

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Estimating Short-Term Returns with Volatilities for High Frequency Stock Trades in Emerging Economies Using Gaussian Processes (GPs)

Leonard Mushunje, Maxwell Mashasha and Edina Chandiwana

Abstract

Fundamental theorem behind financial markets is that stock prices are intrinsically complex and stochastic in nature. One of the complexities is the volatilities associated with stock prices. Price volatility is often detrimental to the return economics and thus investors should factor it in when making investment decisions, choices, and temporal or permanent moves. It is therefore crucial to make necessary and regular stock price volatility forecasts for the safety and economics of investors' returns. These forecasts should be accurate and not misleading. Different traditional models and methods such as ARCH, GARCH have been intuitively implemented to make such forecasts, however they fail to effectively capture the short-term volatility forecasts. In this paper we investigate and implement a combination of numeric and probabilistic models towards short-term volatility and return forecasting for high frequency trades. The essence is that: one-day-ahead volatility forecasts were made with Gaussian Processes (GPs) applied to the outputs of a numerical market prediction (NMP) model. Firstly, the stock price data from NMP was corrected by a GP. Since it not easy to set price limits in a market due to its free nature, and randomness of the prices, a censored GP was used to model the relationship between the corrected stock prices and returns. To validate the proposed approach, forecasting errors were evaluated using the implied and estimated data.

Keywords: short-term volatility, stock prices, stock returns, Gaussian process, GARCH, Numerical market prediction

1. Introduction

Stock prices are towards the determination of investors' portfolio status and their consideration is important not only in stock markets. In South Africa, Stocks listed at Johannesburg Stock Exchange (JSE) are actually going through volatilities whose coefficients are high, and this is most common in almost all emerging economies. Volatility is defined as a tendency for prices to change unexpectedly, Harris [1].

The frequency of the short time scaled volatility hits often poses several investment and operational challenges. Thus volatility prediction is an essential process towards securing the economies of the investment portfolios of investors. Short-term volatility forecasts with a prediction horizon from one hour to several days are critical to optimize stock returns and any associated costs. Broadly speaking, there are two approaches to short-term price volatility forecasting: statistical models and physical models. The former uses only historical stock price data to build statistical models, such as autoregressive integrated moving average (ARIMA), autoregressive conditional heteroscedasticity (ARCH), generalized autoregressive conditional heteroscedasticity (GARCH), artificial neural networks, Kalman filters, support vector machines. The cross-field application of these models appears in wind power generation forecast, (see [2–6]). While Statistical models are robust mainly for very short-term forecasts (between 1 and 3 hours ahead), physical models are good in some ways. They can provide better estimates and predictions for longer horizons (days, weeks, months), because they include (3-Dimension) spatial and temporal factors in a full fluid-dynamics model. However, this type of model has limitations, such as the limited observation set for model calibration. To overcome these limitations, some authors have combined statistical and physical models, Salcedo-Sanz et al. [6] and Al-Yahyai et al. [7] where data from a physical model is used as inputs to a statistical model.

This study proposes a forecast model combined with NMP data in which one-day-ahead price volatility forecasting is realized based on historically recorded close prices, volumes, and other market information. We shall combine our NMP data with the Gaussian Processes (GP). Such integrated methods will be used to tailor make corresponding return predictions. Related to our study, are the works of Ladokhin [8], who examines the accuracy of several of the most popular methods used in volatility forecasting. A comparative approach is employed where historical volatility models such as Exponential Weighted Moving Average, ARMA model and GARCH family of models are compared with Artificial Neural Networks based models. Taylor [9] proposed a simple but less accurate method of estimating volatility where daily squared returns are taken. The jumps associated with intra-day prices are not captured yet these jumps significantly affect volatility.

