

# Optimization Formulation Based on Limited Data and RSM: An Approximation

Mohamed H. Gadallah

**Abstract**—In this study, response surface models (RSMs) based on limited data are developed. Experimental cutting force data for the flat end milling process are employed to build these models. Four RSM models are developed in terms of process variables. The first model is used to build the mean cutting force. Similarly, the second, third and fourth models are used to build models for the variance, skewness and curtosis coefficients of the cutting forces. A bi-objective optimization procedure is developed and solved to generate a set of optimal process settings. An approximation scheme resulted in 6 possible objective combinations. The Pareto set of solutions (or part of) are generated for the six possible combinations. The merit of this study lies in the fact that response surfaces are built from limited data, often experienced in reality.

**Index Terms**—Limited Data, RSM, Multi-objective optimization, Approximations.

## Nomenclature

DOE	Design of experiments
OA	Orthogonal Array
RSM	Response Surface Method
ANOVA	Analysis of variance
A, B, C	Three forms of RSMs
UL8	8 experiments OA, 2-levels
UL27	3-Levels, 27 experiments OA
$\beta_0, \beta_i$	RSM model parameters
$X_1$	Depth of Cut
$X_2$	Rev/min
$X_3$	Feed Rate
$X_4$	Tool Diameter
$F_{\text{mean}}, F_{\text{Std}}$	Mean Force & (Standard Deviation)
$\alpha_3$	Skewness of $F_{\text{mean}}$
$\alpha_4$	Curtosis of $F_{\text{mean}}$
$f_1, \dots, f_4$	First, Second.. objective function
$\tau$	Target value of $F_{\text{mean}}$

## I. INTRODUCTION

Multi-disciplinary optimization (MDO) algorithms especially the dimensionality and complexity issues are addressed [16]. Variable fidelity Response Surface

Manuscript received April 16, 2009. Mohamed H. Gadallah is Associate Professor, Department of Operations Research, Institute of Statistical Studies & Research, Cairo University, Cairo, Egypt 12613 (phone: 202-37484624; fax: 202-37482533; e-mail: [Mohamed@aucegypt.edu](mailto:Mohamed@aucegypt.edu)).

Algorithm (RSA) are used to study the convergence using the Trust region algorithm. A comparative study is given for different sampling strategies based on Design of Experiments (DOE). Insufficient space fitting concepts in relation to response surface approximations are addressed.

The problem of capturing Pareto optimal points on non-convex frontiers with the aid of Aggregate Objective Functions (AOF) are studied [12]. Admissibility, necessary and sufficient conditions are discussed. An efficient method-using surrogate modeling to explore the design space is presented [19]. The method captures the Pareto frontier during multi-objective optimization. Issues related to convexity, concavity and function discontinuity are discussed.

An algorithm based on the Clonal selection principle is presented [1]. Results are compared with other evolutionary algorithms. Evolutionary algorithms are claimed to be less sensitive to the shape or continuity of the Pareto Front. Quality metrics such as the two set coverage; spacing and generational distances are proposed to compare solutions. Physical Programming (PP) can foster the design intent and objectives into mathematical models [10]. The Weighted Sum (WS); The Compromise Programming (CP); and the Physical Programming (PP) are examined. The Normal Boundary method is claimed to generate Pareto Frontier in non-convex regions [3]. A similar approach was followed based on robust design optimization method [11].

The Pareto set of bi-criteria problems is a curve approximated by a hyper ellipse [9]. The approximation is achieved by means of a hyper ellipse to a minimum number of points and the hyper ellipse in explicit analytical description. The  $\mathcal{E}$ -constraint method, one criterion is optimized while the others are the additional constraints. The Tchebycheff scalarization finds the Pareto solution by minimizing a distance between the utopia and Pareto set (the utopia point is a point in the objective space obtained by optimizing each criterion independently). This method is capable of fitting only a sector of the Pareto set. The 2 methods:  $\mathcal{E}$ -constraint and weighted Tchebycheff can be used to auxiliary generate Pareto solutions. The approach is circumvented when the Pareto set is adequately small [2]. An adaptive approximate model (AAM) based on polynomial Genetic Programming with partial interpolation strategy is developed [20]. The AAM is sequentially modified in such a way that the quality of fitting can be gradually enhanced. Various transformation methods are evaluated [18]. The WS approach for MOO is employed to study how well different methods in depicting the Pareto optimal sets. Convex combination of functions is desirable when generating the

Pareto set. Advantages of Normalization technique are also demonstrated. Some improvements to the implicit limit state function method are proposed [15]. The response surface is fitted by a weighted regression technique that allows fitting points weighted to their distance from true surface and estimated design point.

