

크리깅을 이용한 개선된 최적화 알고리즘에 관한 연구

An Improved Algorithm for Practical Optimal Designs by Kriging

노영희*, 임종빈, 박정선(한국항공대학교)

1. 서론

최근 많은 물리적 현상에서 서로 다른 위치에 있는 자료들이 공간적인 상호작용에 의해 특정한 영향을 받는 형상에 대한 모델링이 연구되고 있다. 특히, 통계적 기법을 이용한 근사 모델에 대한 연구가 활발하며, 그 중 비선형성이 강한 모델에 대해 적합하다고 알려진 메타 모델의 구현 방법이 크리깅이다. 하지만 이것은 다른 메타모델에 비해 다소 구현하기 어렵고 오랜 계산시간이 소요되며, 크리깅의 특성상 경험적인 방법에 의한 상관함수의 선택이 근사 모델의 정확성에 영향을 미치게 된다[1]. 이러한 문제점을 해결하기 위해 효율적인 크리깅을 이용한 최적화 알고리즘에 관한 연구가 활발히 진행 중이며, 그 중에서도 설계영역을 이동해가며 최적값을 찾는 반복, 순차적인 근사 최적화 알고리즘에 대한 연구도 그 중 하나이다. 따라서 본 연구에서는 기존의 크리깅에 관한 통계·수학적인 면에서의 접근과 더불어 이를 이용한 개선된 최적화 알고리즘을 제안하여, 단순히 크리깅을 이용한 근사최적화 혹은 크리깅의 통계·수학적인 부분에서만 정확성 향상이 아닌, 전반적으로 개선된 크리깅 메타모델을 이용한 최적화 알고리즘을 구현하는데 본 연구의 목적이 있겠다. 우선 본 연구에서는 적은 실험점으로 가능한 많은 정보를 얻음으로써 크리깅 메타모델의 효율성을 높이기 위하여 고전적인 실험계획법(classical design of experiments) 중 중심합성 계획법(central composite design: CCD)과, 공간충진 실험계획법(space filling experiments) 중 Wing and Fang에 의해 제안된 균일계획법(uniform design: UD)으로 나누어 각각에 대한 결과를 비교·분석하였다[2]. 또한 전역 최적값(global optimum value)과 국부 최적값(local optimum value)을 동시에 만족하는 최적값을 찾기 위한 순차적인 반복에 의한 흥미영역 설정에 있어

확률론적 국부화 기법(stochastic step-reduction method, SSRM)을 이용하여 알고리즘을 구현하였다. 본 연구에서 제안된 최적화 방법에 대한 개선된 알고리즘은 기존의 최적화 예제 중 하나인 pressure vessel(by Sandgren)의 예제에 적용함으로써 그 정확성과 타당성을 비교·검증하고자 하였다.

2. 이론적 배경

2.1 크리깅의 수학적 모형 및 예측 모형

Sacks 등은 크리깅 모형을 다음과 같은 선형 모형(linear model)과 잔차(departure)의 합으로 표현하였다[3].

$$y(x) = f(x) + Z(x) \tag{1}$$

여기서 $y(x)$ 는 우리가 알고자 하는 함수이며, $f(x)$ 는 주어진 실험 데이터를 이용해서 구할 수 있는 전역모델(global model)이다. 또한 $Z(x)$ 는 전역모델로부터 나온 지역적 편차를 생성하여 추출된 데이터 점을 보간하게 된다. 이때 $Z(x)$ 의 공분산(covariance)은 다음과 같이 정의된다.

$$Cov[Z(x^i), Z(x^j)] = \sigma^2 R[R(x^i, x^j)] \tag{2}$$

여기서 σ^2 은 $Z(x)$ 의 분산이며, R 은 상관행렬(correlation matrix), $R(x^i, x^j)$ 는 n 개의 실험점 내 임의의 두 점 사이의 상관함수이다. 이때의 상관함수는 설계자에 의해서 경험적인 방법으로 결정되며, 어떠한 함수를 사용하느냐에 따라 얼마나 유연하게 실험점을 보간해 주는가가 결정된다[3]. 기존의 연구에서는 변수들 간의 관계가 해석적(analytic)이고 반응면이 매끄러운 것으로 알려져 있는 가우스 상관함수(Gaussian correlation function)를 대부분 사용하였으나,

