Best Multiple Non-linear Model Factors for knock Engine (SI) by using ANFIS

[1] Azher Razzaq Hadi Witwit, [2] Azman Yasin, [3] Horizon Gitano, [4] Mohammed Ismael Mahmood

[1] University Utara Malaysia(UUM), Kedah, Malaysia
1University of Babylon, Babylon, Iraq
Corresponding author email: azherwitwit {at} yahoo.com

[2] University Utara Malaysia (UUM),
Kedah, Malaysia

[3] University Kuala Lumpur Malaysian Spanish Institute (UniKL MSI),
Kulim, Kedah, Malaysia

[4] EGMA Garage Equipment Company,
Baghdad, Iraq

ABSTRUCT -- ***Knock Prediction in vehicles is an ideal problem for non-linear regression to deal with, which use many of the factors of information to predict another factor. Training data were collected through a test engine for the Malaysian Proton company and in various states of speed. Selected six influential factors on the knocking (Throttle Position Sensor (TPS), Temperature (TEMP), Revolution Per Minute (RPM), (TORQUE), Ignition Timing (IGN), Acceleration Position (AC_POS)), has been taking data for this study and then applied to a single cylinder, output factor (output variable) to be prediction factor is a knock. We compare the performance of resultant ANFIS and Linear regression to obtain results shows effectiveness ANFIS, as well as three factors were selected from six non-linear factors to get the best model by using Adaptive Neuro-Fuzzy Inference System (ANFIS). Experiments demonstrate that although soft computing methods are somewhat of tolerant of inaccurate inputs, cleaned data results in more robust models for practical problems.***

Keywords -- Knocking, ANFIS, linear regression, Throttle position sensor (TPS), Revolution per minute (RPM).

---

1. INTRODUCTION

Knocking is a process that presents a challenge for many engineers and researchers to achieve the characteristics of quality and to meet customer satisfaction through the achievement of efficiency in the engine. Control systems are designed in modern engines to minimize exhaust emissions while maximizing power and fuel economy. The ability to maximize power and fuel economy by optimizing model contain factors like (Temp, TPS, Rpm, Torque, ….) that effects in engine knocking. There are many factors that affect the occurrence of knocking, including mechanical, electrical, environmental and misuse[1]. The factors can be clarified that influence knocking through analysis of a fishbone diagram as given below in fig. (1):

Measuring the mutual influence of the factors that affect the occurrence of knocking contributes mainly to the design of knocking control systems, in addition to measuring the proportion of influence on the knocking itself. In order to achieve an efficient control system using such a method, a suitable model must be included for the representation of influencing factors in addition to the relationship between them. Many models exist that can represent the relationship between the factors that affect knocking, which can improve the performance of control systems that make the engine work near the border of the knocking phenomenon. Soft computing techniques are known for their efficiency in dealing with complicated problems when conventional analytical methods are infeasible or too expensive, with only sets of operational data available. Some experimental studies have shown that knock intensity does not correlate with most individual cycle parameters like knock occurrence crank angle, burn rate, peak pressure or unburned mass fraction[2]. However, Konig and Sheppard [3] found knock severity to correlate with knock occurrence crank angle. Other experimental studies [4], [5], [6], [7] have characterized knock intensity by analyzing experimental pressure signals, in particular the pressure oscillations.