II. DEVELOPMENT

Interest in engineering mathematical based models is increasing. Models, once developed go through a series of refinements to describe properly the problem under study. Experimentation can alleviate the problem as it allows understanding of different phenomena involved. In this study, limited experimental data are available. The data represents a certain domain of interest. The modeler would start by developing a low order response surface model. These models are developed for the mean, standard deviation, skewness and curtosis coefficients. Analysis of variance is carried on the four models to verify sufficiency of fit. A multi-objective optimization formulation for the four objectives using a unique approximation is proposed and solved via a nonlinear constrained optimization routine developed within Matlab Environment. The four objective functions result in six bi-objective optimization problems. The designer will have several solutions corresponding to the two objectives employed one-at-a-time. Accordingly, the sequence of objectives has no effect on the resulting solution. Further, the Multi-objective problem is always a bi-objective optimization problem. This removes the issue of high number of objectives often experienced in reality. A physical process with four variables and eight experiments is studied. Both skewness and curtosis are calculated from equation 1.

$$\alpha_3 = \frac{E [X - \mu]^3}{\sigma^3}$$

Skewness Coefficient=

$$\alpha_4 = \frac{E [X - \mu]^4}{\sigma^4}$$

Curtosis Coefficient= 1

Where  $x = F_{mean}$ ,  $\mu = \text{mean}$  of eight force measurements,  $\sigma = \text{the standard deviation of the mean force}$ . The  $F_{mean}$ ,  $F_{std}$ ,  $\alpha_3$  and  $\alpha_4$  are used to develop the models in terms of the four variables and their interactions. Three confidence levels are used to assess adequacy of models at 99%, 95% and 90% confidence levels respectively. Table 1 gives the 8 experiments and corresponding mean and force standard deviations. Table 2 gives the significant variables using  $F_{mean}$  and  $F_{std}$  at different confidence levels. At 90% confidence,  $X_1, X_2, X_3, X_2.X_3, X_1.X_4$  are significant using  $F_{mean}$  as a response. Similarly,  $X_1, X_2, X_3, X_1.X_3, X_2.X_3$  are significant at 90% level using  $F_{std}$  as a response. This procedure is repeated similarly at 95% and 99% confidence levels. The maximum error based on  $F_{mean}$  ranges from 13.76% (at 90% confidence) to 16.33% (at 99% confidence). Moreover, the maximum error based on  $F_{std}$  ranges from 35.50% (at 90% confidence) to

66.67% (at 95% confidence).

**Table 1: UL8 and Mean, Fstd, Skewness and Curtosis Coefficients**

Exp	Input Parameters			
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
1	1.5	800	71	8
2	1.5	800	140	12
3	1.5	1600	71	8
4	1.5	1600	140	12
5	3	800	71	12
6	3	800	140	8
7	3	1600	71	12
8	3	1600	140	8

Exp	Output Parameters			
	F mean	F std	$\alpha_3$	$\alpha_4$
1	44.47	31.60	-0.9895	0.9861
2	100.78	98.15	+0.5420	0.4419
3	32.37	20.77	-2.6616	3.6886
4	64.91	55.09	-0.0388	0.01317
5	70.56	59.45	-0.00387	0.000607
6	138.58	108.10	+8.3859	17.0368
7	67.30	65.09	-0.01796	0.004706
8	84.55	65.51	+0.02518	0.007383

**Table 2: Significant variables based on different confidence levels and UL8**

Confidence Level	90%	95%
Max error based on Fmean	13.76%	16.33%
Max Error based on Fstd	35.59%	66.7%
Based on Fmean	$X_1, X_2, X_3, X_2.X_3, X_1.X_4$	$X_1, X_2, X_3, X_2.X_3$
Based on Fstd	$X_1, X_2, X_3, X_1.X_3, X_2.X_3$	$X_1, X_2, X_3, X_2.X_3$