본 논문에서는 가우스 상관함수를 포함하여 Sacks, Guinta 등에 의해 제안된 5개의 상관함수를 사용하였다[3]. 각 상관함수들은 그 형태에 따라 실제함수로 보간 되어질 때 각각의 특징을 가지므로 실제 함수형태에 있어 선형성과 비선형성에 대한 적절한 상관함수의 선택은 정확한 메타모델을 구현하는데 중요한 역할을 한다. 다음으로 실제모델 $y(x)$ 와 오차를 최소화시키는 $\hat{y}(x)$ 가 크리깅 메타모델이 된다.

$$\hat{y} = \hat{\beta} + \gamma^T(x)R^{-1}(y - \hat{\beta}f) \quad (3)$$

여기서 $\hat{\beta} = (f'R^{-1}f)^{-1}f'R^{-1}y$ 은 β 의 일반화된 추정량이며, 식(3)의 우변에 있는 두 항은 상호 연관되지 않으며, 일반화된 최소제곱 예측값을 얻은 다음 회귀모형이 없는 것처럼 잔차를 보간하게 된다. 그리고 실제 함수의 분산에 해당하는 σ^2 에 대한 추정량은 식(4)와 같다.

$$\hat{\sigma}^2 = 1/n(y - f\hat{\beta})R^{-1}(y - f\hat{\beta}) \quad (4)$$

$\gamma(x)$ 는 관측된 지점과 추정하려는 지점에서의 반응값 사이에서의 상관벡터(correlation vector)이며 f 는 길이가 표본의 수이고 모든 요소가 1인 벡터를 의미하고, y 는 표본에서의 반응들로 이루어진 벡터이다. 마지막으로 상관매개변수 θ_k 는 다음의 수식을 최대화함으로서 결정된다.

$$-(\det R)^{\frac{1}{n}} \hat{\sigma}^2 \quad (5)$$

여기서 \det 는 행렬식을 의미하며, n 은 표본의 수, 즉 실험점의 수이다.

본 연구에서는 최적의 크리깅 모델을 구성하기 위해 이산변수에 대한 최적화에 사용되는 simulated annealing을 이용하여 식(5)를 최대화하는 θ_k 를 구하였다.

2.2 순차적 반복의 근사 최적화

설계변수의 영역을 전체영역으로 설정한 후 한번의 근사모델 생성으로 최적화를 시킬 수도 있지만 이는 한번 생성된 크리깅 메타모델이 정확하다는 가정이 필요하며, 만약 그 정확성을 위해 실험점의 수를 늘린다는 것은 효율성의 면에서 오히려 못된 측면이 있다. 그리고 전역적인 최적값(global optimum value)과 동시에 국부적인 최적값(local optimum value)을 만족하는 값을 찾기 위해서, 본 연구에서는 흥미영

역(interest domain)을 반복, 순차적으로 생성해 가며 그 값을 찾는 순차적인 반복의 근사최적화 알고리즘을 구현하였으며 그 알고리즘은 Fig. 1과 같다.

