Proceedings of the International Multiconference on
Computer Science and Information Technology pp. 893–898
ISBN 978-83-60810-14-9
ISSN 1896-7094

Estimating time series future optima using a steepest descent methodology as a backtracker

Eleni G. Lisgara
Department of Business Administration
University of Patras
GR-265.00, Rio, Greece
Email: lisgara@upatras.gr

George S. Androulakis
Department of Business Administration
University of Patras
GR-265.00, Rio, Greece
Email: gandroul@upatras.gr

***Abstract*—Recently it was produced a backtrack technique for the efficient approximation of a time series' future optima. Such an estimation is succeeded based on a selection of sequenced points produced from the repetitive process of the continuous optima finding. Additionally, it is shown that if any time series is treated as an objective function subject to the factors affecting its future values, the use of any optimization technique finally points local optimum and therefore enables accurate prediction making. In this paper the backtrack technique is compiled with a steepest descent methodology towards optimization.**

I. Introduction

PREDICTING future was always a rather challenging task with levels of attractiveness in the case of financial issues. Stock prices prediction, regardless its attractiveness is a very difficult task [1]. The introduction and use of time series provided the task of prediction with data history, decreasing the level of arbitrariness when executing.

Many empirical studies throuhgout literature have discussed the predictability of financial time series in respect woth the data history. In finance the daily stock prices comprise a time series, but also in meteorology, daily maximum or minimum temperatures may report one, too. Agriculture, physics, or geology, as most scientific fields interested in reporting data based on time observations, tend to produce time series reports.

Apart from the data history itself, time series has promoted into a major forecasting tool, based on statistical methodologies that use historical data to predict not future points, but the future prices, of any time series regardless their data content.

However, in the case of financial time series it is of great importance to clarify that the average investor would rather be informed about the time that the lowest and highest values will occur than the next day's– -probably ordinary—value. In mathematical means this could be translated as the time series' local minimum and maximum, respectively. The lack of applications concentrating on when the maximum or minimum would appear regardless the next point's value, led us to the obvious: *Since all known forecasting methodologies are price-oriented, it is essential to focus on a point-oriented one in order to forecast not the value of the time series, but the time that its optima will occur*.

Time series is considered as a sequence of data points arranged according to time. Let t be the time intervals of time T; thus, the time series Y is given by

$$Y = Y_t : t \in T. \tag{1}$$

On the other hand, the phenomenon represented by a time series may be also treated as a mathematical function with m variables. Thus, this phenomenon may be described by an unknown but existable function $F : \mathbb{R}^m \rightarrow \mathbb{R}$ given by:

$$Z = F(x_1, x_2, \ldots, x_m), \tag{2}$$

where $x_1, x_2, \ldots, x_m$ are the m variables' values—each depending on time—that affect the phenomenon. Obviously, the values of the time series in equation (1) are equal to those in equation (2)

$$F(x_{1,t}, x_{2,t}, \ldots, x_{m,t}) = Y_t, \qquad \forall t \in T, \tag{3}$$

where $x_{i,t}$ denotes the values of the x_i in time t.

The unknown function F may be calculated for all possible values of any variable x_i. However, while the phenomenon evolutes the m parameters are assigned by specific arithmetic values that are smoothly modified through time. Therefore, the time series Y is also the graphical representation of discrete points that lead to a curve described by a function $g(x_{1,t}, x_{2,t}, \ldots, x_{m,t})$. So, every time series is the trace of the curve g along the function F; on this basis the time series curve could be treated as a generic function of m variables.

For example, assume that the graphical representation of a random function F is as shown on fig. 1, while equation's g trace on it is represented by the white line. Since this curve may be represented as a generic function, its graphical representation is shown in fig. 2.

So, the time series may be represented as the trace of the curve of g along the function F. On this basis, any time series is equivalent to a curve on the m-dimensional space, thus for the optima's prediction any optimization technique may be applied.

In this paper we propose a backtrack technique that allows any optimization algorithm that obtains "memory" being applied in finding future local optima. Section II includes a brief literature review of most methodologies based on which time series prediction is made. The methodology proposed is decomposed on II-C. Finally in V further research interests and applications are proposed.

