INFORMATION
SYSTEMS IN
MANAGEMENT

Information Systems in Management (2014) Vol. 3 (4) 221–228

# SIMPLE ADAPTIVE FILTER AS A PART OF INFORMATION SYSTEM FOR MARKET DATA ANALYSIS

MACIEJ JANOWICZ, KRYSTYNA KIETLIŃSKA, ANDRZEJ ZEMBRZUSKI

*Department of Computer Science, Faculty of Applied Informatics and Mathematics, Warsaw Univesity of Life Sciences - SGGW*

Application of the simple least mean squares (LMS) adaptive filter of to the Warsaw Exchange Market (GPW) has been analyzed using stocks belonging to WIG20 group as examples. LMS filter has been used as a binary classifier, that is, to forecast the sign of changes in the (normalized) stock values. Two kinds of data has been used, namely, the differenced and double-differenced normalized close values of stocks. It has been shown that while the predictive power of LMS filter is virtually zero for the differenced series, it rises significantly in the case of double-differenced series for all analyzed stocks. We attribute this to the better stationarity properties of the double-differenced time series.

Keywords: Warsaw Exchange Markets, adaptive filters, stationary time series

## 1. Introduction

The problem of optimization of management of investments in financial markets belongs, needless to say, to those of paramount importance for quantitative financial sciences. One can distinguish three major stages in the optimization process. Firstly, one has to gather and organized the market data. It is difficult to overestimate the role of the information systems in performing those tasks. The second stage consists in some form of forecast of prize movement in the market, and the third stage is the development of an investment strategy. Our work can be thought of as a development belonging to the second stage.

Let us here observe that even the strongest orthodox statements condemning technical analysis like that of Graham [Graham 2003] actually do provide some sort of forecast (e.g., respectable stocks of large companies will remain robust and will continue to grow). Thus, such or that predictive tool is always used.

Let us notice here that the predictions of classical theories of time series (please see, e.g., [Hannan 1970, Anderson 1970]) provider forecasts for market data with unsatisfactorily high errors. Therefore, a natural need for other tools to analyze financial time series.

One of the most popular forecasting tool in the realm of discrete stochastic processes are the filters, especially the Wiener (Wiener-Kolmogorov) [Kolmogorov 1941, Wiener 1942] filter for stationary processes and the Kalman filter [Kalman 1960] for non-stationary ones. As is well-known, the celebrated Wiener filter is used to produce an estimate of a target random process by linear time-invariant filtering of an observed noisy process. The most important assumptions which are to be fulfilled for the Wiener filter to work are the stationarity of the signal and noise spectra, and additivity of the noise. The Wiener filter minimizes the mean square error between the estimated random process and the desired process.

The straightforward application of the Wiener filter to any time series is non-trivial as it requires the statistics of the input and target signals to be known. Therefore, a less demanding device in the form of the least-mean-squares (LMS) filter has been developed [Widrow and Stearns 1985, Haykin 2002]. Its particular merit is that it converges to the Wiener filter provided that the investigated

time series is linear and time invariant with stationary noise. What is more, it is very easy to implement numerically even in Excel (in this work we have used a home-made Python code).

Twenty stocks of the Warsaw Exchange Market has been chosen to illustrate the results of the LMS application. Their advantage is that while not all of them belong to the most popular trading stocks, they all form the group of Polish „blue chips", i.e. the WIG20 group.

The main body of our work is organized as follows. In Section 2 we provide a short description of the LMS filter and our time series. Section 3 consists qualitative results regarding the performance of our filter. Section 4 contains the discussion and some concluding remarks.

## 2. Least Mean Squares adaptive filter

The Least Mean Squares filter in its normalized version is defined by the following algorithm.