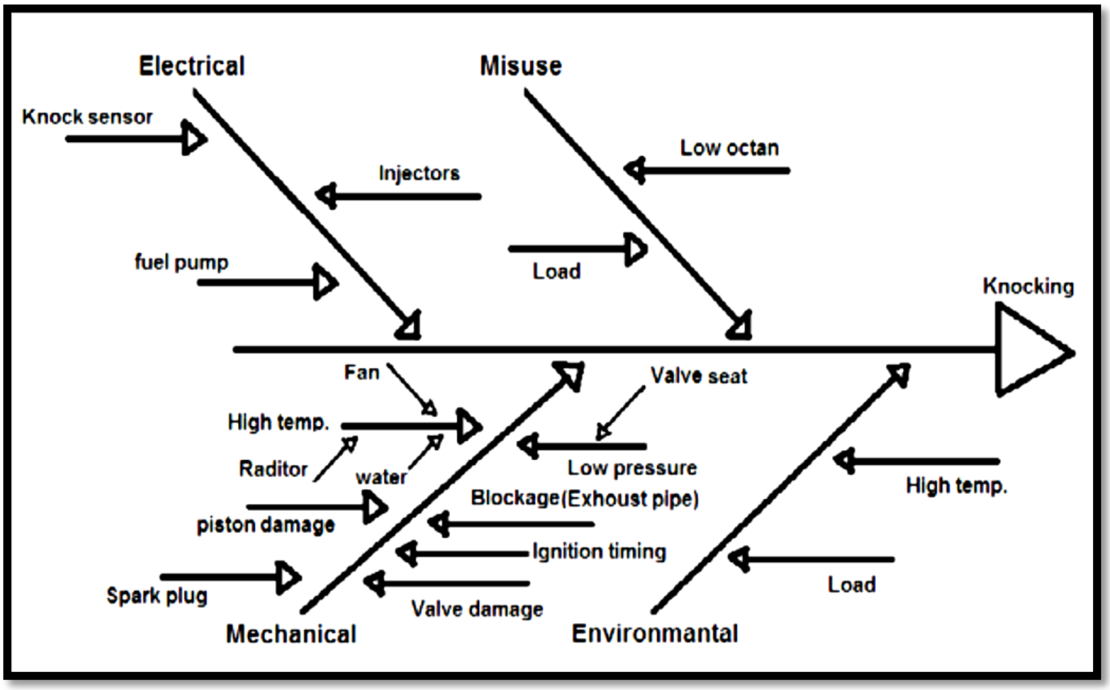


Figure (1): Knock problem

Many researchers adopted the multiple techniques for forecasting results are good, such that was based on both[8] , on curve fitting technology and linear regression method in power Systems load predicted seasonal near-term . The researchers adopted in the analysis of data and results on three factors are temperature, humidity, type of day, where they was studying network load sensitivity to temperature and humidity and the type of day using the concept curve fitting. There was also some researchers[9], deals with mathematical model to assess the performance of photovoltaic (PV) cells. The PV system characteristics are modeled and analyzed by using the curve fitting method referred to the different connections of PV cells and different solar irradiance.

Soft computing methods have been widely applied in many areas in the petroleum industry, such as reservoir description[10], well logging interpretation[11], production prediction[12] and treatment optimization[13].

2. MODEL VALIDATION BY ANFIS

Model validation is the process by which the input vectors from input/output data sets on which the FIS was not trained, are presented to the trained FIS model, to see how well the FIS model predicts the corresponding data set output values. You use a validation data set to check and control the potential for the model overfitting the data. When checking data is presented to ANFIS as well as training data, the FIS model is selected to have parameters associated with the minimum checking data model error. One problem with model validation for models constructed using adaptive techniques is selecting a data set that is both representative of the data the trained model is intended to emulate, yet sufficiently distinct from the training data set so as not to render the validation process trivial. If you have collected a large amount of data, hopefully this data contains all the necessary representative features, so the process of selecting a data set for checking or testing purposes is made easier. However, if you expect to be presenting noisy measurements to your model, it is possible the training data set does not include all of the representative features you want to model. The basic idea behind using a checking data set for model validation is that after a certain point in the training, the model begins overfitting the training data set. In principle, the model error for the checking data set tends to decrease as the training takes place up to the point that overfitting begins, and then the model error for the checking data suddenly increases.

3. PARTITONING DATA

Data set in table .1 were obtained from an original data file (Real data from engine). After that data is fragmented into two parts, a training data set (odd-indexed samples [TPS, RPM, IGN]) and checking data set (even-indexed samples [TEMP, TORQUE, AC_POS]).