Confidence Level	99%
Max error based on Fmean	16.33%
Max Error based on Fstd	-
Based on Fmean	$X_1, X_2, X_3, X_2.X_3$
Based on Fstd	-

Accordingly, models developed based on 90% confidence level are the best using the maximum error. Table 3 gives the RSM model coefficients for UL8 experimental model using 3 models: A, B and C respectively. Error ranges from -4.7% - 15.8% for  $F_{mean}$  and 13.9%-28.5% for  $F_{std}$  respectively.

**Table 3: RSM Model Coefficients for UL8**

Model Coef.	Models	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$
Fmean	A	-6.1551	24.403	-0.024	0.630
	B	-10.657	26.227	-0.024	0.669
	C	-96.155	26.2279	0.0471	1.48
Fstd.	A	24.076	6.9533	-0.0443	0.5433
	B	-35.325	33.3538	-0.0443	1.1063
	C	-127.59	33.3538	0.0326	1.9809

Model Coef.	Models	$\beta_4$	$\beta_5$	$\beta_6$	Error
Fmean	A	-0.003	-	-	15.80%
	B	-0.003	-0.0173	-	14.83%
	C	-0.0039	-0.0173	-0.0007	-4.70%
Fstd.	A	0.0071	-	-	28.52%
	B	0.0071	-0.2502	-	18.66%
	C	0.0071	-0.2502	-0.0007	13.19%

**III Bi-Objective Optimization Formulation**

The optimization problem subject to limits on process variables is stated next as:

Minimize  $f_1 = (F_{mean} - \tau)$

Minimize  $f_2 = F_{std}$

Minimize  $f_3 = \alpha_3$

Minimize  $f_4 = \alpha_4$

Subject to:

$$1.5 \leq X_1 \leq 3.0; 800 \leq X_2 \leq 1600;$$

$$71 \leq X_3 \leq 140; 8 \leq X_4 \leq 12$$

Where:

$$F_{mean} = 96.155 + 26.2279 X_1 + 0.0471 X_2 + 1.48 X_3 + 0.0039 X_1 X_2 + 0.0173 X_1 X_3 + 0.0007 X_2 X_3$$

$$F_{std} = 127.59 + 33.3538 X_1 + 0.0326 X_2 + 1.9809 X_3 + 0.0071 X_1 X_2 + 0.2502 X_1 X_3 + 0.0007 X_2 X_3$$

$$\alpha_3 = -14.7977 + 2.8031 X_1 + 0.0094 X_2 + 0.078 X_3 - 0.0026 X_1 X_2 + 0.0207 X_1 X_3 - 0.0001 X_2 X_3$$

$$\alpha_4 = -21.5231 + 0.8015 X_1 + 0.0328 X_2 + 0.0346 X_3 - 0.0080 X_1 X_2 + 0.1027 X_1 X_3 - 0.0002 X_2 X_3$$

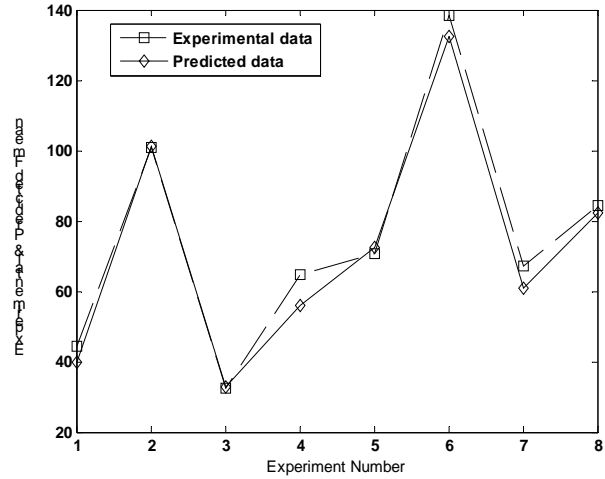
**Approximation:**

A procedure is developed to approximate the sequence of problem approximations. The 4 objectives require  ${}^4C_2 = 6$  approximations as given in Table 4.

