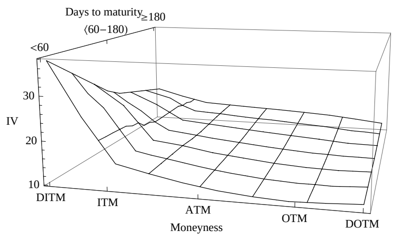

Figure 3: The volatility smirk for call options.

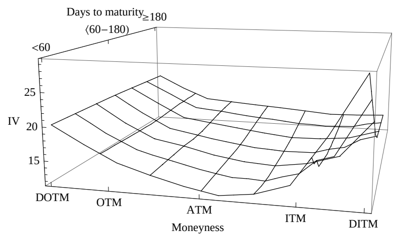

Figure 4: The volatility smile for put options.

On contrary, the ad hoc Black-Scholes RMSE increases rapidly as the days to maturity increase and moneyness deepens. Neural network RMSE is lower than ad hoc Black-Scholes RMSE for all categories. It confirms that neural network price options very well no matter what the moneyness or days to expiration are.

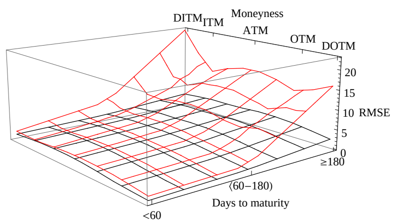

Figure 5: RMSE for call options. Black-Scholes RMSE in red, Neural Network RMSE in black

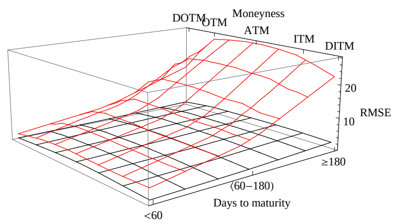

Figure 6: RMSE for put options. Black-Scholes RMSE in red, Neural Network RMSE in black

For call options, Black-Scholes tends to overprice deep in-the-money and deep out-of-the-money options according to the average value of the price, with the highest bias for long-term options. On contrary, Black-Scholes undeprices put options. No such patterns are present for neural networks. Both NMSE and MSE increase heavily as the days to expiration increase for Black Scholes. Again, no such pattern is present for neural network. In absolute values, both measures are much lower for neural network.

Table 3 provides complete results for our out-of-sample performance. We can see that neural network has lower pricing error according to all evaluation criteria. To be rigorous, we also compute the test for the comparison of predictive accuracy of the two tested models. Table 4 summarizes Diebold Mariano (DM) statistic which is approximately normally distributed under the null hypothesis of equal predictive accuracy. For all categories, we strongly reject the null hypothesis of equal predictive accuracy of the neural network and Black-Scholes model. Neural network has significantly lower pricing error than Black-Scholes for all tested call and put options at 1% significance level except for single category of in-the-money call options with expiration less than 60 days. In this single category, neural network produces significantly lower error on 10% level significance. Diebold Mariano test also shows how the difference between Black-Scholes and neural network errors significantly widens with deepness or expiration. The deeper the option in/out-of-the-money, and/or the longer the option has to expiration, the greater the difference between neural network and Black Scholes errors have we found. We thus managed to show that neural network are able to outperform ad hoc Black-Scholes model even without the knowledge of volatility.

6. Conclusion

Since the famous Black-Scholes option pricing formula has been brought into the world of finance, immense volume of option pricing literature has been issued. Soon after the model was proposed, it became heavily criticized for its highly unrealistic assumptions. While Black-Scholes model exhibits strong pricing biases due to these problems, ad hoc Black-Scholes model with frequently updated volatility input performs better, as shown in Section 2 of the paper.

In our paper, we test completely different way of pricing options, which allows us to relax all the restrictive assumptions. Semiparametric neural networks are believed to be able to capture nonlinear dynamic behavior of complex systems, such as stock market. Contributions of the paper are as follow: We show that neural networks learn option pricing formula without the need of volatility as an input. As the benchmark for network, we use generally well-performing ad hoc Black-Scholes model. We train network on fine and wide categorization of moneyness and time to maturity for both call and put options. We show that networks price option within these categories very well.

We evaluate the performance of generalized feed-forward neural network compared to Black-Scholes model on the European style S&P 500 index call and put options. For the Black-Scholes model, we use modified approach. In order to make it more competitive, we use the data that are more likely to be used by practitioners. We use historical volatility which matches exactly the time to maturity day by day, as well as changing interest rates. Generalized feed-forward network with one hidden layer, logsigmoid transformation function and the conjugate gradient learning algorithm is used for comparison. Inputs are the same as to the parametric Black-Scholes model, except for volatility and interest rates inputs, which we relax in the neural network. We do not use volatility at all in order to prove neural network is able to recover it from the real world data.

