2002 추계학술발표회 논문집
한국원자력학회

미국 WIPP 사이트 성능평가를 통해 본
고준위 폐기물 처분장에 대한 확률론적 성능평가의 한 방법론

A Methodology of Uncertainty/Sensitivity Analysis for PA of HLW Repository Learned from 1996 WIPP Performance Assessment

이연명, 김성기, 황용수, 강철형
한국원자력연구소
대전광역시 유성구 덕진동 150

요 약

미국 WIPP 사이트가 1999 년 처분 운영을 개시하였다. 이 WIPP 사이트는 미국 DOE 의 방위프로그램에 의한 부산물로서의 핵폐기물인 초우라늄 방사성 폐기물 (TRU 폐기물)만을 전담 영구 처분하는 처분장으로 처분장 성능에 대해 책임을 지고 있는 미국 에너지성 (DOE)은 이 처분장의 운영허가를 위해 WIPP 운영을 위한 <규제부합 인정신청 (CCA)>을 통하여 EPA 로 부터 허가를 얻어 처분을 시작한 것이다. 이 CCA 를 위한 성능평가 관련연구가 1996 년에 마무리지어짐에 따라, 그 성능평가 연구의 결과나 과정은 비록 우리나라가 고려하고 있는 처분 환경과는 상당히 다른 지질 형태의 것이긴 해도 그 방법론에 대한 고찰을 통하여 이러한 WIPP 사이트의 성능평가에 관련되어 어떻게 어떠한 방법론을 도입하여 불확실성 및 민감도 분석연구가 수행되어졌는지, 그리고 정량적인 불확실성 분석의 결과를 요구하는 EPA 의 규제 기준을 어떠한 논리로 만족시켰는가에 대한 개략적인 검토를 수행하여 보고 우리의 경우 어떤 방법으로 이러한 문제에 접근해 갈 것인지 고찰해 보았다.

Abstract

The WIPP (Waste Isolation Pilot Plant) is a mined repository constructed by the US DOE for the permanent disposal of transuranic (TRU) wastes generated by activities related to defence of the US since 1970. Its historical disposal operation began in March 1999 following receipt of a final permit from the State of NM after a positive certification decision for the WIPP was issued by the EPA in 1998, as the first licensed facility in the US for the deep geologic disposal of radioactive wastes. The CCA (Compliance Certification Application) for the WIPP that the DOE submitted to the EPA in 1966 was supported by an extensive performance assessment (PA) carried out by Sandia National Laboratories (SNL), with so-called 1996 PA. Even though such PA methodologies could be greatly different from the way we consider for HLW disposal in Korea largely due to quite different geologic formations in which repository are likely to be located, a review on lots of works done through the WIPP PA studies could be the most important lessons that we can learn from in view of current situation in Korea where an initial phase of conceptual studies on HLW disposal has been just started. The objective of this work is an overview of the methodology used in the recent WIPP PA to support the US DOE WIPP CCA ans a proposal for Korean case.

