

윤 병 동 서울대학교 기계항공공학부 교수 | e-mail : bdyoun@snu.ac.kr
 윤 현 준 서울대학교 기계항공공학부 박사과정 | e-mail : heonjun@snu.ac.kr
 박 정 호 서울대학교 기계항공공학부 석사과정 | e-mail : hihijung@snu.ac.kr

이 글에서는 컴퓨터 이용 공학(CAE : Computer Aided Engineering) 기술이 발달함에 따라 갈수록 중요해지는 통계적 모델 검증 및 보정(Statistical Model Validation and Calibration)을 수행하는 데 필요한 통계적인 기법들을 각 단계별로 상세하게 소개하고, 실제 제품 개발에 적용하는 데 있어서 예상되는 어려움과 향후 연구방향을 제시하고자 한다.

제품 생산 주기가 짧아지고, 제품의 기능이 다양해짐에 따라 새로운 제품을 개발하고 검증하기 위한 가상 시험(Virtual Testing)의 중요성이 증가하고 있다. 과거 CAE 기술이 발달하기 전에는 제품 개발 단계에서 시제품의 반복 시험을 통해 제품의 성능을 평가하였고, 이로 인해 많은 시간과 비용이 소요되었다. 하지만 CAE 기술의 진보와 컴퓨터 처리 속도의 향상 덕분에 제품의 성능을 CAE 모델을 통하여 사전에 예측하는 것이 가능케 됨으로써, 제품 개발 시간 및 비용이 획기적으로 감소하게 되었다. 그런데 실제 제품에 존재하는 재료 물성, 경계조건, 하중 등에 내포된 불확실성으로 인하여 시험과 해석이 일치하는 정확한 CAE 모델을 개발하는 것은 결코 쉽지 않다. 그렇다면 CAE 모델의 예측 능력(Predictive Capability)에 대한 신뢰도를 보장하기 위한 모델 검증은 어떠한 방식으로 이루어져야 할까? AIAA와 ASME V&V 가이드에서 모델 검증은 '해석모델이 그 모델의 용도에 따른 실제 물리적 거동을 얼마만큼 정확하게 표현하는지를 평가하는 과정'이라고 정의하고 있다. 이를 위해 최근 CAE 모델의 유효성을 통계적으로 검증하고 예측 결과를 향상시킬 수 있는 통계적 모델 검증 및 보정 기술이 큰 관심을 받고 있다. 이 글에서는 통계적 모델 검증 및 보정을 수행하는 데 있어서 필요

한 여러 가지 통계적인 기법들을 소개하고, 몇 가지 사례를 통해 각각의 통계적 기법이 어떻게 적용되는지 구체적으로 알아본다. 또한 이를 통해 실제 제품 개발에 활용하는 데 있어서 예상되는 어려움과 향후 연구방향을 제시하고자 한다.

기존의 결정론적 모델 보정

CAE 모델의 예측 능력을 평가할 수 있는 가장 단순한 방법은 실제 시험 결과와 CAE 해석 결과를 결정론적(Deterministic)으로 비교하는 것이다. 가령 사용자가 임의로 정한 기준(Criterion)보다 두 결과의 차이가 작으면 CAE 모델이 적합하다고 판단을 하게 되고, 기준보다 결과의 차이가 크면 CAE 모델이 부적합하다고 판단을 하게 되는 것이다. 이와 같은 방법을 결정론적 접근(Deterministic Approach)이라고 한다. 하지만 실제 공학적 문제에서는 다양한 불확실성(Uncertainty)과 산포(Variability)의 존재로 인해서 결정론적 모델 보정은 CAE 모델의 예측능력을 오히려 떨어뜨릴 수 있는 오류를 범할 수 있다. 여기서 불확실성은 추론가능(Inferable, 예: 외부온도, 수치에러)하며, 산포는 측정가능(Measurable, 예: 제작공차, 재료 물성치)하다고 정의할

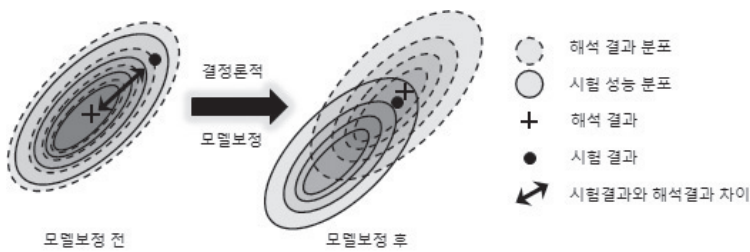


그림 1 결정론적 모델 보정의 그릇된 결과

수 있다.

