Genetic Programming Polynomial Models of Financial Data Series

Hitoshi Iba
Dept. of Inf. and Comm. Engineering
School of Engineering
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-8656, Japan
iba@miv.t.u-tokyo.ac.jp

Nikolay Nikolaev
Dept. of Computer Science
American University in Bulgaria
Blagoevgrad 2700
Bulgaria
nikolaev@nws.aubg.bg

Abstract- The problem of identifying the trend in financial data series in order to forecast them for profit increase is addressed in this paper using genetic programming (GP). We enhance a GP system that searches for polynomial models of financial data series and relate it to a traditional GP manipulating functional models. Two of the key issues in the development are: 1) preprocessing of the series which includes data transformations and embedding; and, 2) design of a proper fitness function that navigates the search by favouring parsimonious, and predictive models. The two GP systems are applied for stock market analysis, and examined with real Tokyo Stock Exchange data. Using statistical and economical measures to estimate the results, we show that the GP could evolve profitable polynomials.

1 Introduction

Recently several applications of Genetic Programming (GP) (Koza, 1992) to financial data analysis and prediction have been reported (Chidambaran et *al.*, 1998), (Allen, and Karjalainen, 1999), (Chen and Lu, 1999), (Iba and Sasaki, 1999), (Zhang, 1999), etc.. Such a problem instance is learning of the trend in series of market prices with the goal to forecast them and to pursue profits increase. The solutions are trend models, which describe systematic price changes for predefined periods of time. There are several motives to process financial data by GP: 1) GP conducts efficient stochastic exploration of the search space with a population of models that progressively adapt to the market conditions; 2) GP discovers automatically the dependencies among the variable factors influencing the market, and thus determines the relevant variables to enter the model; and 3) GP allows to employ discontinuous, non-differentiable objective (fitness) functions for search navigation.

Two of the key issues in the development of GP systems for financial data processing are: 1) how to prepare the raw input observables in order to facilitate their further learning; and, 2) how to control the evolutionary GP search so as to attain less complex and well predicting models. These issues are addressed in this paper by enhancing the STROGANOFF (Iba et *al.*, 1994) system that produces polynomial models. STROGANOFF is extended to use a set of transfer polynomials in order to become more flexible when constructing the models. We present our approaches to the key issues, namely: 1) preprocessing of the financial series with specific data transformation techniques for extracting the significant information from the observables, and next determination of delayed variables for the model; 2) design of a fitness function to navigate the search by favoring parsimonious and smoother polynomials. Using the novel fitness function we achieve better results than in previous experiments (Iba and Sasaki, 1999), which suggests that this fitness function helps to prevent convergence to overfitting solutions which poorly generalize.

The performance of two GP systems is examined on a stock market prediction task: the enhanced version of STROGANOFF, and a traditional Koza-style GP (Koza, 1992). The implementation SGPC-1.1 (Tackett and Carmi, 1994) of Koza-style GP was used, which manipulates expressions formed of elementary functions. The same preprocessing techniques and fitness function are used in the two systems to facilitate the comparisons. The systems are trained and tested on real Tokyo Stock Exchange data, also known as *Nikkei225* price average, which comprises 225 stocks (Iba and Sasaki, 1999). The expected prediction risk from eventual use of the discovered models is evaluated with statistical and economical estimates. The experimental results show that the differential and rational techniques enable discovery of better models by STROGANOFF. Since such transformations capture the economically important information in the financial series, they are expected to be profitable when used for the real market.

This paper is organized as follows. Section two describes the data preprocessing techniques used in the studies. Section three introduces the mechanisms of the enhanced GP system. Thereafter, the experimental results are provided. Finally, a brief discussion is presented and some conclusions are derived.

0-7803-6375-2/00/$10.00

2 Preprocessing

An instance of financial data series is a sequence of market price measurements taken at certain time intervals. The trend modeling task is to identify the regularities among discretely sampled values (observables): $..., v_t, v_{t+1}, v_{t+2}, ...$ *Delayed* (lagged) vectors $\mathbf{v}$ are used to reconstruct the dynamic system that has generated the series:

$$v_{t+1} = f(\mathbf{v}) = f(v_{t-(m-1)\tau}, v_{t-(m-2)\tau}, ..., v_t) \quad (1)$$

where m is *embedding dimension*, and τ is *delay time*. The aim is to find models that map f vectors from independent variables $\mathbf{v}$ to dependent variable values v_{t+1}.