1. 서론

원자력 발전의 필연적인 부산물로서의 방사성 폐기물을 안전하고도 적절한 방법으로 처분하는 일은 아마도 핵연료주기의 전과정을 통틀어 인간환경에 직접적이고도 장기적인 영향을 필연적으로 끼친다는 점에서 가장 중요한 과정일 것이다.
지난 2002 년 7 월 미국 상원은 야카마운틴 사이트를 사용후 핵연료 및 고준위 폐기물 처분장으로 하는 것을 인준하므로써 아직 몇가지 절차가 남아 있긴 해도 아마 미국 최초의 고준위 폐기물 처분장이 될 예정으로 있다.
한편 이에 앞서 미국의 WIPP (Waste Isolation Pilot Plant)사이트에서 1999 년 처분 운영이 처음 개시되었다. 미국 최초로, 그리고 중저준위 폐기물이 아닌 것으로는 세계 최초로 심부 지하 매질내의 동굴형 처분시설로 허가를 받게 된 이 WIPP 사이트는 미국 DOE 의 방위프로그램에 의한 부산물로서의 핵폐기물인 초우라늄 방사성 폐기물 (TRU) 폐기물만을 전담하는 처분장이지만 사실상 고준위 폐기물의 특성으로 볼 수 있는 장수명 핵종을 함유한 폐기물 처분을 역사상 최초로 시작하였다는 점에서 매우 주목할 만한 것이었다.
처분장 성능에 대한 책임을 지고 있는 미국 DOE 는 이 처분장의 운영허가를 위해 WIPP 운영을 위한 <규제부합 인정신청(Compliance Certification Application; CCA)>을 1996 년 10 월 미국 EPA (Environmental Protection Agency)에 요청한 후 1998 년 4 월 EPA 로 부터 긍정적인 결정을 통보받은 다음 그 이듬해 3 월에 첫 폐기물을 들여옴으로써 첫 처분을 시작하게 된 것이다.
그런데 이 CCA 에 요구되는 처분장 성능평가 (PA) 관련 연구는 처분장의 성능, 즉 그 안전성을 보이기 위해 미국 샌디아연구소(Sandia National Laboratories; SNL)를 수행되어 최종 인허가 신청을 위한 장기간의 처분장 성능평가가 1996 년에 최종적으로 마무리된 바 있다.
비록 이 WIPP 사이트 관련 PA 연구의 결과나 과정은 우리나라가 고려하고 있는 처분 환경과는 상당히 다른 지질 형태로서, 지하 심부 불포화 암염층 (Bedded salt)내 사이트에 국한된 것이긴 해도 그 방법론에 대한 고찰은 이제 막 고준위 방사성 폐기물 처분관련 연구를 시작하는 단계에 있는 우리에게 교육적인 효과가 지대할 수가 있다.
이렇게 1996 WIPP PA' 로 통칭되는 이 WIPP 사이트 성능평가 연구는 우선 방대한 부지특성화 작업, 자료 생산을 위한 다양한 실험 계획 프로그램, 그리고 수학적, 개념적 모델링을 통한 평가 모델 개발과 이에 요구되는 대량의 입력자료 도출과 함께, 모델과 입력자료 등에 관련된 불확실성 및 민감도 분석에서 부터 관련 전산코드 개발 및 QA 프로그램까지, 그리고 이미 오래전부터 수행해 온 여러 기초 PA 관련 연구들의 총체적인 결과물로 집약되었다.
이러한 방대한 연구내용중에서 이 보고를 통하여 특별히 주목되는 세부적인 부분은 PA 에 관련된 불확실성 및 민감도 분석 (Uncertainty and Sensitivity analyses)에 대한 방법론이다.
우리가 불확실한 자료에 의해 불확실한 결과를 얻는다고 할 때 언제나 두 개의 질문이 나타나게 되는데, 즉, PA 를 수행하기 위한 모델에 대해 불확실한 입력변수를 사용한다면 그 결과(Consequence)에 대한 불확실성은 어떻게 드러나게 되고, 또 그 불확실한 결과에 기여하는 입력변수들은 도대체 어느 것이 각자 얼마만큼의 비중을 가지고 그 불확실한 결과에 기여하게 되는 것일까?
이러한 질문에 대한 답을 위해서 불확실성 분석과 민감도 분석이 요구되어지는 것이다.
따라서 이 논문을 통해 이러한 WIPP 사이트의 성능평가에 관련되어 어떻게, 그리고 어떠한 방법론이 도입되어 불확실성 및 민감도 분석연구가 수행되어졌는지, 그리고 정량적인 불확실성 분석의 결과를 요구하는 EPA 의 규제 기준을 어떠한 논리로 만족시켰는지, 그리고 나서 우리가 유사하게 수행해 나가야 할 방향은 어떤 것인지에 대한 개략적인 논의를 하고자 한다.

2. 규제치

EPA 는 WIPP 시설내 폐기물 처분에 관련된 규제를 책임지고 있다.
WIPP 처분장의 성능평가의 결과가 만족시켜야 할 EPA 의 제반 규제의 주요한 골자는 40CFR191(b)에 기술된 대로 <사용후핵연료, 고준위 방사성폐기물 및 초우라늄 방사성폐기물의 관리 및 처분에 관한 환경 방사선 방어 기준>과 40CFR194 에 기술된