모든 자연적, 인공적 시스템에는 불확실성과 산포가 존재하며, 공학시스템도 예외는 아니다. 자연적 시스템과 달리, 공학적 시스템에서는 나름의 산포 제어를 통해 불확실성이 상대적으로 적다고 할 수 있다. 공학시스템(또는 제품) 성능해석에 존재하는 불확실성과 산포의 원인은 다음과 같이 세 가지로 분류될 수 있다.

a. 물리적 원인(Physical Source): 공학시스템(이하 시스템)은 설계한 대로 제작생산이 불가능하므로 항상 제작공차와 재료 물성치에 산포(측정 가능하므로)가 존재하며, 시스템을 운영하다보면 하중조건에는 상당한 불확실성과 산포가 동시에 존재함.

b. 모델링 원인(Modeling Source): 시스템의 물리적 지배방정식, 수학적 모델링, 수치적 근사화 과정에서 수반되는 가정들, 재료거동 및 경계조건 단순화, 근사화 오차 등에 의한 모델링 원인이 존재함. 모델링 원인은 대부분 측정이 불가능한 대신에 추론은 가능하므로 불확실성으로 분류됨. 특히 시스템의 복잡도에 따라 모델링 원인이 물리적 원인보다 더 큰 불확실성을 제공하기도 하므로 이를 최소화하는 모델 검증 및 보정이 반드시 필요함.

c. 통계적 원인(Statistical Source): 물리적 산포 및 불확실성 요소를 확률적으로 모델링하기 위해서는 충분한 시료나 관련 정보를 확보하여야 함. 하지만 확률적 모델링에 충분한 수의 시료 또는 정보를 확보하기 위해서는 많은 시간과 비용이 요구되므로, 항상 해당 데이터 부족이 반드시 따르기 마련임. 이 때, 적은 양의 데이터 또는 미지의 변수들에 따른 정보 부족에 기인한

불확실성 요인들을 통계적 원인이라고 함.

위에서 언급한 산포와 불확실성은 무작위 불확실성(Aleatory Uncertainty)과 인식론적 불확실성(Epistemic Uncertainty) 등으로 구분된다. 무작위 불확실성은 충분한 정보하에서 제작공차, 재료 물성치 등과 같이 해당 시스템 자체의 변동에서 오는 것으로,

불확실성의 정도를 더 이상 줄일 수 없다(Irreducible). 반면에 인식론적 불확실성은 정보의 부족에서 오는 것으로, 불확실성의 정도를 줄일 수 있는 여지가 있다.(Reducible)

일반적으로 시스템 성능시험에는 불확실성 요소들이 많이 존재하므로 반복시험을 통한 성능치는 불확실성을 갖는 반면, CAE 모델은 성능을 결정론적으로 예측한다. 이 때, 시험결과는 랜덤하게 얻어지는 반면에 해석결과는 항상 평균값에서 얻어진다. 우선 결정론적인 접근법을 고려해보자. 그림 1과 같이 하나의 시험 결과(●)와 해석 결과(+)를 갖고 있다고 하자. 그림 1의 왼쪽과 같이 모델 보정 전에 해석 및 시험결과가 서로 동일한 분포를 갖는다고 하자. 물론 이러한 통계적 정보는 알려지지 않기 때문에, 두 값의 차이가 있을 경우 그림 1의 오른쪽과 같이 CAE 모델을 보정(Calibration)하여 시험 값으로 맞추는 게 일반적이다. 따라서 결정론적 접근법은 모델의 예측정확도가 오히려 부정확해지는 결과를 초래할 수 있다.

따라서 이러한 결정론적 접근의 한계를 극복하고 CAE 모델 검증 및 보정을 성공적으로 수행하기 위해서는 통계가 반드시 뒷받침 되어야 한다. 이에 다음 장에서는 통계적 모델 검증 및 보정 기술 개념을 간략히 소개하고 각 단계마다 필요한 기법들에 대하여 살펴보도록 하겠다.