The objective of preprocessing is two fold: 1) to isolate the significant information in the observables; and 2) to determine how to express them through independent variables. Isolation of variables that convey essential trend information is crucial for successful learning, since the raw actual data are extremely noisy, which limits their relevance for learning. Learning algorithms, like GP, can find dependencies among the variables, but they can not decide what they have to describe.

Preprocessing of financial series includes two steps: *data transformation* and *embedding*. Normalization should be done in advance to mitigate effects due to the magnitudes of the values. We use linear scaling to convert the input values to the range [0,1]. The lagged vectors $\mathbf{v}$ after normalization become $\mathbf{x}$, i.e. from each v_t derive $x_t = Normalize(v_t)$, which leads to vectors:

$$\mathbf{x} = (x_{t-(m-1)\tau}, x_{t-(m-2)\tau}, ..., x_t) \quad (2)$$

where $x_{t-(m-1)\tau},...,x_t$ are the normalized values.

2.1 Data Transformations

Several techniques for transforming the observables into learnable data are studied to examine how they influence the GP search for good models, and with which GP may yield optimal results.

The *integral techniques* eliminate some noise frequencies from the series viewed as a signal, while preserving the systematic information in it. Such a technique is the moving average filtering implemented as follows:

$$x_a = \frac{1}{l} \sum_{k=t-(l-1)/2}^{t+(l-1)/2} x_k \quad (3)$$

where l is the number of neighbors for averaging, and x_k are the values from the normalized series. The choice of the averaging period l is subtle: a very small l will hardly change the series, while a large l will oversmooth the series. These averaging techniques are widely adopted by specialists in business and economics since by deemphasizing the noise they make the series smoother, and the search for models easier (Brock et *al.*, 1992).

The integral transformations, however, leave too much obscuring information like obvious falling or raising tendencies. This can be alleviated by taking the differences between successive data: $x_t \rightarrow x_t - x_{t-1}$. We consider a *differential transformation* after (Deboeck and Cader, 1994),(Pictet et *al.*, 1995):

$$x_d = x_t - \frac{1}{l} \sum_{k=t-l-1}^{t} x_k \quad (4)$$

where x_d is the difference obtained from its origin x_t and the average from its l neighbors. This formula (4) is a block average which makes the series slightly smoother than the one produced by $x_t \rightarrow x_t - x_{t-1}$.

The influence of the data volume can also be overcome by another technique for smashing large magnitudes. Such an alternative for manipulating raw financial data series is the *rational transformation* (Chen and Lu, 1999):

$$x_r = \ln\left(\frac{x_t}{x_{t-1}}\right) \quad (5)$$

where x_t are x_{t-1} subsequent data from the series. The effect is close to this of differentiating (4) but the rational formula (5) leads to series with no so strong oscillations. The rational manipulations allow to account for the rates of change in the series (Deboeck and Cader, 1994).

2.2 Embedding

The *embedding scheme* prescribes how to pass the series data to the model through variables. In other words, the embedding specifies on which historical data in the series the current time value depends. Embedding by delay vectors with parameters $m = 10$, and $\tau = 1$ is chosen empirically. This scheme is used to make four series after the following transformations:

- *original series*– the given series is processed directly with vectors $\mathbf{x} = (x_{t-m-1}, x_{t-m-2}, ..., x_t)$ in order to realize as to what degree it is difficult to learn;

- *integral series*– each value is replaced by its moving average, computed with a smoothing period $l = 5$ according to (3). Processing is done with delayed vectors $\mathbf{x} = (x_{a-m-1}, x_{a-m-2}, ..., x_a)$. Thus, the averaged values enter the model as variables, not the actual values;

- *differential series*– each value is substituted by its variance from the mean of its neighboring data within a predefined interval $l = 3$ using formula (4), after which delay vectors $\mathbf{x} = (x_{d-m-1}, x_{d-m-2}, ..., x_d)$ are constructed. An advantage of such differencing is that it frees the generate models generated from the requirement to fit very closely the trend;

- *rational series*– applying transformations according to formula (5), a series is obtained which is considered through delay vectors $\mathbf{x} = (x_{r-m-1}, x_{r-m-2}, ..., x_r)$. This allows to incorporate in the model purified information about the data changes.

3 GP Mechanisms

With the employment of GP we attempt to identify the nonlinearities in financial trends. STROGANOFF (Iba et *al.*, 1994) is enhanced to use a set of 10 transfer polynomials, rather than only the complete second-order polynomial, in order to increase the flexibility of GP to fit the data (Nikolaev and Iba, 1999). A fitness function was especially designed with: 1) an accuracy component that reflects the degree of fitting wildly varying series; and 2) a complexity component that favors models with smaller number of coefficients.

