

Development of a Neuro Genetic Algorithm Based Virtual Sensing Platform for the Simultaneous Prediction of NO_x, Opacity and BSFC in a Diesel Engine Operated in Dual Fuel Mode with Hydrogen under Varying EGR Conditions

Rahul Banerjee

Assistant Professor, NIT Agartala, Tripura

Probir Bose

Director, NIT, Agartala.

ABSTRACT

The present study is dedicated to the development of a Coupled Multi Objective Neuro Genetic Algorithm based model to emulate the NO_x, opacity and BSFC (diesel) of an existing single cylinder four stroke CI engine operated in various dual fuel modes with hydrogen operated under EGR of varying thermal signatures. The associated MIMO problem has been transformed into an equivalent MISO model through unique coupling and decoupling algorithms. GA has been used as the search method to optimize the MISO topology. The MISO model substantially reduced the network weight in contrast to its MIMO counterpart thereby ensuring faster convergence and better accuracy with a consequently lower computational footprint and a promising potential to be used as a virtual sensing platform in real time optimization strategies in engines. The MISO model successfully captured the thermal effects of EGR on the desired outputs. The performance of the proposed MISO model has been benchmarked against a conventional equivalent MIMO platform through comprehensive absolute, relative and correlation metrics. Exemplary (R^2) scores of 0.9999 with the experimental values in predicting opacity, NO_x and BSFC simultaneously, together with extremely low Standard Error of Estimate of R^2 , GMRAE and Thiel U2 uncertainty coefficient, consistently, in comparison to its competing MIMO counterpart and was also complimented with an ability to predict each of the desired output with a certainty of 97.92% within a 99.5% confidence interval. The developed model was furthermore capable of mapping the NO_x-Opacity-BSFC trade-off potential of the dual-fuel modes of engine operation, even under EGR, for all cases of actual observations with remarkable accuracy, thereby establishing its inherent proficiency in providing a holistic and robust predictive platform for virtual sensing in real time optimization strategies for such dual fuel operation.

CITATION: Banerjee, R. and Bose, P., "Development of a Neuro Genetic Algorithm Based Virtual Sensing Platform for the Simultaneous Prediction of NO_x, Opacity and BSFC in a Diesel Engine Operated in Dual Fuel Mode with Hydrogen under Varying EGR Conditions," *SAE Int. J. Engines* 5(2):2012, doi:10.4271/2011-01-2472.

INTRODUCTION

1.1. ARTIFICIAL INTELLIGENCE BASED ENGINE MODELING: PERSPECTIVES

With contemporary emission legislations getting more stringent [1,2,3] in order to comply with the responsibilities of environmental obligations, engine manufacturers are turning to explore new avenues to meet the paradox of curtailing particulate matter and NO_x emissions on the one

hand [4,5,6,7] and meeting consumer expectations of reduced BSFC, greater power density and thermal efficiency on the other. In recent years, a variety of technologies have evolved to meet such paradoxical objectives such as common rail fuel injection, electronic engine management, variable turbocharger geometry, variable valve timing, injection timing and combustion chamber geometry optimization. The adoption of such techniques has increased manifold, the number of parameters to be addressed simultaneously for effective engine control strategies. [8,9,10]. Due to the non-linear and complex dependence of the pollutants on engine

