Cognitive Maps for Time Series Modeling

Władysław Homenda

Faculty of Mathematics and Information Science
Warsaw University of Technology, Warsaw, Poland

Faculty of Economics and Informatics in Vilnius
University of Białystok, LT-08221 Vilnius, Lithuania

International Conference
of Artificial Intelligence and Information Technology
Yogyakarta, March 13-15, 2019

The Agenda

The Agenda

Cognitive maps, origin of the term

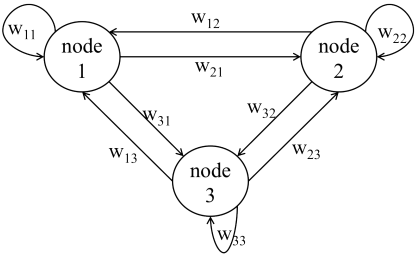
- Cognitive maps:
  - Tolman, E. C. (1948), *Cognitive maps in rats and men*, Psychological Review, 55, 189-208,

    the concepts of hidden learning of animals,
  - Axelrod, R. (1976), *Structure of Decision: the Cognitive Maps of Political Elites*, Princeton, NJ: Princeton University Press,

    causality relations in politics, economics, economy,
  - Kosko, B. (1986), *Fuzzy Cognitive Maps*, International Journal of Man-Machine Studies, 24, 65-75

    causality relations in terms of fuzziness.

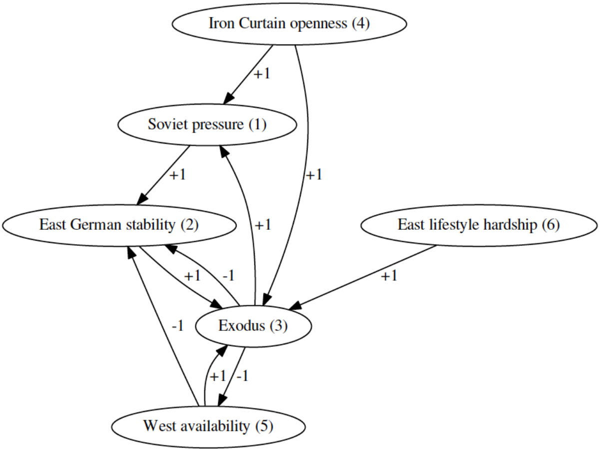
Model



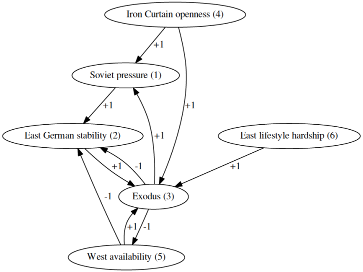
A cognitive map with three concepts

weights: $w_{ij} \in \{-1, 0, 1\}$
states/activations: $node_i \in \{0, 1\}$

Model


Iron Curtain openness (4)
Soviet pressure (1)
East German stability (2)
East lifestyle hardship (6)
Exodus (3)
West availability (5)
+1
+1
+1
+1
+1
-1
+1
-1
+1
-1

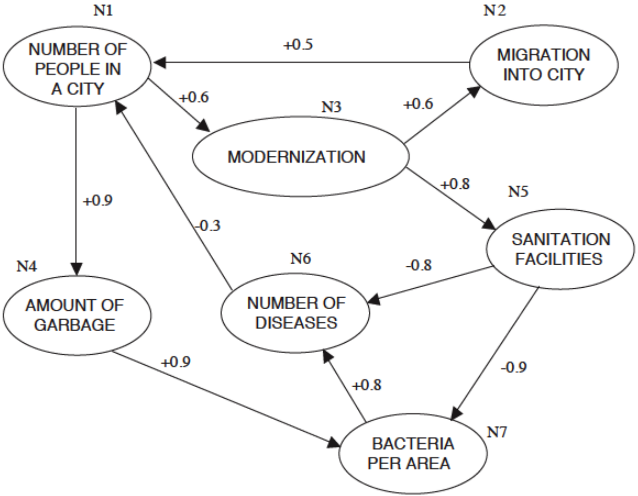
Model



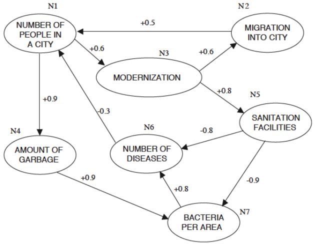
W	1	2	3	4	5	6
1	0	0	+1	+1	0	0
2	+1	0	−1	0	−1	0
3	0	+1	0	+1	+1	+1
4	0	0	0	0	0	0
5	0	0	−1	0	0	0
6	0	0	0	0	0	0