We evaluate the level of fitting with the following *rational average error* (*RAT*):

$$RAT = \frac{\sum_{i=1}^{E}(y_i - f(\mathbf{x}_i))^2}{\sum_{i=1}^{E}(y_i - y_{i-1})^2} + k\sum_{j=1}^{A} a_j^2 \qquad (6)$$

where k is a regularization parameter (Myers, 1990), A is the number of all coefficients in the estimated polynomial $f(\mathbf{x})$, a_j are their values, and E is the number of the data. The first term in this formula (6) shows the improvement over a random walk model. The second term in (6): $k\sum_{j=1}^{A} a_j^2$ is a regularizer that favors models with coefficients of small magnitude. The rationale is that small coefficients indicate a smoother polynomial which does not interpolate tightly the data, and, thus, help to achieve better generalization.

The fitness function relies upon the *generalized cross-validation* (*GCV*) (Wahba, 1990) principle to account for the interpolation *RAT* as well as the model complexity:

$$GCV = \frac{RAT}{(1 - A/E)^2} \qquad (7)$$

where *RAT* is the rational error, A is the number of polynomial coefficients, and E are the examples. The *GCV* formula (7) prefers short size polynomials with small coefficients, that do not overfit too much the data.

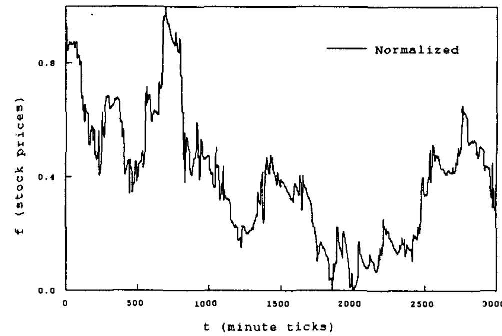

Figure 1. Normalized financial data series.

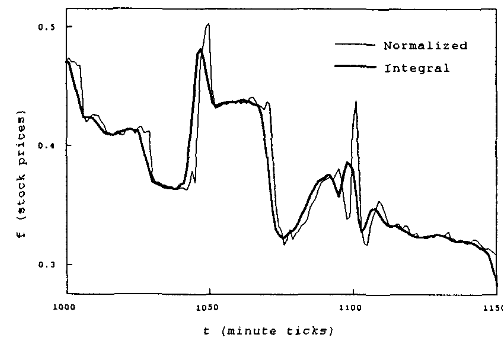

Figure 2. Integral vs. Normalized series.

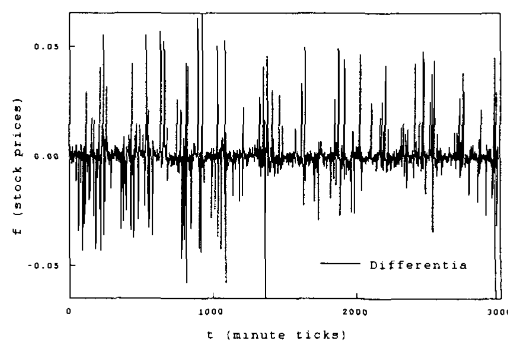

Figure 3. Differential financial data series.

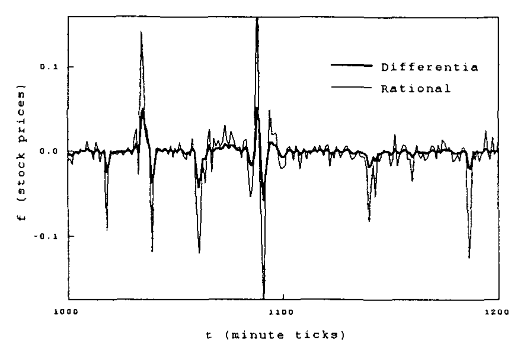

Figure 4. Rational vs. Differential series.

4 Experimental Results

Experiments were conducted with a *Nikkei225* training series of 3,000 points that covers dates during the period April 1st, 1993 through April 17th, 1993. The testing series was of 30,177 data points from the period April 18th, 1993 through September 30th, 1993. These are price averages of 225 representative stocks from the Tokyo Stock Exchange, sampled at every minute from 9:00am to 12:00 pm, and from 1:00pm to 3:00 pm.