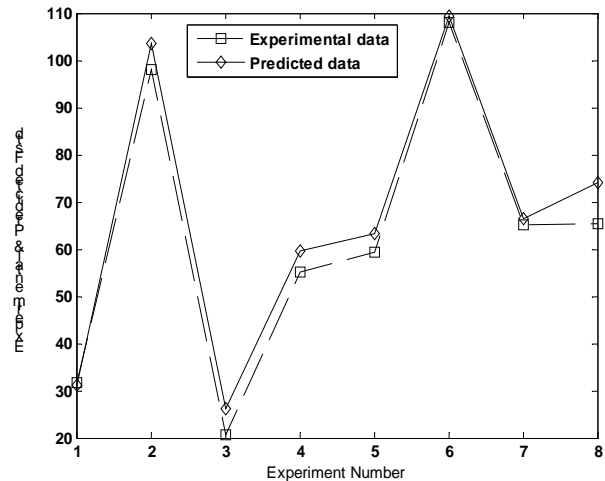
**Table 4: Six possible multi-objective sub-problems.**

Trial	$f_1 = (F_{mean} - \tau)$	$f_2 = F_{std}$	$f_3 = \alpha_3$	$f_4 = \alpha_4$
1	✓	✓	-	-
2	✓	-	✓	-
3	✓	-	-	✓
4	-	✓	✓	-
5	-	✓	-	✓
6	-	-	✓	✓

Figures 1 and 2 give the experimental vs. RSM predicted Fmean and Fstd using UL8. UL8 is an eight trial array and models the linear behavior of the experimental forces. Nonlinearities, once modeled should use three, four, five levels respectively.



**Fig. 1: Exp. versus RSM - Predicted Fmean using UL8 Model**



**Fig. 2: Exp. versus RSM - Predicted Fstd. using UL8 Model**

Different problems are given as:

**Problem # 1:**

Minimize  $f_1 = (F_{mean} - \tau)$

Minimize  $f_2 = F_{std}$

$$1.5 \leq X_1 \leq 3.0; 800 \leq X_2 \leq 1600;$$

Subject to:  $71 \leq X_3 \leq 140; 8 \leq X_4 \leq 12$

and

$$\alpha_3 = 14.7977 + 2.8031 X_1 + 0.0094 X_2 + 0.078 X_3 - 0.0026 X_1 X_2 + 0.0207 X_1 X_3 - 0.0001 X_2 X_3$$

$$\alpha_4 = 21.5231 + 0.8015 X_1 + 0.0328 X_2 + 0.0346 X_3 - 0.0080 X_1 X_2 + 0.1027 X_1 X_3 - 0.0002 X_2 X_3$$

**Problem # 2, 3, 4, 5 and 6 can be similarly stated.**

Figure 3 gives the 6-possible approximations for the bi-objective optimization problem.