통계적 모델 검증 및 보정 개요

통계적 모델 검증(Statistical Model Validation) 및 보

정(Statistical Model Calibration)은 시험 결과와 해석 결과를 통계적으로 비교함으로써 CAE 해석결과의 유효성을 검증하고, 유효하지 않다고 검증된 경우에 모델을 통계적으로 보정하는 기술이다. 통계적 모델 검증 및 보정 절차를 그림 2에서 확인할 수 있다. 이를 위해서는 불확실성 정량화(Uncertainty Quantification; a, b, c), 통계적 모델 검증(d), 통계적 모델 보정(e) 과정이 필수적으로 요구된다. 통계적 모델 검증 및 보정 절차는 다음과 같이 5단계로 구성된다.



그림 2 통계적 모델 검증 및 보정 개략도

a. 불확실성 모델링(Uncertainty Modeling): CAE 모델에 필요한 랜덤입력변수(Random Input Variable)들을 통계적으로 모델링하는 과정이며, 가능한 모든 정보와 전문가 식견을 활용함.

b. 민감도 분석(Sensitivity Analysis): 랜덤입력변수가 많을 경우, 시스템 성능에 큰 영향을 끼치는 주요 입력 변수들을 선별함.

c. 불확실성 확산(Uncertainty Propagation): 유효성 검증을 위해서는 CAE 모델의 해석결과가 통계적으로 표현되어야 하며, 이를 위해 입력변수 내 불확실성이 CAE 모델을 통해 어떻게 성능변수의 불확실성으로 전파되는지를 해석함.

d. 유효성 검증(Validity Check): 해석결과와 시험결과를 통계적으로 비교함으로써 CAE 모델이 유효한지를 검증함.

e. 통계적 모델 보정: CAE 모델이 유효하지 않다고 검증되었을 때, 모델을 통계적으로 보정함.

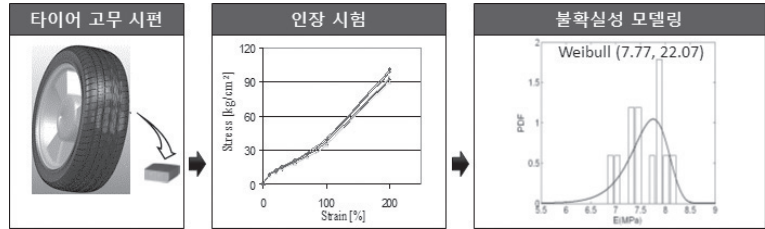


그림 3 자동차 타이어 마찰력 시험에 사용된 타이어 고무 탄성계수의 불확실성 모델링 예

(Probability Theory), Dempster-Shafer 증거 이론(Dempster-Shafer Theory), 최근 가능성 이론(Possibility Theory), 베이즈 이론(Bayesian Theory), 퍼지 이론(Fuzzy Theory) 등이 있다. 이 중에서 통계적 모델 검증과 보정에는 확률 이론과 베이즈 이론이 가장 지배적으로 사용된다. 자동차 타이어 마찰력 시험에 사용되는 타이어 고무의 탄성계수를 확률 이론을 통하여 통계적으로 모델링한 사례를 살펴보자. 결정론적 관점에서는 자동차 타이어 고무의 탄성계수를 하나의 값(7.7 MPa)으로 설정한 후 해석을 수행하여 단일한 마찰력 값을 도출하게 된다. 하지만 확률 이론에서는 자동차 타이어 고무 시험시편을 통해 타이어 고무 탄성계수를 얻은 후 형상모수(Shape Parameter=7.77)와 척도모수(Scale Parameter=22.07)를 갖는 웨이블 분포(Weibull Distribution)로 모델링한다.(그림 3)

불확실성 모델링

불확실성 모델링(Uncertainty Modeling)은 앞에서 서술한 무작위 불확실성과 인식론적 불확실성 등을 고려해서 CAE 모델의 입력변수들을 통계적으로 나타내는 과정이며, 불확실성 정량화 과정의 첫 번째 단계에 해당한다. 불확실성 모델링 기법 종류에는 확률 이론

민감도 분석

복잡한 시스템에는 많은 불확실성 변수들이 존재한다. 모든 입력변수들을 고려한 통계적 해석은 엄청난 해석 횟수와 시간을 요구하게 된다. 민감도 분석은 앞에서 알려진 많은 랜덤입력변수들 중에서 성능 결과에 영향을 크게 미치는 인자를 구분하는 기술이다. 따라서