Other related works were done on the prediction of the stock/index returns by: White [10], Sharda [11], Kimoto [12], Brown [13], Gencay [14] where the Artificial neural networks were used. In addition, Sullivan (n.d) [15] employs a variation of a type of Recurrent Neural Network called Long-Short Term Memory (LSTM) in order to predict stock price volatility in the US equity market. Among their results, they found that greater deal of tuning is required on the deeper network, and in particular, the increased use of dropout layers could help reduce the variance problem associated with the employed model so as to accurately estimate the price volatilities. Recent work on stock market prediction is by Sang and Pierro [16] who focuses on the application of LSTM to predict financial time series in the stock market, using both traditional time series analysis and using technical analysis metrics. This is directly related to the successful application of the traditional (LSTM) to address the problem of volatility prediction in the stock market, Xiong et al. [17] and Sardelicha and Manandhar (n.d) [18]. Bhowmik and Wang [19] on the other hand provides a literature review using a systematic database to examine and cross-reference snowballing where previous studies featuring a generalized autoregressive conditional heteroskedastic (GARCH) family-based model stock market return and volatility are reviewed. They also conduct a content analysis of return and volatility literature reviews over a period of 12 years (2008–2019) and in 50 different papers to see the trends and concentration of volatility linked studies. Their results show that significant studies have been done on volatility. However, a focus on short term trades still lacks. With respect to volatility and deviation

modeling, researchers have proposed different distribution models in order to better describe the thick tail of the daily rate of return. For instance, Engle [20] first proposed an autoregressive conditional heteroscedasticity model (ARCH model) to characterize some possible correlations of the conditional variance of the prediction error. In 1986, Bollerslev extended the ARCH model to form a generalized autoregressive conditional heteroskedastic model (GARCH model). Later, the GARCH model rapidly expanded to other forms such as TARCH, EGARCH, ETARCH to form the so-called the GARCH family. As indicated across the literature, researchers proved that GARCH is the most suitable model to use when one has to analyze the volatility of the returns of stocks with big volumes of observations, (for more see [20–25]).

From the reviewed literature, short-term volatility forecasting has been slimly done and little attention has been paid to jumps in association to these short- timed price swerves. For the volatility studies done, statistical models have been employed as stated earlier in this paper. In this paper, stock prices and some related factors such as returns and volumes' datasets including (NMP) results are analyzed and used to develop volatility forecasting models over a horizon of up to one day, with a (GP) method. The main contributions and thrust of this can be summarized into 4 categories as: 1. The predicted price volatility from an NMP model is corrected using a GP. This process helps to improve performance compared with earlier methods for combining statistical and physical models. 2. To build the relationship between corrected stock prices and stock volumes we employed the censored Gaussian Process (CGP). The method accounts for the probabilistic character of the values that are not known precisely because of censoring. 3. High-stock prices data display different features based on the initial values of the models. As such, we shall be treating one of its subset separately. 4. Past stock price data from the JSE databases is used as an additional input to the forecasting model over the time horizon of 1–3 hours-ahead. This time interval is actually the efficient and supportive to our model. The idea paves a great way for high-frequency trades that are proving to dominate the markets and investment world.

2. Methodology

2.1 Data

The datasets used in this study were extracted from Johannesburg stock exchange (JSE) databases with a time scale from June 2017 to May 2018. The idea is that we shall use a whole year dataset as a training set, and the remainder as an independent test set, from where we will make our suitable inferential conclusions. The missing values were less than 30% and to cater for them, we used the K-nearest neighbor (KNN) approach in R environment; otherwise the obtained data was tied.

2.2 Numerical market prediction model and volatility forecasting

Numerical market prediction uses statistical physics and statistical historical models related to financial markets' mechanics. They are used to predict prices based on certain initial-value and boundary conditions. This study uses the stock price data (SPD) (from JSE) and the NMP model. SPD (stock price data) is extracted from the frequently updated JSE electronic stock price databases. The databases are prepared and well-kept for the interests of investors. In general, short-term price volatility forecasting needs predictions from a NMP model with high spatial resolution. The stock price data from JSE is suitable directly for this application, and hence no needs for extra actions like backward and forward

interpolation. The prediction data is produced once each day, and is usually available at 4:00 PM CAT, where closing valuations are done for most investment assets. The data including stock prices, volumes and returns is provided at an interval of 30 minutes for the following 24 hours. It is no secret that investors and market regulators require accurate stock price and return forecasts: In this study we stress that the forecasting error of 1–3 hours ahead should be less than 10% of the actual recorded figures. Therefore, all the forecast errors contained in this study are calculated using hourly data.