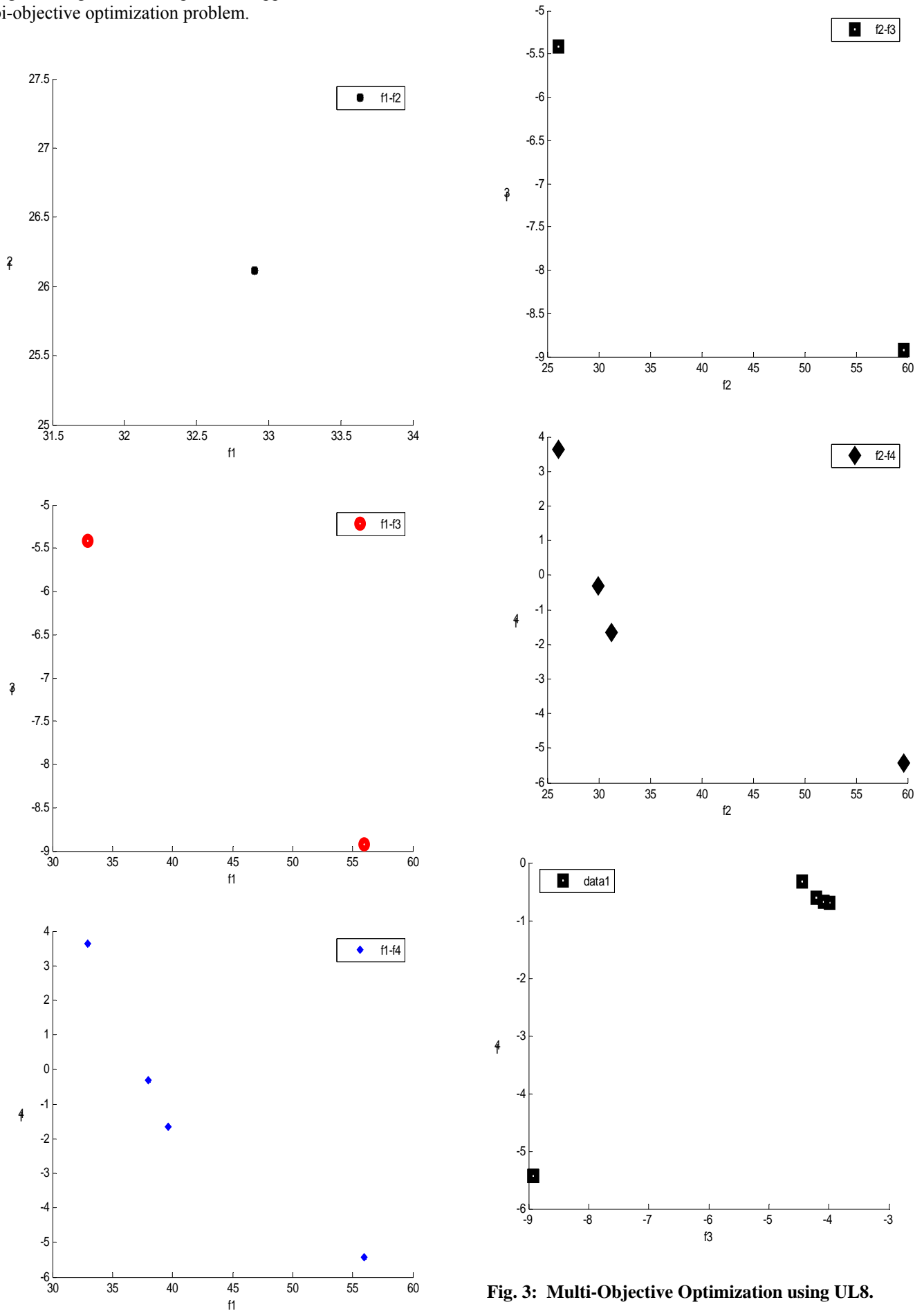


Fig. 3: Multi-Objective Optimization using UL8.

The solutions of the multi-objective problems are shown in Appendix I. Problem 1 is a bi-objective optimization in the f1-f2 domain.  $X^* = (1.5, 1600, 71, 8)$  and  $f_1^* = 32.9044$  and  $f_2^* = 26.1183$ . In problem 2, the f1-f3 domain,  $X^* = (1.5, 1600, 71, 8)$  and  $X^* = (1.5, 1600, 140, 8)$ . The point  $X^* = (1.5, 1600, 71, 8)$  repeats twice regardless of the bi-objective problem solved. A look at the optima generated by the six sub-problems, the point  $X^* = (1.5, 1600, 71, 8)$  is a common solution to all the domains. Hence, this point is certainly a point on the Pareto Frontier. The two points  $(1.5, 1600, 71, 8)$  and  $(1.5, 1600, 140, 8)$  require a bit of attention as the UL8 allows the variation of 4 variables in 2 levels. This means, 16 experiments are needed instead of 8. Since we have used only 50% of what is required, this means that  $X_3$  &  $X_4$  could have been confounded. The consequence could imply that the two points may in fact be one point. Similarly, the point  $(1.5, 800, 71, 8)$  in the f1-f4 domain appears in the f2-f4 domain. The same can be stated for the point  $(1.5, 1000, 71, 8)$  in the f1-f4 and f2-f4 domains respectively. The methodology developed can be depicted graphically in figure 4. Similarly, two UL27 (different domains), UL25 and UL32 are employed to develop the four equations (not shown for brevity). The resulting optima are compared and discussed versus the type and nature of array. These optima are compared via several quality indices such as: cost of solution, stability of solution, uniformity of solution, etc. Appendix II gives measurements and ANOVA for UL8. Appendix III gives the experimental vs. Predicted forces using UL27.

