Department of Mathematics
Faculty of Mechanical Engineering
Slovak University of Technology in Bratislava

3rd International Conference
APLIMAT 2004

MORPHOLOGICAL FILTRATION FOR TIME SERIES

VÁVRA František (CZ), NOVÝ Pavel (CZ), MAŠKOVÁ Hana (CZ), KOTLÍKOVÁ Michala (CZ), NETRVALOVÁ Arnoštka (CZ)

Abstract. In presented article, there will be introduced one possible generalization of using the morphological methods for time series smoothing and filtration. Such methodology is well sophisticated and widely used in the field of digitized image processing. The usage of this method for one dimensional objects is the logical consequence of the usage for the generalization to two dimensional objects. The base of morphological apparatus is the min-max algebra. Presented methodology was developed for the trend analysis in time series of foreign currencies.

Key words. Time series, Morphology, Filtration, Min-Max algebra, Erosion, Dilation

1 Motivation

Smoothing procedure of the two dimensional sets with the help of dilation and erosion operations is described in sources [1,2]. Application of such methods is concentrated mainly into the field of digitized image processing. In sources [4, 5, 6, 7, 8] similar methods are used for dynamic processes (especially discrete events systems). Sources [3, 9, 10, 11] represent the algorithms of joining, forecasting and smoothing. All sources listed above are used as an inspiration for the application of mentioned methods and procedures into the field of the time series.

2 Upper and lower estimation, dilation and erosion, opening and closing of the segment

We will understand the term time series as a general mapping which assign the real value to the time index. The time axis is considered to be the set of integer numbers (without the loss off generality). Therefore time series in our point of view is the following mapping: $x : Z \to R_1;\quad t \mapsto x_t$. The segment of time series $x(t-n, t+n)$ of the length $(2n+1)$, where $(2n+1>1; \mathrm{n} \geq 1)$ with the middle in time point $t \in Z$; will be understand as the sequence $x_{t-n}, ..., x_t, ..., x_{\mathrm{t+n}}$ i.e. the part of time series for discrete time interval $\langle t-n, ..., \mathrm{t+n} \rangle$. Let us adopt the operations of upper and lower segment estimation:

$$U_n(x(t-n,t+n)) \mapsto u \in R_1; \quad U_n : R_{2n+1} \to R_1$$
$$L_n(x(t-n,t+n)) \mapsto l \in R_1; \quad L_n : R_{2n+1} \to R_1 \tag{1}$$
$$L_n(x(t-n,t+n)) \le \min\{x_{t-n},\dots,x_{t+n}\} \le \max\{x_{t-n},\dots,x_{t+n}\} \le U_n(x(t-n,t+n))$$

Upper and lower segment estimation are two numbers representing features of the segment and they satisfy the inequality (1).

Let us define the dilation $D_n(x)$ and erosion $E_n(x)$ of time series x as following mappings:

$$x \to D_n(x) = \ \dots d_{t-1}, d_t, d_{t+1} \ \dots; \quad d_t = U_n(x(t-n,t+n)) \tag{2}$$
$$x \to E_n(x) = \ \dots e_{t-1}, e_t, e_{t+1} \ \dots; \quad e_t = L_n(x(t-n,t+n)) \tag{3}$$

Time series dilation is therefore the series of upper estimations of the original time series. Analogical idea holds for the erosion and lower estimations. Figure 1 demonstrates the dilation and erosion of time series on condition of operations U_1 and L_1 being defined as following:

$d_i = U_1(x(i-1,i+1)) = \max\{x_{i-1}, x_i, x_{i+1}\}$ and $l_i = L_1(x(i-1,i+1)) = \min\{x_{i-1}, x_i, x_{i+1}\}$.

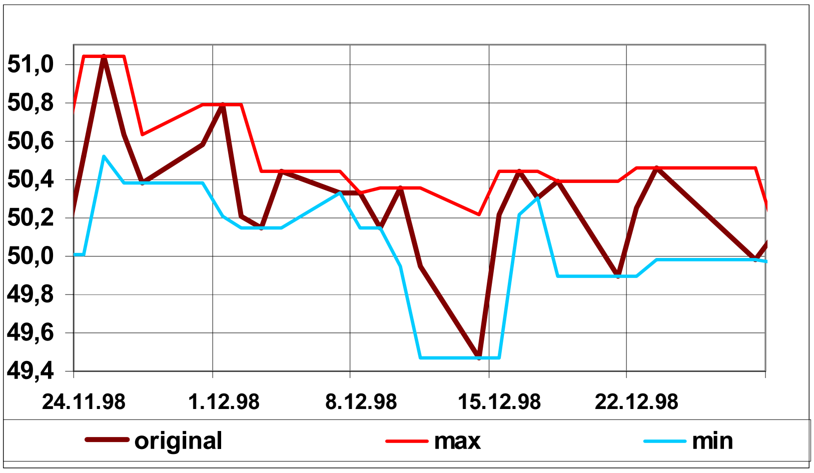

Figure1. Example of dilation end erosion while using the max and min operations

Of course the pair U_1 and L_1 is not the only possible, another pairs could be for example:

$d_i = U_1(x(i-1,i+1)) = \max\{x_{i-1}, x_i, x_{i+1}\} + \alpha(\max\{x_{i-1}, x_i, x_{i+1}\} - \min\{x_{i-1}, x_i, x_{i+1}\}), \textit{where } 0 \le \alpha$

$l_i = L_1(x(i-1,i+1)) \ = \min\{x_{i-1}, x_i, x_{i+1}\} - \alpha(\max\{x_{i-1}, x_i, x_{i+1}\} - \min\{x_{i-1}, x_i, x_{i+1}\}), \textit{where } 0 \le \alpha$

(4)

In compliance with [1, 2] we will adopt the operators of opening $O_n(x)$ and closing $C_n(x)$ as:

$$O_n(x) = E_n(D_n(x)), \ C_n(x) = D_n(E_n(x)) \tag{5}$$

The analogy of the structural element [1, 2] is in several choices of lower and upper estimation operators.

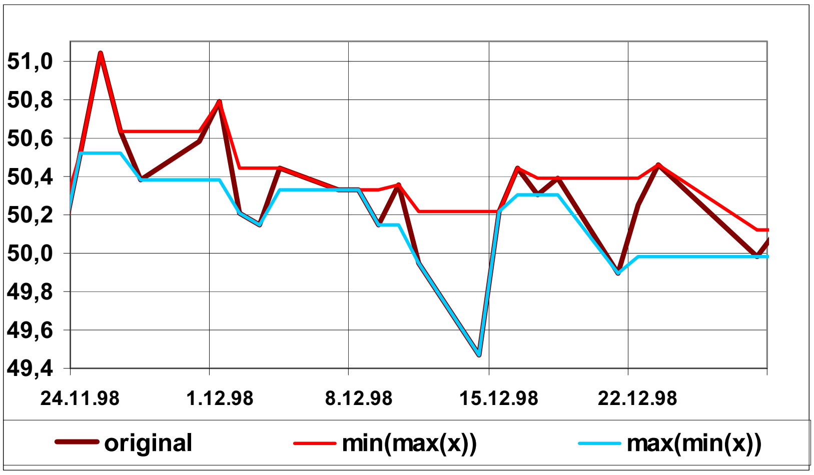

Figure 2.: Example of opening and closing operators (upper hull, lower hull) while using the max, min pair

The "opening" operator can be interpreted as a kind of upper hull of a time series. Analogically the "closing" operator can induce the interpretation of a time series generalized lower hull.

3 Operators properties, idempotency limitation

At first, let us mention some trivial formulas. We will start with the definition of two time series comparison. Let x, y be two time series, then:

$$x \le y \Leftrightarrow \forall t \in Z; x_t \le y_t . \tag{6}$$

Equation (6) does not perform the complete ordering at the set of all time series of the type defined above.
Features:
Obviously: $$E_n(x) \le x \le D_n(x) . \tag{7}$$

Coming out of inequality (7): $$E_n^2(x) = E_n(E_n(x)) \le E_n(x) \le E_n(D_n(x)) \tag{8}$$

$$D_n(E_n(x)) \le D_n(x) \le D_n(D_n(x)) = D_n^2(x) \tag{9}$$

Due to simplicity of proof apparatus we will further use the operation pair (max, min), i.e.:.

$d_i = U_n(x(i-n, i+n)) = \max\{x_{i-n},...,x_{i+n}\}$ and $l_i = L_n(x(i-n, i+n)) = \min\{x_{i-n},...,x_{i+n}\}$ (MM)

For $\forall i : -n \le i \le n$ holds: $x_t \le \max\{x_{t-n-i},...,x_{t+n-i}\}$, therefore $x_t \le \min_{-n \le i \le n} \max\{x_{t-n-i},...,x_{t+n-i}\}$. Due to (MM) this is equivalent to $x \le E_n(D_n(x)) = O_n(x)$. The proof of inequality $x \ge D_n(E_n(x)) = C_n(x)$ is analogical. Therefore:

$$C_n(x) \le x \le O_n(x) \tag{10}$$

Operators of opening and closing do not need to hold (and often do not hold) the feature of closure operators, i.e. following equation does not hold true in general: $O_n^2(x) = O_n(x)$ a $C_n^2(x) = C_n(x)$.