3. The Gaussian process (GP)

The method of Gaussian processes is not new, however less seem to be considerably known on its application to financial data. Moreover, it has been successfully applied to many machine learning tasks. Rasmussen [26] duped a well detailed systematic explanation of Gaussian process regression and Automatic Relevance Determination (ARD). Further, the extension of the Gaussian Processes (GP) to censored data is found in Groot [27].

3.1 Gaussian process model

Let us consider a Gaussian process $f(x)$ for a classic regression problem. Now, assuming that we have training set D with n observations such that $D = \{(x_i, y_i) | i = 1, \dots, n\}$, where x denotes an input vector and y denotes a scalar output, the task is to build a function that satisfies the following multiple linear equation.

$$y_i = f(x_i) + \epsilon_i \quad (1)$$

ϵ_i is the additive noise parameter which is non-observable and is assumed to follow a Gaussian distribution such that $\epsilon_i \sim N(0, \sigma_n^2)$. Note that y is a linear combination of Gaussian variables and hence using the invariant transformation property of linear functions, is itself Gaussian. Therefore, we have $p(y|X, k) = N(0, K + \sigma_n^2 I)$, where $K_{ij} = k(x_i, x_j)$, and the joint distribution for a new input x_* can be written in matrix form as:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim \left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & k(X, x_*) \\ k(x_*, X) & k(x_*, x_*) \end{bmatrix} \right), \quad (2)$$

where, $k(X, x_*) = k(x_*, X)^T = [k(x_1, x_*), \dots, k(x_n, x_*)]$, which we will shortly express as k_* . Consequently, following the properties of joint Gaussian distributions, we predict the distribution of our target variable using the following function:

$$\overline{f_*} = k_*^T (K + \sigma_n^2 I)^{-1} y \quad (3)$$

$$V[f_*] = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_* \quad (4)$$

As a result of the stock price control strategies available in the market, there is always a defined upper limit S_{upper} and lower limit of 0 for the stock prices at JSE. Therefore, in statistics, the true values (unrestricted price output) are ‘censored’ in that they are not observed but are replaced by the threshold value. In our modeling, we assume that Gaussian process (GP) has some latent values $y^* = f(x)$. We then

used the censored GP model by Groot [27] to incorporate our modeling assumption. Expectation propagation is used to approximate the censored distribution of the latent variables. This is accompanied by an exploratory analysis whose results suggest that 4.6% percent is within 5% of the upper limit. Thus, it is in the good range where the noise distribution overlaps significantly with the censored range. This normally guarantees the robustness of our results.

4. Modeling process

Our modeling process follows the same approach used in wind power prediction demonstrated in Chen et al. (n.d). Our proposed forecasting framework used in this paper employs GP models through the incorporation of three additional features and these features are fundamental to our modeling process. The three features are: 1. Automatic Relevance Determination (ARD) which is used to select model data points for inputs; 2. Predicted stock prices from the NMP model are corrected before any volatility forecasting and lastly, 3. detailed adjustments were applied to improve our forecasting accuracy using some adjustments in detail, such as using historical data and a separate model building for high stock prices. The NMP data usually includes several market variables such as trading volumes, stock returns, interest rates and inflation. It is clear that stock returns mainly depend on the actual stock prices, however, we do not know for sure if any other market variables play an important role too, and even if we know, we may fail to know the extent to which the variable can affect the returns. To cater for this case, an ARD is used to investigate the selection of input variables. There are two main ways that can be used to obtain stock returns from NMP data: 1. by learning directly the model between NMP data and stock returns data using a censored GP and correcting the error in NMP stock price prediction and then building a second model for the relationship between stock prices and derived stock returns. 2. Can be obtained qualitatively based on the ideas behind a large body of empirical analysis, which states that there are some systematic and stochastic biases present in the original NMP forecasts. For formality sake, we denote the first way of modeling stock returns as GP-direct, and the second as GP-CPrice (meaning based on corrected price):

For illustration, we give a simple schematic diagram of the modeling process as shown in **Figure 1** below.

