Fuzzy Prediction of Sunspot Number

Dana Majerová, Aleš Procházka, Jaromír Kukal

ICT Prague, Department of Computing and Control Engineering

Abstract. *The paper is devoted to a fuzzy data processing and a time series prediction. We chose a Lukasiewicz algebra with square root function for solving a problem of sunspot number prediction. We tested a lot of fuzzy logic functions. Four criteria were used for evaluating their prediction quality. A collection of fourteen advisable FIR filters was found. A fuzzy network was constructed using this filter set in the second phase. The fuzzy network weights were optimized for each criterion and the results were compared with auto–regressive model. All the fuzzy algorithms were realized in the MATLAB system.*

Keywords: *prediction, fuzzy processing, fuzzy logic function, Łukasiewicz algebra, neural network, Modus Ponens, inductive learning, Matlab.*

1 Introduction

The prediction in time series is a general problem related to many practical tasks. We suppose the fuzzy data processing with constrained sensitivity is a good tool for the prediction. The main idea is to use the simplest possible algebra with the widest abilities. Our algebraic model enables to construct both linear and non-linear filters including a neural network based on Modus Ponens rule.

2 Theoretical Foundation

A background of fuzzy data processing is the Łukasiewicz algebra [2] with square root (ŁA_{sqrt}). The other basic logic algebras like Gougen or Gödel are not good models because of discontinuity of residuum or lack of square root function [7]. The algebraic model is described in [4]. The definition of fuzzy logic function (FLF) and its sensitivity is also included.

The FLF's are used in a Modus Ponens Fuzzy Network (MPFN) [1]. The MPFN consists of four layers containing n input nodes, H FLF nodes, the layer of Modus Ponens law with learnable rule weights and m output nodes. Any MPFN output is FLF of input vector.

THEOREM: Let $H \in \mathbf{N}$ be a number of hidden FLF's in the given Modus Ponens Fuzzy Network. Let λ_j be their sensitivities for $j = 1, \ldots, H$. Then any MPFN output is FLF of input variables with the sensitivity

$$\lambda \leq \max_{j=1,\ldots,H} \lambda_j.$$

3 FIR Filters and Fuzzy Prediction

The FIR filters [6] are the standard tools for time series processing and prediction. Some FIR filters are FLF's.

THEOREM: Let $N \in \mathbf{N}_0$, $n \in \mathbf{N}$ and $m_k \in \mathbf{N}$ for $k = 1, \ldots, n$. Let $\overrightarrow{x} \in \mathbf{L}^n$. Let $w_k = m_k/2^N$ be dyadic weights for $k = 1, \ldots, n$ and $\sum_{k=1}^{n} w_k \leq 1$. Then any FIR filter

$$Y(\overrightarrow{x}) = \sum_{k=1}^{n} w_k \cdot x_k$$

is a FLF in the ŁA_{sqrt} and it satisfies the Lipschitz condition with the sensitivity

$$\lambda = \max_{k=1,\ldots,n} w_k.$$

THEOREM: Let $N \in \mathbf{N}_0$, $n \in \mathbf{N}$ and $m_k \in \mathbf{Z}$ for $k = 1, \ldots, n$. Let $\overrightarrow{x} \in \mathbf{L}^n$. Let $w_k = m_k/2^N$ be dyadic weights for $k = 1, \ldots, n$. Let $\sum_{w_k > 0} w_k \leq 1$, $\sum_{w_k < 0} |w_k| \leq 1$. Then

$$P(\overrightarrow{x}) = \max\left(0, \sum_{k=1}^{n} w_k \cdot x_k\right)$$

is a FLF in the ŁA_{sqrt} and it satisfies the Lipschitz condition with the sensitivity

$$\lambda = \max_{k=1,\ldots,n} |w_k|.$$

4 Experimental Results

A famous problem of sunspot number prediction was solved using the MATLAB [5] environment and year sunspot data from 1700 to 1997. The sunspot number history is described on the Fig. 1. The original sunspot number series $\{s_k\}_{k=1}^{t}$ were transformed using the differential fuzzyfication

$$f_k = \frac{s_k + 1/4}{s_k + 1/4 + s_{k-1} + 1/4}.$$

The f_k truth value corresponds to the proposition: The sunspot number increases (see Fig. 2).

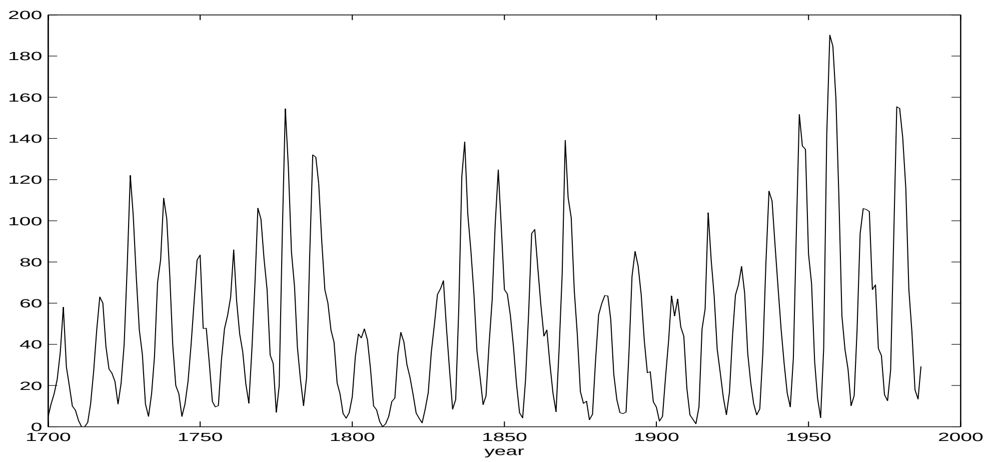

Figure 1: Sunspot number history

The sunspot number prediction from previous four year history was studied for $1 \leq D \leq 4$ where $D \in \mathbf{N}$. The relation between the prediction task and MPFN patterns is defined as $x_i = f_{k-i}$ for $i = 1, \ldots, D$ and $y_1 = \phi(f_{k-1}, \ldots, f_{k-D})$ where ϕ is any FLF.

A lot of experiments with various FLF's which are FIR filters were proceed in the first step. Four criteria were used for evaluated their prediction quality

$$
\begin{aligned}
MAXE &= \max_{i=1,\dots,m}\ \max_{k=1,\dots,p} |y_{ki} - Y_i(\overrightarrow{x}_k)|, \\
STDE &= \sqrt{\frac{1}{m \cdot p} \cdot \sum_{i=1}^{m}\sum_{k=1}^{p} (y_{ki} - Y_i(\overrightarrow{x}_k))^2}, \\
AVGE &= \frac{1}{m \cdot p} \cdot \sum_{i=1}^{m}\sum_{k=1}^{p} |y_{ki} - Y_i(\overrightarrow{x}_k)|, \\
MEDE &= \max_{i=1,\dots,m}\ \operatorname*{median}_{k=1,\dots,p} |y_{ki} - Y_i(\overrightarrow{x}_k)|.
\end{aligned}
$$

Using FLF's $\mu(x) = \max(0, \min(1, x))$, $\nu(x, N) = 2^{-N} \cdot x + (1 - 2^{-N})/2$ the best individual FLF predictors are

$$
\begin{aligned}
h_1 &= \frac{1}{2} \\
h_2 &= \nu(x_1, 1) \\
h_3 &= \nu(x_1, 4) \\
h_4 &= \nu\left(2\left(x_1 \ominus \frac{x_2}{2}\right), 2\right) \\
h_5 &= \nu\left(2\left(x_1 \ominus \frac{x_2}{2}\right), 4\right) \\
h_6 &= \nu\left(4\left(\frac{3 \cdot x_1 + x_3}{4} \ominus \frac{3 \cdot x_2}{4}\right), 5\right) \\
h_7 &= \nu\left(8\left(\frac{x_1 + x_3}{2} \ominus \frac{6 \cdot x_2 + x_4}{8}\right), 5\right) \\
h_8 &= \nu\left(\frac{x_1 + x_2}{2}, 1\right) \\
h_9 &= \nu(x_1 \vee x_2 \vee x_3 \vee x_4, 2) \\
h_{10} &= \nu(x_1 \wedge x_2, 1) \\
h_{11} &= \nu(x_1 \vee x_2, 1) \\
h_{12} &= \nu(x_1 \wedge x_2 \wedge x_3, 1) \\
h_{13} &= \nu\left(\frac{((x_1 \vee x_2) \wedge (x_1 \vee x_3) \wedge (x_2 \vee x_3)) + (x_1 \vee x_2 \vee x_3)}{2}, 3\right) \\
h_{14} &= \nu(x_1 \vee x_2 \vee x_3, 2).
\end{aligned}
$$

The best 14 FLF filters and their negations were used in the MPFN hidden layer of size $H = 28$. The MPFN was realized through two general functions in MATLAB [3]. The MPFN weights were optimized in the second phase. Penalty functions MAXE, STDE, AVGE and MEDE were minimized using the MCRS global optimization technique. The auto–regressive model (AR) of 4–th order was used as prediction standard. The results are described in the Tab. 1. The Figs. 4–7 show selected MPFN outputs and results of simple FLF's.