Let $p$ be a positive integer and let $m$ be a real number – they are the parameters of the algorithm to be chosen by the user. Let the weights $\boldsymbol{w}$ and a part of the observed signal $\boldsymbol{x}$ form vectors of the length $p$. We initialize the weight vector 2 with zeros. Then, at each step of the discrete time $t$, $t = p, p + 1, p + 2, \ldots$ we form the vector $\boldsymbol{x}$ from the observed signals as

$$\boldsymbol{x} = [x(t-1), x(t-2), \ldots, x(t-p)],$$

compute the "predicted" value $y$ with the help of the scalar product of $\boldsymbol{w}$ and $\boldsymbol{x}$, and the error $e$ as the difference between the "desired signal" $d$ and prediction $y$. Finally, at each step $t$ the weights are adjusted according to: $\boldsymbol{w} \to \boldsymbol{w} + m e \boldsymbol{x} / Z$, where $Z$ is the norm of the vector $\boldsymbol{x}$.

In our case, the vector $\boldsymbol{x}$ has been formed for either once or twice-differenced normalized close values of the stocks. The normalization consist of subtracting the overall mean for a given stock and dividing by the standard deviation. Let $u(t)$ denote the stock value normalized in the above way. Then the once-differenced signal value has been computed as $b(t) = u(t + 1) - u(t)$, and the twice-differenced observed signal as $c(t) = b(t + 1) - b(t)$. The vector $\boldsymbol{x}$ has been formed from either $b(t)$ or $c(t)$, and the so-called "desired" value d has, naturally, been equal to $x(t)$.

Let $Tp$ denote the number of true positive results (i.e. both the desired signal and prediction are non-negative), $Tn$ the number of true negative signals (i.e. both the desired signals and prediction are negative), $Fp$ – the number of false positive results ($y$ – non-negative, $d$ – negative) and $Fn$ – the number of false negative results ($y$ – negative, $d$ – non-negative). Then the standard performance measures: accuracy, precision, negative prediction value, sensitivity and specificity are defined as follows:

Accuracy = $(Tp + Tn) / (Tp + Fp + Fn + Tn)$;
Precision (positive predictive value) = $Tp / (Tp + Fp)$;
Negative predictive value = $Tn / (Tn + Fn)$;
Sensitivity = $Tp / (Tp + Fn)$;
Specificity = $Tn / (Tn + Fp)$.

## 3. Results of calculations

We have performed calculations for the securities traded in the Warsaw Exchange Market (GWP) belonging to the groups which has been used to calculate WIG20 index. These are: ALIOR, ASSECOPOL, BOGDANKA, BZWBK, EUROCASH, JSW, KERNEL, KGHM, LOTOS, LPP, MBANK, ORANGEPL, PEKAO, PGE, PGNIG, PKNORLEN, PKOBP, PZU, SYNTHOS, and TAURONPE. The following LMS parameters have been used: $p = 100$, $m = 0.7$ for once-differenced series, and $p = 25$, $m = 0.2$ for twice-differenced series. No attempt has been made to optimized the above parameters (they have been close to the optimal ones computed for the case of a stock - 06MAGNA - not belonging to WIG20). The market data has been obtained from the portal bossa.pl [Bossa 2014].

The results are contained in the following tables and figures. The three tables show examples for three arbitrarily chosen assets while the figures illustrate more comprehensive results for all WIG20 stocks.

1. ALIOR (336 trading sessions)

**Table 1.** Performance of LMS for the case of ALIOR stocks

| | Differenced series | Twice-differenced series |
|---|---|---|
| Accuracy | 0.480 | 0.660 |
| Precision | 0.509 | 0.654 |
| Negative predictive value | 0.454 | 0.667 |
| Sensitivity | 0.476 | 0.658 |
| Specificity | 0.486 | 0.662 |

2. ASSECOPOL (4116 trading sessions)

**Table 2.** Performance of LMS for the case of ASSECOPOL stocks

| | Differenced series | Twice-differenced series |
|---|---|---|
| Accuracy | 0.505 | 0.671 |
| Precision | 0.532 | 0.681 |
| Negative predictive value | 0.478 | 0.662 |
| Sensitivity | 0.517 | 0.667 |
| Specificity | 0.493 | 0.676 |

3. BOGDANKA (1214 trading sessions)

**Table 3.** The same as in Table 1 but for the BOGDANKA stocks

| | Differenced series | Twice-differenced series |
|---|---|---|
| Accuracy | 0.500 | 0.667 |
| Precision | 0.532 | 0.659 |
| Negative predictive value | 0.469 | 0.676 |
| Sensitivity | 0.498 | 0.671 |
| Specificity | 0.503 | 0.663 |

All simulations leading to the above results has been performed using a home-made program written in Python. As the numerics has been rather simple, no special numerical packages except of the module Numpy have been used.