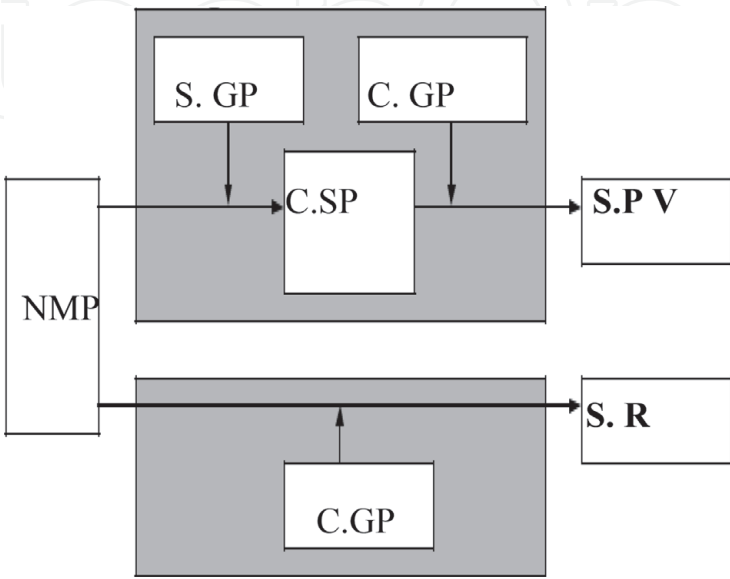


Figure 1.
Model building structure.

The terms in the diagram above are defined as follows:

NMP-Numerical market prediction.

S.GP-Standard Gaussian Process.

C.GP-Censored Gaussian Process.

C.SP-Censored Stock price.

S.PV-Stock price volatility.

S.R-Stock returns.

Further, we then apply our proposed correction model process, taking into account certain constraints to improve the accuracy of modeling. First of all, the process provides forecasts of price volatility and then returns. The main aim is to explore the effect of price volatility on stock returns. As mentioned earlier, the main call of this paper is to develop an efficient price volatility model that can be used to make relevant and frequent volatility estimates. The knowledge of such forecasts and explorations are useful then when modeling stock returns which is the reason behind all investment trades within stock markets.

5. Forecasting accuracy evaluation

Evaluating forecasting accuracy and efficiency can be done using several criteria. This study employed two methods to evaluate our proposed approach and for model evaluation and model comparison: The Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). We defined the error measures as follows:

$$e_t = y_t - \hat{y}_t \quad (5)$$

$$RMSE = \left(\frac{1}{n} \sum_{i=1}^n e_i^2 \right)^{\frac{1}{2}} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (7)$$

Here y_t denotes the actual observation value at time t , \hat{y}_t represents the forecast value for the same period, n is the number of forecasts, and the error is denoted by e_i . The forecasting error threshold for the above specified methods 10%. The accuracy should be less or equal to 10%. Interestingly, our MAE and RMSE are all below 10% as shown in the subsequent tables.

5.1 Experimental validation

Two datasets based on JSE records and estimates are used in this paper to evaluate our approach. We first compared the implied price volatility with the forecasted price volatility. We used the Root Mean Square Error (RMSE) to compute the forecasting error to validate our modeling approach. Secondly, we used the Mean Absolute Error (MAE) method to validate our stock return forecasts where we compared the actual returns and the estimated returns. The two sets of pair wise data are independent from each other as they are extracted differently. Conclusions made are based on that, small values of both MAE and RMSE indicate a high degree of accuracy.