parameters, one dimensional map based strategies has been found to be increasingly insufficient [8,11] to satisfy the evolving multiple distinct parameters being needed to control simultaneously in modern day diesel engines to conform to emission regulations in real time. Advanced engine control systems are being required, needing accurate dynamic models of the combustion process, [12] which are substantially nonlinear. In contrast to traditional nonlinear identification methods, AI platforms are capable of deriving nonlinear models directly from measured input/output data thereby reducing substantially the costs, time, and complexity associated with engine development, performance mapping, and control system development [13]. Advanced on-board diagnostic capability and the ability to reconfigure control in real time has become an essential criterion in implementing complex control paradigms having many more degrees of freedom that need to be addressed simultaneously [16,17]. Artificial intelligence based engine modeling offers a potential [18] for a very fast [19] multidimensional, adaptive model of a system under study along with their inherent robustness to accommodate data with appreciable degrees of uncertainties of observation coupled with low computational cost allowing onboard prediction and interpolations for control decisions [20]. The inherent adaptive strengths of an ANN based artificial intelligence systems to model complex nonlinear systems at high computational speeds even when presented with noisy and incomplete data without any a priori knowledgebase of the physics of the problem under process, have established itself as a robust and sought after system identification and mapping tool [17] in the control of internal combustion engines. Moreover, artificial neural network due to its inherent adaptive capabilities are able to develop robust models simulating accurate engine behavior corresponding to the variables under study and as such are becoming an indispensable and inexpensive basis of a virtual sensing system for on-board measurements of unmeasured or difficult to measure NO_x, PM, TUHC emissions continuously in real time [21,22,23,24,25]. Early researchers in the field of engine emissions [26,27,28,29,30] were quick to recognize the ability of multilayered feed forward ANN networks in embodying nonlinear trends in complex data to fit arbitrarily complex nonlinear models to any desired accuracy [31,32]. Further, with the advent of the era of dual fuel strategies to meet the ever changing emission standards, implementation of conventional multiphysics based simulation platforms have become more computationally demanding in the face of new degrees of freedom [33] that are needed to be addressed thereby increasing model complexity and the associated computational footprint. In addition, there exists no singular universally acceptable multiphysics model applicable to different dual fuel strategies adopted in the various studies [34,35,36,37,38,39,40] each of which is unique in its scope and applicability. The versatility of artificial neural networks to simulate multiple nonlinear objectives simultaneously without the knowledge of the underlying physics of the problem provides an alternative platform to model and

predict the performance and emission spectra of a dual fuelled diesel engines. Contemporary works of [42,43,44,45,46,47,48] have established the indisputable efficacy of ANN platforms to simulate the engine performance and emission spectra in real time along with performance indices of commendable accuracy in dual fuel engines without resorting to complex numerical models.

1.2. MOTIVATION FOR DEVELOPMENT OF A NEW MODEL

It is to be noted that most neural network modeling approaches to dual fuel operation in diesel engines have resorted to the use of a MIMO network topology. Because the model objective function in a MIMO topology is often defined as a weighted sum squared error of the multiple outputs, the model can produce only a compromising optimization of all outputs and cannot guarantee to have the minimum error for each forecasting step [49] thereby increasing the uncertainty when employed in real time control strategies. Furthermore, MIMO architectures due to their inherent nonlinear dynamic coupling and dimensional burden are prone to a higher computational footprint owing to larger number of initial settings and needing considerable training time [50]. The accuracy of MIMO architectures are relatively low because of the complex interactions of the network weights that need to be addressed during each iteration [50]. Moreover, gradient learning rule based back propagation (BP) algorithm has been almost singularly used in the feed forward MIMO topology in such modeling in dual fuel engines as referred in sec [1.1]. But the use of the BP learning/training algorithm being essentially a gradient-based error minimizing search method [51] is often prone to getting trapped in local minima during complex function approximation especially so in multiple output ANN problems as in the present study. The majority of the research works on the application of ANN on emission modeling in engines have evaluated their model against statistical error measures based on mean square error metrics. Empirical studies, [52,53,54,55,56,57,58,59] engaged in evaluating different measures of model accuracy have concluded that mean square errors are inappropriate as model evaluation criteria across different data series as seen in ANN models entrusted with the objective of satisfying several outputs of different scales of observation simultaneously as in the present scope of study and that, for universal acceptance and credibility of any developed model, it is necessary to employ error measures, i.e.; those invariant to scalar transformations. The extent to which model-predicted events approach a corresponding set of independently obtained, observations have been often been quantified in majority of these studies in terms of the correlation coefficient 'r' and co-efficient of determination R²; and is limited in that, they are insensitive to additive and proportional differences between the observed and modeled datasets, such that high scores can be obtained, even if the simulated values are considerably different from

the observed values in terms of magnitude and variability [60].