Four training series were formed from the original series: 1) *normalized version* (Figure 1); 2) *integral series* with a smoothing period $l = 5$, which overall looks like the normalized one but a detailed view reveals the smoothing (Figure 2); 3) *differential series* with $l = 3$ (Figure 3), and 4) *rational series* that looks globally like the differential one (Figure 4).

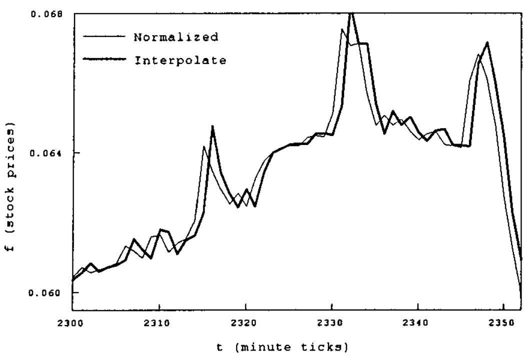

Figure 5. a) Interpolation capacity.

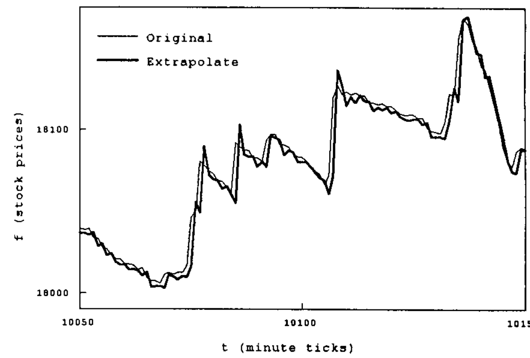

Figure 5. b) Extrapolation capacity of the best (*original*) polynomial found by STROGANOFF.

The measures taken to compare the training sample fit by the derived models are: mean squared error (MSE), ratio of squared errors (RSE), RAE (6), and GCV (7). The *mean squared error* (MSE) is:

$$MSE = \frac{1}{E} \sum_{i=1}^{E} (y_i - f(\mathbf{x}_i))^2 \tag{8}$$

where y_i is the outcome given with the i-th training data, and $f(\mathbf{x}_i)$ is the estimated outcome with this i-th point.

The *ratio of squared errors* (RSE) is the first term from the RAT error (6) without regularization:

$$RSE = \frac{\sum_{i=1}^{E} (y_i - f(\mathbf{x}_i))^2}{\sum_{i=1}^{E} (y_i - y_{i-1})^2} \tag{9}$$

The measures taken to compare the testing performance of the models are: hit percentage, and expected profit gain. The *hit percentage* (*HIT*) shows how accurately the trend directions have been tracked by the model (Iba and Sasaki, 1999):

$$HIT = \frac{N_{up\_up} + N_{down\_down}}{E} \tag{10}$$

where N_{up_up} means number of times when the model outcome and the given outcome exhibit both upward raising tendency, and N_{down_down} means number of times when the model outcome and the given outcome exhibit both falling tendency.

The expected $Profit$ is estimated with a simple algorithm that generates buy/sell trade signals, with which the possible dividends from the model may be calculated. We use a previously defined profit algorithm by Iba and Sasaki (1999), starting from an initial fixed amount of $1,000,000$ Japanese yen committed to each hypothetical trader during a particular test.

Results are reported from 25 runs with each GP, using parameters: $PopulationzeSize = 100$, $Generations = 300$, $MinTreeSize = 5$, and $MaxTreeSize = 35$.

	Accuracy (training)			
	MSE	RSE	RAT	GCV
STROGANOFF	$1.39e-04$	0.8485	0.9153	0.9274
SGPC	$1.83e-06$	0.0107	——	——
	Prediction (testing)			
	HIT		$Profit$	
STROGANOFF	54.02%		$14,714$	
SGPC	62.78%		$31,297$	

Table 1. Estimates of the best polynomials learned by the two GP systems from the *original* financial series.

The regularization parameter k, derived with a statistical technique (Myers, 1990), was $k = 0.005$ in all experiments. The training results with regard to MSE, RSE, RAT, and GCV are estimated using the normalized training series, while all testing results: HITs and $Profits$, are estimated using the original financial series (non-normalized, preprocessed).

Training with the Original Series. The two GP systems showed abilities to evolve well fitting solutions for the original series. The accuracy and prediction of the best results are given in Table 1. The interpolation and extrapolation capacities of the best polynomial discovered by STROGANOFF in arbitrarily selected future intervals are plotted in Figures 5a and 5b.

