References

1. Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via The EM Algorithm. *J. Royal Statistical Society Series B* 39:1-38.
2. Roderick, J., Little, A., and Rubin, D. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.
3. Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. New York: Chapman & Hall/Crc (Software NORM: http://www.stat.psu.edu/~jls/ misoftwa.html #top)
4. Tresp, V., Hofmann, R. (1998). Nonlinear Times-Series Prediction with Missing and Noisy Data. *Neural Computation* 10(3):731-747.
5. Parzen, E., editor. (1984). Time Series Analysis of Irregularly Observed Data. *Lecture Notes in Statistics* Vol. 25, New York: Springer.
6. Ferreio, O. (1987). Methodologies for the Estimation of Missing Observations in Times Series. *Statist. Prob. Lett.* 5(1):65-69.
7. Alexiadis, M.C., Sahsamanoglou, P.S., H.S., Manousaridis, I.M. (1998). Short-Term Forecasting of Wind Speed and Related Electrical Power. *Solar Energy* 63(1):61-68.
8. Pedreira, C. E. and Parente, E. (1995). Neural Networks with Missing Values Attributes. *Proceedings IEEE Internal Conference on Neural Network* 6:3021-3023.
9. Verikas, A.,Gelzinis, A., Malmqvist, K., and Bacauskiene, M. (2001). Using

Unlabelled Data to Train a Multilayer Perceptron. ICAPR 2001, *LNCS* 2013, pp. 40-49.

10. Tanaka, M. (1996). Identification of Nonlinear Systems with Missing Data Using Stochastic Neural Network. *Proc. of the 35th Conf. On Decision and Control*, pp. 933-934. Japan.

11. Miller, D. J., and Uyer, H. S. (1998). Combined Learning and Use for a Mixture Model equivalent to the RBF Classifier. *Neural Computation* 10:281-293.

12. Ghahramani, Z., and Jordan, M. (1994). Supervised Learning from Incomplete Data via an EM Approach. in J. D. Coruan, G. Tesauso, and J. Alspector (ed.), *Advanced in Neural Information Processing System* vol. 6.

13. Hathaway, R. J., and Bezdek, J. C. (2001). Fuzzy c-Means Clustering of Incomplete Data. *IEEE Trans. on SMC part B* 31(5):735-744.

14. Timm, H., Doring, C., Kruse, R. (2004). Different Approaches to Fuzzy Clustering of Incomplete Datasets. *International Journal of Approximate Reasoning* 35:239 249.

15. http://www.csee.ogi.edu/~ericwan/data.html.

16. http://science.nasa.gov/ssl/pad/solar/sunspots.htm

17. Chiewchanwattana, S., and Lursinsap, C. (2002). FI-GEM Network for Incomplete Time-Series Prediction. *Proceeding International Joint Conference on Neural Network* 2:1757-1762, Honolulu.

18. Hughes, D.A., and Smakhtin, V. (2001). Daily Flow Time Series Patching or Extension: A Spatial Interpolation Approach based on Flow Duration Curves. *Hydrol. Sci. J.* 41(6):851-871.

19. Schneider, T. (2001). Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values. *Journal of Climate* 14:853-887.

20 Verbeek, J. J., Vlassis, N., Krose, B. (2002). A k-segments Algorithm for Finding Principal Curves. *Pattern Recognition Letters* 23:8:1009-1017.

21. Chiewchanwattana, S., Lursinsap, C., and Chu, C. H. (2002). Time-Series Data Prediction Based on Reconstruction of Missing Samples and Selective Ensembling of FIR Neural Networks. *Proceedings of The 9th International Conference on Neural Information Processing (ICONIP'02)*, pp. 2152-2156, Singapore.

22. Chiewchanwattana, S., Lursinsap, C., and Chu, C. H. (2004). A Reconstructed Missing Data-Finite Impulse Response Selective Ensemble (RMD-FSE) Network. In Rajapakse, Jagath C.; Wang, Lipo (Eds.), A Chapter in *Neural Information Processing System: Research and Development (Series: Studies in Fuzziness and Soft Computing)* ISBN: 3-540-21123-3, 152:113-127, Springer.

23. Chiewchanwattana, S., Lursinsap, C., and Chu, C. H. (n.d.). The Imputation of Missing Samples based on the Varied Window Clustering of Incomplete Time Series Data. *Pattern Recognition Letter*.

24. Perrone, M. P., and Cooper, L. N. (1992). When Networks Disagree: Ensemble Methods for Hybrid Neural Networks. in *Neural Networks for Speech and Image Processing*. (n.p.): Chapman & Hall.

25. Singhal, S., and Wu, L. (1988). Training Multilayer Perceptrons with the Extended

Kalman Algorithm. in D. S. Touretzky (ed.), *Advances in Neural Information Processing Systems* 1:133-140.

26. Wan, E. A. (1994). Time Series Prediction by Using a Connectionist Network with Internal Delay Lines. in *Time Series Prediction. Forecasting the Future and Understanding the Past, SFI Studies in the Sciences of Complexity*, A. Weigend and N. Gershenfeld (Eds.), Proc. Vol. XVII. (n.p.): Addison-Wesley.

27. Zhou, Z.-H., Wu, J.-X., Tang, W., and Chen, Z.-Q. (2001). Combining Regression Estimators: GA-Based Selective Neural Network Ensemble. *International Journal of Computational Intelligence and Applications* 1(4):341-356.

28. Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., and Kolehmainen, M. (2004). Methods for imputation of missing values in air quality data sets. *Atmospheric Environment* 38:2895-2907.

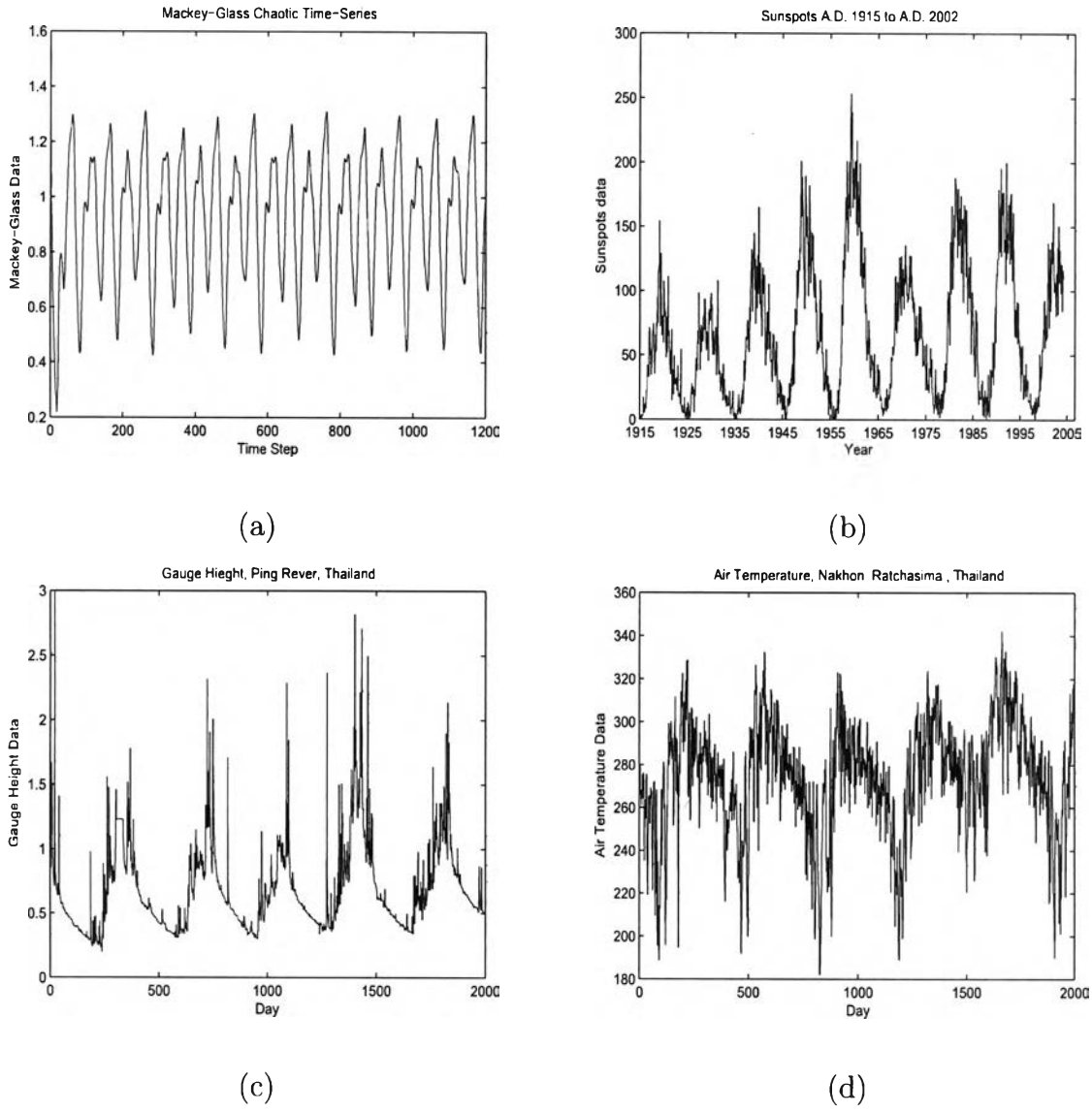


Figure 5.1: The time-series data of (a) Mackey-Glass (b) sunspots (c) gauge height data and (d) the air temperature.

Appendix

Appendix A

Mackey-Glass matlab function

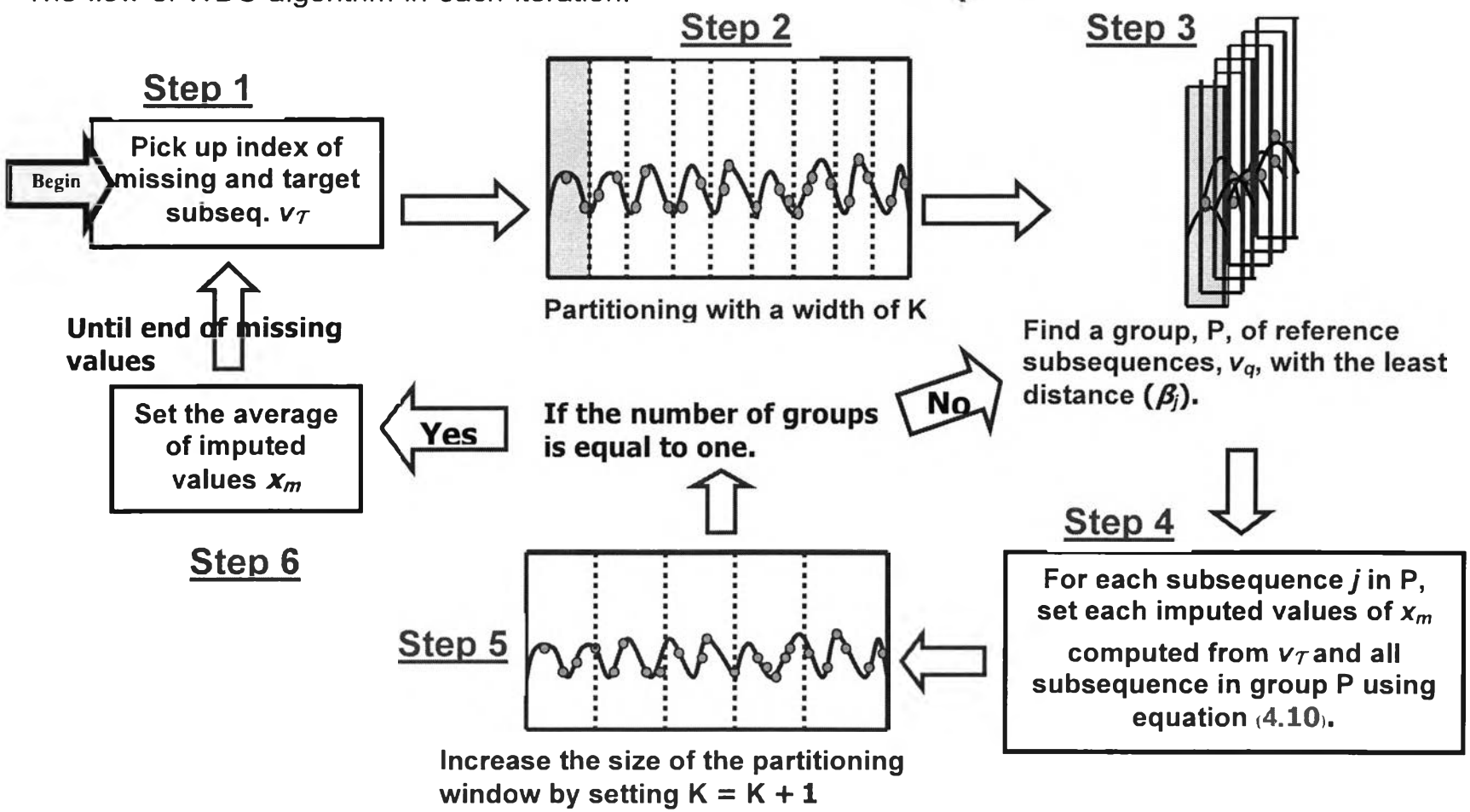
```
function x=mackeyglass(n,level,a,b,c,x0)
%Syntax: x=mackeyglass(n,level,a,b,c,x0)
%_____________________________________
%
% Simulation of the discretized variant of the Mackey-Glass PDE.
%    x(i+1)=x(i)+ax(i-s)/(1+x(i-s)^c)-bx(i)
%
% x is the simulated time series.
% n is the number of the simulated points.
% level is the noise standard deviation divided by the standard
deviation of the
%   noise-free time series. We assume Gaussian noise with zero mean.
% a, b, c, and s are the parameter
% x0 is the initial values vector for x.
%
% Note:
% s=length(x0)
%
%
% Reference:
%
% Mackey M C, Glass L (1977): Oscillation and Chaos in Physiological
% Control Systems. Science 177: 287-289
%
%
% Alexandros Leontitsis
% Department of Education
% University of Ioannina
% 45110 - Dourouti
% Ioannina
% Greece
%
% University e-mail: me00743@cc.uoi.gr
% Lifetime e-mail: leoaleq@yahoo.com
% Homepage: http://www.geocities.com/CapeCanaveral/Lab/1421
%
% 16 Nov 2001

if nargin<1 | isempty(n)==1
    n=500;
else
    % n must be scalar
    if sum(size(n))>2
        error('n must be scalar.');
    end
    % n must be positive
    if n<0
        error('n must be positive.');
    end
    % n must be an integer
    if round(n)-n~=0
        error('n must be an integer.');
```

```
    end
end

if nargin<2 | isempty(level)==1
    level=0;
else
    % level must be scalar
    if sum(size(level))>2
        error('level must be scalar.');
    end
    % level must be positive
    if level<0
        error('level must be positive.');
    end
end

if nargin<3 | isempty(a)==1
    a=0.2;
else
    % a must be scalar
    if sum(size(a))>2
        error('a must be scalar.');
    end
end

if nargin<4 | isempty(b)==1
    b=0.1;
else
    % b must be scalar
    if sum(size(b))>2
        error('b must be scalar.');
    end
end

if nargin<5 | isempty(c)==1
    c=10;
else
    % c must be scalar
    if sum(size(c))>2
        error('c must be scalar.');
    end
end

if nargin<6 | isempty(x0)==1
    x0=0.1*ones(17,1);
else
    % x0 must be either a scalar or a vector
    if max(size(x0))>2
        error('x0 must be either a scalar or a vector.');
    end
end

s=length(x0);
% n must be greater than or equal to s=length(x0)
if n<s
    error('n must be greater than or equal to s=length(x0).');
```

```
end


% Initialize
x(1,1)=x0(s)+a*x0(1)/(1+x0(1)^c)-b*x0(s);
for i=2:length(x0)
    x(i,1)=x(i-1)+a*x0(i)/(1+x0(i)^c)-b*x(i-1);
end

% Simulate
for i=s+1:n
    x(i,1)=x(i-1)+a*x(i-s)/(1+x(i-s)^c)-b*x(i-1);
end

% Add normal white noise
x=x+randn(n,1)*level*std(x);
```

Appendix B

The flow of WDC algorithm in each iteration.



Vita

Mrs. Sirapat Chiewchanwattana was born in October 20, 1962. She received bachelor degree in Statistics from department of Statistics, faculty of Science, Khon Kaen university in 1984. She received master degree in Computer Science from the National Institute of Development Administration, Thailand in 1993. She received a scholarship by The Ministry of University Affairs of Thailand. For pursing the doctorate. Now, she works as a lecturer at department of Computer Science, faculty of Science, Khon Kaen university.