sustainability

MDPI

Article

Modeling and Multi-Objective Optimization of Engine Performance and Hydrocarbon Emissions via the Use of a Computer Aided Engineering Code and the NSGA-II Genetic Algorithm

Richard Fiifi Turkson [1,2,3,*], Fuwu Yan [1,2], Mohamed Kamal Ahmed Ali [1,2,4], Bo Liu [1,2] and Jie Hu [1,2]

Received: 6 November 2015; Accepted: 8 January 2016; Published: 13 January 2016
Academic Editors: Muge Mukaddes Darwish and Marc A. Rosen

[1] School of Automotive Engineering, Wuhan University of Technology, Wuhan 430070, China; yanfuwu@vip.sina.com (F.Y.); Eng.m.kamal@mu.edu.eg (M.K.A.A.); liubo321@whut.edu.cn (B.L.); auto_hj@163.com (J.H.)
[2] Hubei Key Laboratory of Advanced Technology for Automotive Components, Wuhan University of Technology, Wuhan 430070, China
[3] Mechanical Engineering Department, Ho Polytechnic, P.O. Box HP 217, Ho 036, Ghana
[4] Automotive and Tractors Engineering Department, Faculty of Engineering, Minia University, El-Minia 61111, Egypt
* Correspondence: rturkson@hopoly.edu.gh or riturkus@yahoo.com; Tel.: +86-136-2725-0700 or +233-2489-51963

Abstract: It is feared that the increasing population of vehicles in the world and the depletion of fossil-based fuel reserves could render transportation and other activities that rely on fossil fuels unsustainable in the long term. Concerns over environmental pollution issues, the high cost of fossil-based fuels and the increasing demand for fossil fuels has led to the search for environmentally friendly, cheaper and efficient fuels. In the search for these alternatives, liquefied petroleum gas (LPG) has been identified as one of the viable alternatives that could be used in place of gasoline in spark-ignition engines. The objective of the study was to present the modeling and multi-objective optimization of brake mean effective pressure and hydrocarbon emissions for a spark-ignition engine retrofitted to run on LPG. The use of a one-dimensional (1D) GT-Power™ model, together with Group Method of Data Handling (GMDH) neural networks, has been presented. The multi-objective optimization was implemented in MATLAB® using the non-dominating sorting genetic algorithm (NSGA-II). The modeling process generally achieved low mean squared errors (0.0000032 in the case of the hydrocarbon emissions model) for the models developed and was attributed to the collection of a larger training sample data using the 1D engine model. The multi-objective optimization and subsequent decisions for optimal performance have also been presented.

Keywords: engine modeling; NSGA-II genetic algorithm; optimization; emissions

1. Introduction

There is a growing global concern over depleting fossil fuel reserves and climate change issues resulting from industrial activities and the use of automobiles. The environmental footprint of automobiles is a result of an increase in the global fleet of vehicles, creating a high demand for fossil-based fuels. It is therefore feared that if this trend should continue, transportation and other activities that rely on fossil fuels may not be sustainable in the long term. In the search for these alternatives as a result of the concerns mentioned earlier, liquefied petroleum gas (LPG) has been

identified as one of the viable alternatives that could be used in place of gasoline in spark-ignition engines. The use of LPG as fuel leads to the reduction of emissions including Nitrogen oxides (NO_X) and Sulphur dioxide (SO_2). Carbon dioxide emission reduction per kilowatt could also be associated with the use of LPG as fuel. LPG is cheaper to produce than gasoline, reduces the frequency of engine maintenance and facilitates a quick engine start from cold. Against a background of tightening emission legislation around the world coupled with rising fuel costs, the use of LPG as an alternative fuel is becoming widespread [1–9].

The objectives of modeling internal combustion engines may include the design of new engines, design of control systems and the prediction of engine performance without the extensive use of real dynamometer tests which may be expensive in terms of time and resources [10]. According to Tumelaire *et al.* [11], the use of simulation during the engine development process also facilitates the investigation of operational modes that could have a damaging effect on the engine if real dynamometer tests were used. On the other hand, engine calibration and testing is a process of adjusting the input parameters of an engine with the aim of improving driveability and fuel economy to satisfy customer demand or homologation [12–14] to ensure compliance with emission regulations [15,16]. In the calibration of an engine, an engineer is often faced with the problem of maximizing or minimizing certain engine responses, some of which may be conflicting as may be dictated by market demand and/or legislation. For example, two conflicting objectives may be to maximize the brake power while keeping carbon monoxide emissions at a minimum over the entire operating range of the engine. One way of dealing with this problem is the use of multi-objective genetic algorithms. The Group Method of Data Handling (GMDH) polynomial neural networks were used by Atashkari *et al.* [17] to model the effects of variable valve timing and engine speed on the engine torque output and fuel consumption, while the Non-dominating Sorting Genetic Algorithm (NSGA-II) was used for the multi-objective optimization process. In the study [17], an attempt was made to solve the conflicting objective problem of trying to maximize engine output torque while minimizing fuel consumption. In another study [18], a multi-objective optimization of output torque and NO_x emissions was carried out for a single-cylinder gasoline engine for a motorcycle using a 1D fluid-dynamics model. The multi-objective optimization was carried out by comparing the performance of the ε-constraint and the NSGA-II algorithm methods of optimization. It was found in the study [18] that the multi-objective optimization solution for the NSGA-II algorithm was better than that of the ε-constraint method. In the work of Martinez-Morales *et al.* [19], artificial neural networks were used for the non-linear identification of a gasoline engine for the purpose of evaluating objective functions used within an optimization framework involving the use of the NSGA-II algorithm and the Multi-Objective Particle Swarm Optimization (MOPSO) algorithm. The objective of the study [19] was to optimize carbon monoxide (CO), unburnt hydrocarbons (HC) and nitrogen oxides (NO_x) emissions for the gasoline engine used for the study. It was observed, again, that the NSGA-II algorithm performed better than the MOPSO algorithm in the Pareto based optimization process.

In the current study, the modeling and multi-optimization of brake specific hydrocarbon emissions and brake mean effective pressure for a gasoline engine, retrofitted to run on liquefied petroleum gas (LPG), is presented. The study involved the modelling of the target engine using the GT-Power™ 1D engine simulation code. Engine identification via the use of design of experiments (DoE) and trained GMDH polynomial neural networks is presented. The main inputs to the models were intake manifold length and fuel-air ratio (decision variables) and the outputs were brake mean effective pressure and hydrocarbon emissions. The NSGA-II genetic algorithm was chosen for the optimization because it has proven to give better optimal results compared to other optimization methods as observed in references [18,19]. Intake manifold length and fuel-air ratio were chosen as the decision variables because the design and optimization of engine performance relies on the tuning of the intake manifold system and fuel-air ratio [20]. A number of research activities have been conducted to demonstrate the effect of variable intake manifold length on the performance of gasoline engines and are well documented in references [21–24].