We are inclined to think, however, that the two GP systems discover overfitting solutions that fit too closely the available original financial data. An indication for this is the extremely low value $RSE = 0.0107$ achieved by the best SGPC solution. Despite this observation, the evolved solutions especially by SGPC on the original series show satisfactory economical characteristics. The estimated HITs= 62.78% of the best expression found by SGPC are higher than those of the best polynomial from STROGANOFF (Table 1). The computed $Profit = 31,297$ was the second result after the best SGPC percentage from all experiments.

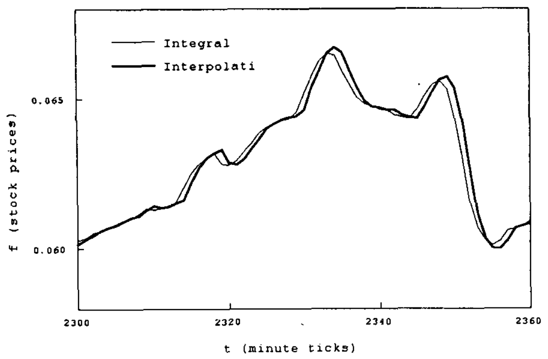

Figure 6. a) Interpolation capacity.

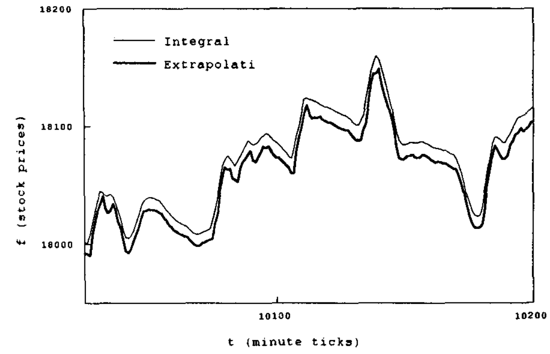

Figure 6. b) Extrapolation capacity of the best (*integral*) polynomial found by STROGANOFF.

	Accuracy (training)			
	MSE	RSE	RAT	GCV
STROGANOFF	$4.09e-05$	0.9328	1.0531	1.0656
SGPC	$1.87e-06$	0.0433	——	——
	Prediction (testing)			
	HIT		$Profit$	
STROGANOFF	53.77%		11,293	
SGPC	62.63%		30,101	

Table 2. Estimates of the best polynomials learned by the two GP systems from the *integral* financial series.

Training with the Integral Series. The results derived with the two GP systems after moving average filtering of the original financial series are given in Table 2. The interpolation and the extrapolation qualities of the best polynomial from STROGANOFF are illustrated in Figures 6.a and 6.b.

The SGPC evolves a slightly worse best solution in mean squared error sense $MSE = 1.87e - 06$ than the one found with the original series with $MSE = 1.83e - 06$, but its generalization is better. We can see in Table 2 that the economical characteristics are also worse: HITs= 62.63% and $Profit = 30,101$. A reason for its worse predictability could be the higher value of the training error $RSE = 0.043$.

The best polynomial from STROGANOFF on the integral series is its worst achievement from economic perspective with lowest HITs= 53.77% and $Profit = 11,293$. Its $RSE = 0.9328$ estimate on the integral series in Table 2 is higher than all other such estimates of its best polynomials from the other series, and this explains why better predictions can not be anticipated. This reasoning is supported also by the higher values of the best STROGANOFF polynomial $RAT = 1.0531$ and $GCV = 1.0656$ in Table 2. The higher RAT and GCV values are due to the fact that the solutions include so many and so accurate coefficients with which a thigh interpolation is attained.

The best SGPC expressions on the original and integral series, which estimates appear in Tables 1 and 2, feature by very high $Profits$, but not compared to the results from the STROGANOFF system (see Tables 3 and 4). Concerning the undetected correlation between the mean squared error MSE and the economic estimates HITs and $Profit$, we think that the moving average filtered data can not protect well the inherent signal dynamics as the theory of signal processing suggests. With respect to the desired economic characteristics, this explains why the integral transformations of the original series did not help the system STROGANOFF to discover polynomial models that could yield optimal hits on future data.

Training with the Differential Series. The highest number of HITs= 74.66% achieved in all runs by the two GP systems exhibits the best polynomial from STROGANOFF using the differential series (Table 3). While the HITs are highest, the $Profit = 54,901$ is only second best result after the one 64,119 achieved with the rational series (Table 4). These estimates are better than the same estimates of the best expression discovered by SGPC (Table 4).