To increase the article clarity we will further on leave out the index of neighbourhood width (we defined the upper and lower estimations for, see (1)). This index will be left out in derived operations and operators as well. In case it will lead to ambiguity we will adopt the index again. Operators $D(x)$, $E(x)$, $O(x)$ and $C(x)$ are isotone considering the basic features of operations max and min. Then:

$$x \le z \Leftrightarrow D(x) \le D(z) \;\; a \;\; x \le z \Leftrightarrow E(x) \le E(z)$$
$$x \le z \Leftrightarrow O(x) \le O(z) \;\; a \;\; x \le z \Leftrightarrow C(x) \le C(z) \,. \qquad (11)$$

Appling the inequalities (10) and (11) we obtain:

$$C(x) \le OC(x) \le O(x) \text{ a } C(x) \le CO(x) \le O(x)\,. \qquad (12)$$

Figures 3, 4 show the progressive construction of operators $OC(x)$ and $CO(x)$.

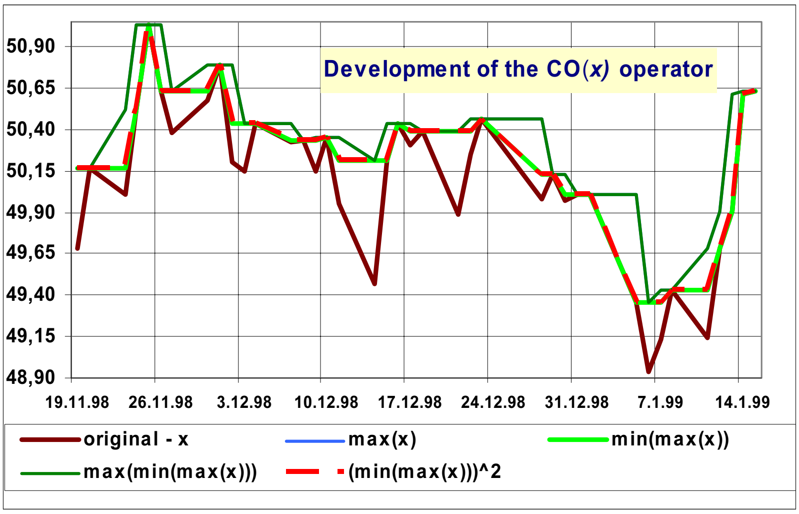

Figure3.: Sequential construction of the $CO(x)$ operator

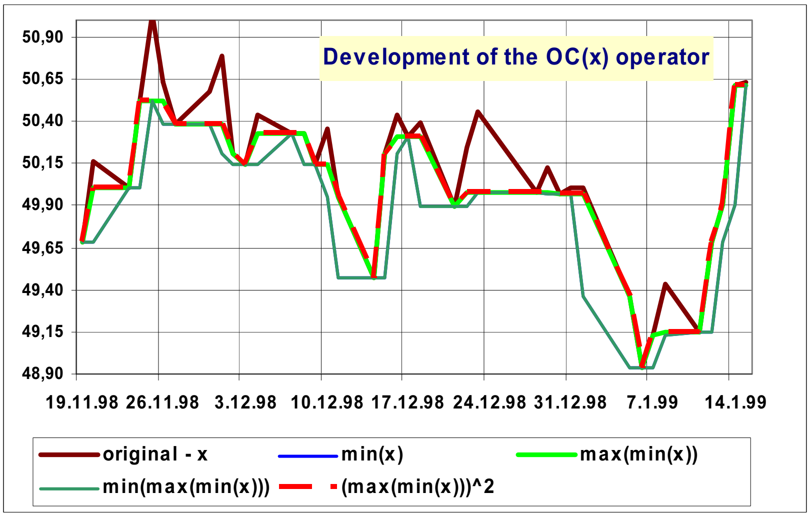

Figure 4.: Sequential construction of the $OC(x)$ operator

Analogically to inequality (10) and using the inequalities (11) and (12) following relation can be proved: (for the proof technique in detail see [2]):

$$OC(OC(x)) = (OC)^2(x) = OC(x) \;\text{ a }\; CO(CO(x)) = (CO)^2(x) = CO(x) \text{ a} \qquad (13)$$
$$OC(x) \le x \le CO(x)\,. \qquad (14)$$

Compound operators $OC(x)$ a $CO(x)$ are then idempotent and therefore they are the closure operators. Compound operators $C(O(O(C(x))))$ and $O(C(C(O(x))))$ would be interesting as well.