2. Experimental Set-Up and Methods

2.1. Data Collection and Engine Modelling

Data was collected for the study by running a four-cylinder, four-stroke Wuling B15 spark-ignition engine retrofitted to run on liquefied petroleum gas (LPG) using an AVL AC dynamometer, under wide-open throttle conditions. The data collected by testing the target engine (Wuling B15 engine) was used to calibrate a GT-Power™ version of the target engine to ensure that the 1D model of the engine exhibited characteristics close to that exhibited by the target engine. Correlation between the model and was carried out by comparing the brake mean effective pressure and brake specific fuel consumption of the engine model and the target engine, and adjusting the engine model's parameters until it exhibited characteristics close to that of the target engine. This was based on the recommendations contained in GT-Power™ engine performance manual. Table 1 contains information that describes the target engine.

Table 1. Target engine description.

Description	Details
Engine type/Number of cylinders	Four-stroke/four cylinders
Number of valves per cylinder	4
Cylinder bore	74.7 mm
Stroke	84.7 mm
Compression ratio	10.2:1
Rated Power/Speed	79 KW@5400 rpm
Idle speed	750 +/− 50 rpm

A comparison between the output of the target engine and the GT Power™ model results are shown in Figure 1. The compared results for brake mean effective pressure, brake specific fuel consumption, brake specific hydrocarbon emissions and brake power are represented by the graphs labeled a, b, c and d, respectively, in Figure 1. The trends indicated in the results for the engine model tracked the results obtained from testing the target engine reasonably well, with the average error between experimental and simulated values being around 3%. This error margin was considered reasonable for the purpose of this study.

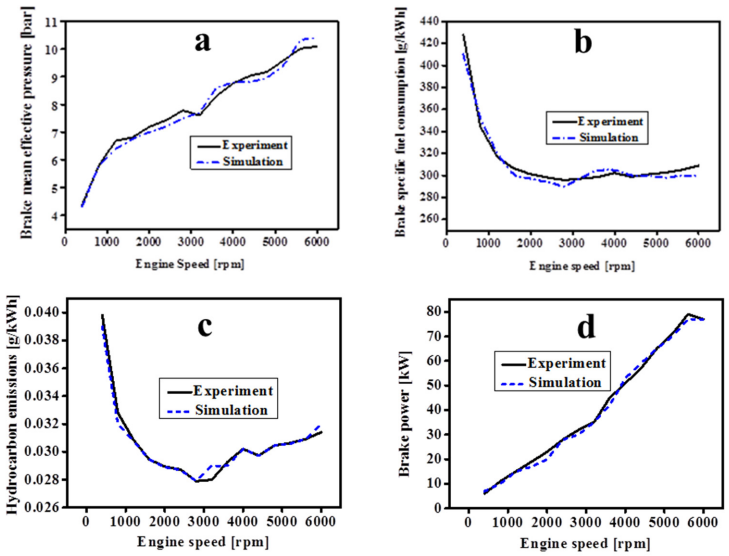

Figure 1. Comparison between experimental and simulation results for (**a**) Brake mean effective pressure; (**b**) Brake specific fuel consumption; (**c**) Brake specific hydrocarbon emissions and; (**d**) Brake power.

2.2. Design of Experiments and System Identification Using GMDH Polynomial Neural Networks

The objective of using a statistical Design of Experiments (DoE) [25–27] is to use a few dynamometer measurements for the purposes of engine identification without losing vital information. The ultimate goal is to facilitate the reduction of the time spent on testing an engine on a dynamometer. This facilitates the collection of data to cover a wide operating range of the engine and decision variables. For the purpose of this study, the GT-Power™ model was used to generate data for 2000 engine operating points. This was considered sufficient for reducing the error of the engine models generated via the training of the GMDH polynomial neural networks.

The primary aim of system identification in the current study was to obtain a function $\hat{f}$ in place of an original function f for estimating the system's output $\hat{y}$ as close as possible to the actual output y for an input vector $\boldsymbol{X} = (x_1, x_2, x_3, \ldots, x_n)$ using trained GMDH polynomial neural networks. The GMDH is the most appropriate method for obtaining the polynomial description of a stochastic system from a small amount of experimental data [28].

In generic terms, the input–output relationship based on the decision variables could be expressed using the Volterra functional series otherwise known as the Kolmogorov-Gabor polynomial [29,30]. The aforementioned functional series could be represented as follows:

$$y = a_0 + \sum_{i=1}^{n} + a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{j=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \ldots \quad (1)$$

Equation (1) can be expressed mathematically using a partial quadratic polynomial involving two decision variables (x_i,x_j) as shown in Equation (2).

$$\hat{y} = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (2)$$

Equation (2) was used to repeatedly build the general mathematical relationship between the inputs (decision variables) and the outputs (brake mean effective pressure and hydrocarbon emissions) in a network of connected neurons in Equation (1). The coefficients $\{a_0, \ldots, a_5\}$ in Equation (2) were determined via regression techniques so that the error between the actual output y and the estimated output for each input pair x_i and x_j is minimized. For the GMDH trained neural network, the coefficients of Equation (2) were determined such that error, E, between the actual outputs and the estimated was given by Equation (3).

$$E = \frac{1}{N}\sum_{i=1}^{N} (y_i \pm \hat{y}_i)^2 \rightarrow \min \quad (3)$$

For a GMDH neural network, a model can be represented by a set of neurons by having different pairs of neurons in each layer connected via a quadratic polynomial for the production of new neurons in the next layer [30]. This could be used in the modeling process to map inputs to outputs. Furthermore, in a GMDH algorithm, it is important to collect data about all the probable combinations of two decision variables (independent variables) for the construction of the regression polynomial Equation (2). For the purpose of this study, the Latin Hypercube Sampling [31–34] method was used to decide on the probable combinations of the decision variables (intake manifold length and fuel-air ratio). A total of 2000 possible combinations of the independent variables were used to generate outputs over the operating range (400–6000 rpm) of the GT-Power™ engine model. The total number of neurons that could be contained in the first layer of the feedforward neural network is given by $\binom{n}{2} = n(n-1)/2$ using collected data in the form $\{(x_{ki},x_{kj},y_k) \mid k = 1,2, \ldots, N\}$ involving the various values for which, $i, j \in \{1,2, \ldots, n\}$. With the collected data, it was possible to form N (2000 for this study) data triples as indicated in Equation (4). Furthermore, in this study, the independent variables

x_{ki} and x_{kj} represented intake manifold length and fuel-air ratio, respectively. The dependent variables y_k considered for the engine models were brake mean effective pressure and hydrocarbon emissions.