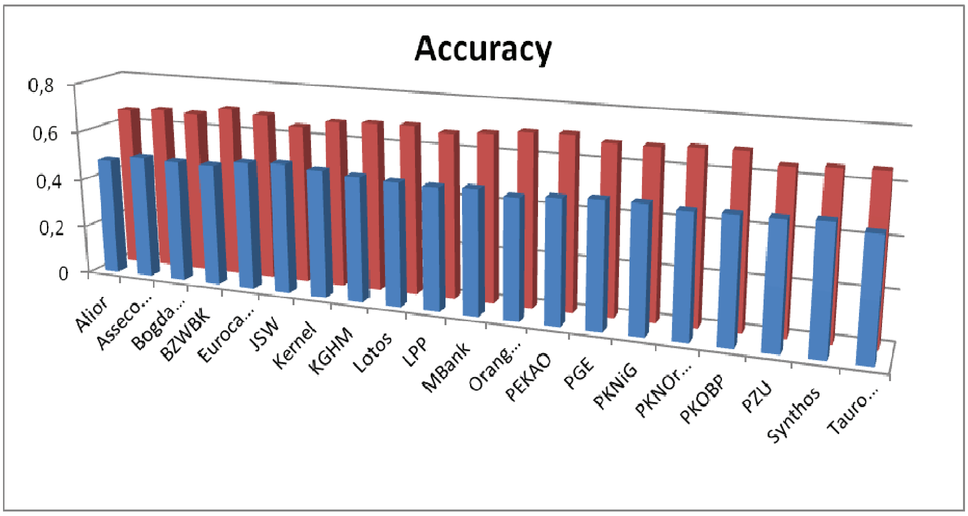


**Figure 1.** Performance of LMS filter for WIG20 stocks: accuracy (please see main text). Shorter column – once-differenced price series; taller column: twice-differenced price series

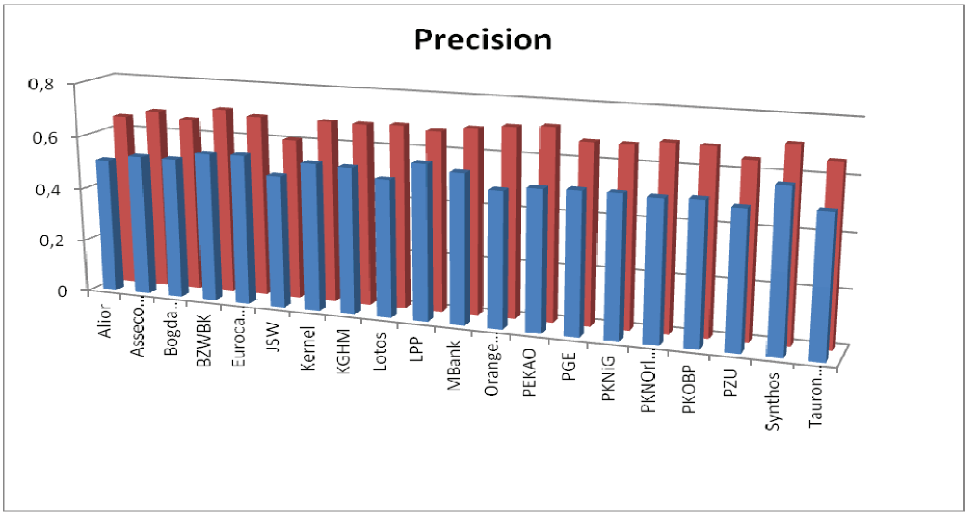


**Figure 2.** Performance of LMS filter for WIG20 stocks: precision (please see main text). Shorter column – once-differenced price series; taller column: twice-differenced price series