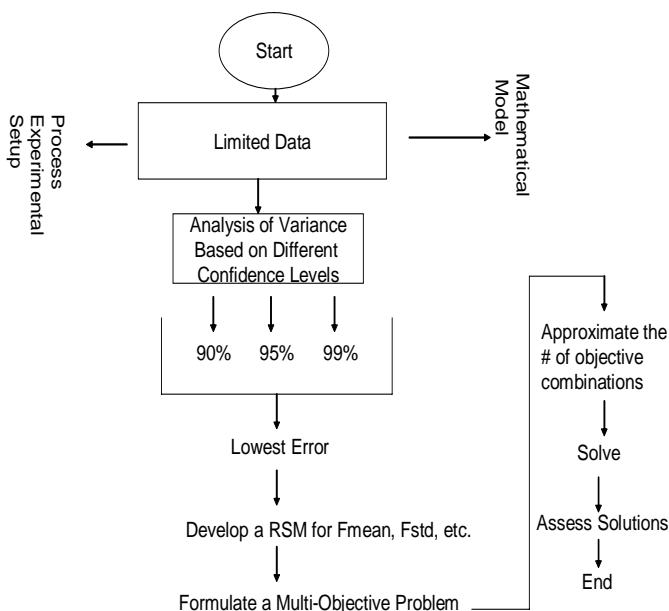


Fig. 4: Flow Chart of Proposed Methodology

#### IV Conclusion

A procedure is given to develop mathematical models from limited data points. RSM is employed to optimize a set of bi-objective problems. The 2-objective functions represent the majority of optimal solutions in their planes according to the approximation given

#### Appendix I: Multi-objective Solutions

Objectives	$X^*$	$f^* = [f_1^*, f_j^*]$	#
$f_1$ & $f_2$	(1.5, 1600, 71, 8)	[32.9044, 26.1183]	1
$f_1$ & $f_3$	(1.5, 1600, 71, 8)	[32.9044, -5.4105]	2
	(1.5, 1600, 140, 8)	[55.9538, -8.9261]	
	(1.5, 1600, 71, 8)	[32.9044, -5.4105]	
$f_1$ & $f_4$	(1.5, 1600, 140, 8)	[55.9538, -5.4298]	3
	(1.5, 800, 71, 8)	[39.6641, -1.6467]	
	(1.5, 1000, 71, 8)	[37.9744, -0.3267]	
	(1.5, 1600, 71, 8)	[32.9044, 3.6333]	
$f_2$ & $f_3$	(1.5, 1600, 140, 8)	[52.6247, -8.9261]	4
	(1.5, 1600, 71, 8)	[26.1183, -5.4105]	
$f_2$ & $f_4$	(1.5, 1600, 140, 8)	[59.6247, -5.4298]	5
	(1.5, 800, 71, 8)	[31.2783, -1.6467]	
	(1.5, 1000, 71, 8)	[29.9883, -0.3267]	
	(1.5, 1600, 71, 8)	[26.1183, 3.6333]	
$f_3$ & $f_4$	(1.5, 1600, 140, 8)	[-8.9261, -5.4298]	6
	(1.5, 965, 103.8, 8)	[-3.9837, -0.7004]	
	(1.5, 972, 98.16, 8)	[-4.0844, -0.6675]	
	(1.5, 976, 89.04, 8)	[-4.2074, -0.6017]	
	(1.5, 1600, 71, 8)	[-4.4505, -0.3267]	
	(1.5, 1600, 140, 8)	[-8.9261, -5.4298]	

#### Appendix II: UL8OA & experimentation using either Fmean and Fstd respectively.

#	$X_1$	$X_2$	$X_1.X_2$	$X_3$	$X_1.X_3$ $X_4$	$X_2.X_3$	$X_1.X_4$
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	1	2
4	1	2	2	2	2	2	1
5	2	1	2	1	2	2	2
6	2	1	2	2	1	1	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