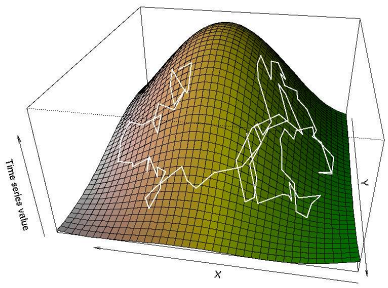

Fig. 1. The time series' general function. The white line traces the time series' real values according to its parameters values.

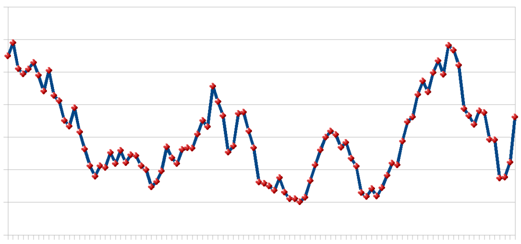

Fig. 2. The function plot of the Y_t using points from the white line of 1.

II. Time Series Forecasting

A. Statistical Based Techniques

The widest used methodologies are based on the knowledge of *some* of the last known prices, i.e.

$$Y_t = \sum_{i=1}^{p} \Phi_i Y_(t-1) + \varepsilon_t \tag{4}$$

where $\Phi_i \in R$, and for the estimation of the new point Y_t are used p known points and a residual ϵ_t that satisfies

$$E[\varepsilon_t | \varepsilon_1, \ldots, \varepsilon_{t-1}] = 0 \tag{5}$$

It was H. Markowitz [2] who applied the mean-variance model on historical data and finally predicted future prices quite accurate in respect to the real ones. Based on his pioneer contribution that future points may be detected through the historical information provided by past data, and statistical assumption including means, variances and covariances, many applications in several subject areas introduced. Several bibliography on time series forecasting for finance incorporated with the probability theory. Sharpre, [3], in his pioneer work in 1970, set basis on the generalized portfolio theory of capital market, introducing the concepts of economics of risk and investment. In 1972, Merton, [4], incorporated with a set of assets by explicitly discussing the characteristics of mean-variance and efficient portfolio frontier. Further extensions added by Pang, [5], Perold ,[6], and extended the level of the parametric methods' usage in large scale selecting problems. In 1990, Best and Grauer, [7], included the general linear programming constraints and in 2000, Best and Hlouskova, [8], also included the concept of the fully and non-risky assets. The use of mean-variance efficient frontiers for the efficient assets' exchange in Jacobs et al, [9], also applied in long positions through the critical line algorithm (CLA); CLA traces out mean-variance efficient sets still including systems of linear equality or inequality constraints.

Furthermore, the distinctive introduction of the exponential smoothing model provided by Brown, [10], and Box and Jenkins, [11], arose new evidence towards predicting time series more efficiently. Such methodologies applied the so called auto-regressive integrated moving average (ARIMA) models to find the best fit of a time series on its own past values; the effectiveness of both methodologies though is a rather controversial issue according to Kuan and Lim, [12].

Most methodologies put aside the issue of the next point on which a research may focus and be interested in, and highlight on forecasting the next value. Thus, Taggart, [13], and Merton, [14], included the concept of the stock market timing in theoretical means, involving financial trends and macroeconomic policies.

B. Artificial Intelligence Techniques

Alternative applications appeared at the late 90's, when, Lee and Jo, [15], and Edwards et al, [16], incorporated with the best time-to-market issue in terms of Artificial Intelligence to predict future stock price movements including weighted factors such as past data and market volatility.

Additionally, Pavlidis et al, [17], [18], incorporated the time series modelling and prediction through the spectrum of feed-forward neural networks as one-step local predictors applied on exchange rates. Grosan and Abraham, [1], and Chen et al, [19], incorporated results from genetic algorithms and neural network applications together, in a single multi-objective algorithm to conclude that the obtained results appear more accurate than the single use of one technique.