1.3. OBJECTIVES OF THE PROPOSED MODEL

Considering the limitations of contemporary ANN platforms as outlined in Sec [1.2], a new Neuro-Genetic MISO model is proposed to satisfy the following objectives:

i. Provide a holistic platform to model the output variables without loss of generalization for a given engine prototype expected to be run in a dual fuel mode with different hydrogen enrichment strategies.

ii. To minimize the dimensional burden of the network to make its computational footprint low enough so that it obtains a possibility of deployment as a fast on-board virtual sensing platform for control and optimization of engine parameters in real time.

iii. To encapsulate and respond to the thermal nature and quantity of the EGR modes of operation. It is pertinent to mention here that as per the knowledge of the authors this is one of the first attempts in its kind to incorporate the thermal characteristics of EGR in ANN based engine modelling which have been found a determining factor on the emission characteristics of an existing diesel engine under dual fuel operation [61,62].

iv. Possess the ability to model both NOx and smoke emissions along with BSFC (diesel) simultaneously to provide a real time trade-off map for effective control strategy. This by itself presents a challenge in the face of the conflicting factors which tend to influence the formation of each of the stated pollutants [63,64,65,66] and make it increasingly difficult to formulate a universal physical model to simultaneously predict them on the same computational platform. The correlation between the chosen input variables and desired outputs in the purview of the present study which has been encapsulated in the correlation matrix {C} (Table 1 & Figure 1) bear testimony to the conflicting influences of the input variables on the desired emission and performance output variables.

v. The model developed should be able to satisfy broad spectra of fitness/performance criteria to qualify for predictive accuracy and quality simultaneously to claim a credible platform of model robustness and integrity.

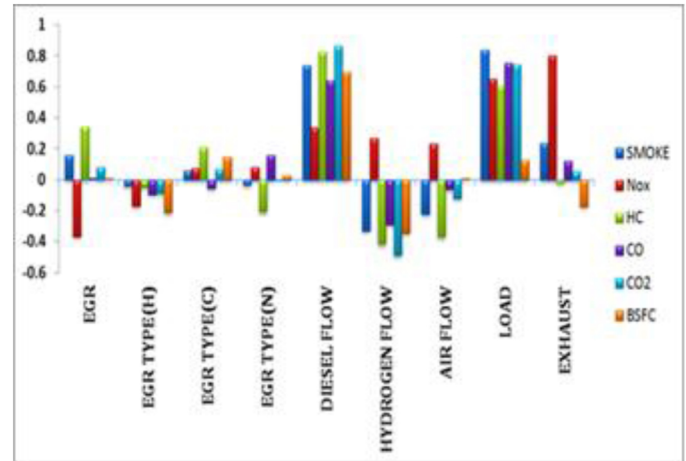


Figure 1. Effect of chosen input parameters on desired output variables as per their respective correlations.

Table 1. Matrix of the Pearson Product Moment Correlation Coefficient of each desired output with chosen input variables.

	θ_{S1}	θ_{NX1}	θ_{BSFC1}
$\dot{m}_E$	0.16	-0.37	0.01
e_H	-0.04	-0.17	-0.21
e_C	0.06	0.07	0.15
e_N	-0.03	0.08	0.03
$\dot{m}_D$	0.74	0.34	0.69
$\dot{m}_H$	-0.33	0.27	-0.35
$\dot{m}_A$	-0.22	0.23	0.01
l	0.84	0.65	0.13
ϕ	0.24	0.80	-0.17