The Agenda

Model


N1
NUMBER OF PEOPLE IN A CITY
N2
MIGRATION INTO CITY
+0.5
+0.6
N3
MODERNIZATION
+0.6
+0.9
-0.3
+0.8
N5
SANITATION FACILITIES
N4
AMOUNT OF GARBAGE
N6
NUMBER OF DISEASES
-0.8
+0.9
+0.8
-0.9
N7
BACTERIA PER AREA

Model



	N1	N2	N3	N4	N5	N6	N7
N1	0	0	0.6	0.9	0	0	0
N2	0.5	0	0	0	0	0	0
N3	0	0.6	0	0	0.8	0	0
N4	0	0	0	0	0	0	0.9
N5	0	0	0	0	0	-0.8	-0.9
N6	-0.3	0	0	0	0	0	0
N7	0	0	0	0	0	0.8	0

weights activations responses goals

$$\begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ & & \dots & \\ w_{n1} & w_{22} & \dots & w_{nn} \end{bmatrix} \text{ with } \begin{bmatrix} x_1. \\ x_2. \\ \dots \\ x_n. \end{bmatrix} \text{ produce } \begin{bmatrix} y_1. \\ y_2. \\ \dots \\ y_n. \end{bmatrix} \text{ modelling } \begin{bmatrix} g_1. \\ g_2. \\ \dots \\ g_n. \end{bmatrix}$$

Model

weights activations responses goals

$$\begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ & & \dots & \\ w_{n1} & w_{22} & \dots & w_{nn} \end{bmatrix} \text{ with } \begin{bmatrix} x_{1.} \\ x_{2.} \\ \dots \\ x_{n.} \end{bmatrix} \text{ produce } \begin{bmatrix} y_{1.} \\ y_{2.} \\ \dots \\ y_{n.} \end{bmatrix} \text{ modelling } \begin{bmatrix} g_{1.} \\ g_{2.} \\ \dots \\ g_{n.} \end{bmatrix}$$

details of computation

$$\begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ & & \dots & \\ w_{n1} & w_{22} & \dots & w_{nn} \end{bmatrix} \star \begin{bmatrix} x_{1.} \\ x_{2.} \\ \dots \\ x_{n.} \end{bmatrix} = \begin{bmatrix} f(w_{11}x_{1.} + w_{12}x_{2.} \dots + w_{1n}x_{n.}) \\ f(w_{21}x_{1.} + w_{22}x_{2.} \dots + w_{2n}x_{n.}) \\ \dots \\ f(w_{n1}x_{1.} + w_{n2}x_{2.} \dots + w_{nn}x_{n.}) \end{bmatrix} = \begin{bmatrix} y_{1.} \\ y_{2.} \\ \dots \\ y_{n.} \end{bmatrix} \approx \begin{bmatrix} g_{1.} \\ g_{2.} \\ \dots \\ g_{n.} \end{bmatrix}$$

Model

activations targets:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{N1} \\ x_{21} & x_{22} & x_{23} & \dots & x_{N2} \\ & \dots & & & \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{Nn} \end{bmatrix} \qquad \begin{bmatrix} g_{11} & g_{12} & g_{13} & \dots & g_{N1} \\ g_{21} & g_{22} & g_{23} & \dots & g_{N2} \\ & \dots & & & \\ g_{n1} & g_{n2} & g_{n3} & \dots & g_{Nn} \end{bmatrix}$$

$$X = \begin{bmatrix} X_1 & X_2 & X_3 & \dots & X_N \end{bmatrix} \qquad \begin{bmatrix} G_1 & G_2 & G_3 & \dots & G_N \end{bmatrix} = G$$

$$W \star \begin{bmatrix} X_1 & X_2 & \dots & X_N \end{bmatrix} = \begin{bmatrix} Y_1 & Y_2 & \dots & Y_N \end{bmatrix} \approx \begin{bmatrix} G_1 & G_2 & \dots & G_N \end{bmatrix}$$

$$W \star X = Y \approx G$$

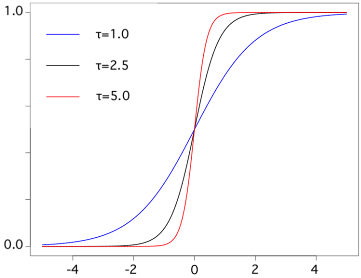
$$y_{ji} = f\left(\sum_{k=1}^{n} w_{jk} \cdot x_{ki}\right)$$