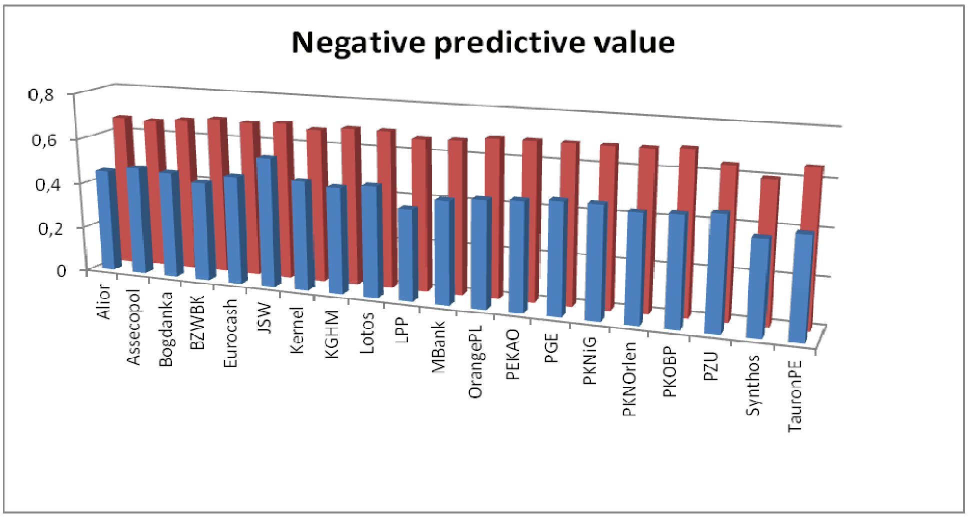


**Figure 3.** Performance of LMS filter for WIG20 stocks: negative predictive value (please see main text). Shorter column – once-differenced price series; taller column: twice-differenced price series

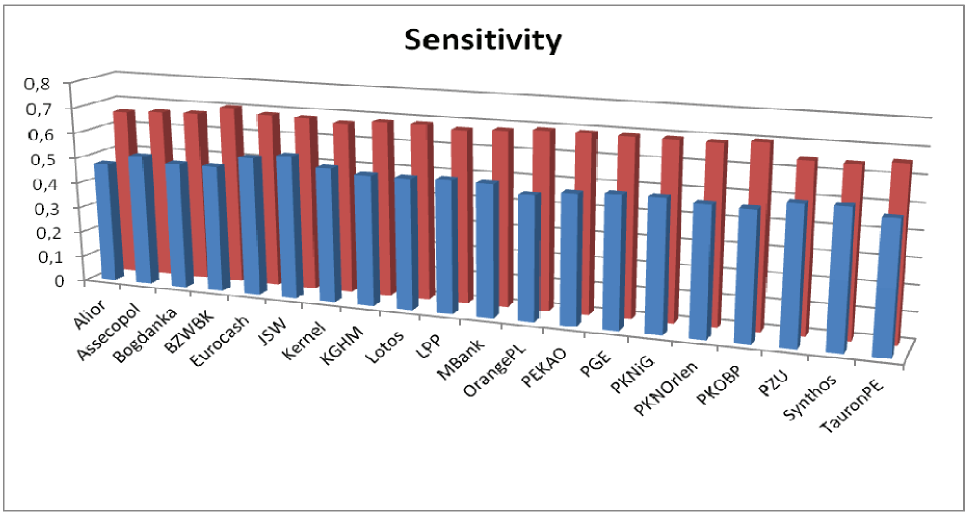


**Figure 4.** Performance of LMS filter for WIG20 stocks: sensitivity (please see main text). Shorter column – once-differenced price series; taller column: twice-differenced price series