2. MODEL DEVELOPMENT

2.1. CHOICE OF INPUT VARIABLES

The 'black box' label notwithstanding, the success of any ANN model depends upon a clear understanding of the problem under consideration with the identification of the most important variables in the process to be modeled. Of the large number of imposing factors influencing the formation of pollutants in dual fuel diesel engines, the final choice of input variables were dictated by the most pertinent contributing factors ascertained after survey of previous similar research studies [61,62,63,64,65,66,67,68,69] as well as the feasibility of on board measurement of the chosen variables in real time. Exhaust gas temperature was chosen as an input variable as a representative of the prevailing in cylinder temperatures during engine operation which has been found to be a dominant contributing factor in NOx

formation [70, 71]. Similarly the rate of air flow has been chosen as an input variable representing the available oxygen concentration during combustion. The choice of considering the flow rates of diesel, hydrogen and inducted air as separate input variables rather than a tempting approach to consider their joint effect by considering the overall equivalence ratio was dictated by the fact that the authors relied on the inherent strength of nonlinear adaptability of an ANN platform to capture the effects and thereby adapt to of each of them separately as the actual equivalence ratio during combustion is highly stochastic and cannot be imposed by a global equivalence ratio [40]. The input and output variables have been addressed in matrices $\{I\}$ as seen in Table 2 and $\{T_{actual}\}$ as shown in (Table 3) respectively. The objective of incorporating the impact of the thermal nature of EGR posed a challenge as it is an implicit categorical variable rather than an explicit numerical one. The present study innovatively attempts to capture the effect with category to numeric transformation as shown in the matrix $\{E\}$ as shown in Table 4.

Table 2. Matrix of the input variables

$$\{I\}_{n \times i} = \begin{bmatrix} \dot{m}_{E1}, e_{N1}, e_{H1}, e_{C1}, \dot{m}_{H1}, \dot{m}_{D1}, \dot{m}_{A1}, l_1, \phi_1 \\ \vdots \\ \dot{m}_{En}, e_{Nn}, e_{Hn}, e_{Cn}, \dot{m}_{Hn}, \dot{m}_{Dn}, \dot{m}_{An}, l_n, \phi_n \end{bmatrix}$$

Table 3. Matrix of the output variables

$$\{T^{actual}\}_{n \times t} = \begin{bmatrix} \theta_{S1}, \theta_{NX1}, \theta_{BSFC1} \\ \vdots \\ \theta_{Sn}, \theta_{NXn}, \theta_{BSFCn} \end{bmatrix}$$

Table 4. Categorical to numerical transformation matrix of the thermal signatures of the imposed EGR.

$$\{E\} = \begin{bmatrix} e_{none} & e_{hot} & e_{cooled} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.2. MATHEMATICAL BASIS OF DEVELOPMENT OF MISO MODEL

In order to reduce the dimensionality burden and its associated disadvantages of a MIMO topology as outlined in Sec [1.2] a MISO architecture has been proposed promising a faster convergence and better accuracy on account of a substantial reduction of the number of weights between hidden and output layers which are much less than that of a convention a MIMO topology, therefore reducing the interaction and complexity. The proposed MISO architecture was built through distinct preprocessing and post processing

algorithms accomplishing the tasks of **Coupling** and **Decoupling** the multiple output variables respectively, thereby attempting to reduce the dimensionality of the MIMO logic by secluding the bulk of the preprocessing outside the main iterative solver engine but simultaneously keeping the traits of the original output vectors as per the weights of their significance. Thus, not only the total network weight is reduced by condensing the multiple output variables to a single weighted output but also the single output variable demands a substantially lower number of hidden neurons given the same number/s of hidden layers in a MIMO architecture for the input variables to adapt itself to the output.