$$D = \begin{bmatrix} x_{1i} & x_{1j} & y_1 \\ x_{2i} & x_{2j} & y_2 \\ \vdots & \vdots & \vdots \\ x_{Ni} & x_{Nj} & y_N \end{bmatrix} \tag{4}$$

It was also imperative to form the N by 6 matrix based on Equation (2) as follows:

$$A = \begin{bmatrix} 1 & x_{1i} & x_{1j} & x_{1i}x_{1j} & x_{1i}^2 & x_{1j}^2 \\ 1 & x_{2i} & x_{2j} & x_{2i}x_{2j} & x_{2i}^2 & x_{2j}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{Ni} & x_{Nj} & x_{Ni}x_{Nj} & x_{Ni}^2 & x_{Nj}^2 \end{bmatrix} \tag{5}$$

The matrix equation based on Equations (2) and (5) could also be represented in Equation (6) as follows:

$$A\boldsymbol{a} = Y \tag{6}$$

where $\boldsymbol{a} = \{a_0, a_1, a_2, a_3, a_4, a_5\}$ is the vector of the coefficients to be determined for the partial quadratic polynomial and $Y = \{y_1, y_2, \ldots, y_N\}^{\mathrm{T}}$. Solving Equation (6) for $\boldsymbol{a}$ leads to the least squares estimation in the form of Equation (7).

$$\boldsymbol{a} = (A^T A)^{-1} A^T Y \tag{7}$$

The use of such an Equation (7) for determining the vector of coefficients for the N data triples is done recursively for each neuron in the hidden layers of the neural network. The downside of using these coefficients directly as obtained from Equation (7) is that it is prone to round-off errors, as well as the existence of singularities.

2.3. The Use of Singular Value Decomposition (SVD) for the Optimal Determination of Coefficients

For linear least squares problems having some singularities in normal equations, the SVD method is normally used for obtaining the optimal coefficients. The singular value decomposition of the matrix $A \in \Re^{N\times 6}$ is the decomposition of the matrix into the product of three matrices. This breakdown comprises an orthogonal matrix $U \in \Re^{N\times 6}$, a diagonal matrix $\sum \in \Re^{6\times 6}$ having singular values and another orthogonal matrix $V \in \Re^{6\times 6}$. The expression for the matrix A based on the SVD is as follows:

$$A = U\sum V^T \tag{8}$$

The optimal determination of the coefficients $(a_0,\ a_1,\ a_3,\ a_4,\ a_5)$ for Equation (2) was achieved by modifying the inversion of the diagonal matrix $\sum$. This involved the setting of the reciprocals of zero or near-zero singulars to zero such that:

$$\boldsymbol{a} = V\left[diag\left(\frac{1}{\sum}\right)\right]U^T \tag{9}$$

Information regarding the use of SVD and the subsequent modification of the diagonal matrix $\sum$ for the optimal determination of coefficients could be found in references [35,36].

Figure 2 represents a summary of the procedure for generating the regression models of the target engine, from the correlation between target engine and 1D model, through Latin Hypercube Sampling/Design of experiments to the formulation of regression models. The training of the GMDH neural network for obtaining the regression models of the engine was implemented in MATLAB®. This involved an initial random generation of a population of individual neural networks represented

as concatenated encoded strings and evaluating the fitness in the form of the inverse of the mean squared error of each network. Genetic operators such as selection, crossover and mutation were then used to improve the fitness from one generation to the other until there was no further improvement in fitness. The equations obtained as a result of training the neural networks were used as objective functions for the multi-objective optimization framework using the NSGA-II genetic algorithm.

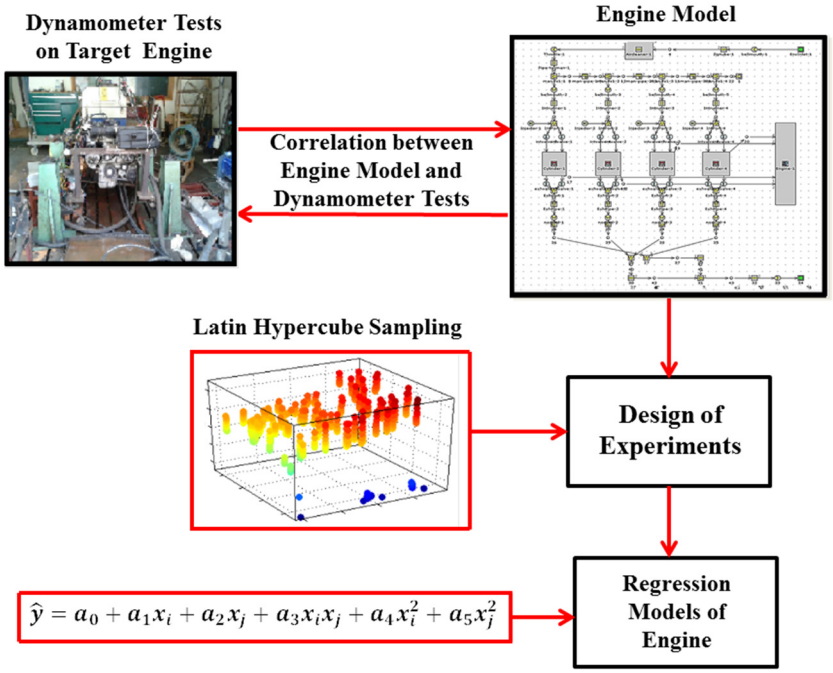

Figure 2. Summary of the modeling process.