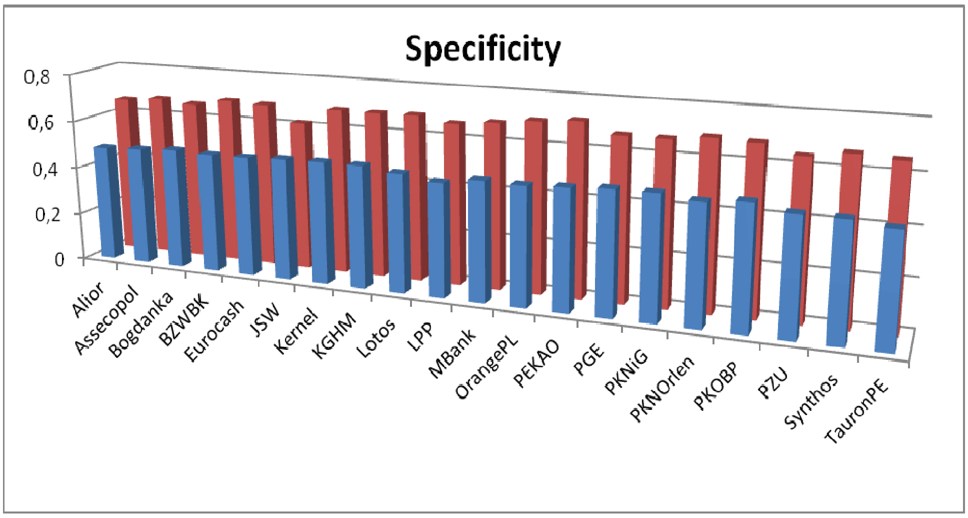


**Figure 5.** Performance of LMS filter for WIG20 stocks: specificity (please see main text). Shorter column – once-differenced price series; taller column: twice-differenced price series

## 4. Discussion

Probably the most remarkable feature of the above results is the striking difference between the performance of LMS for once- and twice-differenced market data. Since all performance characteristics for the once-differenced signals are close to 0.5 (with exception of some values of precision), one can say without any doubts that the most naïve application of the LMS filter to forecast the stock market leads to defeat. On the other hand, the same characteristics for the case of twice-differenced series always exceed 0.6 and, in some cases, even 0.7. This suggests rather strongly that one of the most important conditions of applicability of the Wiener filter and adaptive filters associated with it is better fulfilled. The natural candidate is the time-translation invariance of the series. We believe, however, the quite large values of accuracy of LMS in the case of twice-differenced market time series is still somewhat astonishing (we would not expect it to reach 0.7) and deserves further study.

In the case of once-differenced series we observe that the positive predictive values have been consistently larger than the negative ones while the accuracy being very close to 0.5. This, however, can be ascribed to the natural bias emerging from the fact the market grows and the series is non-stationary.

One can ask a natural question what kind of conclusions follow from the above results concerning the management of portfolio in a stock market. Our answer is this: a good portfolio should be constructed from *predictable* stocks, that is, those for which the performance characteristics of a forecasting method (e.g. LMS filter) are relatively high. This, in particular, eliminates the stocks which are on the market not long enough, for then the accuracy of any type of forecasts is rather low.

In our future work we plan to investigate the performance of other adaptive filters in the stock market and to compare various markets from this point of view. We believe that our work can be further developed as a part of a comprehensive system of gathering and analysis of the market data which would also include extensive databases and intelligent tools creating the investment strategies.

## *REFERENCES*


[1] Anderson, T.W. (1970), *The Statistical Analysis of Time Series*, Wiley, New York.

[2] Bossa (2014): http://bossa.pl/notowania/metastock/

[3] Graham B., Zweig J. (2003), *The Intelligent Investor*, Harper Collins, New York.

[4] Hannan E.J. (1970), *Multiple Time Series*, Wiley, New York.

[5] Haykin S., *Adaptive Filter Theory,* Prentice Hall, 2002.

[6] Kalman, R. E. (1960), *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering 82 (1), 35–45.

[7] Kolmogorov A.N, *Stationary sequences in Hilbert space*, (In Russian) Bull. Moscow Univ. 1941 vol. 2 no. 6, 1–40.

[8] Wiener N., *The interpolation, extrapolation and smoothing of stationary time series*, Report of the Services 19, Research Project DIC-6037 MIT, February 1942.

[9] Widrow B. and Stearns S.D. (1985), *Adaptive Signal Processing,* Prentice Hall, New York.