Further more we can derive that the operators $O(x)$ and $C(x)$ do not commute in general, so following inequality $OC(x) \neq CO(x)$ may (and often does) occur. Equality $OC(x) = CO(x)$ is rare event. (e.g. for x constant).
Figure 5 shows the application of both compound operators:

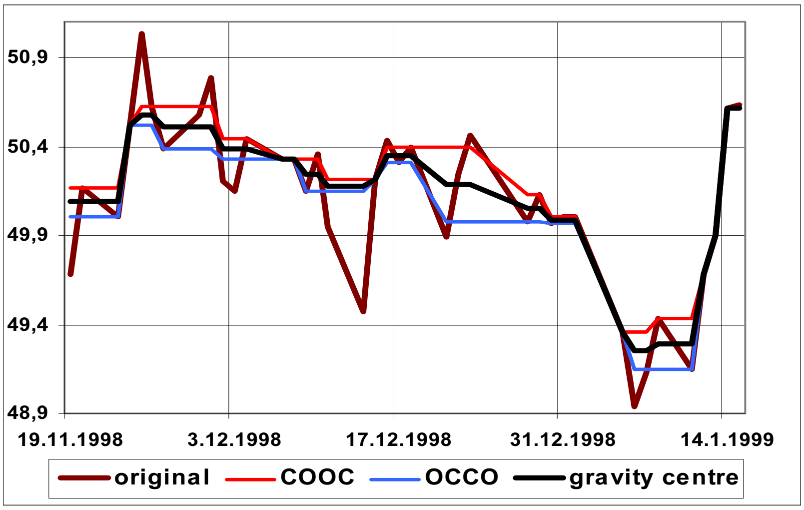

Figure 5.: Application of compound operators, their centre

4 Filtration and centre of time series

While introducing particular morphological operators and related figures, it is obvious, that the progressive smoothing of local extremes takes place. The smoothing is performed from both above and bellow.
Therefore we can adopt following functions:

Opening filtration:

$$F_O(x) = O(C(C(O(x)))) \tag{15}$$

Closing filtration: $$F_C(x) = C(O(C(C(x)))) \tag{16}$$

Further on, it is clear that both filtration operators lead to different results. In situations when such fact happens to be a handicap we can help ourselves with the centre of both curves:

$$T(x) = 0.5(F_O(x) + F_C(x)) \tag{17}$$

Such result could be well interpreted as the trend of analysed time series. Its big benefit is that apart from the last step, the centre holds the mapping of the numbers (i.e. it does not run into other numerical domain than observed values, for the pair of operations min, max). Sometimes it is convenient not to remove the difference between both procedures and we get the interval model of trend (see figures 6, 7).