#	$F_{mean}$	$F_{std}$
1	44.47	31.60
2	100.78	98.15
3	32.37	20.77
4	64.91	55.09
5	70.56	59.45
6	138.58	108.1
7	67.30	65.09
8	84.55	65.51
T	603.52	503.76

#### ANOVA Results based on Fmean & Fstd using UL8.

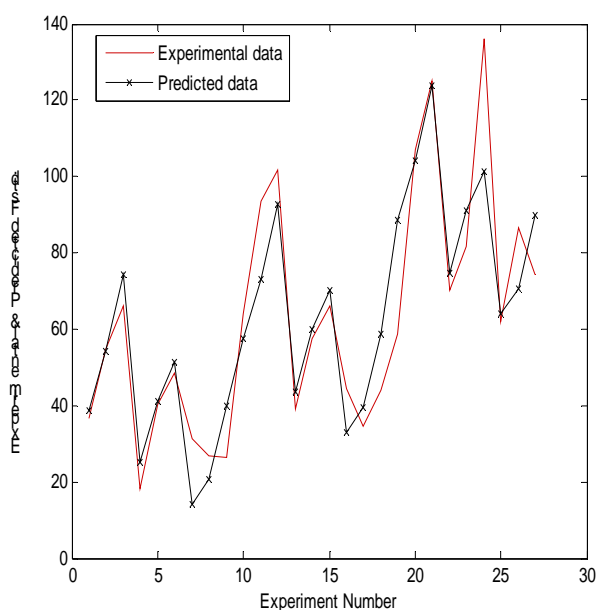
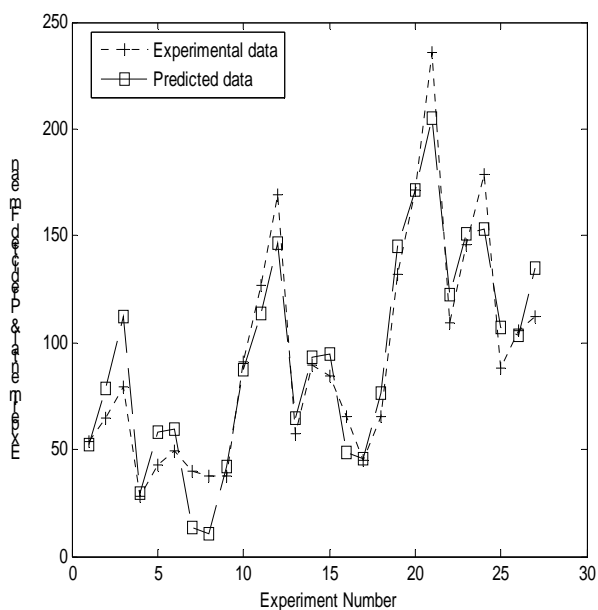
Source	SS	DOF	Variance	Fcal
$X_1$	1754	1	1754	281.2
$X_2$	1384	1	1384	222.03
$X_3$	3789	1	3789	607.57
$X_1.X_4$	91	1	91	14.609
$X_2.X_3$	694	1	694	111.34
Error	12.475	2	6.2375	
SST	7726.88	7		

$F_{1,2,90\%} = 8.53$ ,  $F_{1,2,95\%} = 18.5$ ,  $F_{1,2,99\%} = 93.5$

Source	SS	DOF	Variance	Fcal
X <sub>1</sub>	1070	1	1070	31.52
X <sub>2</sub>	1031	1	1031	30.37
X <sub>3</sub>	2810	1	2810	82.75
X <sub>4</sub> (X <sub>1</sub> X <sub>3</sub> )	335.4	1	335.4	9.877
X <sub>1</sub> .X <sub>4</sub>	32	1	32	~1
X <sub>2</sub> X <sub>3</sub>	809.22	1	809.22	23.83
Error		1		
SST	6124.7	7		

F<sub>1,2,90%</sub> = 8.53 , F<sub>1,2,95%</sub> = 18.5 , F<sub>1,2,99%</sub> = 93.5

### Appendix III: Exp vs. Predicted Fmean and Fstd using UL270A.



### Exp vs. RSM Predicted Force Components using UL27-2 OA

### ACKNOWLEDGMENT

I would like to acknowledge the valuable discussions with Professor M. Bayoumi from Cairo University.