그림 4 불확실성 확산 개념도

민감도 분석(Sensitivity Analysis)을 통해 선정된 주요 랜덤입력변수들만을 고려하면 효율적인 통계적 해석이 가능하다. 가령 앞에서 예로 든 타이어 마찰력 시험의 경우, 타이어 고무의 탄성계수는 마찰력 시험 시 결과에 큰 영향을 미치는 인자라고 할 수 있다. 하지만 고무의 열팽창계수, 두께 등은 마찰력 시험 시 영향이 미미하거나 내재된 불확실성이 상대적으로 매우 작기 때문에 주요 인자에서 제외할 수 있다.

불확실성 확산 해석

민감도 분석을 통해 선정된 주요 랜덤입력인자들은 통계적 CAE 해석을 통해 성능함수를 통계적으로 예측하여야 한다. 이를 위해서는 불확실성 확산(Uncertainty Propagation) 기법을 CAE 해석에 접목해야 한다. 이를 통해 성능함수를 통계적으로 정량화하는 것이 가능케 된다(그림 4). 불확실성 확산 기법은 아래와 같이 다양한 예들이 있다.

a. 랜덤 샘플링 기법: 몬테카를로(Monte Carlo (MC)) 시뮬레이션이 가장 대표적이며, 주로 비교평가대상 연구(Benchmark Study) 목적으로 사용됨. MC 시뮬레이션은 많은 수의 난수(Random Samples)를 생성하여 해석을 수행함으로써 통계적인 결과가 정확하나, 지나치게 높은 계산 비용(Computational Effort)이 요구됨.

b. 전개법(Expansion Method): 1차 또는 저차함수로 근사된 성능함수를 이용하여 통계적 모멘트(Moment)를 근사화하는 방법임. 성능함수의 비선형성이 높거나 랜덤입력변수가 정규분포를 따르지 않는 경우 정확도

가 떨어지며, 근사화 과정에서 성능함수의 미분값들이 필요함.

c. 최적화 기법: 성능함수의 확률분포를 구하기 위해 랜덤입력변수를 표준정규분포로 변형한 후, 성능 값과 해당 확률과의 관계를 최적화 기법을 통해 구하는 방법으로 First-Order Reliability Method (FORM)과 Second-Order Reliability Method (SORM)

등이 있음. 이 방법 또한 최적화 과정에서 성능함수의 미분값들이 필요함.

d. 통계기반 반응함수법: Kriging, Radial Basis Function, Moving Least Squares 방법과 같은 보간(Interpolation) 기법들을 활용하여 성능함수를 근사화한 후에 MC 시뮬레이션을 이용해서 불확실성 확산해석이 가능하며, 통계기반 모델 보정에서 가장 널리 쓰이는 방법임.

e. 통계기반 근사적분법: 불확실성 확산은 랜덤입력변수들의 확률밀도함수를 다차원 적분을 통해 구할 수도 있으나 해석적이나 수치적으로 수행하는 데 한계가 존재함. 이를 극복하기 위하여 차원축소법(Dimension Reduction Method), Polynomial Chaos Expansion, Tensor Product 등의 통계기반 근사적분법 등이 사용됨. 위 방법들 중에서 정확하고 효율적인 불확실성 확산 해석 수행은 가능하나, 프로그래밍이 다소 까다롭다는 단점이 있음.

최근에는 입력변수의 불확실성이 결과의 불확실성으로 확산되는 전방(Forward) 불확실성 확산 해석뿐만 아니라, 결과의 불확실성을 통해 입력변수의 불확실성을 추정하는 역(Inverse) 불확실성 확산 해석에 대한 연구도 활발하게 진행되고 있다.

유효성 검증

유효성 검증(Validity Check)을 수행하기 위해서는 먼저 상기된 불확실성 정량화를 통해서 통계적으로 모델링된 시험 결과와 통계적 해석 결과를 통계적 관점에

서 상호 비교하고 그 유사성 정도를 수치적인 값으로 나타내는 과정이 필요하며, 이를 검증 척도(Validation Metric)라고 한다(그림 5). 확률 분포 특성을 고려하여 검증 척도를 계산하는 방법에는 최대 가능성(Maximal Probability), 쿨백-라이블러 발산(Kullback-Leibler Divergence), 면적 척도(Area Metric), 거리 척도(Distance Metric), 그리고 베이즈 인수(Bayes Factor) 등이 있다. 이렇게 얻어진 검증 척도는 가설 검정(Hypothesis Test)을 통해 최종적으로 CAE 모델의 유효성을 검증하게 된다.

