Full Length Research Paper

Chaotic time series prediction and Mackey-Glass simulation with fuzzy logic

Yasmin Zahra Jafri¹, Amir Waseem²*, Lalarukh Kamal³, S. Masood Raza⁴, Muhammad Sami⁵ and Syed Haider Shah¹

 ¹Departments of Statistics, University of Balochistan, Quetta, Pakistan.
 ²Department of Bio Sciences, COMSATS Institute of Information Technology, Islamabad, Pakistan.
 ³Departments of Mathematics, University of Balochistan, Quetta, Pakistan.
 ⁴Department of Physics, Federal Urdu University of Arts, Science and Technology, Gulshan-e-Iqbal Campus, Karachi, Pakistan.
 ⁵Emerging Sector Wing, Ministry of Production, Islamabad, Pakistan.

Accepted 10 April, 2012

The present study was performed with fuzzy logic (FL) time series prediction modeling on a twenty years hourly averaged wind data, that is, 1985 to 2004 for Quetta, Pakistan. A free fuzzy logic design was followed and hourly wind data for spring prediction were obtained (February, March and April). It was found that the prediction is reliable and precise. Non-stationarity or random walk in wind data exists but it does not influence prediction. Mackey-Glass (MG) simulation of wind data indicated chaos or non periodicity in time series. Moreover, stable attractors were observed in MG-time series, in which the origin is yet unknown. The attractors seen in MG simulation do not influence FL time series prediction.

Key words: Fuzzy logic, artificial neural networks, antecedents, the adaptive neural fuzzy inference system, autoregressive integrated moving average.

INTRODUCTION

The original fuzzy logic pioneered by Zadeh (1965) has been around forty years, and yet it is unable to handle uncertainties. Zadeh introduced the concept of a fuzzy set, a set whose boundary is not sharp or precise. This concept contrasts with the classical concept of a set recently called a crisp set, whose boundary is required to be precise. Probability and fuzzy sets describe different kind of uncertainty. The probability is the theory of sets with random elements. It deals with the likelihood of relevant events or with the expectation of a future event based on something now known (outcome of a random event) while the fuzziness is not the uncertainty expectation. Fuzzy set theory, on the other hand, is not concerned with events. It is concerned with concepts. Rule based fuzzy logic system (FLS), is a powerful design methodology used to minimize the effect of uncertainty (Mendel, 2001; Jafri et al., 2012a, b, c). Model free designs are artificial neural networks (ANN) and fuzzy logic. The fuzzy logic rules are extracted from numerical data and are then combined with linguistic knowledge. The richness of fuzzy logic is that, there are enormous members of possibilities that lead to lot of nonlinear mappings of an input data vector into a scalar output. In model free approaches, the associated model is a representation of architecture to solve a specific problem. With model approach in fuzzy logic, one can endeavor the truth or close approximation theory. FLSs employ 500 rules for one pass and sixteen rules for back propagation steepest descent method of designs, respectively. We followed a model free approach, that is, fuzzy logic on hourly wind speed data to predict future values, that is, consequents from antecedents (past values). A single stage forecasting for a chaotic time series wind data will be used.

^{*}Corresponding author. E-mail: waseemq2000@hotmail.com.

THEORY

Let s(t) be a time series, where t=1,2,...,N; for hourly wind data of Quetta, Pakistan. Measured values of s(t) are denoted by x(t),

 $\mathbf{x}(t) = \mathbf{s}(t) + \mathbf{n}(t)$

Where x(t) denotes measurement error-noise.

Given a window of p past measurements of s(t), namely x(t-p+1), x(t-p+2),...,x(t), to determine an estimate of a future value of s, that is, s(t+l) where p and l are fixed positive integers. For noise free measurements, x(t-p+1), x(t-p+2),...,x(t) are replaced by s(t-p+1), s(tp+2),...,s(t). When l=1, we obtain the single stage forecaster of s. There are l-stage forecasters in time series. We used a single stage forecaster of fuzzy logic.