6. Root mean square error (RMSE) results

As shown in **Table 1**, the implied volatility coefficients are not significantly different the estimated coefficients. As implied volatility measures the realized volatility associated with price changes (short and long term), our estimated volatilities proved to be more reliable. This is well indicated by small RMSE values. Another interesting outcome is that, during holidays and weekends, volatility is high due to less market stability and reduced liquidity levels (**Table 2**).

	Implied Volatility (%)	Estimated Volatility (%)	Errors
Time (month)	(1)	(2)	(3)
1	0.280	0.279	0.001
2	0.23	0.223	0.007
3	0.296	0.283	0.013
4	0.178	0.170	0.008
5	0.312	0.291	0.021
6	0.337	0.32	0.017
7	0.117	0.112	0.005
8	0.49	0.48	0.01
9	0.413	0.419	(0.006)
10	0.60	0.60	0.000
11	0.556	0.532	0.024
12	0.80	0.798	0.003

***RMSE = 0.012168.

Table 1.
Implied volatility versus estimated volatility.

	Actual returns (%)	Forecasted returns (%)	Errors
Time (month)	(1)	(2)	(3)
1	0.360	0.379	(0.019)**
2	0.655	0.635	0.02
3	0.698	0.6901	0.0079
4	0.738	0.738	0.000**
5	0.712	0.691	0.021
6	0.831	0.832	(0.001)*
7	0.8273	0.8121	0.0152
8	0.749	0.748	0.001
9	0.713	0.409	0.304**
10	0.635	0.62	0.015
11	0.756	0.732	0.024
12	0.57	0.568	0.002

MAE = 0.03584 = 3.58%. The footnote ** is representing the extreme values obtained from the analysis. This can either be too low or too high and such numbers are of interest in our analysis.

Table 2.
Actual stock returns versus forecasted stock returns.

Mean absolute error (MAE) results: As depicted in the above presented table. Our model proved to be more accurate as indicated by small MAE values. The forecasts errors are not significant, indicating small deviations of our estimated returns from the true (observed returns).

6.1 Model evaluation

As a preliminary step, the ARD was applied to determine which NMP variables should be included as inputs to the correction model. For clarity, we tabulated the measured conditional stock price values as our target variable in the presence of other selected variables-trading volumes, insider news, inflation, exchange rates and stock returns (**Table 3**).

The intensity and effect of the variables on our prediction accuracy for both stock returns and volatility is at all different. However, we noted that stock prices and trading volumes do impact our prediction much than the rest of factors and at the same time, volatility is mostly influenced by trading volumes and insider news in the market. Therefore, we use trading volumes and stock prices (historic) to predict our stock returns and associated short-term volatility as inputs in the GP correction process. This should be kept in mind that in some market environments and set ups, all of the above variables can be used though their intensity factors vary from market to market.

7. Simulation results

This section presents the results of our stock return-prediction framework and price volatility against some benchmarks. We employed the persistence model and a multi-layer perceptron (MLP) neural network model. We used the approach used by Chen et al. (n.d) and Amjady *et al.* [28], where they apply the MLP method to forecast wind power generation. The idea behind the persistence method is that it simply uses the current value as the forecast, which means that at time, t , the prediction, $\hat{y}_{t+1} = \hat{y}_{t+2} = \dots \hat{y}_{t+30} = y_t$.

Since stock data-prices are Markovian, we excluded the historical data in this model at the pre-processing stage), The Markov property states that the future values of, a stock or its price is well explained by the current/present values than its past. As such, we use the current stock price and volume data in this model, and calculate both the stock returns and price volatilities by stock price-yield curve and volatility smile functions respectively, which can be obtained by training historical

	Modeling period	Modeling period
Variable	(1)	(2)
Stock prices	0.380	0.313
Trading volumes	0.33	0.364
Trade Frequency	0.206	0.21
Interest rates	0.10	0.08
Inflation	0.08	0.121
Insider news	0.27	0.29

Variable effect in percentages: Higher percentage higher effect.

Table 3.
ARD results for stock prices at JSE.