기준으로서 <WIPP Compliance criteria (WIPP 사이트에 대한 인증 기준)>이다 (40CFR191, 1985; Howard, 2000).
미국 샌디아 연구소는 처분시스템이 그 운영을 끝내고 폐쇄된 후 일만년동안에 발생 가능한 핵종 유출에 대한 시나리오를 도출한 후, 처분시스템을 모사하기 위한 코드, 즉 핵종의 이동 및 흡착, 그리고 위해물질이 처분장으로부터 누출되어 여러가지 지하매질을 통해 인간환경으로 나오기까지, 또는 인간의 침입에 의한 가능한 핵종의 유출 등을 모사하기 위한 개념적, 수학적 모델을 개발한 후 이에 대한 계산을 통해 처분장에서 유출된 핵종에 의한 환경에 대한 결과 (Consequence)를 구해낸 후, 1996 년 최종적으로 EPA 의 40CFR191 의 191.13 항에 제시된 유출한계치와 비교하는 것으로 WIPP 성능평가에 대한 정량적 결과물로서 제시하였다.
이 40CFR191 에 의하면, 사실 이는 곧 PA 의 정의이기도 하지만, 처분장의 안전성을 보이기 위해서는 처분장을 나와 접근 가능한 환경으로 유출되는 방사성 핵종의 누적 유출량 $C(\mathbf{x})$를 평가하는 것으로 되어 있고, 이러한 유출에 대한 평가의 결과는 반드시 확률론적인 분포값으로 나타나야 하는 것으로 되어 있다.
즉, $C(\bullet)$가 변수 $\mathbf{x}=[x_1, x_2, \Lambda\ x_{nP}]$에 대한 함수이고, 첨자 nP가 불확실성을 갖는 모델 변수 (이 때 각각의 변수값은 ‘범위’와 ‘통계적인 분포’로서 기술되어지게 된다.) 의 총 갯수가 된다면 그 결과 $C(\mathbf{x})$도 범위와 분포로 나타낼 수 있게 될 것이다. 불확실성 분석이란 바로 이렇게 이 $C(\mathbf{x})$의 분포가 어떠한 형태를 갖는가 하는 것을 결정하는 과정을 포함하게 된다.
이렇게 불확실성 분석이 이루어지면, 위해도 (risk)에 대한 측정이 결과의 ‘기대값’이나 ‘분산’이라는 단일하고 명료한 값 대신 결과값이 전체적으로 가질 수 있는 ‘분포’로 나타낼 수 있게 되어, 보다 많은 불확실성에 관한 정보를 갖는 결과를 보여줄 수 있는 결과를 얻을 수 있게 된다. 기대값이나 분산이 비록 단일하고 알기 쉬운 결과를 나타낼 수 있을지는 몰라도 상세한 불확실성에 대한 정보는 보여줄 수 없다. 이렇게 처분장의 성능평가, 즉 PA 는 원자력발전소에서 활용되는 확률론적 위해도 평가 (Probabilistic Risk Assessments; PRAs)와 동일한 방법론임을 알 수 있다.
언급한대로 EPA 의 40CFR191 에는 폐기물이 처분된 이후 1 만년이 지난 후까지 핵종이 접근 가능한 인간 환경 (Accessible environment)으로 유출되는 누적유출 (Cumulative release)을 정량적인 기준으로 하여 아래의 두가지 항목과 같이 규정한다.

- 누적유출이 제한된 양 (L_i)을 넘을 확률이 1/10 보다 작아야 한다.
- 누적유출이 제한된 양의 10 배 (10 L_i)가 넘을 확률이 1/1000 보다 작아야 한다.

처분장의 성능평가의 결과는 모두 집약되어 CCDF (Complementary Cumulative Distribution Function)로 나타내게 되는데, 이 CCDF 가 모든 중대한 과정이나 사건에 따라 핵종이 유출되는데 따른 누적 유출이 어떤 수준을 넘게 될 확률을 기술해 주는 것이 되고 그 결과를 바로 EPA 의 규제와 비교하게 되는 것이다.

3. 성능평가 방법론

40CFR191.13 과 191.34 에 의하면 그 골자는 다음과 같다:

- EN1: 향후 1 만년간 WIPP 에서 일어날 수 있는 여러 상이한 양상 (feature)이 발생할 가능성을 확률적으로 특성화하는 것으로, WIPP 에 대한 시나리오 개발과정을 통해 이 EN1 이 도출되어지게 되고, 이는 결국 40CFR191 에서 요구하는 대로 WIPP 에서 향후 1 만년간 일어날 수 있는 여러 양상에 대한 발생 가능성을 확률론적으로 기술할 수 있도록 한다.

