DEVELOPMENT OF ANFIS MODEL WITH OPTIMISED INPUTS TO REDUCE THE COMPUTATIONAL COST AND TIME FOR GROUND LEVEL OZONE FORECASTING

Amit Kumar Gorai
Department of Mining Engineering, National Institute of Technology, Rourkela, Odisha 769008, India
Presenting author: Email: amit_gorai@yahoo.co.uk

Summary

This study aims to develop adaptive neuro-fuzzy inference system (ANFIS) for forecasting of daily ozone (O_3) concentrations in the atmosphere of a mega city. The ANFIS model predictor considers the value of seven meteorological factors (pressure, temperature, relative humidity, dew point, visibility, wind speed, and precipitation), NO_2 concentration, and previous day's ozone concentration in different combinations as the inputs to predict the 1-day advance and same day ozone concentration. Collinearity tests were conducted to eliminate the redundant input variables. A forward selection (FS) method is used for selecting the different subsets of input variables. The method reduces the computational cost and time. The performances of the models were evaluated on the basis of four statistical indices [(coefficient of determination (R^2), normalized mean square error (NMSE), index of agreement (IOA), and fractional bias (FB)].

Introduction

Environmental data are typically very complex to model due to the underlying correlation among several variables of different type which yields an intricate mesh of relationships (Marino et al, 2001). Standard statistical techniques may fail to adequately model complex non-linear phenomena (Moussiopoulos et al., 1995). In contrast, expert knowledge is becoming widely used because they showed ability to model non-linear data and their non-reliance on previously assumed equations (Marino et al., 2001).

Methodology and Results

A flowchart of ANFIS model development is presented is Fig. 1. The multi-collinearity test result shown in Table 1 clearly indicates that the VIF values exceeded the recommended value (10) for both the temperature and humidity before removing one of the variables. Also, the tolerance level of temperature and humidity are below the recommended limit (0.2) in each case. After removing of redundant variables, the VIF values and tolerance level are well within the recommended limit for each variable. Applying FS algorithm, all the explanatory variables (input variables) are ordered according to their correlation with the dependent variable (from the most to the least correlated variable). The statistical indices shown in Table 2 indicate that there is no significant influence of the number of input variables on different statistical indices (IOA, FB, NMSE, and R^2). There was a very slight change observed in the output in different input conditions in both the cases (same day and 1-day advance forecasting). Thus, the best model can be selected on the basis of training error and testing error levels. M4 showed the transition or stable input conditions where the error level is stable and thus suitable for forecasting the ozone concentration.

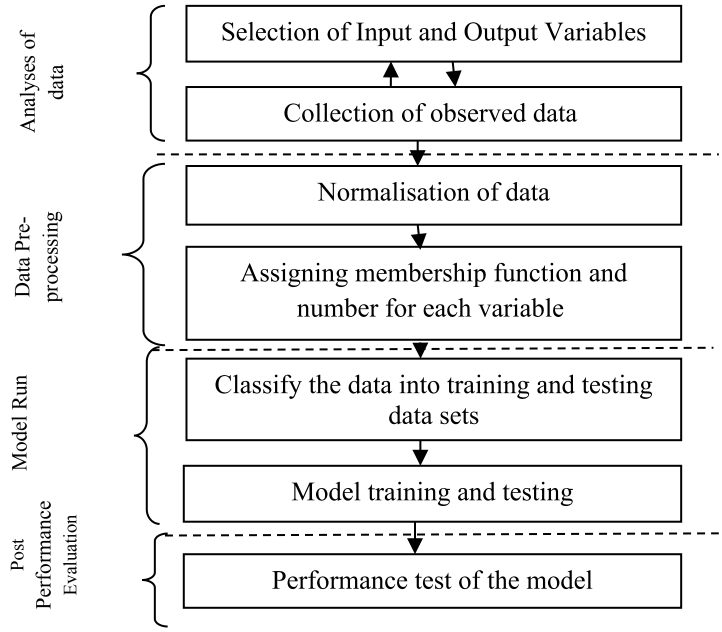

Fig. 1: Flowchart of ANFIS model development

Table 1: Collinearity Statistics

Before Removal of variables			**After removal of variables**		
Variables	**Signif.**	**VIF**	**Variables**	**Signif.**	**VIF**
P- O_3	0.847	1.180	P-O_3	0.861	1.161
NO_2	0.363	2.753	NO_2	0.365	2.740
PRESS	0.404	2.472	PRESS	0.415	2.411
RH	0.015	65.55	RH	0.486	2.058
TEMP	0.005	197.35	TEMP	0.329	3.036
DP	0.003	299.23	VISI	0.717	1.395
VISI	0.711	1.407	WS	0.856	1.169
WS	0.853	1.173	PRECI	0.809	1.237
PRECI	0.764	1.309			

Note: Variables Information
PRESS- Pressure, TEMP- Temperature, RH-Relative Humidity, DP-Dew point, VISI-Visibility, WS-Wind speed, PRECI-Precipitation, P-O_3- Previous day Ozone.

Table 2: ANFIS model results

Model	R^2	NMSE	FB	IOA	R^2	NMSE	FB	IOA
	Same day forecasting				1-day advance forecasting			
M1	0.80	0.02	0.014	0.95	0.68	0.08	0.012	0.91
M2	0.80	0.01	0.013	0.95	0.68	0.09	0.011	0.92
M3	0.81	0.01	0.014	0.95	0.68	0.07	0.010	0.91
M4	0.79	0.05	0.020	0.95	0.72	0.07	0.015	0.92
M5	0.82	0.05	0.017	0.95	0.73	0.07	0.019	0.92
M6	0.82	0.05	0.017	0.95	0.72	0.08	0.017	0.92
M7	0.82	0.05	0.016	0.95	0.72	0.08	0.014	0.92
M8	0.82	0.05	0.006	0.95	0.72	0.08	0.007	0.92

Note: Input combination of models
M1: P-O_3, WS, Visi, RH, Temp, Preci, Press, NO_2; **M2**: P-O_3, WS, Visi, RH, Temp, Preci, Press; **M3**: P-O_3, WS, Visi, RH, Temp, Preci; **M4**: P-O_3, WS, Visi, RH, Temp; **M5**: P-O_3, WS, Visi, RH; **M6**: P-O_3,WS,Visi; **M7**: P-O_3,WS; **M8**: P-O_3

Conclusions

Neuro-fuzzy logic is a stochastic approach which is used in the present study for forecasting of ozone concentrations. It was seen that the statistical performance was good for all the models, indicating that a lesser number of input variables can be considered in order to forecast the model. The IOA ranges from 91% to 95% indicating a good model, and closer to ideal value. The present study helps in finding out the input combination which would lower the computational cost.

References

Marino, D.; Morabito, F.C. (1999) A fuzzy neural network for urban environment monitoring system: the Villa San Giovanni study case, neural nets. Procd. of the WIRN workshop 323-328.

Moussiopoulos, N.; Sahm, P.; Kessler, C.; Kunz, R. (1995) Numerical Simulation of Photochemical Smog Formation in Athens—A Case Study. Atmospheric Environment 29(24): 3619-3632.