Model	RMSE	MAE	NMAPE
	(%)	(%)	(%)
CReturns	14.59	12.79	10.45%
GP-Direct	12.40	12.09	9.53%
GP-CReturns	8.36	7.29	5.73%

Table 4.
Stock returns forecast error.

Model	RMSE	MAE	NMAPE
	(%)	(%)	(%)
MLP-CPrice	15.59	12.79	9.26%
GP-Direct	13.10	11.49	7.53%
GP-CPrice	9.36	8.69	4.73%

Table 5.
Price volatility forecasts error.

dataset. MLP networks are seen in application to short-term wind power forecasting than to stock markets data, for example Amjady *et al.* [28]. This study is making use of the networks, that is, an MLP based model which first corrects stock prices and then predicts stock returns is chosen for comparison. Using the empirical results (model comparison on a validation set), the first MLP model, which corrects stock prices, used NMP stock prices, trading volumes and exchange rates as input variables, and measured returns as the output variable, with a 11-neuron hidden layer. The second part of the MLP-Stock price model used corrected stock prices as input, and has 8-neuron hidden layer, then outputs the final prediction of stock returns. From our empirical results, we can conclude that one well-trained forecast model can be applied to other financial assets data of the same type at JSE. The results of applying the proposed model to the test datasets are shown in **Tables 4** and **5**.

From **Tables 4** and **5**, we note that the proposed GP-Stock Price model has better performance than the other models, and especially presents an outstanding performance in 1–3 hours forecast horizon. In terms of MAE, the improvement of accuracy is 17.98% is required. If comparing to MLP-Stock Price model, the improvement would be 11.61%. The normalized mean absolute percentage error (NMAPE) is the best measure of the forecasting error in our study. This is supported by its ability to provide non-deviating estimates. For easy reference the NMAPE is calculated as:

$$\frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{M} \right| \times 100,$$

where n the number of sample items, M is the market type, in this case we have the stock market (Johannesburg stock exchange).

8. Conclusions

Stock markets are by no means easy to predict. Both stock prices and returns are very stochastic. This poses some difficulties in forecasting their behavior. Volatility

is a key aspect that needs always to be paid attention in the investment world. In this paper, we investigated short-term volatilities associated with stock prices using another approach. We employed the combination of numeric and probabilistic models to forecast one-day interval ahead returns and prices. We used the Gaussian Process (GP) and a Numerical market prediction (NMP) model. To improve the model prediction accuracy, we used a step-wise approach where predicted stock prices are firstly corrected by GP before it is used to forecast stock returns. A censored GP is applied to build the price-return model, mainly to cater for unobserved or missing price records; ARD is used to choose effective NMP variables as inputs to each model; for very short-term forecasts, historical data is added into modeling process; and a high stock prices subset is treated separately by building a single forecast model as we considered it as a special case. The simulation results show that, compared to an MLP-Stock Price model, the proposed model has around 11% improvement of forecasting accuracy, hence the effectiveness and performance of the GP-Stock Price model is proved. Precisely, we proved that the GP performs well in volatility forecasting based on the robustness test done using the forecasting error measures like the RMSE and MAE. Therefore, this paper suggests future works to be carried out on high-frequency trades using the proposed model to make informative forecasts on short-term volatilities.