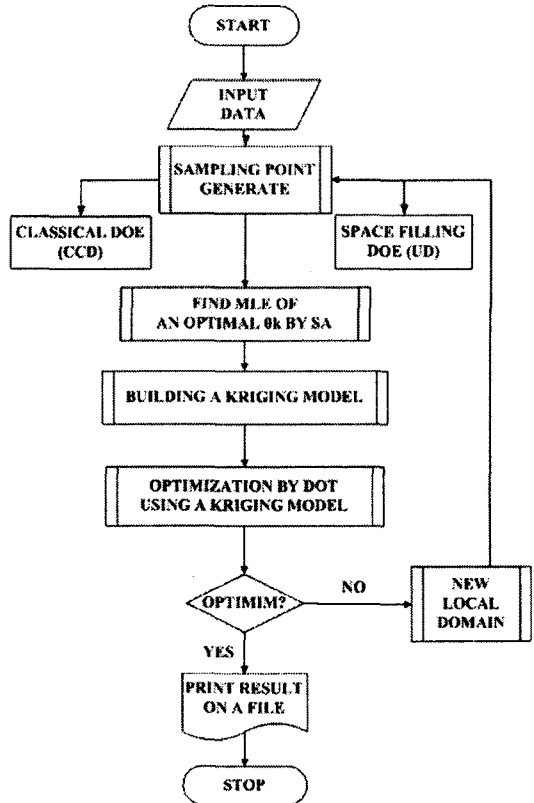


Fig. 1. The flow chart of a sequential approximation optimization

이때 흥미영역을 반복 생성하여 이동해가는 방법에 있어 어떠한 방법을 선택하느냐가 얼마나 정확하고 효율적인 최적값을 찾느냐의 문제와 직결되어 있다. 이에 기존의 연구에서는 초기값과 국부 최적값의 차이를 일정한 비율로 줄여가며 흥미영역을 설정하는 근사화 방법, 주관적인 반응 상한치에 대한 구속을 주어 근사화 하는 방법, 혹은 벌칙함수를 이용한 유용영역의 탐색에 따른 근사화 방법 등이 제안되었다[4]. 이에 본 연구에서는 확률분포를 이용한 통계 추론적(stochastic) 방법을 제안하고 이를 확률론적 국부화 기법(stochastic step-reduction method, SSRM)이라 부르기로 한다.

2.2.1 확률론적 국부화 기법

순차적 반복의 근사 최적화에서 국부적인 설계영역을 만들어가는 방법에 있어 본 연구에서는 확률론적 국부화 기법을 사용하여 설계영역을 순차적으로 반복하여 생성하며, 그 개념은 위의 fig. 2에 있다, 초기 주어진 설계영역에서 실험계획법에 의한 실험점을 선정하며, 이를 바탕으로 크리깅 메타모델을 생성한다.

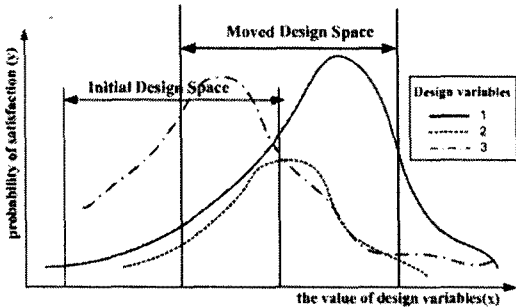


Fig. 2. The concept of a probabilistic design space

그리고 설계영역에서 각 설계변수들의 발생 확률이 균일분포(uniform distribution)를 따르게 약 10,000개의 실험점을 설계영역 내에서 크리깅 메타모델에 의해 선정한 후 제한조건식을 동시에 만족시킬 확률(probability of success: POS)을 구한다. 이때 성공확률은 이항분포(binomial distribution)를 따르게 되며, 이때 이항분포는 다음과 같이 정의된다. 시행횟수 n 이고 성공확률 p 일 때, 이항확률분포(binomial probability distribution)의 확률질량함수는

$$P\{X=x\} = b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \quad (6)$$

$$x = 0, 1, \dots, n$$

이다. 그리고 성공확률 p 는 이항분포의 정규근사 성질에 의해 이는 표준정규 분포를 따르게 되며 이때의 평균 성공확률이 생성된 크리깅 메타모델에 대한 평균 성공확률이라 말할 수 있다. 다음은 이항분포의 정규근사에 관한 이론으로 이항분포 $B(n, p)$ 를 따르는 확률변수 X 에 대해, n 이 충분히 크고 p 가 0 또는 1에 가깝지 않다면 표준화된 확률변수 $\frac{X-np}{\sqrt{npq}}$ 의 분포는 근사적으로 표준정규분포 $N(0, 1)$ 을 따른다.

$$P\{a \leq X \leq b\} \approx P\left\{\frac{a-np}{\sqrt{npq}} \leq Z \leq \frac{b-np}{\sqrt{npq}}\right\} \quad (7)$$

(연속성 수정을 않는 경우)

$$\approx P\left\{\frac{a-0.5-np}{\sqrt{npq}} \leq Z \leq \frac{b+0.5-np}{\sqrt{npq}}\right\} \quad (8)$$

(연속성 수정을 하는 경우)

다음으로 체비셰프 부등식(Chebyshev's Inequality)을 이용하여 설계영역을 이동 시킨 후 이제까지의 과정을 반복하며 설계영역의 변화를 수렴시킨다. 이때 체비셰프 부등식은 X 가 유한평균 μ 와 유한분산 σ^2 을 갖는 확률변수이면, 임의의 $k>0$ 에 대해 다음이 성립됨을 의미한다.