Explanatory power for both models is sufficiently high, as we compare well performing models. Errors of the ad hoc Black-Scholes model increasessignificantly with increasing moneyness and time to maturity while neural network errors surface stay flat. For both call options and put options, the errors surface of neural network lies below the error surface of Black-Scholes. We use Diebold Mariano statistic which tests the equality of predictive accuracy of the models and we find that neural network produces significantly lower error than Black-Scholes model at 1% significance level except for single category of in-the-money call options with expiration less than 60 days. In this category, neural network produce significantly lower error on 10% level significance.

We managed to show that neural networks are able to compete with an ad hoc Black-Scholes model at wide number of categories even without the knowledge of volatility.

Acknowledgements

The support from the Czech Science Foundation under Grants 402/09/0965, 402/09/H045, GA402/09/0732 and Ministry of Education MSMT 0021620841 is gratefully acknowledged.

References

Amilon, H. A. (2003). Neural network versus black-scholes: A comparison of pricing and hedging performances. *Journal of Forecasting 22*, 317–335.

Anders, U., O. Korn, and C. Schmitt (1998). Improving the pricing of options: A neural network approach. *Journal of Forecasting* (17), 369–388.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of Finance 52*(5), 2003–2048.

Baruník, J. (2008). How does neural networks enhance the predictability of central european stock returns? *Czech Journal of Economics and Finance - Finance a Úvěr 58*(7-8), 359–376.

Bates, D. (2000). Post 87 crash fears in the s&p 500 futures option markets. *Journal of Econometrics 94*(1-2), 181–238.

Bates, D. (2003). Empirical option pricing: a retrospection. *Journal of Econometrics* (116), 337–404.

Bennell, J. and C. Sutcliffe (2003). Black-scholes versus artificial neural networks in pricing ftse 100 option. *Intelligent systems in Accounting 12*(4), 243–260.

Berkowitz, J. (2010). On justifications for the ad hoc black-scholes method of option pricing. *Studies in Nonlinear Dynamics & Econometrics 14*(1).

Bernstein, P. L. (1998). *Against the Gods: The Remarkable Story of Against the Gods: The Remarkable Story of Risk.* New York: John Wiley and Sons.

Black, A. J. and D. G. McMillan (2004). Nonlinear predictability of value and growth stocks and economic activity. *Journal of Business Finance & Accounting 31*, 439–474.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637 – 659.

Chesney, M. and L. Scott (1989). Pricing european currency options: A comparison of the modified black-scholes model and a random variance model. *Journal of Financial and Quantitative Analysis 24*, 267–284.

Christoffersen, P. and K. Jacobs (2004). Which garch model for option valuation? *Management Science 50*(9), 1204–1221.

Corrado, C. and T. Su (1997). Implied volatility skews and stock index skewness and kurtosis implied by s&p 500 index option prices. *Journal of Derivatives 4*, 8–19.

Diebold, F. X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 3*, 253–263.

Dumas, B., J. Fleming, and R. Whaley (1996). Implied volatility functions: Empirical tests. *NBER Working papers series*.

Garcia, R. and R. Gencay (2000). Pricing and hedging derivative securities with neural networks and a homogeneity hint. *Journal of Econometrics 94*(1-2), 93–115.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* (6), 327–343.

Heston, S. and S. Nandi (2000). A closed-for garch option valuation model. *The Review of Financial Studies 13*(3), 585–625.

Hornik, K., M. Stinchcombe, and H. White (1989). Multifactor feedforward networks are universal approximators. *Neural Networks 2*(5), 359–366.

Hutchinson, J., A. Lo, and T. Poggio (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance 3*, 851–889.

Jasic, T. and D. Wood (2004). The profitability of daily stock market indices trades based on neural network predictions: Case study for the s&p 500, the dax, the topix and the ftse in the period 1965 -1999. *Applied Financial Economics 14*, 285–297.

Macbeth, J. and L. Merville (1979). An empirical examination of the black-scholes call option pricing model. *The Journal of Finance 34*(5), 1173–1186.

McNelis, P. (2005). *Neural Networks in Finance: Gaining predictive edge in the market.* advanced finance series. Elsevier Academic Press.

Medeiros, M. C., T. Terasvirta, and G. Rech (2005). Building neural network models for time series: A statistical approach. *Journal of Econometrics 25*, 49–75.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science 4*(1), 141–183.

Rapach, D. E. and M. E. Wohar (2005). Valuation ratios and long-horizon stock price predictability. *Journal of Applied Econometrics 20*, 327–344.

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, through august 31,1978. *The Journal of Finance 40*(2), 455–480.