2.2.1. Coupling Phase

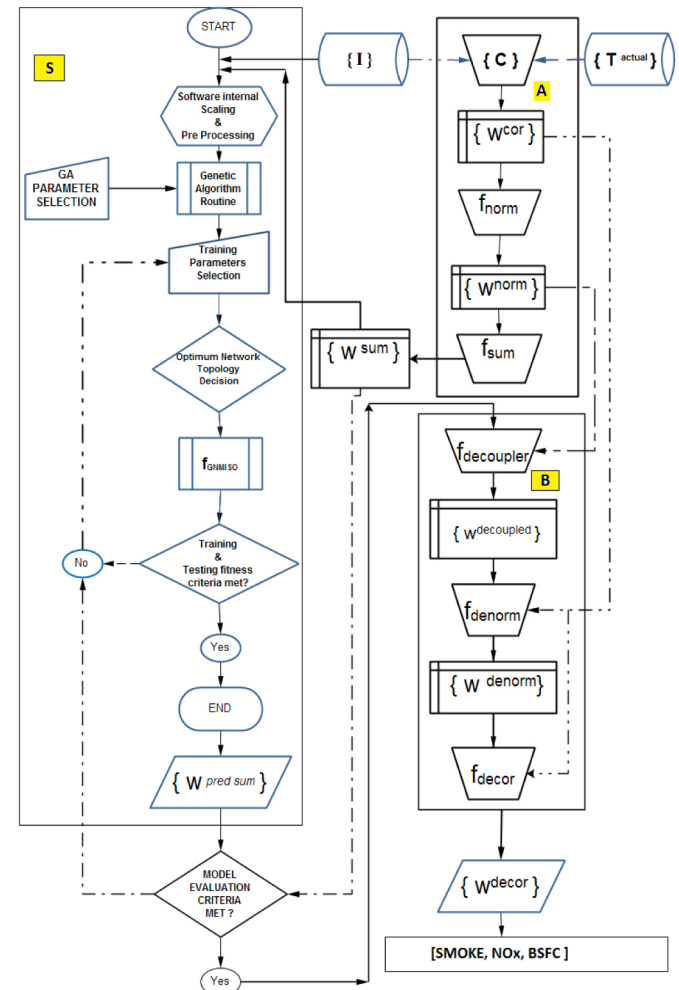


Figure 2. Workflow of MISO model

At the very outset keeping in mind the conflicting impact of the participating input variables on the output variables, a matrix $\{C\}_{corr}$ (Table 1) comprising the respective (Pearson) correlations of each of the input variables with each of the output variables for each observation has been ascertained. This effectively captures the essence of the impacts explicitly where negative values signify inverse relationship among the

variables under consideration. Next, the objective of condensing the multiple outputs/target variables into a single pseudo output vector has been carried out by employing transformation functions f_{norm} , f_{sum} successively on the matrix $[w]^{cor}$ to obtain the single pseudo output variable $[W]^{sum}$. The mathematical basis of the coupling sequence has been enumerated through the following Eq. 1, Eq. 2, Eq. 3, Eq. 4, Eq. 5 and the associated workflow has been encapsulated in **Block A** of Figure. 2.

Coupling Sequence:

$$\{I\}_{n \times i} \times \{C\}_{i \times t} = \{w^{cor}\}_{n \times t} \quad (1)$$

Next:

$$\{w^{norm}\}_{n \times t} = f_{norm} \left[\{w^{cor}\}_{n \times t} \right] \quad (2)$$

Wherein f_{norm} is the **Denormalizing Function** such that the jk^{th} element of $\{w^{norm}\}_{n \times t}$ is given as:

$$\{w_{jk}^{norm}\} = \frac{\{w_{jk}^{cor}\} - \min \left[\{w^{cor}\}_{* \times k} \right]}{\max \left[\{w^{cor}\}_{* \times k} \right] - \min \left[\{w^{cor}\}_{* \times k} \right]} \quad (3)$$

Next:

$$\{w^{sum}\}_{n \times 1} = f_{sum} \left[\{w^{norm}\}_{n \times t} \right] \quad (4)$$

Wherein f_{sum} is the **Summation Function** such that the jk^{th} element of $\{w^{sum}\}_{n \times 1}$ is given as:

$$\{w_{jk}^{sum}\}_{n \times 1} = \left\{ \sum_{k=1}^t w_{jk}^{norm} \right\} \quad (5)$$