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2} \quad (9)$$

예를 들어 확률변수 X 가 평균으로부터 3σ 안에 있을 확률은 약 90%이며 만약 확률변수 X 가 정규분포를 따른다면 위의 확률은 약 96%가 된다. 이때 체비셰프 부등식의 중요성은 확률분포의 평균만 알거나 평균과 분산을 모두 알고 있을 때 확률의 한계를 유도할 수 있다는 것이며, 만일 분포를 알고 있다면 구하려는 확률을 정확히 계산할 수 있으므로 확률의 한계에 연연할 필요가 없다. 또한 체비셰프 부등식이 확률변수 X 의 모든 분포에 유효하지만, 확률의 한계가 대부분의 경우에 실제 확률에 매우 근사하리라고는 기대할 수 없다. 본 연구에서는 확률변수, 즉 실제 설계변수의 분포를 정규분포로 가정하고, 그 설계영역의 정확한 경계값을 사용했던 기존의 연구와는 달리, 설계자에 의해 임의로 정해진 설계영역에 있어 그 분포의 정규성을 가정할 수 없다고 판단하여 경계의 범위가 포괄적으로 적용할 수 있는 체비셰프 부등식을 이용하여 설계영역을 이동시켰다. 이는 본 연구에 앞서 임의로 정해진 설계영역에서의 설계변수의 분포를 왜도와 첨도를 이용한 Jarque-Bera통계량으로 검정한 결과 정규분포성을 기각하는 결과를 얻을 수 있었음을 반영한 것이다. 그리고 확률론적 국부화 기법에 의한 설계영역 이동 시, 매 순차적으로 흥미영역의 생성 때 크리깅 메타모델에 의한 반응값의 분포에 대한 한계치를 두어 설정된 설계영

역이 제한조건을 동시에 만족할 확률이 높을 뿐 아니라, 최적값을 가질 확률이 높은 설계영역으로 이동하도록 하였다.

3. 최적화 예제

본 연구에서 제안한 크리깅 메타모델을 이용한 개선된 최적화 알고리즘을 실질적인 공학문제의 해결에 적용해보기 위하여 4개의 변수를 포함하는 압력용기(pressure vessel)의 최적화 예제에 적용해 보았다[6]. Fig. 3에서의 압력용기는 Sandgren에 의해 1990년에 continuous solver에 의해 최적화 되어진 문제이며, 이를 본 연구의 알고리즘에 적용한 결과와 비교하여 개선된 알고리즘의 신뢰성 및 정확성을 비교·분석한다.

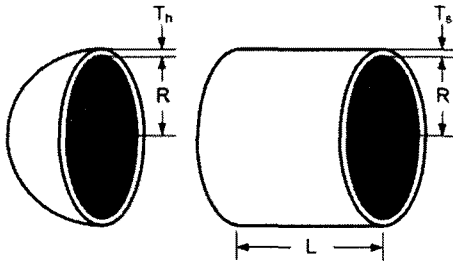


Fig. 3. Pressure vessel

설계변수는 반지름(R), 원통의 길이(L) 및 두께(T_s), 그리고 구형의 상단부분 두께(T_h)이며, 전체 시스템의 비용을 최소화 하는데 그 목적이 있으므로 용접, 재료 등의 비용 등에 대해 전체 시스템의 비용을 정식화 하여 나타낸 식(10)을 최소화 하는 것이다. 식(11)에서 G_1 과 G_2 는 T_s 와 T_h 에 대한 미국기계학회 ASME의 규정에 따른 제한조건식 이며, G_3 는 압력용기 부피에 대한 제한조건식이다. 이 압력용기 재료는 carbon steel ASME SA 203 grade B이며, 3,000psi의 압력에 의해 압축된 공기 $750ft^3$ 을 저장한다[6].

Minimize :

$$OBJ = 0.6224 T_s R L + 1.7781 T_h R^2 + 3.1661 T_s^2 L + 19.84 T_s^2 R \quad (10)$$

Subject to :

$$\begin{aligned} G_1(x) &= T_s - 0.0193R \geq 0 \\ G_2(x) &= T_h - 0.00954R \geq 0 \\ G_3(x) &= \pi R^2 L + (4/3)\pi R^3 - 1.296E6 \geq 0 \end{aligned} \quad (11)$$

$$\begin{aligned} 25 \text{ in} &\leq R \leq 150 \text{ in} \\ 1.0 \text{ in} &\leq T_s \leq 1.375 \text{ in} \\ 25 \text{ in} &\leq L \leq 240 \text{ in} \\ 0.625 \text{ in} &\leq T_h \leq 1.0 \text{ in} \end{aligned} \quad (12)$$