Tables

		Calls					Puts			
		Days to expiration					Days to expiration			
Moneyness, S/X		< 60	60-180	≥ 180	Subtotal		< 60	60-180	≥ 180	Subtotal
OTM	< 0.94	\$0.96	\$3.69	\$23.57		ITM	\$140.01	\$120.54	\$123.52	
		(753)	(4937)	(21467)	(27157)		(1497)	(1462)	(3887)	(6846)
	0.94-0.97	\$2.52	\$11.30	\$61.21			\$53.72	\$59.97	\$83.35	
		(5750)	(4615)	(4936)	(15301)		(2639)	(2714)	(4751)	(10104)
ATM	0.97-1	\$9.51	\$26.05	\$86.70		ATM	\$25.48	\$36.89	\$67.05	
		(9767)	(5608)	(5504)	(20879)		(8875)	(5583)	(5500)	(19958)
	1-1.03	\$30.70	\$48.93	\$110.56			\$9.78	\$22.49	\$53.92	
		(9330)	(4963)	(5543)	(19836)		(9334)	(4963)	(5536)	(19833)
ITM	1.03-1.06	\$61.71	\$78.33	\$133.20		OTM	\$4.35	\$14.91	\$43.13	
		(7942)	(3980)	(5105)	(17027)		(7711)	(3980)	(5105)	(16796)
	≥ 1.06	\$221.67	\$272.46	\$333.07			\$1.86	\$4.82	\$14.99	
		(31386)	(25272)	(46816)	(103474)		(13560)	(17929)	(45161)	(76650)
	Subtotal	(64928)	(49375)	(89371)	(203674)		(43616)	(36631)	(69940)	(150187)

Table 1: Sample properties of the S&P Index options. The reported values are respectively the dollar value of the average bid-offer midpoint price and the number of observations in parentheses within each category defined according to the moneyness and the days to expiration.

Moneyness, S/X			Calls				Puts		
			Days to expiration				Days to expiration		
			< 60	60-180	≥ 180		< 60	60-180	≥ 180
OTM	< 0.94	BS IV	10,456%	10,235%	10,470%	ITM	31,471%	18,176%	16,983%
		Hist.	10,859%	10,592%	13,208%		10,655%	11,354%	11,482%
	0.94-0.97	BS IV	9,5767%	10,230%	11,091%		15,577%	13,991%	16,001%
		Hist.	10,381%	10,290%	11,472%		10,806%	10,641%	11,055%
ATM	0.97-1	BS IV	9,9019%	10,822%	11,176%	ATM	11,892%	13,273%	16,508%
		Hist.	10,067%	10,352%	10,625%		10,310%	10,366%	12,703%
	1-1.03	BS IV	10,941%	11,491%	10,931%		12,870%	14,182%	16,88%
		Hist.	10,107%	10,276%	10,513%		10,008%	10,301%	10,582%
ITM	1.03-1.06	BS IV	10,501%	11,927%	10,627%	OTM	15,232%	15,601%	17,401%
		Hist.	9,9847%	10,256%	10,616%		10,189%	10,292%	10,524%
	≥ 1.06	BS IV	12,427%	6,5197%	4,9808%		19,873%	20,595%	21,502%
		Hist.	9,8070%	10,353%	11,0218%		10,389%	10,451%	13,897%

Table 2: Black-Scholes implied volatility (BS IV), Historical volatility. The reported values are the averaged volatilities for each of the moneyness and days to maturity categories for both calls and puts.

Calls																			
		OTM						ATM						ITM					
Moneyness, S/X		< 0.94			0.94-0.97			0.97-1			1-1.03			1.03-1.06			≥ 1.06		
Days to expiration		< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180
RMSE	NN	0.305	0.831	2.630	0.796	1.680	3.960	1.720	2.580	4.550	2.110	2.790	4.050	1.920	2.640	4.140	1.310	1.840	3.460
	(BS)	1.260	1.870	16.000	2.330	3.190	12.300	2.800	4.030	15.400	2.690	4.590	15.400	1.990	4.060	12.100	2.100	5.520	22.800
MAPE	NN	0.210	0.194	0.392	0.243	0.129	0.058	0.184	0.081	0.043	0.055	0.044	0.030	0.024	0.027	0.026	0.007	0.007	0.010
	(BS)	0.696	0.372	0.268	0.447	0.254	0.103	0.303	0.149	0.091	0.078	0.086	0.073	0.025	0.044	0.053	0.009	0.018	0.048
MAE	NN	0.211	0.524	1.760	0.519	1.190	2.900	1.200	1.920	3.190	1.580	2.070	3.000	1.440	2.010	3.180	1.000	1.350	2.530
	(BS)	0.820	1.350	8.710	1.480	2.440	6.950	1.990	3.270	9.060	2.110	3.700	9.220	1.480	3.160	7.580	1.670	4.620	17.700