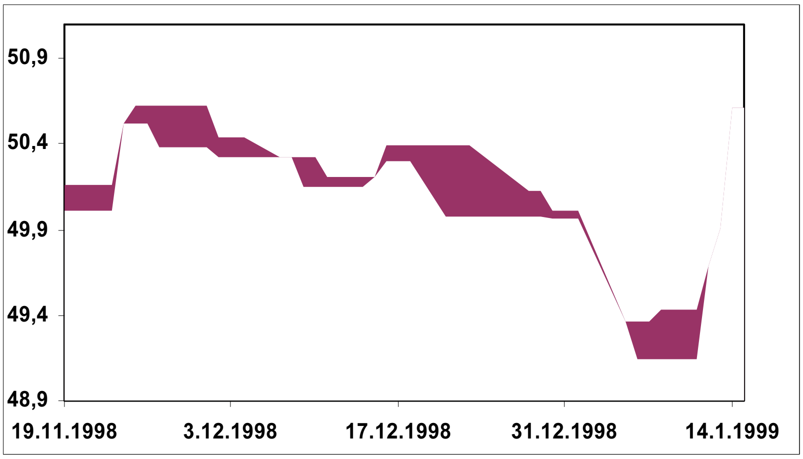

Figure 6.: Intermediate trend model, increased volatility areas

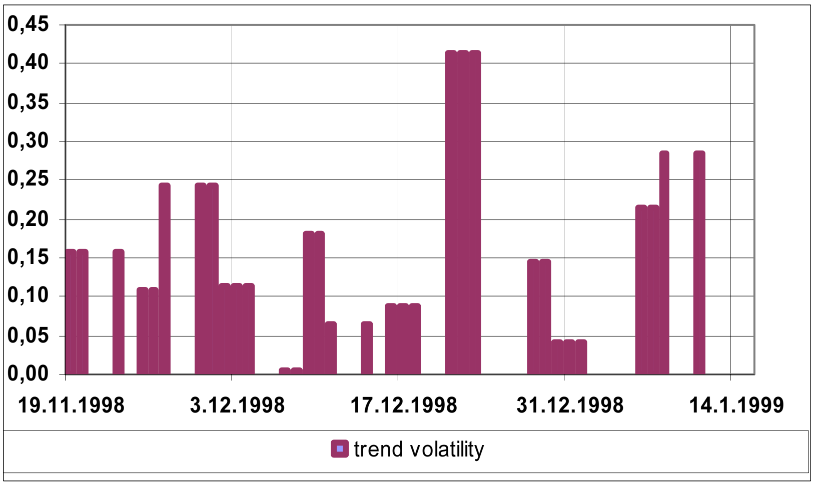

Figure 7.: Trend volatility dimension, bandwidth

In our work we assumed the time series to be boundless from both sides. In real situation this assumption never holds true. Therefore it is important to be particular about the modifications of the operations of lower and upper estimations at the time series bounds to avoid adverse boundary events. The modification for min and max operations is clear, we predefine the time series bounds in required length by its first or last value.

5 Future works

Apart from the trend analysis we can use proposed methods for forecasting. We slightly change the basic operators of dilation and erosion:

$$x \rightarrow D_n(x) = \; \ldots d_{t-1}, d_t, d_{t+1} \; \ldots; \quad d_{t+n+1} = U_n(x(t-n, t+n))$$
$$x \rightarrow E_n(x) = \; \ldots e_{t-1}, e_t, e_{t+1} \; \ldots; \quad e_{t+n+1} = L_n(x(t-n, t+n)$$

The choice of the parameter n (width of basic neighbourhood) is an interesting question for both filtration and forecasting activities especially according to smoothing ability (which local extremes will be smoothed while working with the particular n). Obvious problem is the analysis of multiple filtration operators F_O^k, F_C^k, T^k and their theoretical and application features. Another interesting question performs the detailed study of idempotency of selected operators introduced above. We can not also miss out the problems of algebraic structure of operators space (according to their composition). From the application point of view we will need to study the problems of possible effects according to selection of parameter n and selection of required power of that which operator.

This work was supported by the grant of the Ministry of Education of the Czech Republic No: MSM-235200005 - Information Systems and Technologies.

References

[1] Cohen, G., Gaunery, S., Quadrat J.P.: *Kernels, Images And Projections In Dioids*, WODES' 96. Edinburgh, Scotland, UK,1996.
[2] Granger, C. W.J., Hyung N.: *Introduction to M-M Processes*, Discussion Paper 98-18, Department of Economics University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508, July 1998.
[3] Gunawardena J.: *Idempotency*, Cambridge University Press, Publications of the Newton Institute,1996.
[4] Gunawardena J.: *Cycle Times and Fixed Points of Min-Max Functions*, Stanford University, Stanford USA, 11th International Conference on Analysis and Optimization of Systems, Sophia-Antipolis, France, June 1994.
[5] Gunawardena J.: *Min-Max Functions*, Stanford University, Stanford USA, 1994.
[6] Kiselman, O. C.: *Digital Geometry and Mathematical Morphology*, Lecture Notes, Spring Semester 2002, Uppsala University, Uppsala, Sweden,2002. www.math.uu.se/~kiselman.
[7] Kotlíková, M.: *Knowledge mining – some critical notes for time series*, MOSIS 2003, Brno, CR, 2003.
[8] Nelsen, R.B.: *An Introduction to Copulas*. Springer-Verlag New York, Inc, 1999.
[9] Serra, J.: *Image Analysis and Mathematiacal Morphology*. Academic Press, London 1982.
[10] Vávra F., Nový P., Kotlíková M, Mašková H.: *Simple Forecasting Model of Share Rates and Currencies.* Modelling and SImulation of Systems, pp. 249-253 Ostrava, 2002.
[11] Zmrhal D.: *Forecasting Model using the Predictor-Corrector Method*, MOSIS 2003, Brno, CR, 2003.

Contact address

Doc. Ing. František Vávra CSc., Ing. Pavel Nový, PhD., Ing. Hana Mašková, Ing. Michala Kotlíková, Ing. Arnoštka Netrvalová, FAV ZČU Plzeň, Department of Computer Science and Engineering, Univerzitní 22, 306 14 Plzeň,
e-mail: vavra@kiv.zcu.cz, novyp@kiv.zcu.cz, maskova@kiv.zcu.cz, kotlikov@kiv.zcu.cz, netrvalo@kiv.zcu.cz.