본 예제에서의 실험계획법으로는 앞에서 언급한 바와 같이 중심합성계획법(CCD)과 균일 계획법(UD)을 사용하였으며, 기존의 연구에서 대부분 사용했던 가우시안 상관함수(Gaussian correlation function: fun2)뿐이 아닌 그 외의 지수 상관함수(exponential correlation function: fun1), 3차 상관함수(cubic correlation function: fun3), Bessel function의 변형 형태인 ($\nu = 3/2, 5/2$) Matern 상관함수(Matern correlation function: fun4, fun5)를 사용하여 결과를 비교하였다. 그리고 기존의 흥미영역 설정 방법이었던 흥미영역에서의 초기값과의 거리를 일정한 비율로 축소해가며 옮겨가던 국부화 기법의 최적화 값, 그리고 Sandgren의 결과값과의 비교를 통해 본 연구의 알고리즘의 효율성을 보이고자 한다. 따라서 CCD에 의한 결과를 Table 1에 정리하였으며, CCD의 이용에 있어 각 상관함수에 따른 변화를 Fig. 4에 정리하였고, 공간충진 실험계획법인 UD에 의한 결과는 Table 2에 정리하였으며, 마찬가지로 상관함수에 따른 변화를 Fig. 5에 정리하였다. 특히 Table 2에서 UD를 사용한 확률론적 국부화 기법의 결과를 제외한 Table 1-2의 결과들은 크리깅 메타모델의 구현에 있어 각 상관함수의

Table 1. The optimal results using CCD

| 중심합성법 (CCD) | Sandgren's optimum values | kriging metamodel | |
|-------------|---------------------------|-------------------|-----------------|
| | | 기존의 흥미영역 이동 방법 | SSRM의 흥미영역 이동방법 |
| 목적 | model | 7867.0 | 7552.2 |
| 함수 오차(%) | 실제값 | - | 7549.8 |
| | | - | 0.03 |
| 설계 변수 (in) | R | 47.708 | 50.946 |
| | T_s | 1.100 | 1.000 |
| | L | 117.701 | 108.097 |
| | T_h | 0.600 | 0.600 |

결과를 비교하여 가장 좋은 결과값을 얻을 수 있었던 Matern 상관함수($\nu = 5/2$)에 대한 결과이며, 이는 기존의 연구가 가우시안 상관함수에 치중해 있던 결과와는 다른 결과를 보인다. 이는 Matern 상관함수의 형상은 폭이 좁고 경

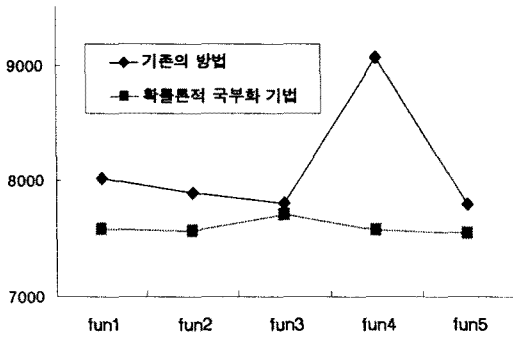


Fig. 4. Comparison the results

사진 포물선의 형태를 잘 보간해 주는 특징이 있으므로 본 연구에서의 목적함수 형태를 잘 근사시켜 주고 있으며, 상관함수의 적절한 선택이 크리깅 메타모델의 정확성과 효율성에 영향을 미침을 알 수 있다. 마찬가지로 Table 1-2의 결과값에서 알 수 있듯이, Sandgren에 의해 구해진 최적값과 비교하였을 때 기존의 흥미영역을 이동해 가는 방법과 확률론적 국부화 기법에 의한 설계영역 이동의 두 알고리즘 모두에서 더 좋은 목적함수의 값을 얻을 수 있었으며, 기존의 흥미영역 이동방법과 확률론적 국부화 기법 둘만을 비교한 결과에서는 모든 상관함수에 대해 확률론적 국부화 기법의 흥미영역 이동 알고리즘의 사용이 더 좋은 결과 값을 얻었음을 Fig. 4-5를 통해 확인할 수 있다.