- EN2: EN1 에서 정의한 대로 향후 1 만년간 WIPP 에서 발생할 수 있는 양상 각각에 관련되어, 핵종 유출에 대한 추정을 의미하는 것으로, 이는 WIPP 에 대한 핵종 이동 모델과 입력자료를 통하여 각각의 상이한 양상에 대해 환경으로의 핵종유출을 평가하는 것을 의미한다.

- EN3: EN1 과 EN2 를 정의하는데 사용된 모든 파라미터에 관련된 불확실성에 대한 확률론적 특성화를 의미하는 것으로, WIPP 에 대한 입력자료 개발에 대한 결과로 나타나는 입력변수의 불확실성 분석을 수행하여 그 결과를 통계적인 분포로 나타내는 것을 의미한다.

이 세 요소의 결과가 결국 EPA 의 규제 (그림 2) 에 모두 관련되어 있는 것을 알 수 있다. 즉, 이 세가지 요소중 EN1 과 EN2 에 대한 규명을 통해 WIPP 사이트의 성능에 관련된 다음과 같은 3 가지 질문에 대한 답을 얻어 낼 수 있는 것이다.

- Q1: 향후 1 만년간 WIPP 사이트에서 상이한 어떤 일들이 일어날 수 있는가?
- Q2: 향후 1 만년간 WIPP 사이트에서 상이하게 발생할 수 있는 그 일은 얼마나 발생 가능한가?
- Q3: 그리고 향후 1 만년간 WIPP 사이트에서 상이하게 발생할 수 있는 그 일에 대한 결과는 도대체 무엇인가?

그리고 이러한 질문에 한가지를 더 추가하자면,

- Q4: 위 3 가지 질문에 답하는 것이 도대체 얼만큼의 신뢰도를 가지고 있는가?

바로 1996 년의 WIPP PA 는 이러한 질문에 대한 답을 제공한 셈인데, 다시 말하자면, EN1 으로서 Q1 과 Q2 에 대한 답을, EN2 로서 Q3 에 대한 답을, 그리고 EN3 로서 Q4 에 대한 답을 제공하면서 EN1 과 EN2 로는 40CFR191 과 40CFR194 에서 요구하는 대로 그림 1 과 같은 CCDF 를 그려내고, EN3 로서 이 CCDF 의 집합 (또는 군), 즉 CCDF 의 분포를 그리게 되는 것이다.

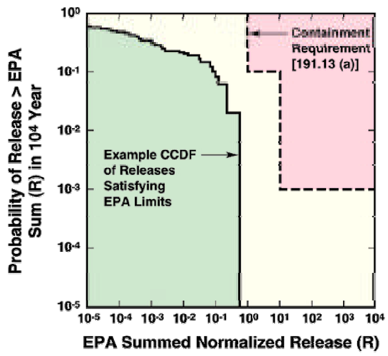


그림 1. 40CFR191 Subpart B 에 기술된 EPA 규제치 경계선 (점선)과 관련 CCDF

4. 불확실성 분석의 방법론 (Helton, 1991; Helton 1993a; Helton 1993b; Helton et al., 2000; Helton 2000)

처분장 성능에 관련된 위해도 (risk)를 일반적인 결과 (Consequences)와 그 결과에 관련된 확률(probability) 또는 빈도(frequency)로만 나타내는 것은 막상 그 위해도를 감지하기에 모호한 느낌도 줄 뿐 아니라 실제 이러한 정량적인 결과 추정을 위한 계산을 어떻게 수행되어야 되는지 막연한 느낌을 주게 된다. 이는 도대체 어떻게 나쁜 일이 진행될 수 있는지, 그리고 얼마만큼의 확률로 그러한 일이 진행될 수 있는지, 그리고 그런 일이 진행된 결과로서 얼마나 정량적인 영향을 주는지에 대한 정보를 정량적으로 명쾌하게 제시하여 주지는 못한다.
처분 시스템의 경우 환경으로의 핵종유출이 가장 중요한 평가 결과가 될텐데, 바로 이 결과, 즉 위해도를 EPA 에서 제시하는 핵종 유출제한 기준치와 비교하여 그 규제기준에 합당한지를 판단하는 것이 처분장의 성능을 평가하는 것이 PA 의 과정이고, 이러한 과정은 결국 불확실한 입력 정보에 의해 얻어진 불확실한 계산 결과에 따른 정보, 즉 불확실성에 관한 내용을 그대로