Puts																			
		ITM						ATM						OTM					
Moneyness, S/X		< 0.94			0.94-0.97			0.97-1			1-1.03			1.03-1.06			≥ 1.06		
Days to expiration		< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180
RMSE	NN	0.914	1.470	2.590	1.380	1.970	2.930	1.880	2.200	2.940	1.800	2.270	2.760	1.350	2.130	2.610	0.756	1.010	1.580
	(BS)	4.060	9.340	22.500	3.790	9.500	26.500	4.010	9.270	27.900	4.330	10.000	27.800	3.660	10.200	26.400	2.100	5.530	14.500
MAPE	NN	0.006	0.010	0.017	0.019	0.025	0.028	0.060	0.046	0.035	0.164	0.083	0.043	0.242	0.126	0.052	0.281	0.187	0.152
	(BS)	0.031	0.088	0.230	0.066	0.164	0.438	0.163	0.302	0.700	2.280	0.828	1.080	151	2.750	1.700	894	283	754
MAE	NN	0.751	1.110	2.020	0.964	1.450	2.260	1.370	1.600	2.240	1.260	1.730	2.140	0.898	1.610	2.020	0.475	0.649	1.120
	(BS)	3.660	8.740	18.800	3.210	8.270	23.600	3.260	7.960	25.300	3.560	8.990	25.400	3.000	9.380	24.500	1.710	4.270	11.400

Table 3: The neural network and BS model performance according to the moneyness and maturity on the S&P Index Call (Put) options. RMSE - root mean square error, MAPE - mean absolute percentage error, MAE - mean absolute error.

Calls																		
	OTM						ATM						ITM					
Moneyness, S/X	< 0.94			0.94-0.97			0.97-1			1-1.03			1.03-1.06			≥ 1.06		
Days to expiration	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180
D-M0	8.87**	21.30**	38.50**	19.90**	18.70**	14.10**	18.60**	18.00**	17.40**	12.60**	19.70**	18.60**	1.30*	12.60**	16.10**	114.00**	79.30**	112.00**
D-M1	9.50**	20.60**	37.90**	18.90**	18.70**	13.90**	18.90**	16.70**	17.20**	12.80**	19.50**	18.10**	1.33*	12.10**	15.80**	113.00**	79.20**	112.00**
D-M2	10.50**	20.40**	37.60**	19.20**	18.00**	13.80**	19.80**	16.80**	17.00**	12.60**	19.60**	18.90**	1.34*	12.30**	15.50**	113.00**	78.70**	112.00**
D-M3	11.70**	19.90**	37.50**	19.60**	18.50**	13.80**	19.90**	17.30**	16.40**	12.50**	19.70**	18.70**	1.34*	12.30**	15.90**	117.00**	78.30**	113.00**
D-M4	11.40**	19.30**	37.90**	19.70**	18.90**	13.70**	18.90**	18.60**	16.70**	12.50**	19.80**	18.70**	1.35*	12.60**	16.50**	116.00**	78.90**	114.00**

Puts																		
	ITM						ATM						OTM					
Moneyness, S/X	< 0.94			0.94-0.97			0.97-1			1-1.03			1.03-1.06			≥ 1.06		
Days to expiration	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180	< 60	60-180	≥ 180
D-M0	34.20**	45.40**	42.10**	31.40**	37.60**	60.70**	36.40**	48.70**	73.10**	44.70**	55.80**	75.40**	47.70**	57.60**	78.30**	70.00**	73.80**	127.00**
D-M1	36.00**	39.10**	41.40**	30.40**	36.40**	63.70**	35.60**	50.00**	72.40**	45.50**	57.70**	78.70**	48.20**	53.90**	78.50**	70.10**	73.60**	129.00**
D-M2	35.40**	38.80**	42.00**	31.30**	36.10**	64.20**	34.90**	49.10**	72.50**	43.70**	56.20**	78.10**	47.60**	53.60**	79.80**	69.60**	74.40**	130.00**
D-M3	34.70**	39.00**	41.90**	32.00**	35.70**	62.80**	35.00**	47.30**	71.90**	44.80**	57.10**	77.90**	47.60**	56.90**	79.60**	68.90**	75.70**	130.00**
D-M4	33.90**	38.30**	43.60**	32.60**	34.70**	64.00**	34.20**	47.60**	73.10**	44.10**	56.90**	80.80**	46.50**	53.90**	80.00**	69.00**	74.60**	129.00**

Table 4: Diebold-Mariano statistics of neural network and Black Scholes errors for lags zero through four (D-M0 to D-M4) for all categories of Call and Puts. *,** are 10%, and 1% significance levels.