The usefulness of the studied differential transformation formula (4) has been acknowledged by specialists in economics and it is not a surprise that good economic results are achieved with it. The way in which we define it with averaging of the three recent price data is suitable for modeling short and medium trends. If the goal is to perform long term prediction it should be adjusted to use for example 10 or more recent price data, depending on the desired prediction interval (Deboeck and Cader, 1994). What such a differential transformation provides is the possibility to filter out noise and avoid close fitting of the price magnitudes which are not essential to forecasting. The results in Table 3 confirm that the learning abilities of the GP systems are to a great degree sensitive to such series transformations.

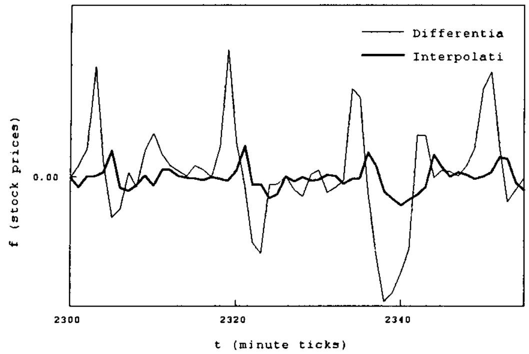

Figure 7. a) Interpolation capacity.

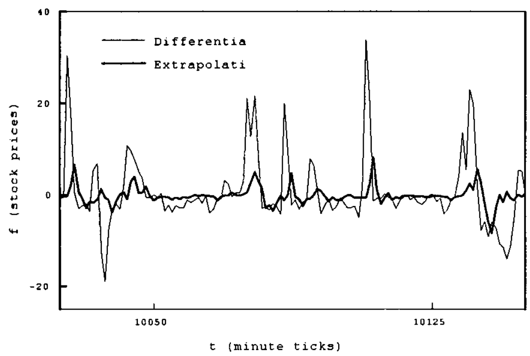

Figure 7. b) Extrapolation capacity of the best (*differential*) polynomial found by STROGANOFF.

	Accuracy (training)			
	MSE	RSE	RAT	GCV
STROGANOFF	$7.86e-05$	0.8066	0.8627	0.8751
SGPC	$1.89e-06$	0.0199	——	——
	Prediction (testing)			
	HIT		$Profit$	
STROGANOFF	74.66%		54,901	
SGPC	56.07%		11,241	

Table 3. Estimates of the best polynomials learned by the two GP systems from the *differential* financial series.

The training estimates: RSE, RAT and GCV in Table 3 of the best STROGANOFF polynomial on the differential series are lower than these on the original and integral series in Tables 1 and 2, which suggests that this polynomial will exhibit higher generalization. The $MSE = 7.86e-05$ is not the best in mean squared error sense (the best $MSE = 4.09e-05$ is on the integral series) among all best polynomials found by STROGANOFF using the other transformations. This could serve as indication that the highest degree of fitting the series is not the only sufficient condition for expecting highest prediction, that is the other criteria: RSE, RAT and GCV, should be also taken into consideration in order to judge which solution is best.

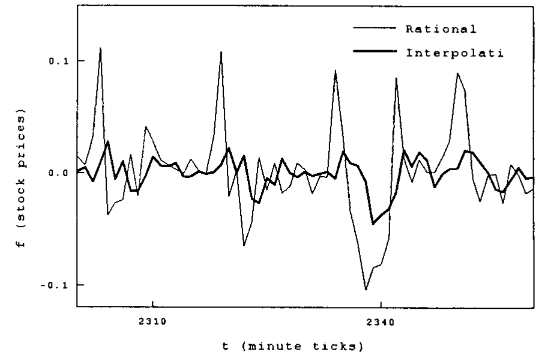

Figure 8. a) Interpolation capacity.

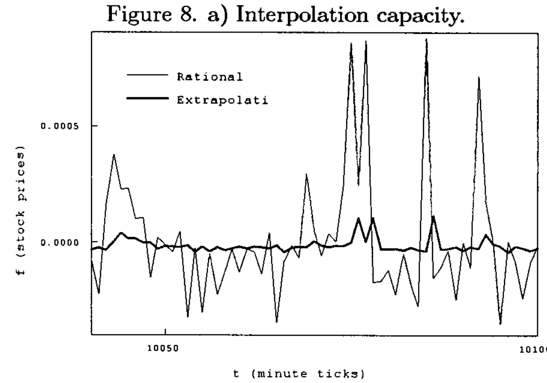

Figure 8. b) Extrapolation capacity of the best (*rational*) polynomial found by STROGANOFF.