### REFERENCES

- [1] Coello, C.A.C., Cortes, N.C., "Solving Multi-Objective Optimization Problems using an Artificial Immune System", Kluwer Academic Publishers, 2002, pp. 1-35.
- [2] Christopher, A.M., Anoop, A.M. and Messac, A., "Smart Pareto Filter: Obtaining A Minimal Representation of Multi-objective Design Space", Engineering Optimization, Vol. 36 no. 6, 2004, pp. 721-740.
- [3] Das, I. and J.E.Dennis, "Normal Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multi-Criteria Optimization Problems", Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1996.
- [4] Fabio, F. and Mauizio, R., "VIS: An Artificial Immune Network for Multi-Objective Optimization", Engineering Optimization, Vol. 38, no.8, 2006, pp.975-996.
- [5] Hongbing, Fang and Mark F. Horstemeyer, "Global Response Approximation with Radial Basis Functions", Engineering Optimization, Vol. 38, no. 4, 2006, pp. 407-424.
- [6] Jenn-Long Liu, "Novel Orthogonal Simulated Annealing with Fractional Factorial Analysis to Solve Global Optimization Problems", Engineering Optimization, Vol. 37 no. 5, 2005, pp. 499-523.
- [7] Janga R. and Nagesh, K., "An Efficient Multi-Objective Optimization Algorithm based on Swarm Intelligence for Engineering Design", Engineering Optimization, Vol. 39, no. 1, 2007, pp. 49-68.
- [8] Kok, S. Won and Tapabrata Ray, "A Framework for Design Optimization using Surrogates", Engineering Optimization, Vol. 37 no. 7, 2005, pp. 685-703.
- [9] Li, Y., Fadel, G.M. Wiecek, M., Blouin, V., "Minimum Effort Approximation of the Pareto Space of Convex Bi-Criteria Problems", J. of Optimization and Engineering, Vol. 4, 2003, pp. 231-261.
- [10] Messac, A. and Mattson, C., "Generating Well Distinguished Sets of Pareto Points frontier Engineering Design using Physical Programming", Optimization and Engineering, Vol. 3, 2002, pp. 431-450.
- [11] Messac, A., Yahia, I., "Multi-Objective Robust Design Using Physical Programming", Structural and Multi-Disciplinary Optimization, Vol. 23, No. 5, 2002, pp.357-371.
- [12] Messac, A., Sukam, P.C., Melachrinoudis, E., "Aggregate Objective Functions and Pareto Frontiers: Required Relationships and Practical Implications", Optimization and Engineering, Vol. 1, 2000, pp.171-188.
- [13] Matlab user Manual, Version 7, The Math Works Inc., 2004.
- [14] Narayanan, A., Toropov, V.V., Wood, A.S., Campean, I.F. "Simultaneous Model Building & Validation with Uniform Design of Experiments", Engineering Optimization, Vol. 39, no. 5, 2007, pp. 497-512.
- [15] Nguyen X.S. Sellier, A., Duprat, F., Pons G. "Adaptive Response Surface Method based on a Double Weighted Regression Technique" Probabilistic Engineering Mechanics Vol. 24 2009, pp. 135-143.
- [16] Rodriguez J.F., Perez, V. M. Padmanabhan, D. and Renaud, J.E. "Sequential Approximate Optimization using Variable Fidelity Response Surface Approximations" Structural Multi-Disciplinary Optimization Vol. 22, 2001 pp. 24-34.
- [17] Shin, S., A New Paradigm for Integrated Bi-Objective Robust-Tolerance Design Modeling and Optimization, Ph.D. Dissertations, Clemson University, 2005.
- [18] Timothy, R.M., Arrora, J.S., "Function Transformation Methods for Multi-Objective Optimization", Engineering Optimization, Vol. 37 no. 6, 2005, pp. 551-570.
- [19] Wilson, B., Cappelleri, D., Simpson, T. Frecker, M., "Efficient Pareto Frontier Exploration using Surrogate Approximations", Optimization and Engineering, Vol. 2, pp. 31-50, 2001.
- [20] Yeun, Y.S., Kim B.J. Yang, Y.S. and Ruy, W.S., "Polynomial Genetic Programming for RSM", Technical Report, 2004.