CAE 모델의 유효성 검증을 수행할 때, 시간 및 비용의 제약으로 인하여 한정된 시험 데이터를 사용해야 한다. 따라서 제한된 표본으로도 유효성을 검증할 수 있는 체계적인 가설 검정법이 필요하다. 가설 검정은 일반적으로 ① 귀무가설(Null Hypothesis) 선정, ② 유의수준(Significance Level) 선정으로 이루어진다. 유효성 귀무가설은 CAE 모델이 유효하다고 가정하고, 이의 기각 여부에 따라 그 유효성 여부를 판단한다. 이때 중요한 것은 유의수준 결정이며, 개발하고자 하는 제품이 갖는 위험도(Risk)에 따라 정도를 달리할 수 있다. 위험도가 큰 제품은 엄격한 기준으로, 상대적으로 위험도가 작은 제품은 유연하게 유의수준을 결정한다.

통계적 모델 보정

만약 유효성 검증을 통해 CAE 모델이 적합하지 않다는 결론이 도출되면 어떻게 될까? 이런 경우 CAE 모델의 경계조건, 하중조건, 요소(Element) 선정, 물리적 법칙(예: 지배방정식) 자체에 문제가 있거나, 아니면 입력 변수의 불확실성을 통계적으로 모델링하는 과정에서 해당 변수의 통계량(예: 평균, 분산, 왜도, 첨도 등)이 잘못되었을 가능성이 크다. 가령 초기 CAE 모델에 입력되는 재료 물성치는 많은 경우에 사전 지식을 통해 얻

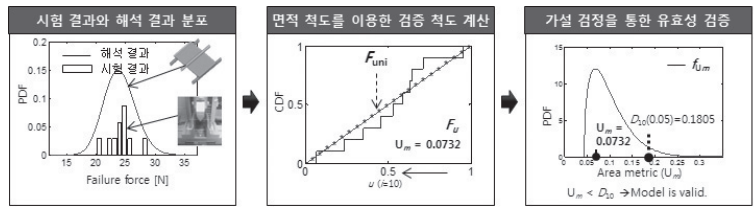


그림 5 면적 척도를 이용한 검증 척도 계산 및 유효성 검증

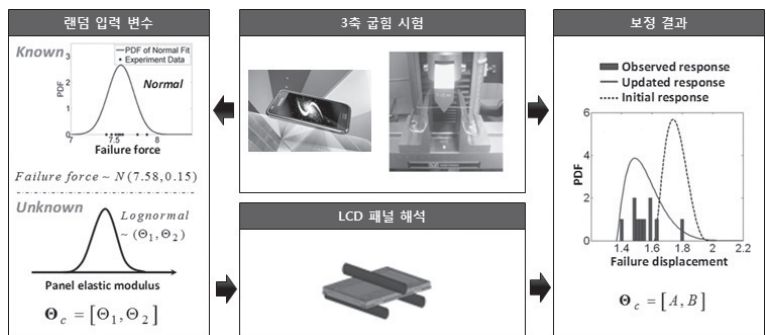


그림 6 최대우도추정법을 이용한 통계적 모델 보정(예: 휴대폰 LCD 패널의 기계적 파손시험 - 3축 굽힘 시험)

거나 임의의 가정으로 선정되기 때문에 부정확한 값을 가질 수 있다. 만약 후자의 경우라면 통계적 모델 보정(Statistical Model Calibration)을 통하여 CAE 모델의 예측 능력을 향상시킬 수 있다.

통계적 모델 보정은 랜덤입력변수들을 미지변수(Unknown)로 설정하고 해석 결과 분포와 시험 결과 분포의 유사도를 최대화하는 미지변수 분포를 찾아내는 것이다. 이러한 통계적 모델 보정 방법에는 최대 우도 추정법(Maximum Likelihood Estimation), 베이저안 보정법(Bayesian Calibration), 편향 보정(Bias Correction) 등이 있다. 예를 들어 최대 우도 추정법은 해당 시험 결과를 가장 잘 표현할 수 있는 확률 분포를 최적화 알고리즘을 이용하여 찾는 방법이다(그림 6). 이러한 보정 과정을 반복함으로써 새롭게 구한 검증 척도가 가설 검정 결과를 만족하게 되면 최종적으로 검증된 모델을 얻게 되는 것이다.