2.2.2. Iteration Phase

2.2.2.1. Selection of Network Architecture

Having thus reduced a MIMO problem to a MISO, the next step was to determine the network architecture. The neural network architecture refers to the arrangement of neurons into layers and the connection patterns between the hidden layers, the choice of activation functions and most pertinently the number of hidden neurons in the hidden layers which are the units of nonlinear adaption wherein an increased number of hidden neurons may be used in order to get more 'degrees of freedom' and allow the network to store

more complex patterns. However in the effort to find an optimum network topology, there must be a balance in increasing the number of hidden layers and the number of hidden neurons therein such that the total network weights do not impose a high computational cost and avoid a generalization loss during the testing phase of the network. Past studies utilizing a MIMO topology in emission modeling in diesel engines and dual fuel strategies have widely used a single hidden layer but with the expense of a correspondingly high number of neurons therein to match the highly complex and nonlinear dependence of the input parameters to the emission outputs. Starting from empirical relations to modern, more exhaustive search methods, genetic algorithm provides a robust alternative method of searching the optimum solution to complex problems like the selection of an optimal network topology where it is difficult or impossible to test for optimality due to the lack of a universally accepted fitness criterion. A genetic or evolutionary algorithm applies the principles of evolution found in nature to the problem of finding an optimal solution from a given set of possible solutions of a problem based on the mechanics of natural genetic and natural selection. The basic elements of natural genetics - reproduction, crossover, and mutation - are used in the genetic search procedure. The basics of GA and its pertinent merits in being employed as credible search routine have already been established [72,73,74,75]. Hence, in the present study genetic algorithm was used to find the optimal architecture of the neural network. The process of choosing the optimum topology in the present study through GA routine was performed on *Alyuda Forecaster v 1.6 Build (784)* which was also chosen as our iteration solver. The details of the GA training parameters have been detailed in Table 5. For the present case study a single hidden layer was proposed by the GA search routine with an optimal 6 hidden neurons therein as schematically represented in Figure 3.

Table 5. MISO Topology selection on iteration solver

Alyuda Forecaster v 1.6 Build (784)		
TOPOLOGY	GFNN (9-6-1)	Genetic algorithms parameters: Population size: 50 Generations: 60 Penalty: 5 Crossover rate: 0.8 Mutation rate: 0.2
Activation Function	Log Sigmoid	
Training Algorithm	Levenberg-Marquardt	

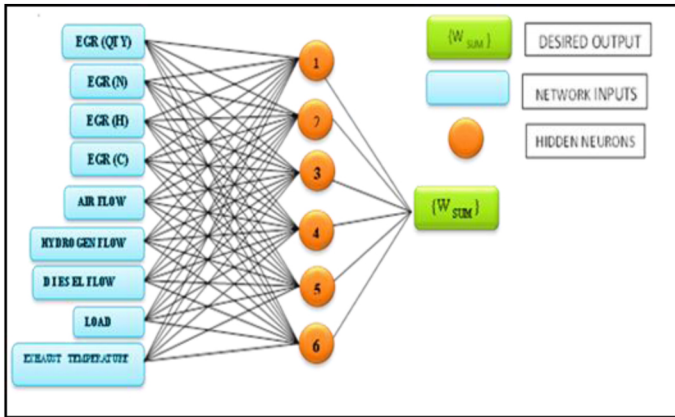


Figure 3. Schematic of MISO network topology

2.2.2.2. Choice of the Learning Rule

Subsequent to having determined the optimum topology, it was then necessary to apply an appropriate learning rule which is essentially a performance criterion involving the simultaneous and incremental adjustment of network weights in such a way that the activation functions gradually assume features that help collectively approximate the desired response. The Levenberg-Marquardt method was chosen as the learning algorithm in the present study. The Levenberg-Marquardt is a robust second order error minimization hybrid algorithm that combines the advantages of the steepest descent and Gauss-Newton methods to produce a more efficient method [76-77] than either of these two methods and help achieve convergence with a minimum number of iterations. The second-order methods such as the Levenberg-Marquardt algorithm provide faster solutions because of the incorporation of an extra second derivative of error information and automatic internal adjustments that are made to the learning parameters. A detailed exposition on the merits and demerits as per its cases of applicability can be had from the works of Shepherd [78].