2.4. Multi-Objective Optimization

A multi-objective optimization problem is concerned with maximizing or minimizing a number of objective functions in the presence of certain inequality and equality constraints, as well as other constraints in the form of lower and upper bounds defining the decision variable space. In the generic sense, a multi-objective optimization problem could be expressed mathematically as follows:

$$\begin{cases} \text{Maximize/Minimize } f_k(x),\ k = 1, 2, \ldots, K \\ \text{Subject to } g_m(x) \geqslant 0,\ m = 1, 2, \ldots, M \\ h_j(x) = 0,\ j = 1, 2, \ldots, J \\ x_i^L \leqslant x_i \leqslant x_i^U,\ i = 1, 2, \ldots, n \end{cases} \tag{10}$$

The solution to a multi-objective problem is to find a vector $\boldsymbol{X}^* = (x_1^*, x_2^*, x_3, \ldots\ x_n^*) \in \Re^n$ of the decision variables that gives an optimal solution to the multi-objective functions $F(x) = \{f_1(x), f_2(x), \ldots, f_x(x)\} \in \Re^k$. The last set of constraints represent decision variable bounds which constrain the decision variables x_i to vary between the lower bound x_i^L and the upper bound x_i^U [37–39].

Prior to the discussion on the multi-objective optimization using the NSGA-II genetic algorithm, it would be important to describe the optimization process in relation to the Pareto dominance and the Pareto optimal solution to a multi-objective optimization problem.

2.5. Pareto Dominance and Optimal Solutions

A vector $\boldsymbol{B} = \{b_1, b_2, \ldots, b_n\} \in \Re^n$ is said to be dominant to another vector, $\boldsymbol{C} = \{c_1, c_2, \ldots, c_n\} \in \Re^n$ (that is $B \prec C$, for a minimization problem) if it satisfies the only condition that for $\forall_i \in \{1,2, \ldots\ n\}$, $b_i \leqslant c_i \wedge \exists_j \in \{1,2, \ldots\ n\}: b_j < c_j$. That is, there should be at least one value such as b_j in the vector $\boldsymbol{B}$

being less than another value such as c_j in the vector $\boldsymbol{C}$, with the rest of the elements in vector $\boldsymbol{B}$ being less or equal to the elements in vector $\boldsymbol{A}$. For the feasible region denoted by Ω and defined by the decision variable space $\Re^n$, an element x^* of the vector X^* is defined as a Pareto optimal if and only if $F(x^*) \prec F(x)$. This could be expressed mathematically as $\forall_i \in \{1,2, \ldots k\}, f_i(x^*) \leqslant f_i(x) \wedge \exists_j \in \{1,2, \ldots n\}: f_j(x^*) < f_j(x)$. It could also be said that a solution x^* could be described as a Pareto optimal solution if there does not exist any other point that dominates x^*.

2.6. The Pareto Set and Optimal Frontier

The Pareto set for a given multi-objective optimization problem could be described as the set of decision variables in the defined variable space, comprising all the Pareto optimal vectors and could be expressed as $P^* = X \in \Omega$ for which there does not exist any vector such as $X' \in \Omega$ such that $F(X') \prec F(X)$. On the other hand, the Pareto frontier $P_f{}^*$ is a set of objective functions and could be expressed mathematically as $P_f{}^* = \{F(X) = (f_1(X), f_2(X), \ldots f_k(X)): X \in P^*\}$.

2.7. Use of the NSGA-II Genetic Algorithm for the Multi-Objective Optimization Framework

In the generic sense, genetic algorithms belong to a class of numerical optimization methods that are inspired by natural selection and genetics, and could be used to solve a broad spectrum of real world problems. The use of genetic algorithms for solving real world problems could be thought of as an extension of Charles Darwin's theory of evolution that the genetic diversity in our world today is as a result of the differential survival and reproduction of individuals due to differences in phenotype [40]. Genetic algorithms have attracted attention because of their ability to solve problems by finding a global optimum in the presence of many local optimums. The basic procedure for executing a genetic algorithm is summarized in Figure 3.

The NSGA-II genetic algorithm was proposed by Deb *et al.* [39] to help reduce the computational complexity based on a certain number of decision variables and a given population of solutions, preserve the elite members of a population of solutions and eliminate the need for a sharing parameter associated with other multi-objective evolutionary algorithms like the Pareto-archived evolutionary strategy (PAES) [41] and the strength-Pareto evolutionary algorithm (SPEA) [42].

The procedure for executing the NSGA-II genetic algorithm based on crowding distance sorting is illustrated in Figure 4. The process starts by forming an initial population, Pr, of candidate solutions for which Br = 0.5 Pr, Ar = 0.5 Pr and Pr = Ar $\cup$ Br. The non-dominated sorting is performed on the initial population, Pr, in a manner that ensures that elitism from the previous generation of solutions is preserved. For the situation where the members of the non-dominated set M_1 are smaller than the number of members in Ar, all the members of M_1 are used for the formation of the subsequent population A_{r+1}. Other members of the new population A_{r+1} are selected from other non-dominated sets such as M_2 and M_3. In the generic sense, the number of non-dominated sorting solutions from M_1 to the last front M_L may be larger than the size of Ar. To ensure that the new population A_{r+1} has the same size as Ar, an operator known as the crowded-comparison operator is used for selecting the best solutions to occupy the available spaces for forming the new population. The new population A_{r+1} is then processed by applying the genetic operators of selection (analogous to the natural selection in biological systems), crossover (analogous to the sexual reproduction in humans) and mutation (for preserving the genetic diversity) for producing a new generation of solutions A_{r+2}.

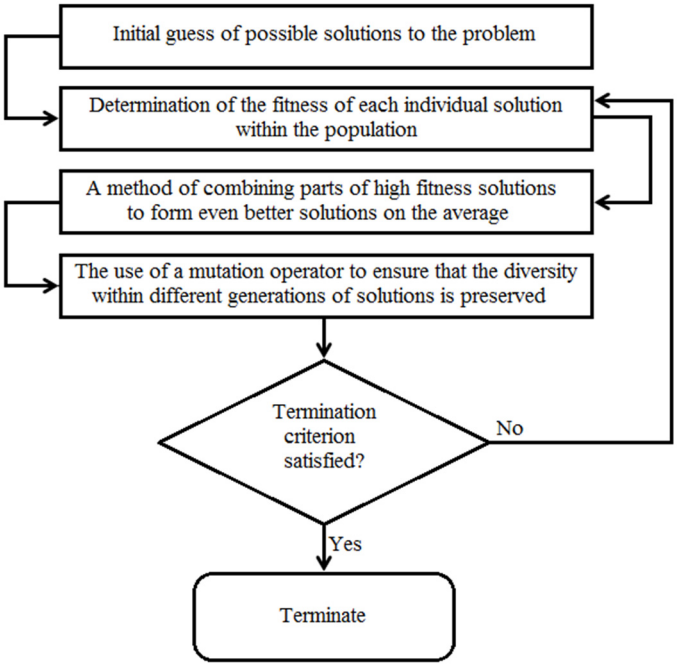

Figure 3. General procedure for executing a genetic algorithm.