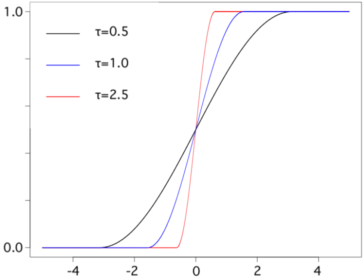
Model

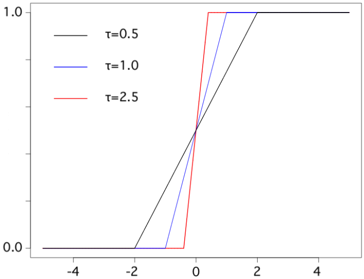
- Examples of squashing functions:

$$f : (-\infty, +\infty) \rightarrow [0, 1]$$

  - sigmoidal: $f(x) = \dfrac{1}{1 + e^{-\tau \cdot x}}$,
  - sinus: $f(x) = \begin{cases} 0 & \tau \cdot x < -\pi/2, \\ \frac{1}{2}(\sin(\tau \cdot x) + 1) & -\pi/2 \leqslant \tau \cdot x \leqslant pi/2 \\ 1 & \tau \cdot x > \pi/2 \end{cases}$
  - linear: $f(x) = \begin{cases} 0 & \tau \cdot x < -1, \\ \frac{1}{2}(\tau \cdot x + 1) & -1 \leqslant \tau \cdot x \leqslant 1 \\ 1 & \tau \cdot x > 1 \end{cases}$


1.0
0.0
τ=1.0
τ=2.5
τ=5.0
-4
-2
0
2
4


1.0
0.0
τ=0.5
τ=1.0
τ=2.5
-4
-2
0
2
4


1.0
0.0
τ=0.5
τ=1.0
τ=2.5
-4
-2
0
2
4

A problem

Given

- activations X,
- targets G

design a fuzzy cognitive map W modelling a given problem, i.e. such that:

$$W \star X \;=\; G$$

or at least such W' that

$$W' \star X = Y$$

where Y is as close to G as possible

Evaluations

$$W' \star X = Y \approx G$$

$$\begin{array}{cc} \text{responses} & \text{targets} \\ \begin{bmatrix} y_{11} & y_{12} & y_{13} & \cdots & y_{1N} \\ y_{21} & y_{22} & y_{23} & \cdots & y_{2N} \\ & \cdots & & & \\ y_{n1} & y_{n2} & y_{n3} & \cdots & y_{nN} \end{bmatrix} & \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1N} \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2N} \\ & \cdots & & & \\ g_{n1} & g_{n2} & g_{n3} & \cdots & g_{nN} \end{bmatrix} \end{array}$$

Fundamental quality measures

$$W' \star X = Y \approx G$$

responses targets

$$\begin{bmatrix} y_{11} & y_{12} & y_{13} & \cdots & y_{1N} \\ y_{21} & y_{22} & y_{23} & \cdots & y_{2N} \\ & \cdots & & & \\ y_{n1} & y_{n2} & y_{n3} & \cdots & y_{nN} \end{bmatrix} \quad \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1N} \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2N} \\ & \cdots & & & \\ g_{n1} & g_{n2} & g_{n3} & \cdots & g_{nN} \end{bmatrix}$$

Response evaluation, fundamentals:

- $MSE = \frac{1}{N \cdot n} \sum_{j=1}^{N} \sum_{i=1}^{n} (y_{ij} - g_{ij})^2$
- $MAXSE = \max\{(y_{ij} - g_{ij})^2 : i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, N\}\}$

Model construction

Map reconstruction from data:

- Data:
  - $X = [X_1, X_2, \ldots, X_N]$ - activations,
  - $G = [G_1, G_2, \ldots, G_N]$ - targets,
  - W - random weights, uniform distribution.
- Map reconstruction:
  - $W' \star X = Y$
    $Y = [Y_1, Y_2, \ldots, Y_N]$ as close as possible to $G = [G_1, G_2, \ldots, G_N]$
  - minimization of objective function,
- Results:
  - objective function is a choice of a quality measure (MSE, MSEk, etc.),
  - gradient, PSO or other optimization methods,
  - "as close as possible" in terms of a quality measure.