Table 1: Real Data Sets

Input Factors						Output
'TPS'	'TEMP'	'RPM'	'TORQUE'	'IGN'	'AC_POS'	'KNOCK'
80.023	90.1	1000.5	65.8	-6.375	40.2	0.275
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
80.021	89.9	4999.5	134.3	21.000	77.2	0.854

The table shown above is several observations or samples from the real data set. The six input attributes are 'TPS', 'TEMP', 'RPM', 'TORQUE', 'IGN', 'AC_POS'. The output variable to be predicted is the 'Knocking'.

4. INPUT SELECTION

Implementation of an exhaustive search within the available input to determine a set of inputs that affect the process of knocking. First Parameter been identified is the number of combinations to be trained during the search. After that, it was built "ANFIS" model for each combination and trained for one epoch and give a report on the performance achieved (case). In the beginning, has been determining the most input factor that effect in the process to predict of the output (Knock).

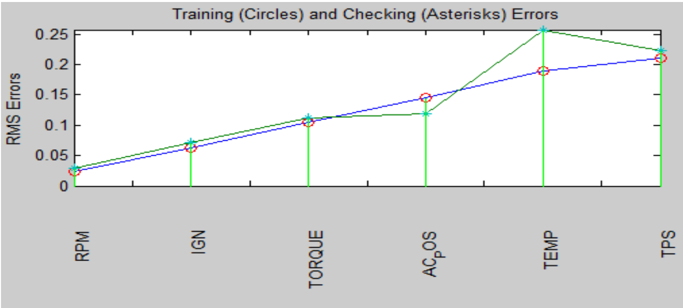


Figure 1: Every input variable's influence on knocking

Variable (factor) in the far left in the figure has less error, or in other words more important in relation to outputs. From the drawing, and the results (ANFIS model 3: RPM ---> trn=0.0241 , chk=0.0298) , it is clear that factor (RPM) is the most effect, which implies that there is no overfitting. This means that we can push beyond it a little and explore whether we can identify more than one factor in building a model ANFIS.

'TPS' 'TEMP' 'RPM' 'TORQUE' 'IGN' 'AC_POS'

Train 6 ANFIS models, each with 1 input selected from 6 candidates...where:

Trn=training data error.

Chk=Checking data error.

ANFIS model 1: TPS --> trn=0.2106, chk=0.2227

ANFIS model 2: TEMP --> trn=0.1896, chk=0.2572

ANFIS model 3: RPM --> trn=0.0241, chk=0.0298

ANFIS model 4: TORQUE --> trn=0.1047, chk=0.1112

ANFIS model 5: IGN --> trn=0.0627, chk=0.0717

ANFIS model 6: AC_POS --> trn=0.1456, chk=0.1196

Intuitively, we can simply select 'RPM' and 'IGN' directly since they have the least errors. This will not necessarily be the optimal combination of two input factors that lead to a minimum of training errors. Then searching for the perfect gathering for two of the input factors and the results were as follows:

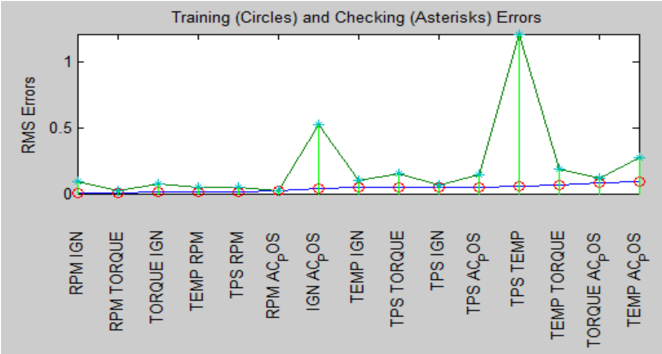


Figure 2: All two input variable combinations and their influence on knocking

Train 15 ANFIS models, each with 2 inputs selected from 6 candidates...