In the current study, an initial solutions population of 200 and a maximum number of generations of 500 were used for the multi-objective optimization. The tournament selection method was used to select the best solutions in a particular population for creating a mating pool for producing child solutions. Probabilities of 0.9 and 0.1 were used for the crossover and mutation operators, respectively. The lower and upper bounds for the intake manifold length decision variable were 231 mm and 673 mm, respectively. Similarly, the lower and upper bounds for the fuel-air ratio decision variable was 0.056 and 0.083, respectively, corresponding to an air-fuel ratio in the range of 18:1 to 12:1. The multi-objective optimization had two conflicting objective functions: brake mean effective pressure and hydrocarbon emissions, derived from the training of the GMDH polynomial neural networks. The multi-objective optimization was implemented in MATLAB®.

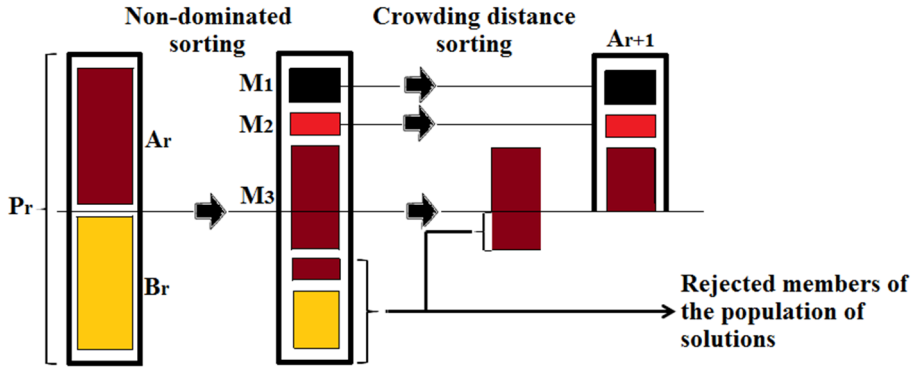

Figure 4. Non-dominated sorting for an NSGA-II genetic algorithm based on crowding distance.

3. Results and Discussion

3.1. Results for Training GMDH Neural Networks

The quadratic polynomial equations resulting from using the GMDH neural networks for engine identification are given in Equations (11) and (12).

$$y_{bmep} = 6.523 - 5.232I_L + 6.12F_{air} + 9.63I_L F_{air} + 3.25I_L^2 - 5.45F_{air}^2 \quad (11)$$

$$y_{HC} = 125 + 0.026I_L - 1.8F_{air} + 0.0035I_L F_{air} - 0.04I_L^2 + 0.0033F_{air}^2 \quad (12)$$

where y_{bmep} = Brake mean effective pressure (bar); y_{HC} = Brake specific hydrocarbon emissions (g/kWh); I_L = Intake manifold length (mm); F_{air} = Fuel-air ratio.

Equations (11) and (12) were used as objective functions for the optimization framework using the NSGA-II genetic algorithm. The mean squared error and other statistics regarding the brake specific hydrocarbon emissions and brake mean effective pressure models are shown in Table 2.

Table 2. Statistics for brake specific hydrocarbon emissions and brake mean effective pressure models.

Errors and Standard Deviation	Brake Specific HC Emissions		Brake Mean Effective Pressure	
	Training	**Test**	**Training**	**Test**
Mean Squared Error (MSE)	0.0000032	0.000032	0.297	0.263
Root Mean Squared Error (RMSE)	0.0018	0.0018	0.545	0.513
Mean Error (Absolute)	0.0000039	0.000078	0.00399	0.0135
Standard Deviation (StD)	0.0018	0.0018	0.545	0.545

The training and testing mean squared errors for the brake specific hydrocarbon emissions model were the same at 0.0000032. This indicated that apart from the models good ability to map training input data to target outputs, the model will also generalize well when presented with data not used for the training process. In similar research [17], higher values for training (0.027) and testing (0.035) mean squared errors were reported (a case of the best of errors reported). In a more recent study [43], the best training mean squared error reported was 0.1718 for a similar case of system identification. The main reason for the lower mean squared errors achieved in this research for the brake specific hydrocarbon model may be due to an increased number of data (2000) collected using a GT-Power™ model of the target engine. The data for the study [17] was collected by using dynamometer tests that limited the number data samples, hence a higher model error. The same could be said for reference [43] because experimental data was used directly for the study. The evidence of this is given in Equation (13).

$$Model\ error \approx \sqrt{\sigma^2}\frac{\sqrt{n_{eff}}}{\sqrt{N}} \quad (13)$$

where σ^2 = variance; n_{eff} = Effective neural network parameters (weights and biases) that are optimized during training and testing; N = Number of data samples.

For an optimized neural network structure, the model error could be approximated by Equation (13). It could be seen that the model error given by the aforementioned equation will be decreased with an increase in the number of data samples, N [44].

On the contrary, a higher level of mean squared errors for training and testing was achieved for the brake mean effective pressure model. This could be due to an increase in the variance error component of the model which is basically due to the deviation of network parameters from their optimal values as a consequence of the noisy nature of collected data. The model error could be decomposed into the bias and variance error components, with the bias error being mainly due to

model flexibility in terms of structure. A higher order polynomial model, say a cubic one, could yield a lower bias error [44].

Figures 5 and 6 show the training and test statistics for the brake specific hydrocarbon emissions model, respectively. Figure 5 shows the comparison between the target outputs and model predictions against training data indices on the horizontal axis. In all, 70% of the 2000 data samples obtained from the simulation were used for training. Also shown in Figure 5 is a plot of the errors and a fitted normal distribution curve for the model errors. The statistics for the testing shown in Figure 6 show similar results for 30% of the total data samples, corresponding to 600 samples selected at random from the 2000 data samples. Similar statistics are also shown graphically for the brake mean effective pressure model in Figures 7 and 8.

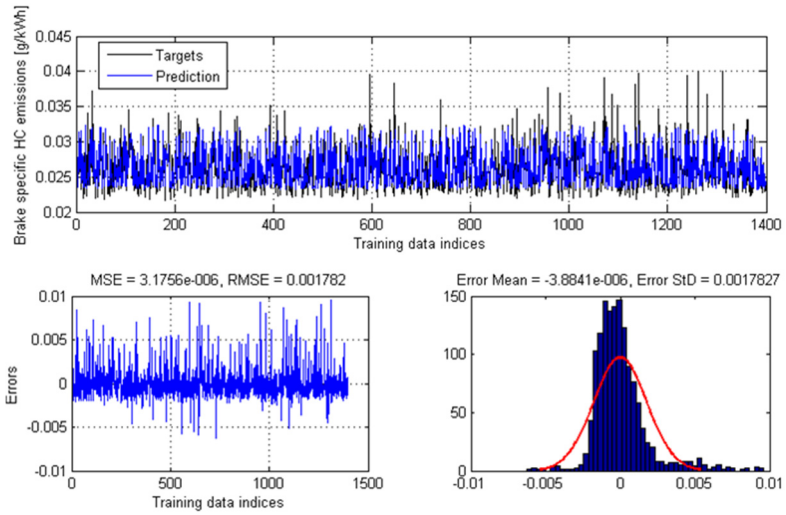

Figure 5. Training statistics for the brake specific hydrocarbon emissions.

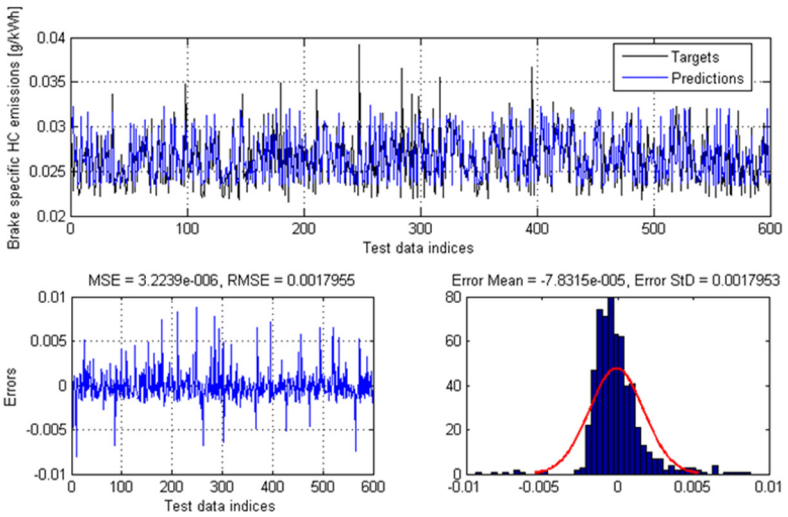

Figure 6. Statistics for testing the brake specific hydrocarbon emissions model.

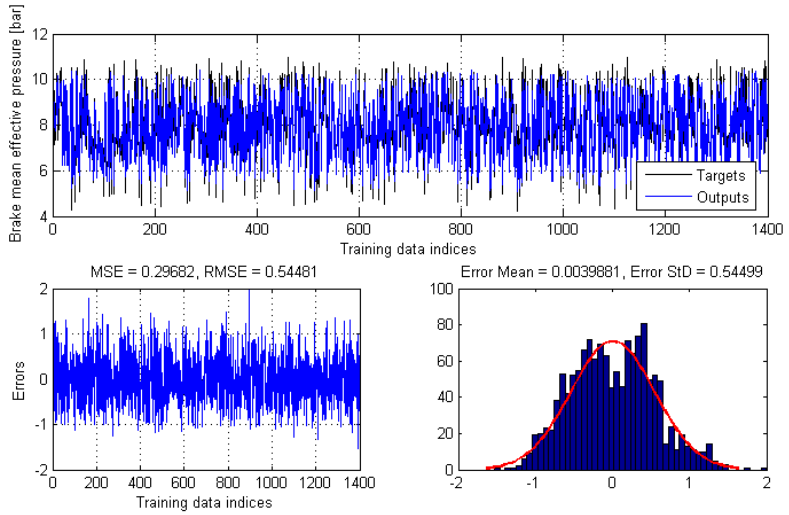

Figure 7. Statistics for the training of the brake mean effective pressure model.

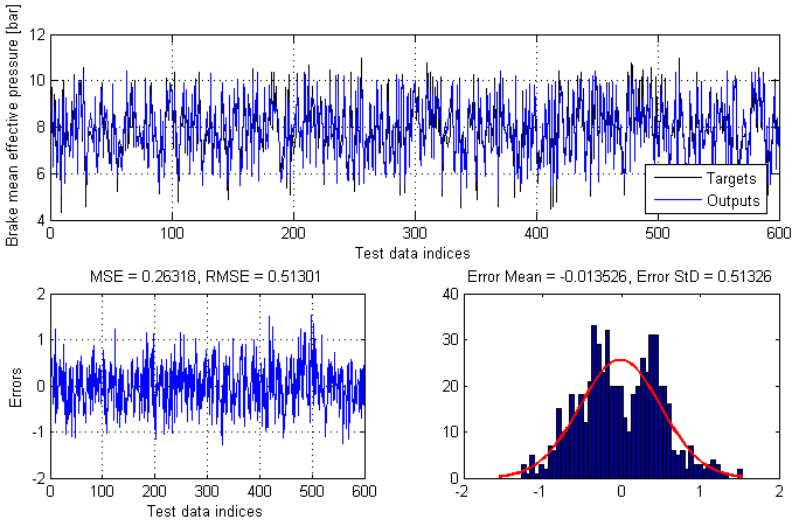

Figure 8. Statistics for the testing of the brake mean effective pressure model.

3.2. NSGA-II Multi-Objective Optimization Results

Figure 9 shows the final 200 Pareto optimal solutions for the objective functions brake mean effective pressure and brake specific hydrocarbon emissions. The points C and D (Figure 9) represent the minimum and maximum optimal values, respectively, for the brake mean effective pressure and brake specific hydrocarbon emissions. It could be seen from the graph that it is not possible to improve (increase) brake mean effective pressure without worsening (increasing) the brake specific hydrocarbon emissions. This trend generally characterizes a conflicting multi-objective optimization problem. Similar results were obtained in a study [17] involving the multi-objective optimization of engine output torque (an engine performance metric) and fuel consumption based on engine speed and variable valve timing as decision variables.

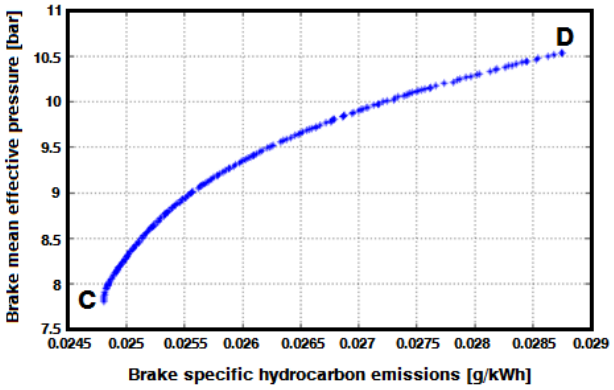

Figure 9. Pareto-based optimal solutions to the multi-objective optimization problem.

Figure 10 shows a two-dimensional scatter plot (involving 200 points) of the decision variables used for the multi-objective optimization. The scatter plot shows that the decision variable (fuel-air ratio) sat at around 0.083, except for a few outliers, with the intake manifold length I_L decision variable varying such that 231 mm $\leqslant I_L \leqslant$ 323.53 mm, between the points marked C and D. For this particular multi-objective optimization problem, if the designer wishes to obtain an optimal solution to the conflicting objectives of maximizing brake mean effective pressure while minimizing the brake specific hydrocarbon emissions, the fuel-air ratio must be kept close to 0.083 while varying the intake manifold between 231 mm and 323.53 mm.

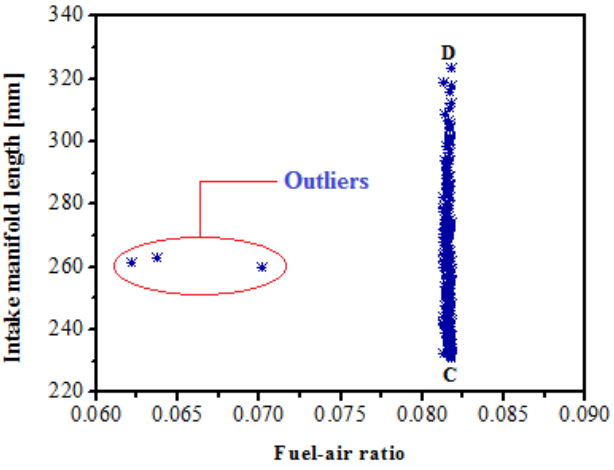

Figure 10. Scatter plot for decision variables.

4. Conclusions

The modeling and multi-objective optimization of brake mean effective pressure and brake specific hydrocarbons have been presented in this study. The modeling process involved the use of a GT-Power™ engine model of a target gasoline engine retrofitted to run on LPG. The multi-objective optimization of the conflicting objectives considered for the study, based on the decision variables of intake manifold length and fuel-air ratio, has also been presented in this work. Based on the current study, the following conclusions could be drawn:

- The prediction accuracy of the engine models could be improved by using GMDH polynomial neural networks. However, the prediction accuracy could be even better if a larger number of training and test data sets were generated from 1D engine models (through Latin Hypercube Sampling) rather the collection of fewer data through experiments.

- For the purpose of achieving optimal results, the relationship between the two decision variables must be such that the fuel-air ratio lies close to 0.083 while the intake manifold length varied between 231 mm and 323.53 mm.

The multi-objective optimization procedure adopted in this study involved the simple case of two decision variables (fuel-air ratio and intake manifold length) and two objective functions. While the current study presented a basic approach for solving a multi-objective optimization problem, future research activities should focus on carrying out similar multi-objective optimizations involving more than two decision variables as well as a higher number of objective functions than considered in the current study.

Acknowledgments: The authors would like to express their deep appreciation to the Hubei Key Laboratory of Advanced Technology for Automotive Components (Wuhan University of Technology) for financial support. Furthermore, R.F. Turkson and M.K.A. Ali wish to acknowledge the Chinese Scholarship Council (CSC) for financial support for their PhD studies in the form of CSC Grant No. 2013GXZ993 and 2014GF032 respectively. M.K.A. Ali also wishes to appreciate the additional financial support from the Egyptian Government.

Author Contributions: This is to certify that all the authors have contributed significantly in diverse ways including design of the study, data collection, analysis/interpretation of the results, and the final preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tums, S.R. *An Introduction to Combustion: Concepts and Applications*, 2nd ed.; McGraw-Hill Education: New Delhi, India, 2000.
2. Babu, M.G.; Subramanian, K. *Alternative Transportation Fuels: Utilisation in Combustion Engines*; CRC Press: Boca Raton, FL, USA, 2013.
3. Lee, E.; Park, J.; Huh, K.Y.; Choi, J.; Bae, C. *Simulation of Fuel/Air Mixture Formation for Heavy Duty Liquid Phase LPG Injection (LPLI) Engines*; SAE International: Warrendale, PA, USA, 2003.
4. Agostinelli, D.; Carter, N.; Fang, G.; Hamori, F. *Optimization of a Mono-Fuel Liquid Phase LPG MPI Fuel System*; SAE International: Warrendale, PA, USA, 2011.
5. Helin, X.; Yusheng, Z.; Huiya, Z. *Experimental and Numerical Study on the Characteristics of Liquid Phase LPG and Diesel Fuel Sprays*; SAE International: Warrendale, PA, USA, 2006.
6. Ramadhas, A.S. *Alternative Fuels for Transportation*; CRC Press: Boca Raton, FL, USA, 2012.
7. Cipollone, R.; Villante, C. *Model-Based A/F Control for LPG Liquid-Phase Injected SI ICEs*; SAE International: Warrendale, PA, USA, 2004.
8. Teene, E.A.A. *Development and Modeling of a Liquid Phase LPG Fuel Injection System*; ProQuest: Ann Arbor, MI, USA, 2007.
9. Dill, S.; Anderson, E.; Taylor, B.; Dunbar, B.; Baker, M.; Tischler, T.; Wandrie, H.; Cheroudi, B.; Pinhas, B. *Conversion of Engine Fuel System from Gasoline Injection to Liquid Propane Injection*; SAE1997 International Congress & Exposition: Detroit, MI, USA, 1997.
10. Shamekhi, A.-M.; Shamekhi, A.H. A new approach in improvement of mean value models for spark ignition engines using neural networks. *Expert Syst. Appl.* **2015**, *42*, 5192–5218. [CrossRef]
11. Tumelaire, C.-F.; Gurney, D.; Cains, T.; Milovanovic, N.; Warth, M. Flexible ECU Function Development Calibration and Engine Performance Assessment Based on Co-Simulation. Available online: http://papers.sae.org/2013-01-0342/ (accessed on 22 December 2015).
12. Wallner, T.; Lohse-Busch, H.; Gurski, S.; Duoba, M.; Thiel, W.; Martin, D.; Korn, T. Fuel economy and emissions evaluation of BMW hydrogen 7 mono-fuel demonstration vehicles. *Int. J. Hydrog. Energy* **2008**, *33*, 7607–7618. [CrossRef]
13. Merkisz, J.; Pielecha, J.; Gis, W. Gaseous and Particle Emissions Results from Light Duty Vehicle with Diesel Particle Filter. Available online: http://papers.sae.org/2009-01-2630/ (accessed on 22 December 2015).
14. Merkisz, J.; Pielecha, J.; Pielecha, I. Road test emissions using on-board measuring method for light duty diesel vehicles. *Jordan J. Mech. Ind. Eng.* **2011**, *5*, 89–96.
15. Lumsden, G.; Browett, C.; Taylor, J.; Kennedy, G. *Mapping Complex Engines*; SAE International: Warrendale, PA, USA, 2004.