Author details


Leonard Mushunje^{1,2*}, Maxwell Mashasha¹ and Edina Chandiwana¹

1 Department of Applied Mathematics and Statistics, Midlands State University, Zimbabwe

2 Commercial Bank of Zimbabwe

*Address all correspondence to: leonsmushunje@gmail.com

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Hong, Chang, and Chiu. Hour-ahead wind power and speed forecasting using simultaneous perturbation stochastic approximation (spsa) algorithm and neural network with fuzzy inputs. *Energy*, 35(9):3870–3876, 2010.
- [2] Hussain, Murthy, Singh, (2019) Stock market volatility: A review of the empirical literature. *IUJ J. Manag.*, 7, 96–105.
- [3] Liu, Erdem, and Shi, (2011). Comprehensive evaluation of arma arch (–m) approaches for modeling the mean and volatility of wind speed. *Applied Energy*, 88(3):724–732.
- [4] Li and Shi, (2010). On comparing three artificial neural networks for wind speed forecasting. *Applied Energy*, 87 (7):2313–2320.
- [5] Louka, Galanis, Siebert, Kariniotakis, Katsafados, Pytharoulis, and Kallos. Improvements in wind speed forecasts for wind power prediction purposes using kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(12):2348–2362, 2008.
- [6] Sancho Salcedo-Sanz, ngel, Bellido, Emilio, Garca, Figueras, Prieto, and Paredes, (2009). Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. *Renewable Energy*, 34(6):1451–1457.
- [7] Al-Yahyai, Charabi, and Gastli (2010). Review of the use of numerical weather prediction (nwp) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, 14(9): 3192–3198.
- [8] Leung, M. T., Chen, A., & Daouko, H. (2000). Forecasting exchange rates using general regression neural networks. *Computers and Operations Research*, 27(11–12), 1093–1110.
- [9] Taylor, S. (1986). *Modelling Financial Time Series*. Chichester: John Wiley & Sons Ltd.
- [10] White, H. (1988). Economic Prediction Using Neural Networks: The Case of IBM Daily Stock. *Proceedings of the IEEE International Conference on Neural Networks*, 451–458.
- [11] Sharda, S. A. (1990). Neural Networks as Forecasting Experts: An Empirical test. *Proceedings of the International Joint Conference on Neural Networks*, 491–494.
- [12] Ladokhin, (2009). *Forecasting Volatility in the Stock Market*. BMI Paper
- [13] Brown, S. G. (1998). The Dow theory: William Peter Hamilton's track Record Reconsidered. *Journal of Finance*, 1311–1333.
- [14] Groot, (2012). Gaussian process regression with censored data using expectation propagation. In *Sixth European Workshop on Probabilistic Graphical Models*.
- [15] Sullivan. *Stock Price Volatility Prediction with Long Short-Term Memory Neural Networks*. Department of Computer Science Stanford University Stanford, CA (unpublished paper)
- [16] Sang, C., & Di Pierro, M. (2018). Improving trading technical analysis with tensor flow long short-term memory (lstm) neural network. *The Journal of Finance and Data Science*.
- [17] Xiong Ruoxuan, Eric P. Nichols, Yuan Shen, *Deep Learning Stock Volatility with Google Domestic Trends*, 2015.
- [18] Sardelicha and Manandhar, "Multimodal deep learning for

short-term stock volatility prediction,”
<https://arxiv.org/abs/1812.10479>.

[19] Bhowmik and Wang, (2020). Stock Market Volatility and Return Analysis: A Systematic Literature Review. *Entropy* 2020, 22, 522; doi:10.3390/e22050522, www.mdpi.com/journal/entropy

[20] Gencay, R. (1998). The Predictability of Security Returns with Simple Technical Trading Rules. *Journal of Empirical Finance*, 347–359.

[21] Bollerslev, T (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

[22] Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. Inflation. *Economics* 1982, 50, 987–1008.

[23] Kimoto, T. A. (1990). Stock Market Prediction System with Modular Neural Networks. *Proceedings of the International Joint Conference on*.

[24] Leung, M.T.; Daouk, H.; Chen, A.S. Forecasting stock indices: A comparison of classification and level estimation models. *Int. J. Forecast.* 2000, 16, 173–190

[25] Scott, L.O. Financial market volatility: A survey. *Staff Pap. Int. Monet. Fund* **1991**, 38, 582–625.

[26] Rasmussen, Williams, (2006). *Gaussian processes for machine learning*. Cambridge: MIT Press.

[27] Harris, L. *Trading and Exchanges: Market Microstructure for Practitioners*; Oxford University Press: Hong Kong, China, 2003.

[28] Amjady, Keynia, and Zareipour, (2011). Short-term wind power forecasting using ridge let neural network. *Electric Power Systems Research*, 81(12):2099–2107.