The Agenda

1. Cognitive maps
   - Origin of the term
   - Model
2. Fuzzy cognitive maps
   - Model
   - Evaluating
   - Fundamental quality measures
3. Time series models
   - Model with fuzzy cognitive map
   - Empirical evaluation
4. Processing time series at concept level
   - Concept level of time series
   - Understanding time series

Fuzzy cognitive map as a model of time series

Time series:

$$c_1, c_2, c_3, c_4, \ldots$$

Activations:

$$X_{\cdot 1} = [c_1, c_2, \ldots, c_n]^T$$

$$X_{\cdot 2} = [c_2, c_3, \ldots, c_{n+1}]^T$$

$$\ldots$$

$$X_{\cdot N} = [c_N, c_{N+1}, \ldots, c_{N+n-1}]^T$$

Targets:

$$G_{\cdot 1} = [c_2, c_3, \ldots, c_{n+1}]^T$$

$$G_{\cdot 2} = [c_3, c_4, \ldots, c_{n+2}]^T$$

$$\ldots$$

$$G_{\cdot N} = [c_{N+1}, c_{N+2}, \ldots, c_{N+n}]^T$$

Fuzzy cognitive map as a model of time series

X	$X_{\cdot 1}$	$X_{\cdot 2}$	$X_{\cdot 3}$	$\dots$	$X_{\cdot N-1}$	$X_{\cdot N}$
$X_{1\cdot}$	c_1	c_2	c_3	$\cdots$	c_{N-1}	c_N
$X_{2\cdot}$	c_2	c_3	c_4	$\cdots$	c_N	c_{N+1}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$X_{n-1\cdot}$	c_{n-1}	c_n	c_{n+1}	$\cdots$	c_{N+n-1}	c_{N+n-2}
$X_{n\cdot}$	c_n	c_{n+1}	c_{n+2}	$\cdots$	c_{N+n-2}	c_{N+n-1}

G	$G_{\cdot 1}$	$G_{\cdot 2}$	$G_{\cdot 3}$	$\dots$	$G_{\cdot N-1}$	$G_{\cdot N}$
$G_{1\cdot}$	c_2	c_3	c_4	$\cdots$	c_N	c_{N+1}
$G_{2\cdot}$	c_3	c_4	c_5	$\cdots$	c_{N+1}	c_{N+2}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$G_{n-1\cdot}$	c_n	c_{n+1}	c_{n+2}	$\cdots$	c_{N+n}	c_{N+n-1}
$G_{n\cdot}$	c_{n+1}	c_{n+2}	c_{n+3}	$\cdots$	c_{N+n-1}	c_{N+n}

Fuzzy cognitive map as a model of time series

Modeling considered so far:

$$Y_{\cdot j} = f(W \cdot X_{\cdot j})$$

$$y_{ij} = f(W_{i\cdot} * X_{\cdot j}) = f\Big(\sum_{k=1}^{n} w_{ik} \cdot x_{kj}\Big)$$

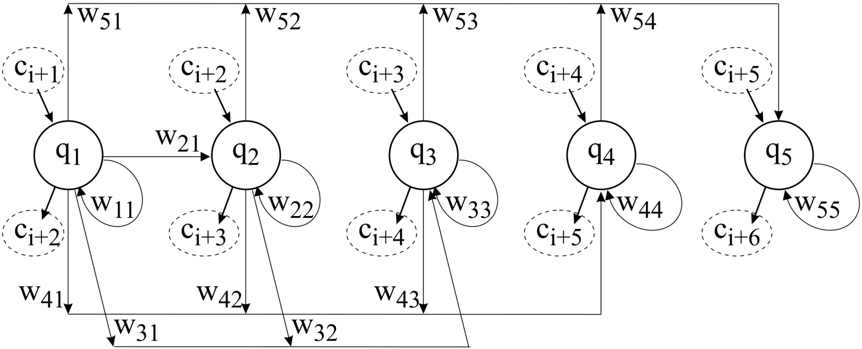
$$w_{ij} = 0 \quad \text{dla} \quad i > j$$

Adding bias:

$$Y_{\cdot j} = f(B + W \cdot X_{\cdot j})$$

$$y_{ij} = f(b_i + W_{i\cdot} * X_{\cdot j}) = f\Big(b_i + \sum_{k=1}^{n} w_{ik} \cdot x_{kj}\Big)$$