2.2.2.3. Choice of Miso Network Parameters

For the chosen feed forward network, a continuous differentiable sigmoid activation function [51] was chosen both for the hidden as well as the output neurons. The advantages of employing Log-sigmoid activation function in the feed forward neural architecture is that it provides a form of automatic gain control and its output cannot grow infinitely large or small. Mean Square Error was chosen as the loss function to be minimized. The virtues of the metric of MSE due to its inherent properties of convexity, differentiability and symmetry in the context of optimization have been discussed in the works of Wang [79]. Subsequent to the selection of the network topology, the neural network iterations were performed subject to the stop training and loss function criteria as detailed in Table 6 on the data set segregated (Table 7) by the iteration solver for network training. Data subsets were internally rescaled by the iteration

solver as shown in Table 8 and retained blind by the solver for testing of model training performance and cross-validation purposes to prevent over-training during the training phase.

Table 6. Stopping Criteria of iteration set on both cases.

Solver Iteration	
1.	MSE on training subset must drop below 0.00001.
Additional Training Stop Criteria	
2.	Early stopping method, i.e. the training process is stopped when the MSE on a validation set of data increases or remains the same for an imposed number of epochs (=100).
3.	Attainment of an imposed maximum number of epochs [=10 exp (6)].
4.	Early stopping on 5% generalization loss.

Table 7. Data Subset selection by iteration solver.

	Data Subset
Training	66.67% randomly selected
Validation	16.67% randomly selected
Testing	16.67% randomly selected

Table 8. Internal scaling factor of MISO solver platform

$\dot{m}_E$	e_H	e_C	e_N	$\dot{m}_D$	$\dot{m}_H$	$\dot{m}_A$	l	ϕ	$\{W\}_{sum}$
0.04	0.8	0.8	0.8	0.63573	3.2	0.07032	0.19231	0.0018	0.37505

2.2.2.4. Miso Network Iteration Phase

The iteration sequence, encapsulated symbolically in Eq. 6, mainly consisted of network weight updates through the Levenberg- Marquardt optimization routine between the neurons of the input, hidden and output layers through Eq.7 and Eq. 8, sequentially for each step of network update until the desired loss function 'e_k' (Eq. 9) is achieved whence the network is assumed to be sufficiently trained and ready for estimation and comparison with the test data subset hitherto retained blind to the iteration. The details of the network update iteration mechanism of a GFNN can be had from the works of Sudhir [80]. The entire work flow in the iteration solver has been encapsulated in **Block S** of Figure 2 and that of the iteration routine in Figure 4 while the iteration convergence results for the chosen stopping criteria set has been detailed in Table 9.

$$\left\{w^{pred\ sum}\right\}_{n \times 1} = f_{GNMISO} \left[f_{GA} \left(f^{nnprep} \left(\left\{I\right\}_{i \times n} \left\{w^{sum}\right\}_{n \times 1} \right) \right) \right] \quad (6)$$

Iteration Sequence:

$$z_j = f_{hidden} \left(\sum_{i=1}^n w_{ij} x_i + \mu_j \right) \quad (7)$$

$$y_m = f_{out} \left(\sum_{j=1}^h z_j b_{jk} + \mu_k \right) \quad (8)$$

$$e_k \equiv MSE_k = \frac{1}{n_{obs} \times n_{out}} \sum_{\lambda=1}^{n_{out}} \sum_{j=1}^{n_{obs}} (t_{\lambda j} - y_{\lambda j})^2 \quad (9)$$