5 Conclusions

The Łukasiewicz algebra enriched by square root enables to realize operators approaching to weighted averages. It is appropriate for construct fuzzy network with limiting sensitivity. The

k	h_k–name	N	MAXE	STDE	AVGE	MEDE
1	constant	0	0.4287	0.1559	0.1259	0.1058
2	"linear" prediction	1	0.5768	0.1408	0.1012	0.0644
3	"linear" prediction	4	0.4150	0.1520	0.1214	0.1031
4	"quadratic" prediction	2	0.4934	0.1395	0.1035	0.0713
5	"quadratic" prediction	4	0.4112	0.1498	0.1193	0.1001
6	"cubic" prediction	5	0.4130	0.1540	0.1240	0.1069
7	"biquadratic" prediction	5	0.4130	0.1554	0.1252	0.1080
8	average of two	1	0.5357	0.1516	0.1125	0.0860
9	maximum of four	2	0.4151	0.1529	0.1278	0.1151
10	minimum of two	1	0.6143	0.1692	0.1183	0.0704
11	maximum of two	1	0.4946	0.1440	0.1117	0.0975
12	minimum of three	1	0.6143	0.1857	0.1338	0.0852
13	2.5th of three	3	0.4153	0.1527	0.1226	0.1011
14	maximum of three	2	0.4383	0.1488	0.1216	0.1037
	MPFN for minimum MAXE		0.4040	0.1547	0.1263	0.1094
	MPFN for minimum STDE		0.4598	0.1345	0.0981	0.0650
	MPFN for minimum AVGE		0.5255	0.1389	0.0956	0.0648
	MPFN for minimum MEDE		0.5834	0.1473	0.1007	0.0578
	AR of fourth order		0.4880	0.1338	0.0963	0.0641