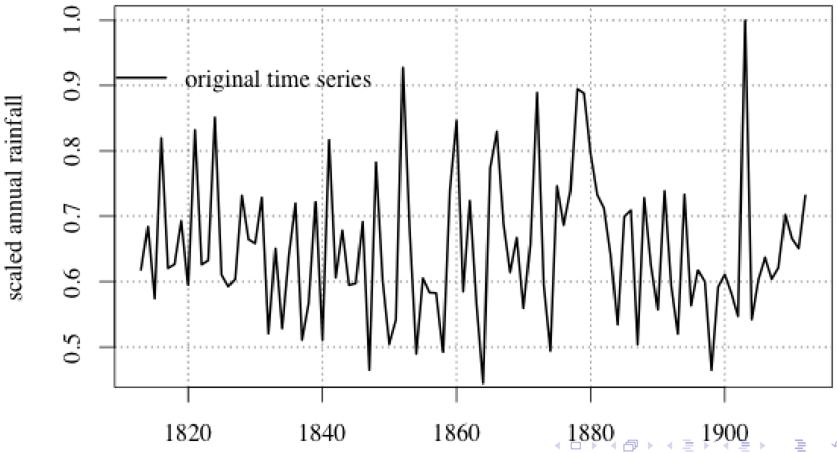
Fuzzy cognitive map as a model of time series



Prediction - MSE as objective function
Evaluating results - MSE

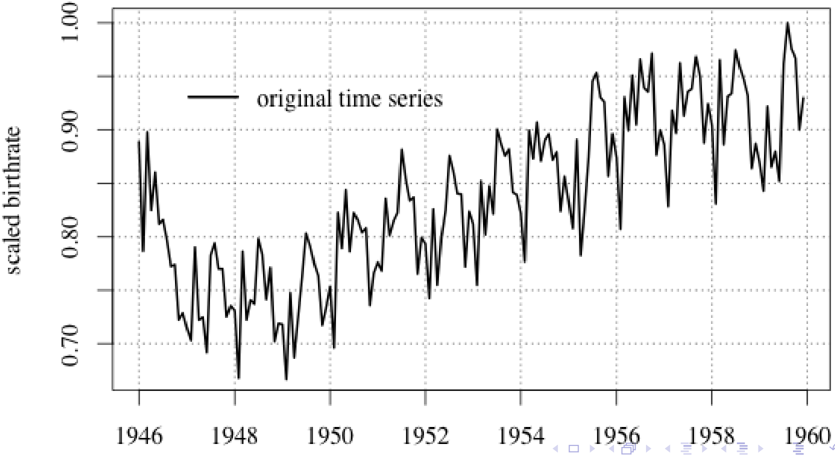
Fuzzy cognitive map as a model of time series

Annual rainfall in London from 1813-1912



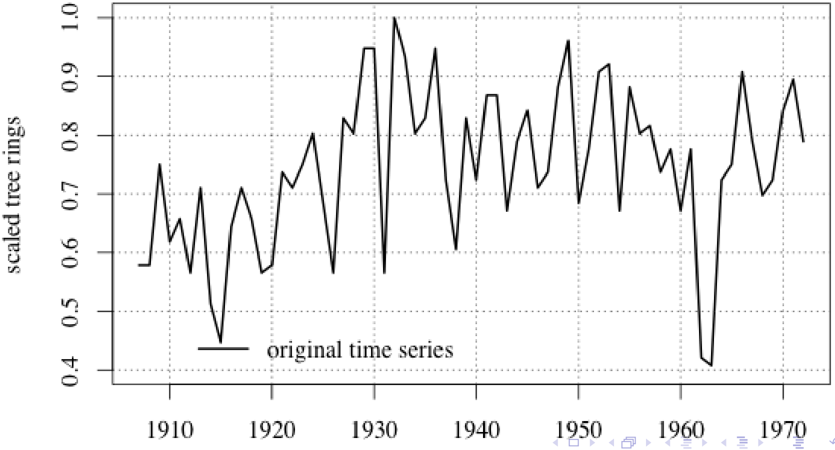
Fuzzy cognitive map as a model of time series