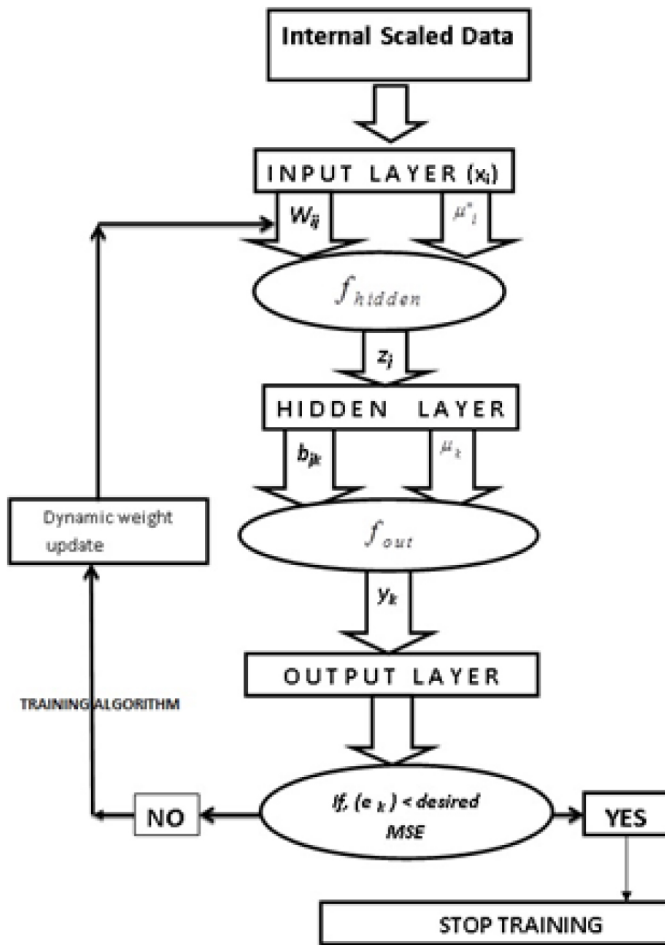


Figure 4. Workflow of the iteration sequence

Table 9. Iteration results of MISO and MIMO-BP-MLP platforms

CPU:		INTEL® CORE 2 DUO®T9300@ 2.5 Ghz	
System Memory:		2 GB @ 667 MHZ	
OS:		WINDOWS 7® SP1	
MISO TOPOLOGY		MIMO TOPOLOGY (BP-MLP)	
Training subset		Training subset	
Maximum absolute error	0.00027	Maximum absolute error	0.0063
Minimum absolute error	0.00003	Minimum absolute error	0.0008
Average absolute error	0.00014	Average absolute error	0.0036
Average MSE	2.32E-07	Average MSE	8.6E-03
Test subset		Test subset	
Maximum absolute error	0.00235	Maximum absolute error	0.045
Minimum absolute error	0.00011	Minimum absolute error	0.0033
Average absolute error	0.00102	Average absolute error	0.025
Average MSE	1.68E-06	Average MSE	3.7E-02
Typical Iteration Step interval	0.134 s	Typical Iteration Step interval	NA
Total Iteration Time For Convergence	4m 52s	Total Iteration Time For Convergence	9 m 03s

2.2.3. Decoupling Phase

Once the iterative solver had simulated the pseudo output $\{w^{pred\ sum}\}_{n \times 1}$ within the acceptable range of the imposed MSE loss function of the learning rule, the pseudo output vector was decoupled to its respective predicted output components $\{w^{pred\ out}\}_{n \times t}$ via the decoupling sequences enumerated in detail through Eq.10, Eq.11, Eq.12, Eq.13, Eq.14, Eq.15 through the transformation functions $f_{decoupler}$, f_{denorm} and f_{decor} respectively, which were then compared with the actual output variables by various measures enlisted in sec [4]. The workflow of the decoupling phase has been has been encapsulated in **Block B** of Figure 2.

Decoupling Phase:

$$\left\{w^{decoupled}\right\}_{n \times t} = f_{decoupler} \left(\left\{w^{pred\ sum}\right\}_{n \times 1} \right) \quad (10)$$

Wherein $f_{decoupler}$ is the **Decoupling Function** such that the jk^{th} element of $\left\{w^{decoupled}\right\}_{n \times t}$ is given by: