Uncertainty quantification using Auto-tuned Surrogates of CFD model Simulating Supersonic flow over tactical missile body

Srinivas Soumitri Miriyala
Department of Chemical Engineering
Indian Institute of Technology, Hyderabad
Hyderabad, India
ch13m15p000002@iith.ac.in

Raja Banerjee
Department of Mechanical and Aerospace Engineering
Indian Institute of Technology, Hyderabad
Hyderabad, India
rajabanerjee@mae.iith.ac.in

Kishalay Mitra
Department of Chemical Engineering
Indian Institute of Technology, Hyderabad,
Hyderabad, India
kishalay@che.iith.ac.in

***Abstract*—In contemporary practices, Computational Fluid Dynamics (CFD) based tools are increasingly applied to build high fidelity First Principles based Models (FPMs) for designing tactical missile systems. However, optimization, sensitivity analysis and uncertainty quantification using such models still remain to be extremely tedious and, hence, are performed offline. Artificial Neural Networks (ANNs) are known to be efficient machine learning based models capable of modelling large amount of nonlinearities in the data. However, heuristics involved in their modelling prevent their application as surrogate models to computationally intensive FPMs. In this work, a novel algorithm, aimed at simultaneous optimal estimation of architecture (number of hidden layers and nodes in each layer), training sample size and activation function in ANNs is proposed. The proposed algorithm classifies as a generic multi-objective evolutionary neural architecture search strategy to design ANNs. It is solved using the population based evolutionary optimization algorithm called Non-dominated Sorting Genetic Algorithm-II. The high fidelity data to train the ANNs are obtained from CFD model for simulating supersonic flow over a tactical missile body in ANSYS Fluent©. The obtained ANNs with a test set accuracy of around 99% are then used for quantifying uncertainties, due to order of discretization, flux and Spalart-Allmaras turbulence model in CFD simulations, on the coefficients of lift, drag and rolling moment using the method of analysis of variance (ANOVA). The applicability of optimally designed MLP surrogates for performing ANOVA based study presents novel insights into the epistemic uncertainties arising due to the CFD model which can help in realistic designs of tactical missiles.**

Keywords—Surrogate models, Neural Architecture Search, Uncertainty quantification, ANOVA, CFD based Epistemic Uncertainties, Missile design, NSGA-II

I. Introduction

Contemporary engineering practices have been witnessing a tremendous surge in the implementation of complex mathematical First Principles based Models (FPMs) with the advent of high performance computers. However, optimization, control, sensitivity analysis and uncertainty studies with these robust models still remain to be extremely tedious, mainly due to iterative nature of these algorithms involving repetitive time-expensive function evaluations.

978-1-7281-2547-3/20/$31.00 ©2020 IEEE

The situation aggravates further, when the considered model involves large number of coupled differential equations or detailed Computational Fluid Dynamics (CFD) based solvers whose single simulation consumes large computational time [1]. It is, therefore, nearly impossible to practically implement the optimization, control and uncertainty analysis using such high fidelity models [2]. However, these activities can be accelerated by streamlining them with data-driven models, which work by obscuring the time consuming physics based models, thereby acting as surrogates to FPMs [2].

One classic example of surrogate models is the Artificial Neural Networks (ANN), capable of handling large amount of nonlinearities in the data. ANNs are advantageous over several other prominent classes of surrogates such as Kriging, Response Surface methods, etc., which are used extensively in studies related to uncertainty quantification [3 to 10]. Kwon et al. 2020 proposed a method for surrogate driven uncertainty propagation in complex vehicular design where, a novel adaptive sampling approach was used for minimizing the sample points for training the surrogate models. In this work, the parameters causing the uncertainty were ranked using ANOVA based method [3]. In another recent work, the authors proposed ANOVA based a decomposition scheme to quantify the uncertainties due to physiological inputs. Here, ANNs, SVMs and Adaptive Neuro Fuzzy Inference Systems (ANFIS) are built for predicting the 10-day forecast of a reservoir inflow in China [4]. They concluded that the effect of uncertainties due to interaction between the inputs and data-driven models is more than the individual effects implying the need for designing the surrogate models with high accuracy. Feng et al. 2019 developed a two-stage sensitivity and optimization study for design of plastic injection moulding and its subsequent application to industrial plastic products [5]. In the first stage, the effect of operating conditions on the product characteristics is studied using ANOVA, resulting in reduction of decision variable space. Subsequently, ANN surrogate model was built to map the reduced decision search space with properties and optimization study was conducted using Genetic algorithms. The parametric uncertainty due to Spalart-Allmaras (SA) model for predicting two dimensional aerodynamic stall and compressor stall was studied by He et al. in 2019 using surrogate model and Monte Carlo Simulations [6]. This study

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

provided with rules of thumb for users implementing the turbulence model in CFD simulations through surrogate assisted uncertainty analysis. Similarly, several works from a variety of domains, for example, Carbon dioxide sequestration [7], multi-scale modelling using finite volume methods [8], aerogel glazing system design [9] and so on, are reported in recent literature, where various machine learning models are implemented as alternatives to time consuming first principles based models and uncertainty quantification is performed using these inexpensive surrogate models. However, as detailed in this recent *Nature*© article [10], which describes the necessity for uncertainty quantification while using Artificial Intelligence (AI) and Machine Learning (ML) based models such as ANNs, there is strong need to ensure that the uncertainties and inaccuracies arising from surrogate models be minimized. For instance, ANNs suffer with following major disadvantages:

- ANN design is based on heuristics
- No proper guidelines for activation function choice.
- Training sample size is allocated randomly.
- ANNs are data-greedy and often get Over-fitted

These issues with ANNs prevent them from qualifying as efficient surrogate models. An optimal ANN construction framework capable of simultaneous estimation of all the hyper-parameters is thus, the immediate need of the hour.

With the growth and applicability of ML in various fields, significant amount of research is carried out in developing automated machine learning models capable of smart and intelligent estimation of hyper-parameters [11]. Neural Architecture Search (NAS) is one such topic which aims at eliminating the heuristic based approach for determining the architecture of ANNs [12]. Broadly, the NAS strategies are classified into three categories: a) NAS based on reinforcement learning [13], b) NAS based on unsupervised learning [14] and c) optimization formulation based NAS. While the first two categories are applied to Deep Neural Networks (DNNs), the third category is generally applied for hyper-parameter optimization of shallow neural networks (where the size of network is much smaller compared to DNNs). In [15] a mixed integer nonlinear programming problem (MINLP) is solved for optimal design of ANN. In another work, conflicting scenarios of prediction accuracy maximization for each output are utilized to build neural networks [16]. In this method, the problem scales as a function of number of outputs and thus becomes difficult to generalize. In Carvalho et al. 2011, combination of training and validation error was used as a single objective to design ANNs [17]. However, instead of weighted sum of training and validation error, a more generic Akaike Information Criteria (AIC) [18], might serve a better objective as it penalizes the model for over-fitting. Moreover, the it has been shown in literature that weighted sum approach to solve multi-objective optimization problems (MOOPs) is least efficient among other methods [19]. While most of these NAS strategies address the issue of optimal design of ANNs, none of them speak about other problems faced by ANN modellers. For example, the training size determination and activation function choice is generally governed by heuristics leading to the sub-optimal fitting of ANNs. In [20] in addition to using different surrogates, the problem of sample size estimation is also addressed, however, this method used random sampling plans rather than uniform space-filling sampling strategies. We summarize the existing challenges associated with the methods aimed at optimal design of ANN based surrogate models below:

- Problem of Over-fitting not being addressed.
- Large Computational time of algorithms.
- Not being able to generalize to high-dimensional problems (with respect to number of inputs in ANN).
- No effort to solve the problems of train sample size determination (SSD) and optimal ANN design, simultaneously.

In this paper, we, therefore, aim to present a holistic approach for neural network design along with other hyper-parameter estimation using a multi-objective evolutionary optimization framework. This framework is designed to balance the trade-off between accuracy and parsimony of the multi-layered perceptron networks (MLPs). Further, to prevent the over-fitting of MLP models due to sample size, optimal training sample size estimation algorithm is also developed. The proposed framework, being multi-objective in nature, results in a Pareto front (a list of equally significant solutions – MLP models) from which a single MLP network is selected based on AIC, a critically acclaimed model evaluation criteria known for penalizing the models for over-fitting [18]. Thus, apart from the common techniques to prevent over-fitting such as the early-stopping and regularization, the proposed framework incorporates additional mechanisms (SSD algorithm and AIC) to prevent over-fitting of MLPs. The objectives of this paper are a) to utilize the proposed algorithm for developing accurate MLP surrogate models to emulate highly non-linear and time expensive CFD model used for simulating supersonic flow over a tactical missile body and then b) utilize the developed optimal MLPs to quantify epistemic uncertainties arising due to order of discretization, flux and turbulence models used in CFD simulations on the coefficients of drag (C_D), lift (C_L) and rolling moment (C_M) using decomposition based approach of ANOVA. The coefficients of drag, lift and rolling moment are parameters of prime importance in the design and manufacturing studies of tactical missiles. Since these are always estimated using CFD simulations, it becomes imperative to model the uncertainties in these parameters to provide realistic estimations.

The rest of the paper is organized as follows. Section II describes the CFD model in brief followed by the description about proposed multi-objective evolutionary NAS algorithm and a short description of applying surrogate assisted ANOVA for uncertainty quantification of CFD model parameters. Section III presents the results and discussions before concluding the work in Section IV.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

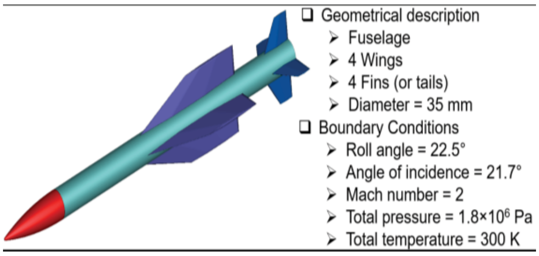

Figure 1. A typical picture of cruciform tactical missile with geometrical description and boundary conditions obtained from literature [21].

II. FORMULATION

A. CFD model for supersonic flow over tactical missile body

A cruciform tactical missile system is considered in the current work. The configuration of the missile is given by ogive-cylinder-fuselage-wings-fins, as shown in Fig. 1.

The missile geometry is taken from Khalil et al. 2019 [21]. While undergoing flight, the missile undergoes through several forces and moments. The predominant ones are the lift forces (F_L) and drag (F_D) generated due to pressure differences and skin friction, respectively. The pressure and viscous forces are defined using Eq. 1 and Eq. 2, respectively, where P_F is pressure, τ_w is the wall shear stress, which is estimated using the law of the wall functions and $d\vec{A}$ is the surface area vector.

$$\vec{F}_P = \int -P_F d\vec{A} \tag{1}$$

$$\vec{F}_V = \int \tau_w d\vec{A} \tag{2}$$

The missile body is subjected to rolling moment, generated due to the interaction between the vortices and wings, which increases with angle of attack and reaches a critical value at a roll angle of 22.5°. Rolling moment $\vec{M}$ is determined using $\vec{F}_P$ and $\vec{F}_V$. The parameters, which effect the missile due to these forces and moment are C_L, C_D and C_M defined in Eqs. 3 to 5, respectively, where ρ is fluid density, A is cross-sectional area of fuselage, u is velocity and L is wing span.

$$C_L = \frac{|\vec{F}_L|}{\frac{1}{2}\rho A u^2} \tag{3}$$

$$C_D = \frac{|\vec{F}_D|}{\frac{1}{2}\rho A u^2} \tag{4}$$

$$C_M = \frac{|\vec{M}|}{\frac{1}{2}\rho u^2 L} \tag{5}$$

A system of governing equations for a single-component fluid, is considered to describe the mean flow properties in integral Cartesian form for an arbitrary control volume V with differential surface area dA using laws of conservation of energy, mass and momentum. The governing equations are solved in CFD to determine C_L, C_D and C_M . Uncertainties arising in C_L, C_D and C_M due to three different sources in CFD model are considered for study in this work, i.e. spatial discretization, flux and turbulence modelling.

Advancement in high performance computing has resulted in two prominent types of spatial discretization, i.e. Flux Vector Splitting (FVS) or Flux Difference Splitting (FDS). Comparatively, FVS methods are robust, but they are not accurate in shear layer regions due to excessive numerical dissipation error. In contrast, FDS methods such as variants of Roe's scheme are accurate in shear layer regions. A hybrid method called Advection Upstream Splitting Method (AUSM) is suggested to resolve the issues with FVS and FDM. The AUSM based scheme is used to simulate flow for the base case.

The default conditions in CFD solver ANSYS FLUENT® enable the evaluation of discrete values of a scalar quantity at the cell centers. However, face values are required for convection terms in governing Naiver Stokes (NS) equation. This requires interpolation of the face values using an upwind scheme where the face values are interpolated from upstream cell values relative to the direction of the normal velocity in NS equation. In this work, second order accurate upwind scheme was considered for the base case, while first and third order were considered for evaluating the uncertainties.

Spalart-Allmaras (SA) model was considered for determining transport equation for turbulent kinematic viscosity owing to the compressible flow with adverse pressure gradients in shock region. The SA Model is one-equation model, which is applicable in wall-bounded flows. Of all the constants in SA model, C_{b1} was found to be most sensitive. Thus, the uncertainty analysis for SA model is performed by varying C_{b1} and assuming the base case value of 0.1355.

The missile geometry is enclosed within a computational domain of hemispherical inlet and cylindrical far-field. The radius of hemisphere was set to 80D where D is the diameter of missile fuselage. The tip of nose was situated at the global coordinate position of (0,0,0). A body-fitted non-orthogonal mesh was created for the missile body with 96, 38 and 266 nodes in radial direction, axial direction in nose and axial direction in fuselage, respectively. A total of 5,316,960 mesh elements was created for simulating the nominal case in

Table 1. A list of all possible sources of uncertainty in the CFD model for supersonic flow over the missile body as studied in literature and their base values.

No.	Sources of Uncertainty	Nominal Case Value
1	Geometrical Parameter	23D/3
2	Turbulence Model	Spalart-Allmaras
3	Viscosity Model	Sutherland
4	Thermal Conductivity	0.0242 W/m K
5	Specific Heat	1006.43 J/kg K
6	Static Temperature	166.67 K
7	Turbulence Intensity	1%
8	Inviscid Flux Type	AUSM
9	Discretization Scheme Order	2nd Order

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

ANSYS Fluent® V13 with base values mentioned in Table 1. Of all the inputs governing the CFD model (as shown in Table 6.1), inviscid flux type, turbulence model, and order of discretization were found to be major sources of uncertainty using a preliminary sensitivity analysis. In the rest of the work, the uncertainty analysis is performed to quantify the effect of these variables on C_L, C_D and C_M defined in Eqs. 3 to 5.

B. Algorithm for Optimal design of MLPs

Geometrically, the hidden layer in neural network represents a hyper-plane capable of linearly segregating the given data. Thus it is evident that more than one hidden layer is needed when considered data is nonlinear. Since the behaviour of data is unknown prior to modelling, it is justified to explore multi-layered perceptron networks. However, a large number of hidden layers leads to several fold increase in number of parameters. This eventually results in over-fitting of the network. Thus, there exists a trade-off between predictability and parameters in the network, which in many terms is similar to the well-known bias-variance trade-off in ANNs.

On the other hand, for an ANN with fixed complexity, the predictability can also be improved by providing with sufficient number of training samples. However, large dataset for training is another reason behind over-fitting of ANNs. Thus, another trade-off between predictability and training size can be observed. The aforementioned two trade-offs are used to design the multi-objective optimization framework proposed in the current work. The proposed algorithm for ANN design is based on this conflict between predictability, parsimony and sample size. A three objective formulation is thus designed based on conflicting objectives of maximizing the ANN predictability on the test set, minimizing the total parameters in the ANN model and minimizing corresponding training points as shown in Table 2.

Table 2. Multi-objective formulation for optimal design of MLPs along with simultaneous estimation of activation function choice and training sample size.

Objectives	Decision Variables
Maximize R^2 Minimize N Minimize P	Nodes in hidden layer 1: $1 \leq N1 \leq 8$
	Nodes in hidden layer 2: $0 \leq N2 \leq 7$
	Nodes in hidden layer 3: $0 \leq N3 \leq 7$
	Activation function N_TF: 1 or 2

Although any number of hidden layers in the MLP can be explored through the proposed algorithm, in this work, MLP architectures with maximum of three hidden layers were evaluated. Therefore, number of nodes in three hidden layers represent the first three decision variables for the MOOP in Table 2 (N1, N2 and N3). Fourth variable is binary representing the activation function choice (1 means log-sigmoidal and 2 means tan-sigmoidal). 0 as lower bounds for number of nodes in second and third hidden layer ensures the emergence of one to three layered networks during optimization. For example, an architecture presented as 3-2-7-1-1 signifies an ANN with 3 inputs and 1 output and three hidden layers with 2, 7 and 1 nodes, respectively. The proposed MOOP formulation is solved using non dominated sorting genetic algorithm, NSGA-II [19]. Each population will correspond to an architecture. Given the architecture, the number of parameters P can be evaluated and after determining the optimal sample size N (as explained in next section), the MLP model can be trained using Levenberg Marquardt algorithm and tested on an unseen dataset to evaluate the accuracy in terms of R^2, shown in Eq. 6, where y is the original output and $\hat{y}$ is the predicted output.

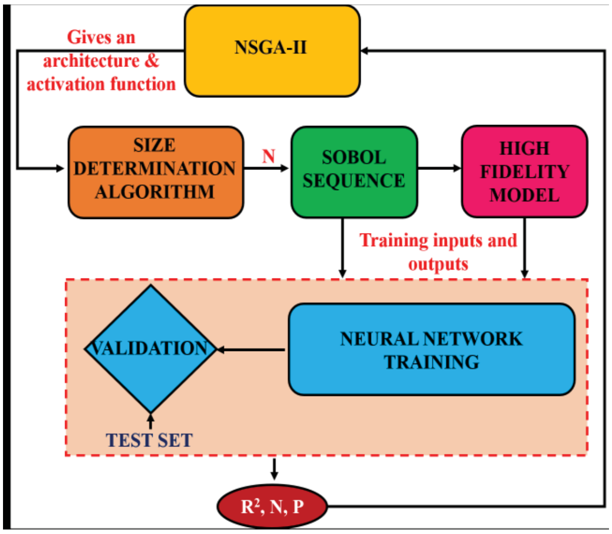

Figure 2. Flowsheet of proposed algorithm for optimal design, activation function estimation and optimal training sample size estimation in MLPs.

$$R^2 = \left(\frac{\mathrm{cov}(y,\hat{y})}{\sqrt{\mathrm{var}(y)\mathrm{var}(\hat{y})}}\right)^2 \quad (6)$$

$$\mathrm{cov}(y,\hat{y}) = n\sum_{i=0}^{n} y^{(i)}\hat{y}^{(i)} - \sum_{i=0}^{n}\hat{y}^{(i)}\sum_{i=0}^{n} y^{(i)}$$

$$\mathrm{var}(y) = n\sum_{i=0}^{n} y^{(i)2} - \left(\sum_{i=0}^{n} y^{(i)}\right)^2$$

This completes the evaluation of all objectives for the given set of decision variables. This procedure is repeated for all population in the generation of NSGA-II. Subsequently, the operations of cross-over, mutation and selection results in the creation of a new generation [19]. This cycle is repeated till a suitable termination criterion (100 generations and 100 populations) of NSGA-II algorithm is met, as shown in the flowsheet in Fig. 2.

C. Algortihm for optimal training sample size estimation

The optimal training sample size determination algorithm proposed in this work draws inspiration from the versatile K-fold Cross Validation approach, a well-established technique to identify the least over-fitted model among a set of alternatives. The methodology of the SSD algorithm using K-fold method is explained briefly below:

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

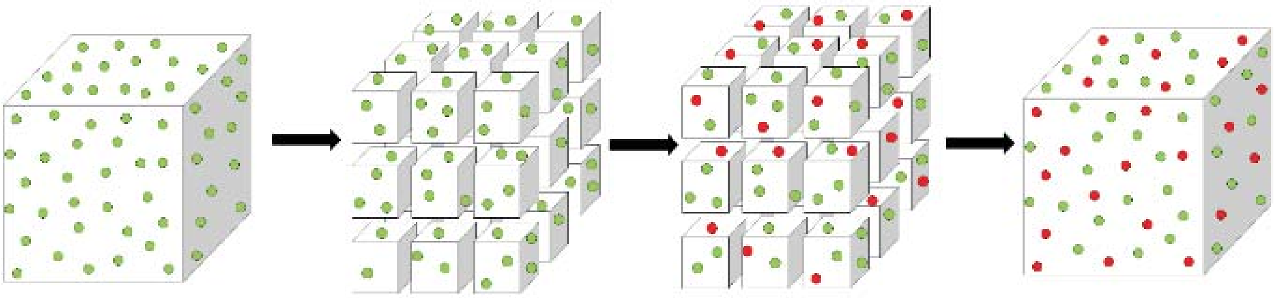

Figure 3. Pictorial representation of hyper-cube method implemented as alternative to expensive K-fold method in proposed sample size determination algorithm.

a) Given an architecture [3-N1-N2-N3-1] the algorithm starts by generating suitable number (K times 3, where K is number of folds and 3 is number of inputs) of data points using Sobol sampling scheme, for which the corresponding output is generated from the first principle model.
b) It then divides this data set into K folds, out of which the architecture is trained with samples from K-1 folds and validated with samples from left over fold to generate the validation error. Since the validation set can be obtained equally likely from all K folds, K validation errors are generated by K times training and validation. The mean of these validation errors is reported as the cross validation error for the architecture and considered sample size.
c) Next the sample size is incremented by a suitable value and steps a) and b) are repeated to generate the corresponding cross validation error for the same architecture.
d) The algorithm is continued iteratively till a sample size is obtained beyond which the increment in sample size doesn't show an improvement in cross validation error.

The repetitive training and validation of the model in the K-fold based approach prevents overfitting. In the aforementioned algorithm, the K-fold method is called repeatedly till convergence. This results in significant increase in computational time. Thus, we propose a simpler and inexpensive alternative to evaluate the Cross validation error in the Step (b) of the aforementioned algorithm. Since the idea of K-fold method is to test the model across all the regions of input space, in the alternative approach, the input domain determined by the given dataset, is divided into hyper-cubes of equal volume and a representative from each hyper-cube is randomly selected to constitute the validation set. The rest of the points are used to train the MLP model and the error obtained by validating this model is considered as an approximation to the cross validation error obtained by the extensive K-fold method. This approach is pictorially represented in Fig. 3.

In this manner, in the algorithm for optimal design of ANNs described in previous section, when the architecture is given by NSGA-II, the optimal sample size is determined. After this, the ANN model is trained with this optimal sample size and tested with unseen data to generate the R^2 metric of Eq. 6.

D. Uncertainty quantification

In this work, Three-way ANOVA [22] was performed to systematically analyze the effect of uncertainty due to the three inputs (X1 = SA model, X2 = Order of Discretization and X3 = Flux) on the outputs Y: Coefficient of Lift, Coefficient of Drag and Coefficient of Rolling moment. Three-way ANOVA specifically studies the influence of second and third order terms – X1X2, …, and X1X2X3 on each of the outputs, i.e., the interaction effects along with main effects. In order to perform ANOVA, the inputs should be categorical in nature. In the present study, order of discretization and flux are categorical in nature – order of discretization has three levels: 1st order, 2nd order and 3rd order and flux has two levels: Roe's scheme and AUSM. However, the SA turbulence model parameter C_{b1} is continuous in nature. In order to perform ANOVA, the range of C_{b1} is divided into three levels, thereby allowing us to consider SA turbulence model parameter as a categorical variable with three levels. The Design of Experiments (DoE) for Three-way ANOVA is shown in Table 3. Using the high fidelity CFD model, sufficiently large number of samples in each cell shown in Table 3, i.e., n cannot be generated due to the extreme computational cost. Thus, in this work, first an optimal MLP based surrogate is built to emulate the high fidelity CFD model accurately and then the Three-way ANOVA is performed using the samples generated from optimal MLP. In this work, multiple input single output MLP models are constructed. Thus, corresponding to the three outputs, three MLPs are designed using proposed algorithm.

Table 3. DoE table for Three-Way ANOVA where the main effects of 3 inputs along with their second and third order interactions are evaluated on the desired output Y. Number of points in each cell = n and number of levels for SA model, flux and order is denoted by a, b and c, respectively.

	$\mathbf{C_{b1j}}$		
$\mathbf{Flux_q}$	1	Y_{1jqk}	$\mathbf{Order_k}$
	2	Y_{2jqk}	
	i	Y_{ijqk}	
	n	Y_{njqk}	

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

III. RESULTS AND DISCUSSIONS

Using the uniform low-discrepancy sampling plan Sobol [23] first a set of 100 3-dimensional input points corresponding to the 3 inputs – C_{b1}, order and flux are generated. These points are utilized to generate high fidelity values of C_D, C_L, and C_M using the CFD model. Out of this data, 20 points were set aside to test the MLP models and the remaining were used in the proposed algorithm. The training size estimation algorithm tried to find out the optimal number of points from the remaining 80 points to train and validate the MLP model. To honor the space constraints results are shown only for one output – C_D.

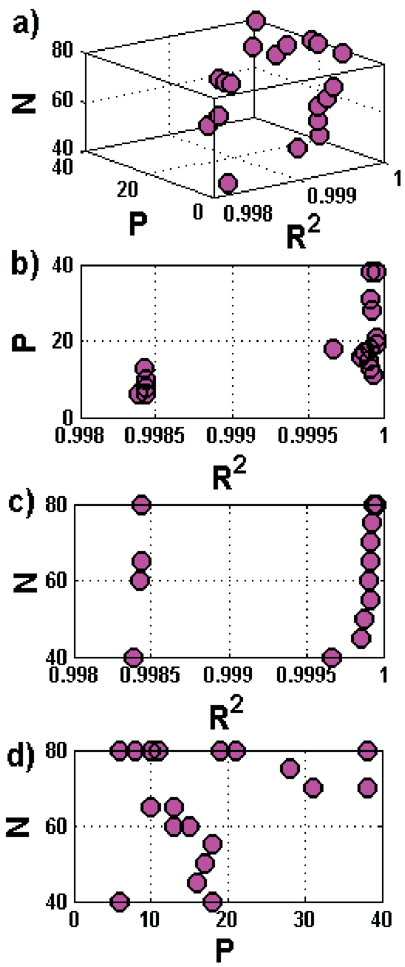

Figure 4. a) 3 dimensional Pareto front (maximize R^2, minimize N and minimize P) obtained from the proposed algorithm. Subfigures b to d present the 2 dimensional representations of the 3 dimensional plot in subfigure a.

Fig. 4 presents 3 dimensional Pareto front obtained from the proposed algorithm. NSGA-II was run for sufficiently large number of generations beyond 100 (~500) to ensure convergence of Pareto front. Each point in this front is an MLP model as shown in Table 4 along with the AIC values. The model with minimum AIC value is considered to be least over-fitted. The selected models for all the three outputs are shown in Table 5.

Table 4. List of Pareto solutions presented in Fig. 4.

N1	N2	N3	N_TF	R^2	P	N	AIC
1	0	0	1	0.998385	6	40	-283.561
1	0	0	2	0.998434	6	80	-688.868
1	1	0	1	0.998434	8	80	-573.975
1	1	0	2	0.998434	8	80	-684.871
1	1	1	1	0.998436	10	80	-570.074
1	1	1	2	0.998431	10	65	-550.869
1	2	1	2	0.998425	13	60	-500.495
2	0	0	2	0.999937	11	80	-936.072
2	1	0	2	0.999912	13	65	-730.999
2	1	1	2	0.999906	15	60	-663.406
2	1	3	1	0.999953	21	80	-833.376
2	2	0	2	0.999871	17	50	-528.579
2	2	1	2	0.999951	19	80	-946.119
2	5	2	1	0.999954	38	80	-800.452
3	0	0	1	0.999854	16	45	-408.081
3	1	0	1	0.999914	18	55	-531.792
3	1	0	2	0.999669	18	40	-376.664
3	3	0	1	0.999924	28	75	-729.482
3	5	0	2	0.999918	38	70	-745.239
6	0	0	1	0.999917	31	70	-664.857

Table 5. List of optimal MLP models selected for emulating the high fidelity CFD model.

Output	N1	N2	N3	N_TF	R^2	P	N
C_D	2	2	1	2	0.999951	19	80
C_L	2	0	0	2	0.999999	11	80
C_M	6	0	0	2	0.991923	31	80

The emergence of multi-layered architectures against the common belief of a single layered architecture for feed forward neural networks justifies the need for the neural architecture search algorithm. The densely populated and widely spread Pareto front shown in Fig. 4 speaks about the importance of the multi-objective evolutionary nature of the proposed NAS algorithm. Finally, even though the optimal sample size estimation algorithm was implemented, it was found out that all 80 points are needed for training the MLP models in case of all three outputs. R^2 values, evaluated on the test set, being close to 1 indicate the high accuracy of the optimally designed surrogate models with least scope for over-fitting. Thus, the proposed method for hyper-parameter optimization of MLP models has been successful in designing optimal and highly accurate auto-tuned MLPs for emulating the high fidelity CFD model. The results of Three-way ANOVA using these models corresponding to the three outputs of CFD model are performed next and presented in Table 6.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

Table 6. Results of optimally designed auto-tuned MLP surrogate assisted Three-way ANOVA for epistemic uncertainty quantification of CFD model on supersonic flow over a tactical missile body. In this table, alpha = significance level, Fcalculated = Test-statistic for F-test = $\frac{\text{Between groups variance}}{\text{Within group variance}}$, Fcritical: obtained from F-distribution table,

If Fcalculated < Fcritical, then the alternate hypothesis is accepted or the null hypothesis is rejected.

Input	Output	alpha	Fcalculated	Fcritical	Decision
Flux	C_D	0.01 (99%)	13.68	6.90	Accept alternate hypothesis
Flux	C_L	0.01 (99%)	17.90	6.90	Accept alternate hypothesis
Flux	C_M	0.01 (99%)	2.38	6.90	Accept null hypothesis
Order	C_D	0.01 (99%)	135.37	4.83	Accept alternate hypothesis
Order	C_L	0.01 (99%)	110.03	4.83	Accept alternate hypothesis
Order	C_M	0.01 (99%)	499.09	4.83	Accept alternate hypothesis
C_{b1}	C_D	0.01 (99%)	0.08	4.83	Accept null hypothesis
C_{b1}	C_L	0.01 (99%)	0.08	4.83	Accept null hypothesis
C_{b1}	C_M	0.01 (99%)	0.12	4.83	Accept null hypothesis
Flux and Order	C_D	0.01 (99%)	64293.84	4.84	Accept alternate hypothesis
Flux and Order	C_L	0.01 (99%)	142634.87	4.84	Accept alternate hypothesis
Flux and Order	C_M	0.01 (99%)	117.66	4.84	Accept alternate hypothesis
Cb1 and Flux	C_D	0.01 (99%)	0.02	4.84	Accept null hypothesis
Cb1 and Flux	C_L	0.01 (99%)	0.02	4.84	Accept null hypothesis
Cb1 and Flux	C_M	0.01 (99%)	0.17	4.84	Accept null hypothesis
Cb1 and Order	C_D	0.01 (99%)	0.10	3.53	Accept null hypothesis
Cb1 and Order	C_L	0.01 (99%)	0.10	3.53	Accept null hypothesis
Cb1 and Order	C_M	0.01 (99%)	0.10	3.53	Accept null hypothesis
Cb1, Flux and Order	C_D	0.01 (99%)	1.71	3.56	Accept null hypothesis
Cb1, Flux and Order	C_L	0.01 (99%)	1.19	3.56	Accept null hypothesis
Cb1, Flux and Order	C_M	0.01 (99%)	1.94	3.56	Accept null hypothesis

If X1, X2 and X3 correspond to the three inputs in the aforementioned table, then for the main effects, i.e., effect of X1 on the uncertainty related to a particular output, the Null hypothesis is, the means of all the groups/levels are statistically equal to each other and the Alternate hypothesis is, at least one of the means differs from other group/level means.

In case of the interaction effects, for X1X2, the Null hypothesis states: there is no statistically significant interaction between X1 and X2 with respect to considered output. Similarly, for X1X2X3, the three-way interaction, the Null hypothesis states: there is no statistically significant interaction among X1, X2 and X3 with respect to the considered output.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.

Following inferences can be drawn from the results of uncertainty analysis reported in Table 6.

- It was found that the variation in Flux had statistically signficant main effect on the outputs C_D and C_L but it had no effect on C_M.
- The variaiton in order had statistically significant effects on all the outputs.
- However, changing only C_{b1}, did not effect any of the outputs, i.e., there was no main effect found due to the variation in C_{b1}.
- This means, the uncertainty in C_D and C_L due to flux and order is significantly high. While the uncertainty in C_M is only due to order of descretization.
- By observing the two-way interactions, it was found that the interactions between Flux and order have significant effect on C_D, C_L and C_M . However, the interactions between C_{b1} and Flux & C_{b1} and order do not effect the outputs.
- Finally there was no significant effect of interactions between all three inputs on any of the outputs C_D or C_L or C_M.

IV. Conclusions

In the present work, the significance of optimally designed MLP networks for uncertainty quantification of computationally intensive CFD model on design parameters associated with supersonic flow of tactical missile body is studied. MLPs were designed optimally using a novel multi-objective optimization framework by balancing the trade-off between accuracy and parsimony in the network. The optimal MLPs were able to emulate the CFD model with R^2 value of 0.99. Subsequently, three-way ANOVA was performed using the optimally designed MLPs to determine the uncertainties arising in coefficients of drag, lift and rolling moment due to order of discretization, type of flux model and turbulence model used in the CFD simulations. The corresponding results were quantified, which is first of its kind, and thus provided unique insights into the design aspects of tactical missile systems.

V. References

[1] A. Mogilicharla, T. Chugh, S. Majumdar, and K. Mitra, “Multi-objective optimization of bulk vinyl acetate polymerization with branching,” Mater. Manuf. Process., vol. 29, no. 2, pp. 210–217, 2014.

[2] Y. Jin, “Surrogate-assisted evolutionary computation: Recent advances and future challenges,” Swarm Evol. Comput., vol. 1, no. 2, pp. 61–70, 2011.

[3] K. Kwon, N. Ryu, M. Seo, S. Kim, T. H. Lee, and S. Min, “Efficient uncertainty quantification for integrated performance of complex vehicle system,” Mech. Syst. Signal Process., vol. 139, p. 106601, 2020.

[4] X. Zhang, H. Wang, A. Peng, W. Wang, B. Li, and X. Huang, “Quantifying the Uncertainties in Data-Driven Models for Reservoir Inflow Prediction,” Water Resour. Manag., vol. 34, no. 4, pp. 1479–1493, 2020.

[5] Q. Q. Feng, L. Liu, and X. Zhou, “Automated multi-objective optimization for thin-walled plastic products using Taguchi, ANOVA, and hybrid ANN-MOGA,” Int. J. Adv. Manuf. Technol., vol. 106, no. 1–2, pp. 559–575, 2020.

[6] X. He, F. Zhao, and M. Vahdati, “Uncertainty Quantification of Spalart–Allmaras Turbulence Model Coefficients for Simplified Compressor Flow Features,” J. Fluids Eng., vol. 142, no. 9, 2020.

[7] B. Chen, D. R. Harp, Y. Lin, E. H. Keating, and R. J. Pawar, “Geologic CO2 sequestration monitoring design: A machine learning and uncertainty quantification based approach,” Appl. Energy, vol. 225, no. May, pp. 332–345, 2018.

[8] S. Chan and A. H. Elsheikh, “A machine learning approach for efficient uncertainty quantification using multiscale methods,” J. Comput. Phys., vol. 354, pp. 493–511, 2018.

[9] Y. Zhou and S. Zheng, “Uncertainty study on thermal and energy performances of a deterministic parameters based optimal aerogel glazing system using machine-learning method,” Energy, vol. 193, p. 116718, 2020.

[10] E. Begoli, T. Bhattacharya, & D. Kusnezov, “The need for uncertainty quantification in machine-assisted medical decision making,” Nature Machine Intelligence, vol. 1, no. 1, pp. 20-23, 2019.

[11] F. Hutter, L. Kotthoff, & J. Vanschoren, “Automated machine learning: methods, systems, challenges,” Springer Nature, pp. 219 2019.

[12] T. Elsken, J. H. Metzen, & F. Hutter, “Neural architecture search: A survey, ” 2018 arXiv preprint arXiv:1808.05377.

[13] Y. Jaafra, J. Luc Laurent, A. Deruyver, and M. Saber Naceur, “Reinforcement learning for neural architecture search: A review,” Image Vis. Comput., vol. 89, pp. 57–66, 2019.

[14] N. Nayman, A. Noy, T. Ridnik, I. Friedman, R. Jin, & L. Zelnik, “Xnas: Neural architecture search with expert advice,” In Advances in Neural Information Processing Systems, pp. 1977-1987, 2019.

[15] V. Dua, “A mixed-integer programming approach for optimal configuration of artificial neural networks,” Chem. Eng. Res. Des., vol. 88, no. 1, pp. 55–60, 2010.

[16] F. Boithias, M. El Mankibi, and P. Michel, “Genetic algorithms based optimization of artificial neural network architecture for buildings’ indoor discomfort and energy consumption prediction,” Build. Simul., vol. 5, no. 2, pp. 95–106, 2012.

[17] A. R. Carvalho, F. M. Ramos, and A. A. Chaves, “Metaheuristics for the feedforward artificial neural network (ANN) architecture optimization problem,” Neural Comput. Appl., vol. 20, no. 8, pp. 1273–1284, 2011.

[18] H. Akaike, “Information theory and an extension of the maximum likelihood principle,” In Proc. 2nd Int. Symposium on Inf. Theory, pp. 267-281, 1973.

[19] K. Deb. Multi-objective optimization using evolutionary algorithms, John Wiley & Sons, 2001.

[20] J. Müller and C. A. Shoemaker, “Influence of ensemble surrogate models and sampling strategy on the solution quality of algorithms for computationally expensive black-box global optimization problems,” J. Glob. Optim., vol. 60, no. 2, pp. 123–144, 2014.

[21] M. Khalil, A. Hashish, and H. M. Abdalla, “A preliminary multidisciplinary design procedure for tactical missiles,” Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng., vol. 233, no. 9, pp. 3445–3458, 2019.

[22] T. C. Urdan, Statistics in plain English, Taylor & Francis, 2016.

[23] I. J. Forrester, A. Sobester, A. J. Keane, Engineering Design via Surrogate Modelling, A practical guide, Wiley, 2008.

Authorized licensed use limited to: Hebei University of Technology. Downloaded on February 02,2023 at 17:00:47 UTC from IEEE Xplore. Restrictions apply.