The common characteristic amongst these forecasting methodologies is that they neglect the aspect of *optima's* prediction in favor of that of *value*'s prediction.

C. Optimization Techniques

"Optimization is the act of obtaining the best result under given circumstances" [20]. Mathematically this may be

$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & g_j(X) \leq 0, j = 1, 2, \ldots, m \\ & l_j(X) = 0, j = 1, 2, \ldots, n \end{aligned} \tag{6}$$

where $X = \{x_1, x_2, \ldots, x_n\}$ is an n-dimensional vector, $f(X)$ is called the objective function and $g_j(X)$ and $l_j(X)$ are, respectively, the inequality and equality constraints. This is the basic form of the constrained optimization problem. In the case of the lack of constraints, the problem is the unconstrained optimization one. It is common, by applying appropriate modifications, to transform any constrained problem into an unconstrained one.

A well known class of algorithms for unconstrained optimization is the Steepest Descent methods firstly introduced by Cauchy, [22],

$$x^{i+1} = x^i - \lambda_i^* \nabla F_i. \quad (7)$$

where λ_i^* is the optimal step length along the search direction $-\nabla F_i$.

Vrahatis et al, [21], proposed a modification of the steepest descent algorithm model, called Steepest Descent with Adaptive Stepsize (SDAS-2) algorithm, based on the the Armijo's method [23].

The SDAS-2 is decomposed, using the parameters x^0 for the initial point, λ^0 as the randomly large initial stepsize, MIT the maximum number of iterations required and ε as the predetermined desired accuracy.

Algorithm 1 The SDAS-2 algorithm, [21]

Require: $\{F; x^0; \lambda^0; MIT; \varepsilon\}$
  Set $k = -1$
  while $k < MIT$ and $\|\nabla F(x^k)\| > \varepsilon$ **do**
    $k = k + 1$
    if $k \geq 1$ **then**
      Set $\Lambda^k = \frac{\|\nabla f(x^k) - \nabla F(x^{k-1})\|}{\|x^k - x^{k-1}\|}$
      if $\Lambda^k \neq 0$ **then**
        $\lambda = 0.5/\Lambda^k$
      else
        Set $\lambda = \lambda^0$
      end if
    else
      Set $\lambda = \lambda^0$
    end if
    Tune λ by means of a stepsize tuning subprocedure.
    Set $x^{k+1} = x^k - \lambda \nabla F(x^k)$
  end while
  return $\{x^k; F(x^k); \nabla F(x^k)\}$

The following theorem—provided by Vrahatis et al, [21]—concerns the iterative scheme's convergence of the Algorithm 1:

Theorem 1: [21] Suppose that the objective function $F : \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable and bounded below on $\mathbb{R}^m$ and assume that ∇F is Lipschitz continuous on $\mathbb{R}^m$. Then, given any $x^0 \in \mathbb{R}^m$, for any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 1, satisfying the Wolfe's conditions [24], [25], [26] implies that $\lim_{k\to\infty} \nabla F(x^k) = 0$.

III. Decomposition of the Backtrack Procedure

As already mentioned in I the time series Y is also the trace of a curve along function F of m variables. Thus, the problem of time series' optima search is equivalent to the constrained optimization problem (equation 6).

Using the notation of equation (6) lots of the unconstrained minimization methods are iterative in nature and hence they start from an initial random solution, $X^0 \in \mathbb{R}^m$ and proceed towards the minimum point in a sequential manner, $x^1, x^2, \ldots x^{\min}$. According to the time series notation in equation (1), point t_i is denoted by $t_i = (x_1^i, x_2^i, \ldots, x_m^i)$.

Without lack of generality it is supposed that the prediction of future maximum is in question; the procedure is decomposed in two stages:

1) **Backward search for minimum** When applying any optimization technique, it is concluded that most of them in order to generate the new point use prior knowledge collected from the process data including points, function values, gradient values, matrix approximations etc. In our case, let t_n be the last known point of time series, and by using the SDAS-2 algorithm the next estimation of local minimum is calculated when applying Algorithm 1. Obviously, in order to calculate an approximation of the ∇F we have used finite differences. Therefore, by applying the SDAS-2, the sequence of points that leads to a local minimum $[t_n, t_{k_1}, t_{k_2}, \ldots, t_{k_{m-2}}, t_{\min}]$, where $n > k_1 > k_2 > \cdots > k_{m-2} > \min$ is estimated.
2) **Forward search for minimum** When this "sequence" of points is viewed as a forward process, it appears as a sequence that starts from the minimum past point $t_{\min}$, crisscrosses the last known point t_n and probably leads to a maximum future one $t_{\max}$. In this forward process the stepsize Λ_i^* and the sequence of points are known from the previous stage 1. The constructed sequence of points provides us with all the information needed to proceed in estimating a future maximum. So, by applying one step of Algorithm 1 from the initial point t_n an approximation of $t_{\max}$ is obtained.

Thus, the previously described process gives the following backtrack algorithm for the estimation of future local maximum; note that ε is a very small positive number that reassures desired accuracy.

Algorithm 2 The backtrack algorithm

Require: $\{t_n, \varepsilon\}$
  repeat
    Run stage 1 to compute a sequence of points t_n, t_{k_1}, t_{k_2}, $\ldots$, $t_{k_{m-2}}$, $t_{\min}$ that leads to "past" local minimum.
    Calculate "future" point t_{n+k} by applying stage 2 using points $t_{k_{m-2}}$, $\ldots$, t_{k_2}, t_{k_1}, t_n.
    Set $t_n = t_{n+k}$
  until $Y'_{t_n} < \varepsilon$
  Set $t_{\max} = t_{n+k}$.
  return t_{max}

Note that the procedure described in Algorithm (2) is equivalent to the optimization problem given by equation (6). Therefore, since all assumptions of Theorem 1 are satisfied, the Algorithm 2 converges to an optimum future point.

IV. Numerical Applications

To investigate the proposed method's reliability, the method was applied on two different time series samples; one fi-

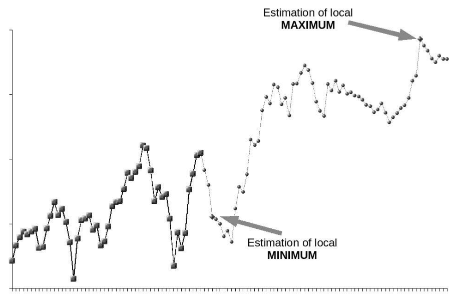

Fig. 3. An application of the backtrack algorithm on Athens Stock Market general index on April 14, 1999

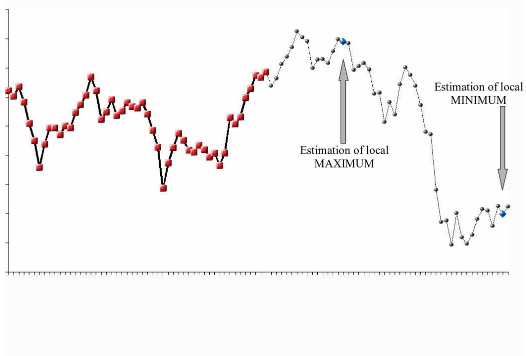

Fig. 4. An application of the SDAS backtrack algorithm on Athens Stock Market general index on July 13, 1998

nancial stock market data and another with meteorological ones.

A. Athens' Stock Market

The proposed application is tested on the daily closing prices of the Athens' Stock Market. The data consists of the daily closing prices of 18 years—from 1985 until 2002. In Figure 3 is presented an application of the backtrack algorithm towards predicting Athens' Stock Market general index for the randomly chosen date of April 14, 1999. The last 50 known values of general index are used; i.e. in the case of April 14, 1999, the 50 last known indexes are from January 29, 1999 until April 14, 1999. These points are represented on Figure 3 with the square symbol. In Figure 3, again, the gray circles stand for the index's actual values index for the exchanging period from April 15, 1999 till July 14, 1999. Future values are connected together with the discontinuous line. The points depicted from the backtrack algorithm are symbolized with the rhomb symbol.

In Figure 4 is presented an application of the SDAS-2 backtrack algorithm towards predicting Athens' Stock Market general index for the randomly chosen date of July 13, 1998; then Algorithm 2 is applied to approximate future local minima and maxima. The last 50 known values of the general index are used; i.e. in the case of July 13, 1998, the 50 last known indexes are from May 4, 1998 until July 13, 1998. These points are represented on Figure 4 with the square symbol. In Figure 4, again, the gray circles stand for the index's actual values for the exchanging period from July 14, 1998 till September 16, 1998. Future values are connected together with the discontinuous line. The points depicted from the SDAS backtrack algorithm are symbolized with the rhomb symbol.

To investigate the SDAS backtrack algorithm's reliability, the algorithm was applied on 1000 randomly chosen prices to approximate the dates t_{min} and t_{max} that produce local minima and local maxima, respectively. Since we assume that the investor seeks for assets that are relatively cheap to buy and their price increase, the table only includes the cases that the local minima emerges before the local maxima.

TABLE I
STATISTICAL INDEXES APPLIED ON RANDOMLY CHOSEN ASSETS

Quantity	**Days**	**Return**
Min	1.00	-0.2320
Q_1	4.00	0.0020
Median	9.00	0.0520
Mean	15.54	0.0662
Q_3	21.25	0.1162
Max	55.00	0.4020
s.d.	15.65	0.1343

The difference between the estimations for local minima and local maxima is extracted in terms of time, that is days of market activity. Furthermore, each asset's return was calculated based on the function

$$R = \frac{Y_{t_{\max}} - Y_{t_{\min}}}{Y_{t_{\min}}} \qquad (8)$$

In Table I are briefly reported the statistical indexes that characterize the variables "days" and "return" obtained from equation 8. Specifically, the raw "days" includes indexes regarding the difference in market days between local minima and local optima, $t_{\max} - t_{\min}$; likewise, raw named "return" illustrates the indexes that result from variable R estimated from equation (8). The abbreviations used as columns are referred to the local minimum, first quarter Q_1, mean value, median, third quarter Q_3 and standard deviation.

It is therefore, observed that the mean between the predicted local minima and local maxima is 15 days approximately. During this period the mean return is about 6.6%.

B. Athens Temperature

As widely known, the time series derived from meteorological data appear to have several local optima due to weather's dependence from sensitive factors. We have constructed a time series based on the data collected from the National Technical University of Athens' meteorological station, that includes information about the temperature for every 10 minutes from February 14, 1999 until today.

If the time series' graph is closely observed, for instance for the Christmas Day of 2006 (Fig. 5), it is obvious that

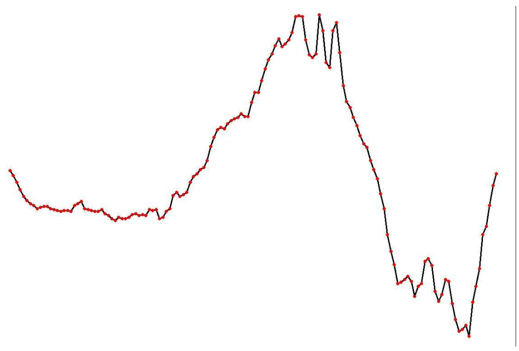

Fig. 5. Athens' temprature during Christmas Day of 2006.

TABLE II
STATISTICAL INDEXES APPLIED ON RANDOMLY CHOSEN ASCENDING TEMPERATURES

Quantity	Forecasted Max	Forecasted Min
Min	-1.16	-8.93
Q_1	-0.20	-0.21
Median	0.03	0.01
Mean	0.20	-0.08
Q_3	0.24	0.28
Max	4.64	3.62
s.d.	0.99	1.39

there is a high number of optima. Furthermore, there are areas of continuous changes from minima or maxima, especially around the global maxima—occurred in 15.00—and the global minima—occurred in 22.30.

The proposed algorithm was tested on the time series data for the year 2006. In order to avoid the effects derived from the continuous changes we have randomly chosen points that appear to follow an ascent or a descent stream. It is supposed that a time series follows an ascent stream when its m last known points are in an ascent order, too. Furthermore, when the time series appeared following an ascent stream, the backtrack algorithm was applied for the detection of the past known minima; the SDAS-2 algorithm was applied only for the prediction of the future local maxima. Samewise, if the time series appeared following a descent stream the backtrack algorithm was applied for the detection of the past maxima, and the SDAS-2 for the prediction of the local minima.

The results obtained from ascending temperature data are shown on Table II while those obtained from descending temperature are shown on Table III. In order to appraise the statement that for the ascending stream case the level of successfulness in predicting local maxima is greater than this for the local minima and vice-verca, table IV was constructed, it contains the successfulness level for each case.

The values shown on both "Forecasted Max" and "Forecasted Min" columns conclude from the difference between the value of the predicted local optimum point and the value of the initial known one.

The column "Direction" illustrates the timeseries' direction, in means of an "ascending" or a "descending" stream. In column "Min" is shown the level of successfulness in predicting local minima; same wise, the column "Max" demonstrated the successfulness level when predicting local maxima. It is observed that the local maxima's prediction while the timeseries follows an ascending stream is quite better than this of the local minima, and vice-versa. However, such behavior is rather expected due to the usage of the SDAS-2 algorithm that belongs to the category of the steepest descent algorithms. According to [27] this category of techniques tends to converge to the closest optima.

TABLE III
STATISTICAL INDEXES APPLIED ON RANDOMLY CHOSEN DESCENDING TEMPERATURES

Quantity	Forecasted Max	Forecasted Min
Min	-2.24	-5.36
Q_1	-0.43	-0.38
Median	-0.11	-0.07
Mean	0.16	-0.37
Q_3	0.34	0.10
Max	4.00	0.56
s.d.	1.24	1.12

TABLE IV
LEVELS OF SUCCESSFULNESS IN PREDICTING MAXIMUM AND MINIMUM IN CO-RELATION WITH THE TIME SERIES' STREAM

Direction	Min	Max
Ascending	47%	52%
Descending	55%	38%

V. Conclusions and Further Research Interests

This paper is structured based on the rationalization that the financial time series may be treated as a function subjected to those that represent all different factors affecting its values during time, thus we incorporated with optimization techniques instead of statistical ones. Obviously, such a rationalization provides strong enough evidence towards applying any optimization technique. Due to this, the proposed backtrack technique was applied using the SDAS-2 algorithm incorporated in [21]. The results obtained provide strong evidence regarding predictions' accuracy. Further, it is observed that the local maxima's prediction—while the timeseries follows an ascending stream—is quite better than this of the local minima, and vice-versa.

The SDAS-2 algorithm was applied since

(a) uses prior knowledge (calculates an approximation of Lipschitz constant using all sequence points), and
(b) as proposed by [27] the steepest descent methods are the most reliable as far as it concerns the convergence to the closest optimum; both characteristics could applied in portfolio optimization problems.

At this point, three directions could be proposed based on the backtrack technique: (a) applications of the backtrack technique in other minimization algorithms that use information collected from sequence points, such as quasi-Newton methods, conjugate gradient, etc, (b) using these techniques towards approximate a future minimum and a future maximum, as well, of an individual asset, so as to formulate an asset management technique for its behavior in an asset portfolio and (c) using backtrack technique with a multi-objective optimization method for managing a given portfolio.

Further research could be focused on these directions, due to the interesting results the technique provides; our research concentrates towards these directions, as well.

REFERENCES

[1] C. Grosan and A. Abraham, *Stock Market Modeling Using Genetic Programming Ensembles*. Berlin/Heidelberg: Springer, 2006.

[2] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, pp. 77–91, 1952.

[3] W. Sharpe, *Portfolio Theory and Capital Markets*. McGraw–Hill, New York, 1970.

[4] R. Merton, "An analytic derivation of the efficient frontier," *Journal of Finance and Quantitative Analysis*, vol. 9, pp. 1851–1872, 1972.

[5] J. Pang, "A new efficient algorithm for a class of portfolio selection problems," *Operational Research*, vol. 28, pp. 754–767, 1980.

[6] A. Perold, "Large-scale portfolio optimization," *Management Science*, vol. 30, pp. 1143–1160, 1984.

[7] M. Best and R. Grauer, "The efficient set mathematics when mean-variance problems are subject to general linear constrains," *Journal of Economics and Business*, vol. 42, pp. 105–120, 1990.

[8] M. Best and J. Hlouskova, "The efficient frontier for bounded assets," *Mathematical Methods of Operations Research*, vol. 52, pp. 195–212, 2000.

[9] B. Jacobs, K. Levy, and H. Markowitz, "Portfolio optimization with factors, scenarios, and realistic short positions," *Operations Research*, vol. 53, pp. 586–599, 2005.

[10] R. G. Brown, *Statistical Forecasting for Inventory Control*. McGraw-Hill, New York, 1959.

[11] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. 2nd Edition, Holden–Day, San Francisco, 1976.

[12] C. M. Kuan and T. Lim, "Forecasting exchange rates using feedforward and recurrent neural networks," *Journal of Applied Econometrics*, vol. 10, pp. 347–364, 1994.

[13] R. A. Jr Taggart, "A model of corporate financing decisions," *The Journal of Finance*, vol. 32, pp. 1467–1484, 1977.

[14] R. C. Merton, "On market timing and investment performance. i. an equilibrium theory of value for market forecasts," *The Journal of Business*, vol. 54, pp. 363–406, 1981.

[15] K. Lee and G. Jo, "Expert system for predicting stock market timing using a candlestick chart," *Expert Systems with Applications*, vol. 16, pp. 357–364, 1999.

[16] R. D. Edwards, J. Magee, and W. H. Bassetti, *Technical Analysis of Stock Trends*. AMACOM, Division of the American Management Associations, 2001.

[17] N. G. Pavlidis, V. P. Plagianakos, D. K. Tasoulis, and M. N. Vrahatis, "Financial forecasting through unsupervised clustering and neural networks," *Operational Research: An International Journal*, vol. 6, no. 2, pp. 103–127, 2006.

[18] N. G. Pavlidis, D. K. Tasoulis, V. P. Plagianakos, and M. N. Vrahatis, "Computational intelligence methods for financial time series modeling," *International Journal of Bifurcation and Chaos*, vol. 16, no. 7, pp. 2053 – 2062, 2006.

[19] Y. Chen, L. Peng, and A. Abraham, "Stock index modeling using hierarchical radial basis function networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4253 LNAI - III, pp. 398–405, 2006.

[20] S. S. Rao, *Optimization, Theory and Applications*. New Delhi: Wiley Eastern Limited, 1984.

[21] M. N. Vrahatis, G. S. Androulakis, J. N. Lambrinos, and G. D. Magoulas, "A class of gradient unconstrained minimization algorithms with adaptive stepsize," *Journal of Computational and Applied Mathematics*, vol. 114, pp. 367 – 386, 2000.

[22] A. Cauchy, "Méthode générale pour la résolution des systémes d'équations simultanées," *Comp. Rend. Acad. Sci. Paris*, vol. 25, pp. 536–538, 1847.

[23] L. Armijo, "Minimization of function having Lipschitz continous first partial derivatives," *Pacific Journal of Mathematics*, vol. 16, pp. 1–3, 1966.

[24] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, pp. 226–235, 1969.

[25] ——, "Convergence conditions for ascent methods II: some corrections," *SIAM Review*, vol. 13, pp. 185–188, 1971.

[26] J. Nocedal, "Theory of algorithms for unconstrained optimization," in *Acta Numerica 1991*, A. Iserles, Ed. Cambridge University Press, Cambridge, 1991, pp. 199–242.

[27] G. S. Androulakis and M. N. Vrahatis, "OPTAC: A portable software package for analyzing and comparing optimization methods by visualization," *Journal of Computational and Applied Mathematics*, vol. 72, pp. 41–62, 1996.