Table 1: FLF filter efficiency

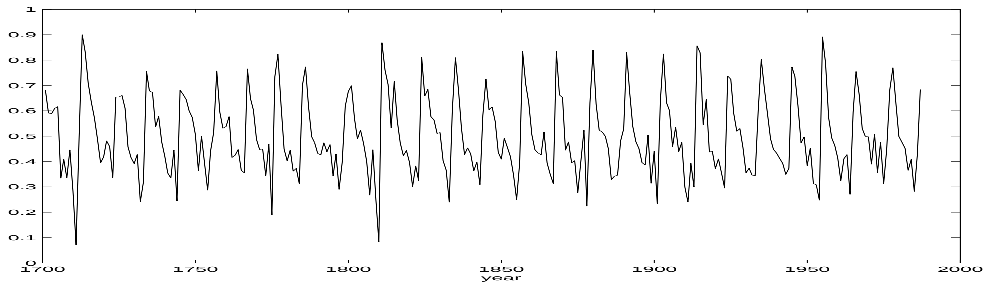

Figure 2: Fuzzyfication: sunspot number increases

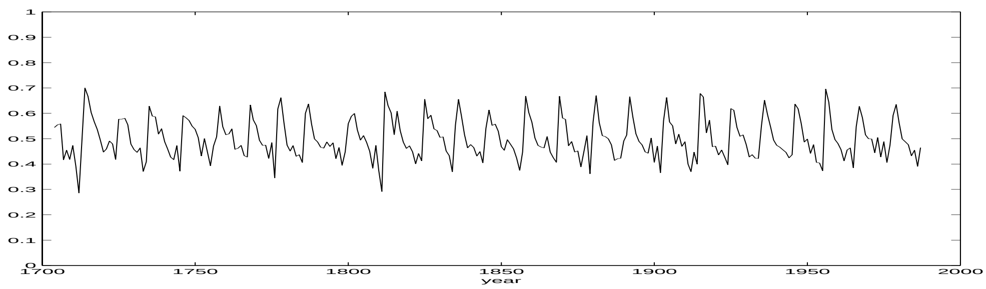

Figure 3: "Linear" prediction for $N = 1$

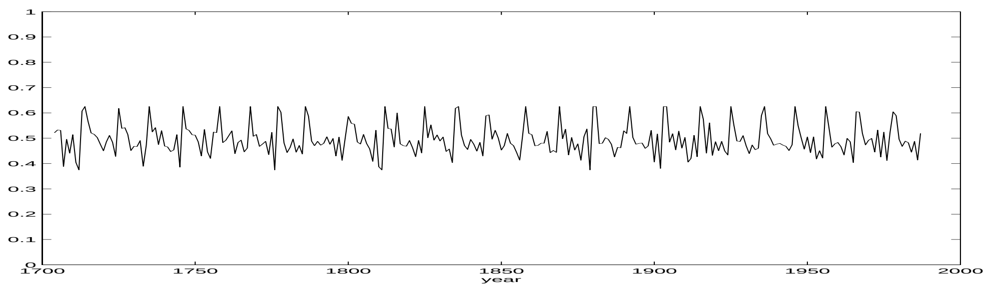

Figure 4: "Quadratic" prediction for $N = 2$

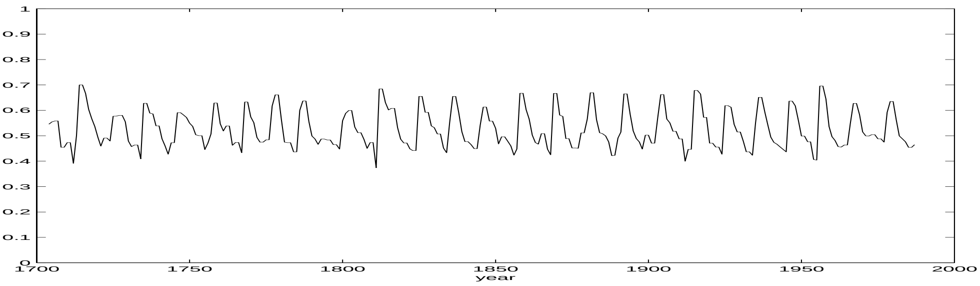

Figure 5: Maximum of two for $N = 1$

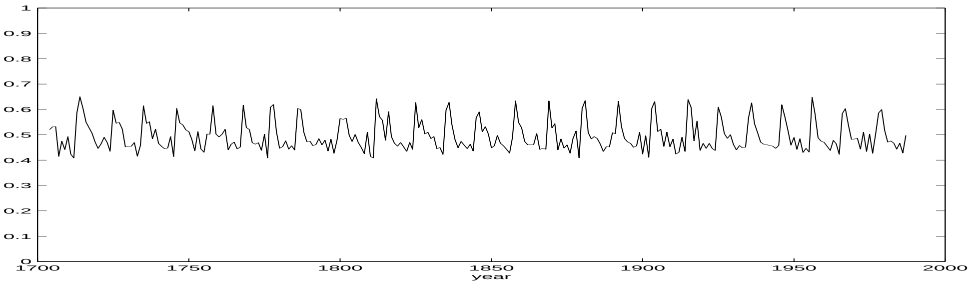

Figure 6: MPFN for optimum STDE

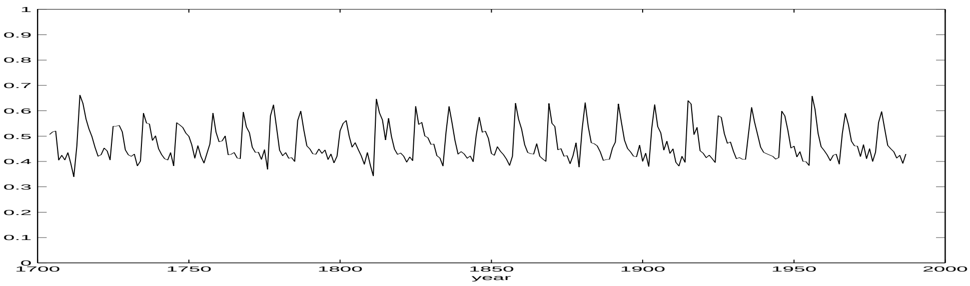

Figure 7: MPFN for optimum AVGE

Modus Ponens Fuzzy Network is applicable tool for fuzzy prediction. The optimization of fuzzy network structure will be a subject of our future research.

Acknowledgement

The work has been supported by the Research Intention of the Faculty of Chemical Engineering of the Institute of Chemical Technology in Prague. This support is very gratefully acknowledged.

References

[1] Kukal J. *Modus Ponens Fuzzy Network.* In: MENDEL 2000. Brno. 2000.

[2] Łukasiewicz J. *Selected Works.* Amsterdam. 1970.

[3] Majerová D., Kukal J. *Modus Ponens Fuzzy Network in Matlab.* In: MATLAB 2000. Prague. 2000.

[4] Majerová D., Kukal J., Procházka A. *2D Image Fuzzy Filters in Matlab.* In: MATLAB 2001. Prague. 2001.

[5] *Matlab 5.3.* The MathWorks, Inc. 1999.

[6] Mitra S. K., Kaiser J. F. *Handbook for Digital Signal Processing.* John Wiley & Sons. New York. 1993.

[7] Novák V., Perfilieva I., Močkoř J. *Mathematical Principles of Fuzzy Logic.* Kluwer Academic Publishers. Boston. 1999.

Contact

Dana Majerová, Aleš Procházka, Jaromír Kukal

INSTITUTE OF CHEMICAL TECHNOLOGY, Prague
Department of Computing and Control Engineering
Technická 5, 166 28 Prague 6 Dejvice
Phone: 420-2-2435 4170, fax: 420-2-2435 5053
E-mails: {Dana.Majerova, Ales.Prochazka, Jaromir.Kukal}@vscht.cz
WWW addresses: http://www.vscht.cz/ (university),
http://uprt.vscht.cz/ (department),
http://phobos.vscht.cz/ (research group)