	Accuracy (training)			
	MSE	RSE	RAT	GCV
STROGANOFF	$1.62e-03$	0.4676	0.9354	0.9473
SGPC	$2.05e-04$	0.0068	——	——
	Prediction (testing)			
	HIT		$Profit$	
STROGANOFF	68.33%		64,119	
SGPC	62.81%		31,571	

Table 4. Estimates of the best polynomials learned by the two GP systems from the *rational* series.

Training with the Rational Series. The expectation was that training with the rational series has to yield most profitable results (Chen and Lu, 1999). Our experimental work reveals that SGPC is really most more successful than all its runs on the rational series from economical point of view. Support for this claim are its lowest $RSE = 0.0068$ on the training rational series, which has lead to HITs= 62.81% despite the highest mean squared error $MSE = 2.05e-04$.

The rational series, however, is much more difficult for learning financial forecasts than the differential one by the STROGANOFF system. We attribute this to a great degree to the slightly higher frequency oscillations of the rational series compared to the oscillations of the differential series. The rational series curve exhibits more sporadic characteristics in the sense of having more frequent occurrences of sharp spikes with higher amplitudes than the differential one (Figure 4), which do not lend themselves easily to accurate descriptions. The sharper spikes can be represented by models with higher curve amplitudes, but trying to identify such function models usually leads to models which overfit the inherent curvature in the data. That is why, the sharper character of the rational series curve is an argument that hints why a good predictive polynomial is difficult to find by STROGANOFF on such series transformations.

The estimates of the best solutions evolved by the two GP systems are given in Table 4. The series interpolation and approximation by the best STROGANOFF polynomial are illustrated in Figures 8.a and 8.b.

The lower generalization capacity of the best STROGANOFF polynomial on the rational series related to this on the differential series could be explained with the higher magnitude of its coefficients, which can be seen in Table 4 by the higher RAT and GCV estimates than these in Table 3. The best rational polynomial shows a tendency to fit more accurately the spikes in the series for the expense of loosing some extrapolation power. The best STROGANOFF polynomial on the rational series actually shows a very low $RSE = 0.4676$ (Table 4) compared to the best differential polynomial having $RSE = 0.8066$ (Table 3), which suggests that it will generalize well. The problem, however, seems to be that the coefficient values, accumulated in $RAT = 0.9354$ and $GCV = 0.9473$, also impact the prediction quality in the sense that they have an impact on the modeling of the directional changes in the series.

These accuracy results may be considered as arguments for the economical characteristics of the best solutions in Table 4. The best STROGANOFF polynomial exhibits highest $Profit = 64.119$ than all its best polynomials on the other transformations.

We find that the differential and rational transformations enable to achieve most profitable results from economical point of view by STROGANOFF, which is not the case for SGPC. Although the such transformations produce series of sporadic character, that is they lead to series with much more frequent oscillations, they seem to be more amenable to learning from them highly predictive polynomial models by the system STROGANOFF. These two series transformations, however, feature by irregular curve characteristics that are difficult to capture by the expression models of simple functions evolved by SGPC. Such expression models of simple functions can be useful when rational series are considered. Therefore, when the economical criteria are the objective STROGANOFF should be preferred with differential and rational series, while SGPC should be preferred with original and integral series.

Additionally, we would like to point out that our data were not sampled with a constant time window. That is, we were not given data from the periods from 12pm to 1am, and from 3am to 9am. During these periods the world economy is changing, it does not stop, but such effects are simply not taken into consideration in our present investigations since we lack these data. Further research will address these effects also.

5 Discussion

The presented experimental results allow us roughly to summarize that:

1) the simple moving average transformation techniques may not help the GP of polynomials to evolve profitable polynomial trading models from given price movements, while they are helpful for GP systems with expression models from simple functions like SGPC;

2) the MSE estimate is not sufficient to judge which financial series model to select, additional estimates like the RSE, RAT and GCV that account for the relative fitting, for the number and for the magnitudes of the model parameters should be taken into consideration to prefer not simply accurate but also predictive solutions;

3) the differential and rational financial series transformations are approximated well by genetically programmed polynomials. This confirms the advises of other researchers that such detrended transformations allow to capture the economically important information in the training data, and should be considered for learning (Chen and Lu, 1999);

4) we found that the GP of polynomials may be most beneficial for stock market modeling of differential series. It should be noted that other studies also claim that similar differential series should be subject to learning, because they were found to provide the best linear explanatory model among a dozen examined (Dorsey and Sexton, 1998);

5) the defined fitness function enables to identify more predictive solutions than those in the previous studies (Iba and Sasaki, 1999) employing directly the mean squared error as a fitness function.