ANFIS model 1: TPS TEMP --> trn=0.0538, chk=1.2138

ANFIS model 2: TPS RPM --> trn=0.0155, chk=0.0461

ANFIS model 3: TPS TORQUE --> trn=0.0458, chk=0.1538

ANFIS model 4: TPS IGN --> trn=0.0474, chk=0.0684

ANFIS model 5: TPS AC_POS --> trn=0.0513, chk=0.1424

ANFIS model 6: TEMP RPM --> trn=0.0113, chk=0.0478

ANFIS model 7: TEMP TORQUE --> trn=0.0617, chk=0.1849

ANFIS model 8: TEMP IGN --> trn=0.0453, chk=0.0984

ANFIS model 9: TEMP AC_POS --> trn=0.0924, chk=0.2690

ANFIS model 10: RPM TORQUE --> trn=0.0081, chk=0.0255

ANFIS model 11: RPM IGN --> trn=0.0038, chk=0.0875

ANFIS model 12: RPM AC_POS --> trn=0.0195, chk=0.0210

ANFIS model 13: TORQUE IGN --> trn=0.0099, chk=0.0719

ANFIS model 14: TORQUE AC_POS --> trn=0.0782, chk=0.1146

ANFIS model 15: IGN AC_POS --> trn=0.0432, chk=0.5251

Results of the figure indicates that the model “RPM” and “Torque” combination is ideal for two of the factors affecting the knocking. Training and checking errors were obtained distinction between them, and therefore there is a reference to the beginning of overfitting. It may not be wise to use more than two input to build a model ANFIS, this hypothesis was tested to validate it.

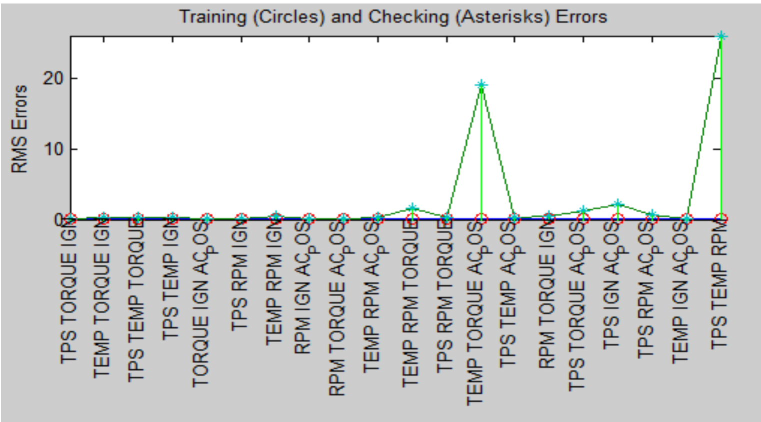


Figure 3: All three input variable combinations and their influence on knocking

Figure shows the test result of three factors enter (TPS, RPM and IGN) (ANFIS model 6: TPS RPM IGN --> trn=0.0000, chk=0.0619) as the best combination of three influential factors on the knocking. However, probably the minimum error of the training and checking does not differ greatly from the best model for the influential factors in the knocking, which shows that the factor newly added does not improve the prediction much. In general, we prefer the model with a simple structure, so stick to two factors in the model ANFIS for further exploration.

Train 20 ANFIS models, each with 3 inputs selected from 6 candidates...