Number of births per month in New York city, 1946-1959



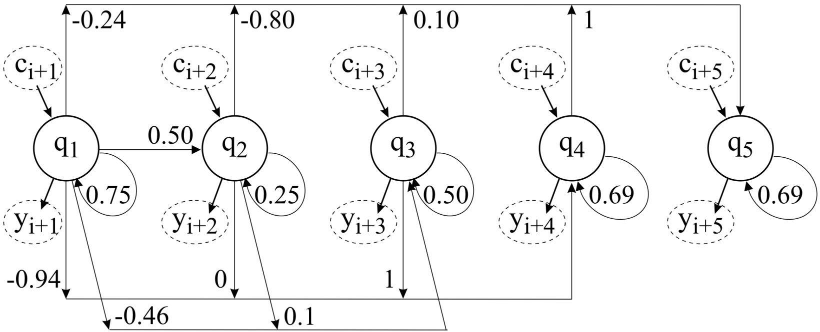
Fuzzy cognitive map as a model of time series

Campito tree rings, which indicate tree growth, 1907-1960



Fuzzy cognitive map as a model of time series

FCM n=5 for *births* time series



Fuzzy cognitive map as a model of time series

MSE*100 for time series forecasts with different map sizes with and without bias.

	rain		birth		tree rings	
map size	no bias	with bias	no bias	with bias	no bias	with bias
3	1.69	1.69	1.00	0.89	1.00	1.36
4	1.70	1.70	0.98	0.86	1.15	1.78
5	1.71	1.72	0.96	0.78	1.80	2.06
6	1.77	1.80	0.92	0.78	2.19	2.33
7	1.82	1.87	0.91	0.75	1.96	2.07
8	1.86	1.89	0.89	0.77	2.66	2.74
9	1.92	1.92	0.89	0.73	2.92	2.97
10	1.81	1.86	0.89	0.68	2.99	3.11
11	1.67	1.71	0.89	0.69	2.89	3.01
12	1.64	1.66	0.89	0.74	2.85	3.05

Fuzzy cognitive map as a model of time series

Comparison of MSE obtained for three time series with different modeling and forecasting methods.

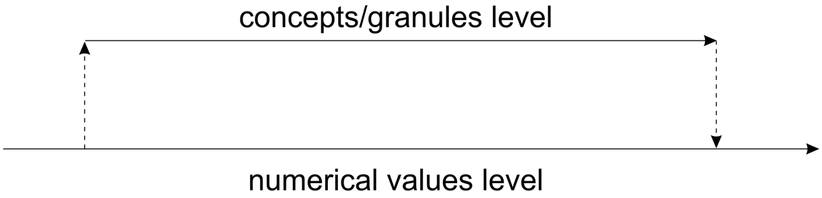
MSE*100	rain		births		tree rings	
method	train	forecast	train	forecast	train	forecast
ARIMA par=(1,0,0)	1.18	2.15	0.13	0.59	1.30	3.51
Holt-Winters	1.27	1.77	0.37	0.18	1.31	4.60
FCM, n=5, no bias	1.20	1.71	0.34	0.96	4.03	3.71
FCM, n=5 with bias	1.17	1.72	0.21	0.78	5.44	2.33

The Agenda

1. Cognitive maps
   - Origin of the term
   - Model
2. Fuzzy cognitive maps
   - Model
   - Evaluating
   - Fundamental quality measures
3. Time series models
   - Model with fuzzy cognitive map
   - Empirical evaluation
4. Processing time series at concept level
   - Concept level of time series
   - Understanding time series

Concept level of time series

- weather will be very/moderately hot
  in south part of Poland tomorrow
- level of rivers are in lower part of the scale and is equal:
  - 452 centimetres of Vistula river in Sandomierz
    its increase is 23 centimetres during 24 hours
  - ...



Handling time series

Time series representation in the space
Amplitude/Amplitude change/Change of amplitude change

- $a_1, a_2, a_3, a_4, \ldots$ - amplitude
- $\delta a_2, \delta a_3, \delta a_4, \ldots$ - amplitude change
  - $\delta a_2 = a_2 - a_1,$
  - $\delta a_3 = a_3 - a_2,$
  - $\delta a_4 = a_4 - a_3,$
- $\delta\delta a_3, \delta\delta a_4, \delta\delta a_5, \ldots$ - change of amplitude change
  - $\delta\delta a_3 = \delta a_3 - \delta a_2,$
  - $\delta\delta a_4 = \delta a_4 - \delta a_3,$
  - $\delta\delta a_5 = \delta a_5 - \delta a_4,$