6 Conclusion

This paper presented empirical evidence, which indicates that GP of polynomials seems to be a promising paradigm to searching for non-linear dependencies in financial data sequences. The relevance of such a GP system STROGANOFF was illustrated on a series of stock market price movements and related to a traditional Koza-style GP. The successfull results from STROGANOFF concern specifically the quality of the evolved polynomial solutions, and should be expected to be valid for high-order multivariate polynomials only. We limited ourselves to a few basic transformations only, but additional research should be performed on combinations of various transformed data through hybrid embeddings.

Another topic for further research is the consideration of STROGANOFF for other financial applications, like: option pricing, risk estimation, foreign currency exchange rates, etc., upon the assumption that the historical trends are correlated with the retrospective indicators from technical analysis point of view.

Acknowledgments
The authors are grateful to the anonymous reviewers for the detailed comments that helped us to improve the quality of the paper.

Bibliography

Allen,F. and Karjalainen,R. (1999). Using Genetic Algorithms to Find Technical Trading Rules. *Journal of Financial Economics*, vol.51, N: 2, pp. 245-271.

Brock,W.A., Lakonishok,J. and LeBaron,B. (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns, *Journal of Finance*, vol. 47, pp. 1731-1764.

Chen,S.-H. and Lu,C.-F. (1999). Would Evolutionary Computation Help in the Design of ANNs in Forecasting Foreign Exchange Rates? In: (Eds.), *Proc. 1999 Congress on Evolutionary Computation, CEC99*, IEEE Press, NJ, 267-273.

Chidambaran,N.K., Jevons Lee,C.J. and Trigueros,J.R. (1998). An Adaptive Evolutionary Approach to Option Pricing with GP. In J.R.Koza, W.Banzhaf, K.Chellapilla, K.Deb, M.Dorigo, D.B.Fogel, M.Garzon, D.Goldberg, H.Iba and R.L.Riolo (Eds.), *Genetic Programming 1998: Proc. of the Second Annual Conf. (GP'98)*, Morgan Kaufmann Publ., CA, 38-41.

Deboeck,G.J. and Cader,M. (1994). Pre and Postprocessing of Financial Data, In: G.J.Deboeck (Ed.), *Trading on the Edge. Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets.* John Wiley & Sons, New York, 27-44.

Dorsey,R.E. and Sexton,R.S. (1998). The Use of Parsimonious Neural Networks for Forecasting Financial Time Series. *Journal of Computational Intelligence in Finance*, vol.6, N: 1, pp. 24-31.

Iba,H., Sato,T. and de Garis,H. (1994). System Identification Approach to Genetic Programming. In *Proc. of the First IEEE Conf. on Evolutionary Computation*, vol.1, New York, NY, 401-406.

Iba,H. and Sasaki,T. (1999). Using Genetic Programming to Predict Financial Data. In: (Eds.), In: *Proc. 1999 Congress on Evolutionary Computation, CEC99,* IEEE Press, NJ, 244-251.

Koza,J.R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* The MIT Press, Cambridge, MA.

Myers,R.H. (1990). *Classical and Modern Regression with Applications.* PWS-KENT Publ., Belmont, CA: Duxbury Press.

Nikolaev,N. and Iba,H. (1999). Automated Discovery of Polynomials by Inductive Genetic Programming, In: J.M.Zytkow and J.Rauch (Eds.), *Principles of Data Mining and Knowledge Discovery. Third European Conf. PKDD '99*, LNAI-1704, Springer, Berlin, 451-456.

Pictet,O.V., Dacorogna,M.M., Chopard,B., Oudsaidene,M., Schirru,R. and Tomassini,M. (1995) Using Genetic Algorithm for Robust Optimization in Financial Applications, *Neural Network World*, vol. 5, N: 4, 573-587.

Tackett,W.A. and Carmi,A. (1994). The Donut Problem: Scalability and Generalization in Genetic Programming. In: K.E.Kinnear Jr. (Ed.), *Advances in Genetic Programming* vol.I, The MIT Press, Cambridge, MA, 143-176.

Wahba,G. (1990). *Spline Models for Observational Data.* CBMS-NSF Regional Conf. Series 59, SIAM Press, Philadelphia, Pennsylvania.

Zhang,B.-T. (1999). Forecasting High-frequency Financial Time Series with Evolutionary Neural Trees: The Case of Heng-Sheng Stock Market. In: *Proceedings ICAI'99.*