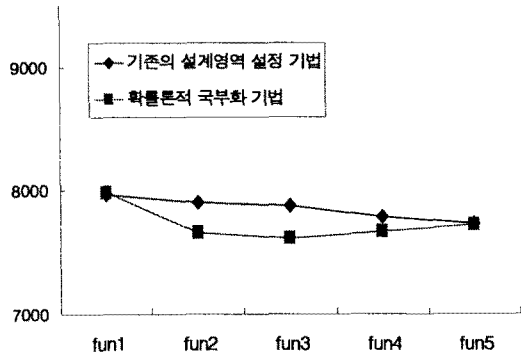


Fig. 5. Comparison the results

4. 결론 및 향후과제

본 연구에서는 근사모델을 이용한 최적화 방법 중 크리깅 메타모델을 이용한 최적화 방법에 있어 개선된 알고리즘을 제안하고, 그 정확성과 신뢰성을 비교·검증 하였다. 우선 고전적 실험계획법과 공간충진 실험계획법의 활용과 반복·순차적인 근사화 기법에 의한 최적화를 구현하였으며, 순차적인 반복 근사 최적화에 있어 흥미영역 설정 시 확률·통계추론적인 방법인 확률국부화 기법(SSRM)에 의한 설계영역 이동을 구현하였다. 그 결과, 본 알고리즘에 기존의 최적화 문제였던 pressure vessel에 적용해 봄으로써 더 향상된 결과를 얻을 수 있었으며, 정확성에 있어서도 좋은 결과를 얻을 수 있었다. 하지만 기존의 연구에서 배제되었던 상관함수에 관한 심층적인 연구의 필요성을 알 수 있었으며, 이 연구는 지금 진행 중에 있다.

후 기

본 연구는 스마트무인기기술개발사업 및 국방과학연구소 그리고 서울대학교 비행체특화연구센터의 연구비 지원에 의한 연구결과입니다. 이에 감사드립니다.

Table 2. The optimal results using UD

| 균일계획법 (UD) | Sandgren's optimum values | kriging metamodel | |
|------------|---------------------------|-------------------|-----------------|
| | | 기존의 흥미영역 이동 방법 | SSRM의 흥미영역 이동방법 |
| 목적 | model | 7867.0 | 7610.3 |
| 함수 | 실제값 | - | 7605.9 |
| | 오차(%) | - | 0.05 |
| 설계 변수 | R | 47.708 | 49.109 |
| | T_s | 1.100 | 1.000 |
| (in) | L | 117.701 | 121.308 |
| | T_h | 0.600 | 0.600 |

이때 Fig. 4-5에서 가로축은 상관함수의 종류, 세로축은 목적함수의 값이 된다. 또한 크리깅 메타모델에 의해 구해진 설계변수의 최적값, 그리고 그에 대한 목적함수의 값과 구해진 설계 변수에서의 실제 결과값과의 오차에서도 0% - 0.25%로 본 연구에서 사용된 알고리즘의 신뢰성과 정확성을 확인할 수 있다.

참고문헌

1) Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P., "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, 1989, pp. 409-435

- 2) Zhao, S., Jiao, L., Zhao, J., Wang, Y., "Evolutionary Design of Analog Circuits with a Uniform-Design Based Multi-Objective Adaptive Genetic Algorithm," *NASA/DoD Conference of Evolution Hardware(EH'05)* 2005, pp. 1-4.
- 3) Simpson, T. W., Mauerv, T. M., Korete, J. J., and Mistree, F., "Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization, Vol.1*. 1998, pp. 381-391.
- 4) Lin, Y., Mistress, F., Allen, J. K., Tsui, K., "Sequential Metamodeling in Engineering Design," 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2004, pp. 1-13.
- 5) Giunta, A., Waton, L., "A Comparison of Approximation Modeling Technique Polynomial Versus Interpolatin Models", *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Vol. 1, AIAA, Reston, VA*, 1998, pp. 392-404.
- 6) Simpson, T. W., "A Concept Exploration Method for Product Family Design," *PH.D. Thesis, Georgia Institute of Technology*, 1998, pp. 280-281.
- 7) Haldar, A., Mahadevan, S., "Probability, Random Variables and Stochastic Methods in engineering Design," *John Wiley & Sons*, 200, pp. 35-48.
- 8) Krige, D. G., "Two dimensional weighted moving average trend surfaces for one valuation," *special edition of Journal of South African Institute of Mining and metallurgy*, 1966, pp. 13-38
- 9) Koehler, J. R. and Owen, A. B., "Computer Experiments," *Handbook of Statistics (Ghosh, S. and Rao, C. R, eds.)*, Elsevier Science, New York, 1996, pp. 261-308.