ANFIS model 1: TPS TEMP RPM --> trn=0.1399, chk=25.9267

ANFIS model 2: TPS TEMP TORQUE --> trn=0.0000, chk=0.3306

ANFIS model 3: TPS TEMP IGN --> trn=0.0000, chk=0.3497

ANFIS model 4: TPS TEMP AC_POS --> trn=0.0002, chk=0.2610

ANFIS model 5: TPS RPM TORQUE --> trn=0.0001, chk=0.2223

ANFIS model 6: TPS RPM IGN --> trn=0.0000, chk=0.0619

ANFIS model 7: TPS RPM AC_POS --> trn=0.0012, chk=0.6813

ANFIS model 8: TPS TORQUE IGN --> trn=0.0000, chk=0.0909

ANFIS model 9: TPS TORQUE AC_POS --> trn=0.0005, chk=1.1865

ANFIS model 10: TPS IGN AC_POS --> trn=0.0005, chk=2.2198

ANFIS model 11: TEMP RPM TORQUE --> trn=0.0001, chk=1.5055

ANFIS model 12: TEMP RPM IGN --> trn=0.0000, chk=0.4297

ANFIS model 13: TEMP RPM AC_POS --> trn=0.0001, chk=0.2767

ANFIS model 14: TEMP TORQUE IGN --> trn=0.0000, chk=0.3673

ANFIS model 15: TEMP TORQUE AC_POS --> trn=0.0001, chk=18.9690

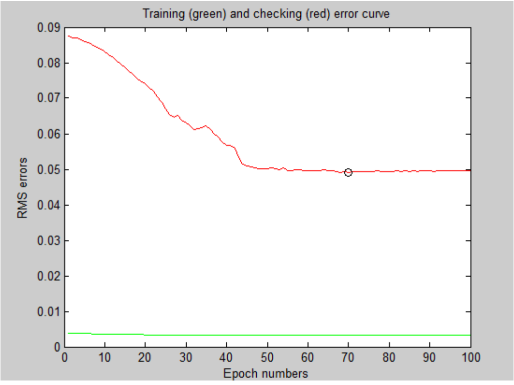
ANFIS model 16: TEMP IGN AC_POS --> trn=0.0397, chk=0.1454

ANFIS model 17: RPM TORQUE IGN --> trn=0.0003, chk=0.5385

ANFIS model 18: RPM TORQUE AC_POS --> trn=0.0000, chk=0.0313

ANFIS model 19: RPM IGN AC_POS --> trn=0.0000, chk=0.0941

ANFIS model 20: TORQUE IGN AC_POS --> trn=0.0000, chk=0.1725



ANFIS training and checking errors

The plot above shows the error curves for 100 epochs of ANFIS training. The green curve gives the training errors and the red curve gives the checking errors. The minimal checking error (0.0491) occurs at about epoch 70, which is indicated by a circle. Notice that the checking error curve tend to stability after 70 epochs, indicating that further training, that means more training is useless and there is no more generalization.

errorValue = 0.0491

epochNo = 70

5. ANFIS VS LINEAR REGRESSION

At this point would be to check the performance of the ANFIS model with a linear regression model.

The ANFIS prediction can be compared against a linear regression model by comparing their respective RMSE (Root Mean Square Error) values against checking data.

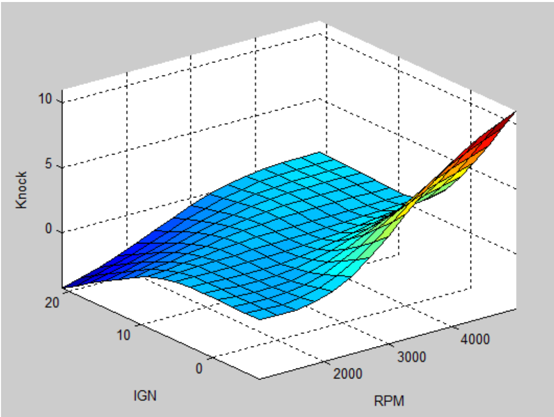
RMSE against checking data:

ANFIS: 0.049 **-** Linear Regression: 0.062

It can be seen that the ANFIS model outperforms the linear regression model.

6. ANALYSING THE ANFIS MODEL

After calculating the lowest point of error for the check data for model ANFIS through the training process, note in the input-output surface of the model in figure below:



Input-Output surface for trained FIS

The input-output surface is a nonlinear and monotonic surface, we will note the output (knocking) in the ANFIS model will respond to varying values of 'RPM' and 'IGN'. Results obtained through the application of the new model is a low level knocking with increasing Ignition timing (IGN) at the same time increasing in the Revolution Per Minute (RPM), which shows the effectiveness of the new model with non-linear behavior of the factors affecting the knock.