16. Martyr, A.J.; Plint, M.A. *Engine Testing: The Design, Building, Modification and Use of Powertrain Test Facilities*; Elsevier: Philadelphia, PA, USA, 2012.
17. Atashkari, K.; Nariman-Zadeh, N.; Gölcü, M.; Khalkhali, A.; Jamali, A. Modelling and multi-objective optimization of a variable valve-timing spark-ignition engine using polynomial neural networks and evolutionary algorithms. *Energy Convers. Manag.* **2007**, *48*, 1029–1041. [CrossRef]
18. D'Errico, G.; Cerri, T.; Pertusi, G. Multi-objective optimization of internal combustion engine by means of 1d fluid-dynamic models. *Appl. Energy* **2011**, *88*, 767–777. [CrossRef]
19. Martínez-Morales, J.D.; Palacios-Hernández, E.R.; Velázquez-Carrillo, G.A. Modeling and multi-objective optimization of a gasoline engine using neural networks and evolutionary algorithms. *J. Zhejiang Univ. Sci. A* **2013**, *14*, 657–670. [CrossRef]
20. Heywood, J.B. *Internal Combustion Engine Fundamentals*; McGraw-Hill: New York, NY, USA, 1988; Volume 930.
21. Saravanan, D.; Gokhale, A.; Karthikeyan, N. Design and Development of a Novel Charge Boosting System for a Single Cylinder SI Engine. Available online: http://papers.sae.org/2014-01-1707/ (accessed on 22 December 2015).
22. Taylor, J.; Gurney, D.; Freeland, P.; Dingelstadt, R.; Stehlig, J.; Bruggesser, V. Intake Manifold Length Effects on Turbocharged Gasoline Downsizing Engine Performance and Fuel Economy. Available online: http://papers.sae.org/2012-01-0714/ (accessed on 22 December 2015).
23. Malkhede, D.N.; Khalane, H. Maximizing Volumetric Efficiency of IC Engine through Intake Manifold Tuning. Available online: http://papers.sae.org/2015-01-1738/ (accessed on 22 December 2015).
24. Vaughan, A.; Delagrammatikas, G.J. A High Performance, Continuously Variable Engine Intake Manifold. Available online: http://papers.sae.org/2011-01-0420/ (accessed on 22 December 2015).
25. Anderson, M.J.; Whitcomb, P.J. Design of experiments. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons: Hoboken, NJ, USA, 2000.
26. Hockman, K.K.; Berengut, D. Design of experiments. *Chem. Eng.* **1995**, *102*, 142–148.
27. Park, G.-J. Design of experiments. In *Analytic Methods for Design Practice*; Springer: London, UK, 2007; pp. 309–391.
28. Ivakhnenko, A. Polynomial theory of complex systems. *IEEE Trans. Syst. Man Cybern.* **1971**, *4*, 364–378. [CrossRef]
29. Farlow, S.J. *Self-Organizing Methods in Modeling: GMDH Type Algorithms*; CRC Press: Boca Raton, FL, USA, 1984; Volume 54.
30. Mueller, J.-A.; Lemke, F. Self-organising data mining: An intelligent approach to extract knowledge from data. Pub. Libri: Hamburg, Germany, 2000.
31. Iman, R.L. Latin hypercube sampling. In *Encyclopedia of Quantitative Risk Analysis and Assessment*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
32. Stein, M. Large sample properties of simulations using latin hypercube sampling. *Technometrics* **1987**, *29*, 143–151. [CrossRef]
33. Huntington, D.; Lyrintzis, C. Improvements to and limitations of latin hypercube sampling. *Probab. Eng. Mech.* **1998**, *13*, 245–253. [CrossRef]
34. Florian, A. An efficient sampling scheme: Updated Latin hypercube sampling. *Probab. Eng. Mech.* **1992**, *7*, 123–130. [CrossRef]
35. Golub, G.H.; Reinsch, C. Singular value decomposition and least squares solutions. *Numer. Math.* **1970**, *14*, 403–420. [CrossRef]
36. Vetterling, W.T.; Teukolsky, S.A.; Press, W.H. *Numerical Recipes: Example Book (C)*; Press Syndicate of the University of Cambridge: Cambridge, UK, 1992.
37. Coello, C.; Christiansen, A.D. Multiobjective optimization of trusses using genetic algorithms. *Comput. Struct.* **2000**, *75*, 647–660. [CrossRef]
38. Osyczka, A. Multicriteria optimization for engineering design. *Des. Optim.* **1985**, *1*, 193–227.
39. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Hoboken, NJ, USA, 2001; Volume 16.
40. Coley, D.A. *An Introduction to Genetic Algorithms for Scientists and Engineers*; World Scientific: Singapore City, Singapore, 1999.

41. Knowles, J.; Corne, D. *The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Pareto Multiobjective Optimisation, Proceedings of the IEEE Congress on Evolutionary Computation*; Washington, DC, USA, 6–9 July 1999, CEC: Washington, DC, USA, 1999.
42. Zitzler, E. *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*; Citeseer: Princeton, NJ, USA, 1999; Volume 63.
43. Ahmadi, M.H.; Ahmadi, M.-A.; Mehrpooya, M.; Rosen, M.A. Using GMDH neural networks to model the power and torque of a stirling engine. *Sustainability* **2015**, *7*, 2243–2255. [CrossRef]
44. Nelles, O. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*; Springer Science & Business Media: Berlin, Germany, 2001.