Suppose we are given a collection of N data points, x(1), x(2),...x(N). Then as it is commonly done when neural networks are used to forecast a time series, we shall partition this data set into two subset: a training data and a testing data subset with N-D data points, x(D+1), x(D+2),...,x(N). Because we will use a window of p data points to forecast the next data point, there are at most D-p training pairs, $x^{(1)}$, $x^{(2)}$,..., $x^{(D-P)}$, where:

$$x^{(1)} = [x(1), x(2), ..., x(p), x(p+1)]^T$$

 $x^{(2)} = [x(2), x(3), ...x(p+1), x(p+2)]^T$

$$x^{(D-p)} = [x(D-p), x(D-p+1), ..., x(D-1), x(D)]^{T}$$
 (1)

In Equation 1, the first p elements of $x^{(t)}$ are the inputs to the forecaster and the last element of $x^{(t)}$ is the desired output of the forecaster, that is:

-

where t =1,2,...D-p and T is over the complete time series.

The training data are used in a fuzzy logic system by forecaster to establish its rule. There are at least three ways to extracting rules from the numerical training data:

1. Let the data establish the centers of the fuzzy sets that appear in the antecedent and consequents of the rules. For single stage forecasting, there are D-p rules that one can extract from the (D-p) training pairs, $x^{(1)}$, $x^{(2)}$,..., $x^{(D-p)}$ (Mendel 1995).

2. Pre-specify fuzzy sets for the antecedents and consequents and then associate the data with these fuzzy sets. Fuzzy rules are generated from the given data pairs (Wang and Mendel, 1992). Initially, fuzzy sets are established for all the antecedents and the consequents (Wang, 1994).

3. Establish the architecture for a FLS and use the data to optimize its parameters. In this approach, we fix the architecture of the FLS ahead of time; we fix the number of rules, the number of antecedents, the shapes of the antecedents and consequents membership functions, the inference method and the kind of fuzzification.

Since we are dealing with chaotic time series, therefore, it is imperative to talk on defuzzifier chaos simulation. Chaos now-a-days is having an impact on diverse disciplines of knowledge including physics, biology, chemistry, economics and medicine (Farmer, 1982; Rasband, 1990). Very briefly, chaotic behavior can be described as bounded fluctuations of the output of a nonlinear system with high degree of sensitivity to initial conditions (Casdogli and Stephen, 1992). The observation in a system exhibiting chaotic dynamics appears to be uncorrelated, thus making forecasts difficult. A nonlinear delay differential equation developed by Mackey and Glass (1977) of the form:

$$\frac{ds}{dt} = \frac{0.2 \ s(t-T)}{1+s^{10}(t-T)} - 0.1s(t)$$
(3)

can be used for time-series prediction in a deterministic manner.

In Equation 3, for $\tau > 17$ is known to exhibit chaos in both neural networks and fuzzy logic (Wang, 1994; Casdogli and Stephen, 1992; Farmer, 1982; Rasband, 1990; Mackey and Glass, 1977). When $\tau < 17$ in Equation 3, we observe non chaotic, that is, periodic time series (Mendel, 1965). Equation 3 can be converted to a discrete time series by using Euler's method with a step size equal to 1.

RESULTS AND DISCUSSION

We used mean hourly wind data of the years 1985 to 2004 to gain more reliability about model validation and indeed better prediction. We follow the adaptive neural fuzzy inference system (ANFIS) which constructs a fuzzy inference system for input/output data set and whose membership function parameters are tuned using either back propagation algorithms alone or in combination with least squares estimation method. This is how fuzzy inference system (FIS) structure and parameter adjustment were accomplished. The parameters are associated with membership functions and will change during the course of learning from the data that are to be modeled.

The model validation is dependent on trained FIS model which predicts the corresponding data set output values from the input vectors. We considered 600 hourly wind speed data for trainee and the remaining 120 or 144 for checking. With FIS modeling we checked the trainee data and found its learning fit with checking data sets to adjust the membership functions. We also viewed FIS structure after adjusting membership functions for our

Epoch	Month	Epoch error	Average checking error on FIS output	FIS Data structure
5 10	January	0.55325 0.54888	0.51091	Same for all months
5 10	February	0.56150 0.55817	13.242	-
5 10	March	1.11460 1.10300	6.6941	-
5 10	April	1.01030 1.00770	5.1694	-
5 10	Мау	1.42290 1.41970	4.7963	-
5 10	June	1.54530 1.51620	1.6881	-
5 10	July	1.7249 1.7063	11.9788	-
5 10	August	1.2153 1.2123	2.0175	-
5 10	September	0.79920 0.79590	2.6837	-
5 10	October	0.78743 0.77950	1.0969	-
5 10	November	1.11760 1.10920	5.9558	-
5 10	December	0.24367 0.24215	0.82591	-

 Table 1. Shows model validation for 5 epochs, four input membership functions and for FLS in case of hourly averaged wind data for the period of 1985 to 2004 over months. No of epochs do not influence FIS output or its associated testing error.

monthly data which looked same for all months (Jafri, 2008). We then performed ANFIS training and test the data against the trained FIS. The validation of model is true only if the beginning gap between the training and checking error would have a minimum value; checking error data would decrease in the initial epoch and would go on decreasing even for larger epoch along with the training error. When the checking error goes down to a minimum value at a certain epoch and rises suddenly from its minimum value, we get model over fitting. Model over fitting is avoided as much as possible. Tables 1 and 2 show model validations for months and seasons respectively, which are found compromising and suitable for five epochs and four membership function for FLS. The number of epochs do not influence FIS output or its associated checking error. The checking data does not always validate FIS model. In such a case, we train the trained FIS model against the checking data.

We have been able to set indices in MATLAB 5.3 for our wind data to execute prediction of hourly wind speed data for the period of 1985 to 2004 with window interval of 6 and with four antecedent membership functions, that is:

for t = 59:658
Data (t-58,:) = [x(t-18) x(t-12) x(t-6) x(t) x(t+6)];
end
$$(4)$$

Where x(t+6) is the predicted value of wind speed and x(t+6) is also referred to as consequent. We obtained forecast for wind speed ahead of 6 h for each hours of

S #	Season	Epoch	Epoch error	Average checking error from FIS output	FIS Data structure
4	Winter	5	0.6109	0.02440	
1	(November, December, January)	10	0.5966	0.63416	Same for all seasons
2	Spring	5	1.3675	2.0746	
Ζ	(February, March, April)	10	1.3495	2.0746	-
0	Summer	5	0.79207	0,7000	
3	(May, June, July)	10	0.7879	2.7368	-
4	Autumn(fall)	5	0.4847	2 1022	
	(August, September, October)	10	0.4797	2.1922	-

Table 2. Shows model validation for 5 and 10 epochs respectively, four input membership functions and for FLS in all seasons.

Table 3. Predicted values of averaged wind data for the month of February (1985 to 2004).

S #	x(t-18)	x(t-12)	x(t-6)	x(t)	x(t+6)
1	1.0288	8.2304	0	9.2592	7.2016
2	0	8.2304	9.2592	9.2592	8.2704
3	0	8.2304	8.2304	9.2592	4.1152
4	0	6.1728	9.2592	8.2304	3.0864
5	1.0288	2.0576	7.2016	7.2016	2.0576
6	1.0288	3.0864	7.2016	8.2304	3.0864
7	8.2304	0	9.2592	7.2016	2.0576
8	8.2304	9.2592	9.2592	8.2304	2.0576
9	8.2304	8.2304	9.2592	4.1152	9.2592
10	6.1728	9.2592	8.2304	3.0864	7.2016
11	2.0576	7.2016	7.2016	2.0576	2.0576
12	3.0864	7.2016	8.2304	3.0864	3.0864
13	0	9.2592	7.2016	2.0576	2.0576
14	9.2592	9.2592	8.2304	2.0576	6.1728
15	8.2304	9.2592	4.1152.	9.2592	7.2016.
591	1.0288	9.2592	1.0288	1.0288	7.2016
592	6.1728	7.2016	2.0576	2.0576	8.2304
593	8.2304	4.1152	2.0576	3.0864	8.2304.
594	10.288	1.0288	1.0288	0	8.2304
595	9.2592	2.0576	1.0288	2.0576	8.2304
596	9.2592	2.0576	1.0288	1.0288	8.2304
597	9.2592	1.0288	1.0288	7.2016	6.1728
598	7.2016	2.0576	2.0576	8.2304	4.1152
599	4.1152	2.0576	3.0864	8.2304	2.0576
600	1.0288	1.0288	0	8.2304	3.0864

the month. These four antecedent membership functions are selected over a day with a window of six hours. The results of prediction are heavy and cumbersome (Jafri, 2008), therefore, it cannot be reproduced either in a research paper or in a thesis. Therefore, we produced time series prediction of wind data only for months of spring (February, March, April) as shown in Tables 3, 4 and 5. We must remember that the antecedent and the consequent membership functions are initially tuned and modeled through ANFIS. We also calculated root square mean error value (RMSE) for both trainee and checking data sets (Jafri, 2008).

S #	x(t-18)	x(t-12)	x(t-6)	x(t)	x(t+6)
1	1.0288	0	1.0288	4.1152	5.1440
2	2.0576	0	3.0864	5.1440	2.0576
3	2.0576	0	3.0864	7.2016	1.0288
4	1.0288	0	1.0288	8.2304	1.0288
5	0	1.0288	0	5.1440	2.0576
6	1.0288	1.0288	0	7.7160	1.0288
7	0	1.0288	4.1152	5.1440	0
8	0	3.0864	5.1440	2.0576	0
9	0	3.0864	7.2016	1.0288	0
10	0	1.0288	8.2304	1.0288	0
11	1.0288	0	5.144	2.0576	0
12	1.0288	0	7.7160	1.0288	0
13	1.0288	4.1152	5.144	0	0
14	3.0864	5.144	2.0576	0	0
15	3.0864	7.2016	1.0288	0	0
591	5.1440	9.2592	8.2304	2.0576	0
592	5.1440	9.2592	6.1728	2.0576.	0
593	4.1152	10.288	5.1440	2.0576	1.0288
594	4.1152	1.0288	1.0288	2.0576	2.0576
595	10.288	2.0576	1.0288	2.0576	3.0864
596	9.2592	2.0576	1.0288	3.0864	2.0576
597	9.2592	1.0288	1.0288	0	3.0864
598	9.2592	2.0576	2.0576	0	7.2016
599	10.288	2.0576	3.0864	1.0288	2.0576
600	9.2092	1.0288	0	2.0576	8.2304

Table 4. Predicted values of averaged wind data for the month of March (1985 to 2004).

Table 5. Predicted values of averaged wind data for the month of April (1985 to 2004).

S #	x(t-18)	x(t-12)	x(t-6)	x(t)	x(t+6)
1	8.2304	5.144	5.144	5.144	2.0576
2	6.1723	5.144	5.144	4.1152	2.0576
3	8.2304	8.1728	5.144	2.5720	10.200
4	8.2304	2.0576	5.1440	2.5720	9.2592
5	6.1728	3.0864	6.1728	2.5720	2.0576
6	5.1440	3.0864	5.144	2.5720	3.0864
7	5.1440	5.1440	6.1728	2.5720	3.0864
8	5.1443	5.1440	8.2304	2.0576	3.0864
9	6.1728	5.1440	5.1440	2.5720	3.0864
10	2.0576	5.1440	2.0576	10.288	5.1440
11	3.0864	6.1728	.2.0576	9.2592	5.144
12	3.0864	8.2304	2.5720	3.0864	5.144
13	5.144	5.144	2.5720	3.0864	9.2592
14	5.144	4.1152	2.5720	3.0864	10.288
15	5.144	2.5720	8.0576	3.0864	8.2304
•					

590	0	6.1728	3.0864	1.0288	2.0576
591	1.0288	5.1440	2.0576	0	0
592	1.0288	4.1152	2.0576	2.0576	0
593	1.0288	2.0576	2.0576	0	0
594	2.0576	3.0864	1.0288	0	2.0576
595	5.1440	2.0576	1.0288	0	6.1728
596	6.1728	2.0576	1.0288	2.0576	0
597	5.1440	2.0576	0	0	1.0288
598	4.1152	1.0288	2.0576	0	1.0288
599	2.0576	1.0288	0	0	1.0288
600	3.0864	1.0288	0	2.0576	5.144

Table 5. Cont.d

We performed chaotic time series prediction on wind data using Equation 3. To obtain the time series values over the hours of the month, we applied a Runge-Kutta method to find the numerical solution of Mackey-Glass equation; the result we saved in the file mgdata.dat. Here we assume x(0) = 1.2 for $\tau = 17$ and x(t) = 0 for t<0. Surprisingly enough, our data exhibited chaos even for τ = 17. To plot the MG-time series we type

$$\left.\begin{array}{c} \text{load mgdata .dat} \\ t = \text{mgdata}(:,1); x = \text{mgdata}(:,2); \text{plot}(t,x); \\ \text{end} \end{array}\right\}_{(5)}$$
$$trnData = Data(1:600, :); \\ \text{chkData = Data(601:end,:);} \end{array}\right\}_{(6)}$$

To start the training, we need FIS structure that specifies the structure and initial parameters of the FIS for learning. This is the task of genfis 1.

fismat =genfis1(trn Data);

The generated FIS structure contains fuzzy rules with 104 parameters. In order to get good generalization capability, the ratio between trainee data sets and parameters is about five, that is, in our case it is about 600/120.

Figures 1, 2, 3, 4, 5, 5, 6, 7, 8, 9, 10, 11 and 12 show MG-simulation of wind data. We observe in these figures line or loops and triangular chaotic attractors. These chaotic attractors to our knowledge have not been seen in MG-simulation in any literature. The chaos or non periodicity in wind speed time series data is evident. We checked the position of attractors on time axis from where it took the start and then compared the nearby consequent or predicted values. We found that the predicted values were not affected. It appears as if chaotic attractors are inherited characteristics of wind speed data.

In addition to chaotic attractors, there is a chaos in the data which supposedly to our conjecture is nonstationary. We confirmed the non-stationarity of wind speed data by determining the standard deviation of every hour over the days of the month (Jafri, 2008). We know that a pure integrated (I) process is a random walk. Integrated process remembers where it was and then moves at random (Anderson et al., 2011). A random walk is said to be a non-stationary process because over time it tends to move further and further away from where it was. In contrast, the autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) model each, represents a stationary process because they tend to behave similarly over long periods, staying relatively close to their long time periods and staying relatively close to their long run mean. The autoregressive integrated moving average (ARIMA) process remembers its changes. The Box-Jenkins ARIMA processes form a family of linear statistical model based on the normal distribution that have the flexibility to imitate many different real time series by combining AR process, I-process and MA-process. The result is a parsimonious model. Stochastic simulation and time series models were studied (Sami et al., 2012; Jafri et al., 2012a, b, c) and developed to forecast synthetic sequence of wind speed and global solar radiations, respectively (Kamal and Jafri, 1997). The daytime and night time sleep patterns of a new born baby defined a random biological process with long range power law correlation (Canessa and Calmetta, 1994). The chaos may behave almost linearly in some part of the phase space and highly non-linearly in other parts. We are familiar with approximately two chaotic non-linear dynamical systems (Kosko, 1997). But the recognition of attractors in a chaotic system is indeed difficult. The behavior of chaos in wind speed time series data due to presence of attractors seems to be almost non-linear. The use of neural networks and the practical considerations in controlling chaos (Alsing et al., 1994; Bayly and Virgin, 1994) are based on the observation that a chaotic attractor has embedded within it an infinite



Figure 1. MG-simulation for January (1985 to 2004).



Figure 2. MG-simulation for February (1985 to 2004).



Figure 3. MG-simulation for March (1985 to 2004).



Figure 4. MG-simulation for April (1985 to 2004).



Figure 5. MG-simulation for May (1985 to 2004).



Figure 6. MG-simulation for June (1985 to 2004).



Figure 7. MG-simulation for July (1985 to 2004).



Figure 8. MG-simulation for August (1985 to 2004).



Figure 9. MG-simulation for September (1985 to 2004).



Figure 10. MG-simulation for October (1985 to 2004).



Figure 11. MG-simulation for November (1985 to 2004).



Figure 12. MG-simulation for December (1985 to 2004).

number of unstable periodic orbits. But we did not observe periodic orbits in the attractors. It may happen that attractors which we have seen are stable. Recursive partitioning algorithms feed forward back propagating neural network and small perturbations of a control parameter can generate accurate models for chaotic systems. Rules explosion is the biggest problems in FLSs. Therefore, the feedback fuzzy system may offer a more efficient way to approximate dynamical systems with a fixed and a small number of rules. An open research problem is to find tractable learning laws for controlling and tuning such FLSs. Moreover, chaotic attractors should be viewed in terms of long range time series power law correlation or with cascade correlation algorithms.

CONCLUSIONS

We infer from this study the following conclusions:

1- The chaotic time series prediction with fuzzy logic is reliable and precise.

2- Non-stationarity in wind data exists. The nonstationarity in wind data does not influence FL prediction.

3- The line/loop and triangular attractors are seen in Mackey-Glass simulation for chaotic time series data. But these attractors do not affect FL prediction.

4- The recognition of attractors whether stable or unstable in a non linear dynamical system such as occurrence of wind and its speed is essential to deciphering an appropriate time series statistical correlation. This will provide a rationale solution to learning procedures for tuning and controlling fuzzy logic.

REFERENCES

- Alsing PM, Gavrielides A, Kovanis V (1994). Using neural networks for controlling chaos, Phys. Rev. E., 49(2): 1225-1231.
- Anderson DR, Sweeney DJ, Williams TA (2011). Statistics for Business and Econs, 11th Edition, South-western Cengage Learning, USA.
- Bayly PV, Virgin NL (1994). Practical considerations in the control of chaos. Phys. Rev. E., 50(1): 604 -607.
- Canessa E, Calmetta A (1994). Physics of a random biological process. Phys. Rev. E., 50(1): R47-R49.
- Casdogli M, Stephen E (1992). Nonlinear modeling and forecasting, Santa Fe Institute Series, Addison-Wesley Pub. Co, USA.
- Farmer JD (1982). Chaotic attractors of infinite dimensional dynamical systems. Physica, 4D: 366-393.
- Jafri YZ (2008). Recent trends in time series modeling, simulation and prediction: Statistical and fuzzy logic approach Ph.D. Thesis, University of Balochistan Quetta Pakistan.
- Jafri YZ, Kamal L, Waseem A, Raza SM (2012b). Rule Based Fuzzy Logic (FL) Time Series Prediction, Int. J. Phys. Sci., 7(11): 1862-1873.
- Jafri YZ, Sami M, Waseem A, Murtaza G, Akbar S (2012a). Stochastic approaches for time series forecasting of rate of dust fall: A case study of North West of Balochistan, Pakistan. Int. J. Phys. Sci., 7(4): 676–686.

- Jafri YZ, Waseem A, Kamal L, Raza SM, Sami M (2012c). Recent trends on artificial neural networks for prediction of wind energy. Sci. Res. Essays, 7(14): 1521-1526.
- Kamal L, Jafri YZ (1997). Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. Solar Energy, 61(1): 23-32.
- Kamal L, Jafri YZ (1999). Stochastic modelling and generation of synthetic sequences of hourly global solar irradiation at Quetta, Pakistan. Renewable Energy, 18: 565-572.
- Kosko B (1997). Fuzzy Engineering. 1st Ed. Prentice Hall, Inc., New Jersey, USA.
- Mackey MC, Glass L (1977). Oscillation and chaos in physiological system. Science, 197: 287-289.
- Mendel JM (1995). Fuzzy logic system for engineering: A tutorial IEEE Proc., 83: 345-377.
- Mendel JM (1995). Lessons in estimation theory for signal processing, communication and control Second edition Prentice Hall, USA.
- Mendel JM (2001). Uncertain rule based fuzzy logic system, Introduction and new direction. Prentice Hall PTR. Inc. NJ. USA.
- Rasband SN (1990). Chaotic dynamics of nonlinear systems Wiley NY.,
- Sami M, Waseem A, Jafri YZ, Shah SH, Khan MA, Akbar S, Siddiqui MA, Murtaza G (2012). Prediction of the rate of dust fall in Quetta city by using seasonal ARIMA (SARIMA) modeling. Int. J. Phys. Sci. 7(10): 1713 – 1725.
- Wang LX (1994). Adaptive Fuzzy System and control: Design and stability analysis, PTR Prentice Hall, Englewood Cliffs, NJ, USA.
- Wang LX, Mendel JM (1992). Generating fuzzy rules by learning from examples. IEEE Transactions on systems man and cybernetics. 22(6): 1414-1427.
- Zadeh LA (1965). Fuzzy sets. Inf. control, 8: 338-353.