7. DISCUSSION AND CONCLUSIONS

The purpose of this study to improve the control systems, these systems are designed in modern engines to minimize exhaust emissions while maximizing power and fuel economy. The ability to maximize power and fuel economy by optimizing model contain factors like (Temp, TPS, Rpm, Torque, ….) that most influential in engine knocking. As well as study the behavior of the factors affecting the problem of knocking in internal combustion engine.

In order to balance the complexity of the model and the amount of error, which is obtained through the training and checking of the data was initially taking a single entry (one factor) and then calculate the error training and checking, and the results show that the RPM factor is the most effect on the output factor "Knock" where the amount of training and checking error is (trn=0.0241, chk=0.0298) (Model 3). After that has been added to the second influential factor on output "Knock" in order to find an explanatory model has a more power than the previous, and an appropriate training and checking error. The results showed that the model (ANFIS Model 10) is ideal for two factors, their impact on the Knocking, where the amount of training and testing error is (trn=0.0081, chk=0.0255), but it is through the drawing notice for overfitting the beginning, so it is possible to test more than two factors. After adding a third factor results showed that (ANFIS Model 6) is the best model for the three factors most influence on output factor “knock”. The plot above shows the error curves for 100 epochs of ANFIS training, the minimal checking error (0.0491) occurs at about epoch 70, Notice that the checking error curve tend to stability after 70 epochs, indicating that further training, that means more training is useless and there is no more generalization.

8. REFERENCES

[1] J. B. Heywood, *Internal combustion engine fundamentals* vol. 930: Mcgraw-hill New York, 1988.

[2] K. M. Chun, "Characterization of knock and prediction of its onset in a spark-ignition engine," Massachusetts Institute of Technology, 1988.

[3] G. Konig and C. Sheppard, "End–gas auto–ignition and knock in a spark ignition engine," *Fuel*, vol. 2013, pp. 09-12, 1990.

[4] C. V. Ferraro, "A KNOCK INTENSITY METER BASED ON (A) KINETIC CRITERION," 1978.

[5] R. Barton, S. Lestz, and L. Duke, "Knock intensity as a function of engine rate of pressure change," *Training*, vol. 2005, pp. 04-01, 1970.

[6] C. V. Ferraro, M. Marzano, and P. Nuccio, "Knock-Limit Measurement in High-Speed S. I. Engines," *Studies*, vol. 2011, pp. 12-14, 1984.

[7] W. Leppard, "Individual-cylinder knock occurrence and intensity in multicylinder engines," 1982.

[8] M. B. Jain, M. K. Nigam, and P. C. Tiwari, "Curve fitting and regression line method based seasonal short term load forecasting," in *Information and Communication Technologies (WICT), 2012 World Congress on*, 2012, pp. 332-337.

[9] H. Andrei, T. Ivanovici, G. Predusca, E. Diaconu, and P. C. Andrei, "Curve fitting method for modeling and analysis of photovoltaic cells characteristics," in *Automation Quality and Testing Robotics (AQTR), 2012 IEEE International Conference on*, 2012, pp. 307-312.

[10] D. Tamhane, P. Wong, F. Aminzadeh, and M. Nikravesh, "Soft computing for intelligent reservoir characterization," in *SPE Asia Pacific Conference on Integrated Modelling for Asset Management*, 2000.

[11] R. Ibatullin, N. Ibragimov, R. Khisamov, E. Podymov, and A. Shutov, "Application and method based on artificial intelligence for selection of structures and screening of technologies for enhanced oil recovery," in *SPE/DOE Improved Oil Recovery Symposium*, 2002.

[12] W. W. Weiss, R. S. Balch, and B. A. Stubbs, "How artificial intelligence methods can forecast oil production," in *SPE/DOE Improved Oil Recovery Symposium*, 2002.

[13] Y. Liu, B. Bai, Y. Li, J. Coste, and X. Guo, "Optimization design for conformance control based on profile modification treatments of multiple injectors in a reservoir," in *International Oil and Gas Conference and Exhibition in China*, 2000.