만약 통계적 모델 보정을 통해서도 검증된 모델을 얻지 못하게 되면, 모델 개선(Model Refinement) 기술을

통하여 물리 지배방정식, 수학적 모델링, CAE 모델 자체를 수정함으로써 검증된 모델을 얻게 된다. 예를 들어 재료의 등/이방성, 비선형 거동 등을 재차 확인하거나, 모델을 단순화하는 과정에서 쓰인 가정이나 근사화를 재검토해야 하며, 하중조건 및 경계조건의 적절성을 판정해야 한다.

결론 및 향후 전망

CAE 모델의 예측 능력을 향상시키고 해석 결과의 유효성을 체계적이고 과학적인 방법으로 검증할 수 있는 통계적 해석모델 검증 및 보정은 새로운 시스템(또는 제품)을 개발하고 성능을 평가하기 위한 가상 시험에 매우 중요한 기술이 될 것으로 기대된다. 특히 검증된 CAE 모델은 산업체의 시스템(또는 제품) 설계 능력을 향상시키는 데 큰 기여를 하게 되어, 향후 점점 짧아지는 제품 수명주기 및 갈수록 치열해지는 국제 시장 등의 현실 속에서 우수한 경쟁력을 확보할 수 있을 것으로 사료된다. 단, 통계적 모델 검증 및 보정 기술 수행하는데 있어서 중요한 점은 반드시 실제 시험에서 나타나는 현상을 정확하게 이해하고, CAE 모델 구축에 필요한 역학 지식들을 숙지한 후에 모델 검증을 수행해야만 한다는 것이다. 만약 데이터들을 무의미하게 통계적으로 단순 비교하는 것에 그친다면, 오히려 물리적으로 타당하지 못한 해석 결과를 초래할 수 있기 때문이다. 본문에서 살펴본 통계적 모델 검증 및 보정을 수행하는데 있어서, 앞에서 열거한 통계적 기법들을 실제 산업체의 제품 개발을 위한 가상 시험에 적용하는 데 예상되는 어려움은 다음과 같다.

a. 다수의 시험 데이터 확보의 어려움: CAE 모델의 유효성 검증을 통계적으로 수행하기 위해서는 다수의 시험 데이터 확보가 요구됨. 그러나 자동차 충돌 시험처럼 높은 시험 비용과 시간이 요구되는 경우에는 시험을 여러 번 반복적으로 수행하는 데 어려움이 있음. 만약 시험 데이터 수가 적은 경우에는 해당 시험 조건에서의 유의미한 통계치를 얻기가 용이하지 않으며, 시험 결과와 CAE 모델의 예측 결과의 차이가 데이터 수가 적어서 생긴 불확실성에 기인한 것인지 아니면 실제로 값의 차이가 있는 것인지 판단하기가 어려움.

b. 임시방편적(Ad hoc) 모델 보정의 위험성: CAE 모델의 경계조건 및 하중조건 등이 물리적으로 타당하지 못한 상황에서 무분별하게 입력변수 값만 보정하면, 오히려 무의미한 해석 결과를 도출하게 됨. 따라서 유효성 검증에서의 무효(Invalid) 판정으로부터 CAE 모델의 무효성 원인이 입력변수 보정으로 해결 가능한 문제인지 정확히 유추하는 기법 개발이 필요함.

c. 다양한 검증 척도의 필요성: 대부분의 검증 척도는 시불변(Time-invariant) 성능 예측에 국한되어 있기 때문에, 시변(Time-variant) 성능 예측 모델에는 적용할 수 없음. 더불어 예측해야 하는 성능이 다수인 경우에는 두 종류 이상의 시험 결과와 해석 결과를 비교해야 되기 때문에, 다차원(Multi-Dimension) 검증 척도 개발이 요구됨.

따라서 향후에는 통계적 모델 검증 및 보정 기술을 위한 학술적 연구뿐만 아니라, 실제 산업체 제품 개발 단계에 활용할 수 있는 실용적이고 효율적인 통계적 기법 개발을 위한 연구가 활발히 진행될 전망이다.