Handling time series

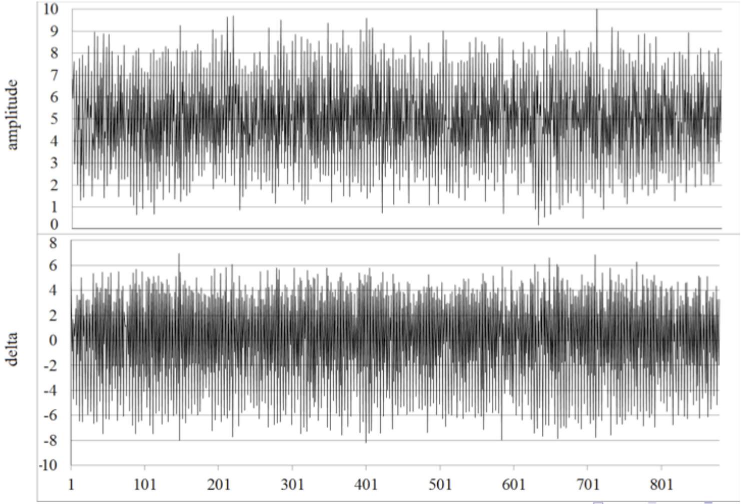
Synthetic time series were constructed according to the following procedure:

- select base sequence
- replicate the base sequence to the total length 3000,
- add a random distortion drawn from the normal (Gaussian) probability distribution with mean equal to 0 and standard deviation equal to 0.7.

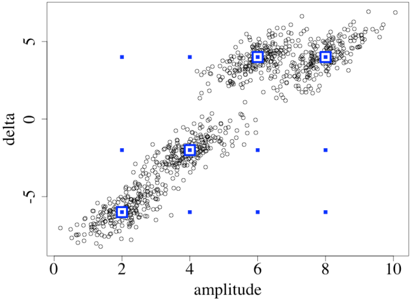
3-dimensional dynamics representation of the 2648-based synthetic time series

time	1	2	3	4	5	6	7	...	2998	2999	3000
amplitude	2	6	4	8	2	6	4	8	...	4	8
amp. change	$\sim$	4	-2	4	-6	4	-2	4	...	-2	4
change of a. c.	$\sim$	$\sim$	-6	6	-10	10	-6	6	...	-6	6

Handling time series

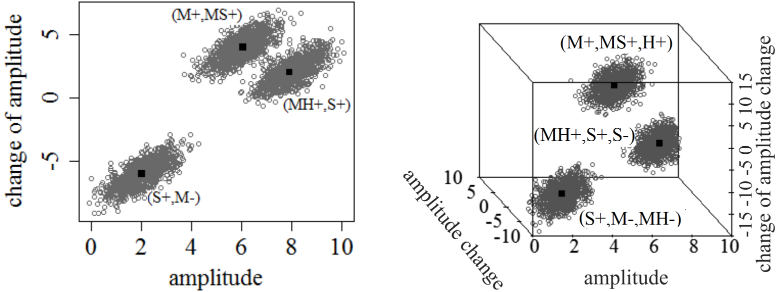


Transformation to concepts space



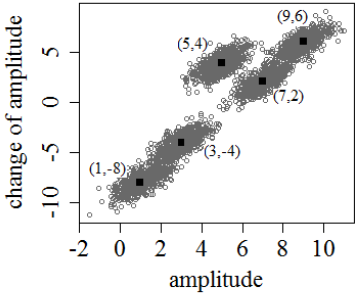
- One dimensional clustering:
  - amplitudes: Small, Moderately Small, Moderately High, High,
  - deltas: High Negative, Small Negative, Moderately Positive.
- Two dimensional clustering:
  (S,HN), (S,SN), (S,MP), (MS,HN), . . . , (H,MP).

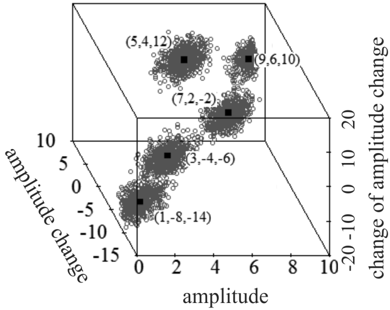
Transformation to concepts space



The 268-based synthetic time series

Transformation to concepts space

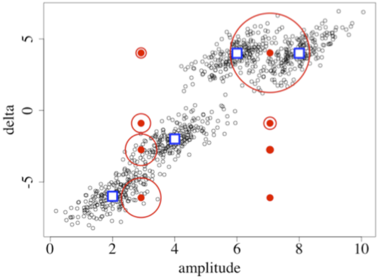


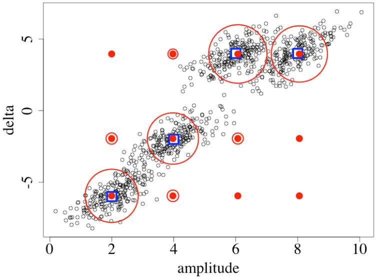


The 15739-based synthetic time series

Transformation to concepts space

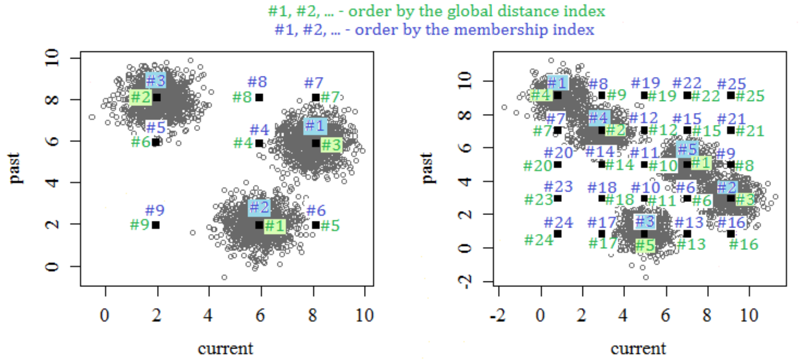
Strength of concepts





Transformation to concepts space

Strength of concepts



Ranking of concepts for the 268 and 15739-based time series. Best concepts are emphasised with coloured background.

Concepts evaluation

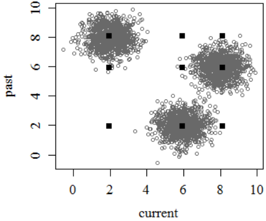
$$M(\mathbf{v}_j) = \sum_{j=1}^{N} x_{ij}$$ sum of membership degrees to the concept

$$GD(\mathbf{v}_j) = \sum_{i=1}^{N} e^{-\|\mathbf{z}_i - \mathbf{v}_j\|}$$ sum of distances to the concept

where:

N - number of points, $\mathbf{v}_j$ - j-th concept, $\mathbf{z}_i$ - i-th data point

x_{ij} - i-th activation corresponds to the j-th concept

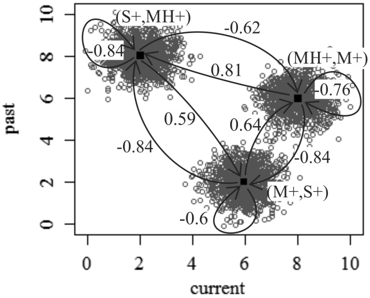


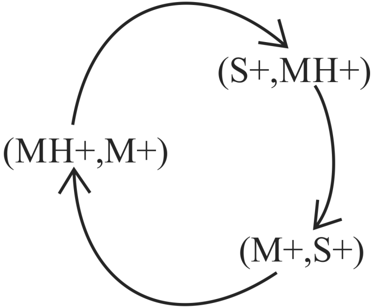
Understanding time series

3-dimensional dynamics representation
of the 268-based synthetic

time	1	2	3	4	5	6	7	...	2998	2999	3000
amplitude	2	6	8	2	6	8	2	...	2	6	8
amp. change	$\sim$	4	2	-6	4	2	-6	...	-6	4	2
change of a. c.	$\sim$	$\sim$	-2	-8	10	-2	-8	...	-8	10	-2

Understanding time series





On the left: FCM with 3 nodes trained for the 268 time series. Nodes are black squares. Labels are their linguistic description.
On the right: the actual cycle of values in the 268 time series. Positive linkages in the map form the same cycle as sketched on the left.

Cognitive Maps for Time Series Modeling

Władysław Homenda

Faculty of Mathematics and Information Science
Warsaw University of Technology, Warsaw, Poland

Faculty of Economics and Informatics in Vilnius
University of Białystok, LT-08221 Vilnius, Lithuania

International Conference
of Artificial Intelligence and Information Technology
Yogyakarta, March 13-15, 2019