CCDF 의 군으로 나타내거나, 또는 CCDF 의 군으로 나타내는 대신 이들을 평균하거나 가공하여 가능한 한 단일한 CCDF 로 알기쉽게 나타내거나 하는 작업이 될텐데, 우선 위해도를 ‘ 순서를 갖는 3 개의 요소 ’ 로서 다음과 같은 집합으로 나타내보자.

$$R = \{(E_i, pE_i, \mathbf{cE}_i), i = 1, \Lambda \ nE\} \tag{1}$$

여기에서

E_i = 사건의 집합 (a set of similar occurrence)

pE_i = 사건집합중 사건 E_i 가 발생할 확률 (probability that an occurrence in set E_i will take place)

$\mathbf{cE}_i$ = 사건 E_i 에 관련된 결과 벡터 (a vector of consequences associated with E_i)

nE = 사건의 갯수 (number of sets selected for consideration)

이렇게 해놓고 보면 위해도는 결국 무엇이 발생하고 (what can happen; E_i), 얼마나 그 일이 가능하며 (how likely things are to happen; pE_i), 이에 대한 결과는 무엇인가 (the consequences of each set of similar occurrences; $\mathbf{cE}_i$)로 분리되는 것을 알 수 있다.
이 식에서 $\mathbf{cE}_i$ 는 아직은 단일한 값이며 아직 그 어떤 입력 파라미터값에 대한 불확실성에 관한 내용에 대해서는 아무런 기술은 하고 있지 않고, 그저 (1)식은 파라미터의 불확실성에는 아무런 관련이 없이 오직 nE 개의 시나리오만 반영된 식이 된다.
한편 R 도 이 식만으로는 무엇을 의미하고 있는지 애매하게만 보여 이 R 을 보다 가시적으로, 보다 의미있는 결과로 만들 필요성이 있는데, 그 중의 하나가 바로 CCDF 를 이용하는 것이다. 그러면 이 CCDF 는 비로소 pE_i 와 $\mathbf{cE}_i$ 에 담긴 정보를 결과로서 고스란히 보여 줄 수 있게 된다.
이제 pE_i 와 $\mathbf{cE}_i$ 를 표현하기 위해서 CCDF 를 이용해 보기로 한다. 어떤 특정한 결과 cE 가 순서를 가지고 발생하여 $cE_i \le cE_{i+1}, \quad i = 1, \Lambda \ , nE - 1$ 이 성립한다면 이 결과에 대한 CCDF, 즉 밀도함수 F 는 다음과 같이 정의될 수 있다.

$$F(x) = cE\text{가 특정한 결과치 x를 넘을확률} = \sum_{j=i}^{nE} pE_j \tag{2}$$

그러면 그림 2 와 같이 이 F 는 x 축에 표시된 값 (cE) 이상을 가질 확률을 나타내는 계단 함수로 나타내지게 된다.

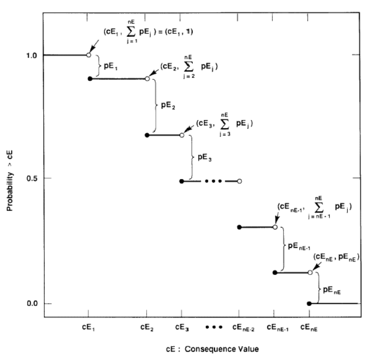


그림 2. 결과 $\mathbf{cE}_i$ 를 나타내기 위해 추정된 CCDF

사실 CCDF 를 만들어 가는 방법은 비교적 단순한데, 모든 발생 가능한 시나리오 E_i 를 $E_1,\Lambda\ ,E_{nE}$ 로 나누어 두고 pE_i 와 $\mathbf{cE}_i$ 에 대해 그림에서 보는 방식을 통해 커브를 작성해 나가면 된다. 이 과정에서 보다 많은 E_i 를 사용할 수만 있으면 계단의 크기가 잘게 줄어들어 보다 부드러운 커브를 얻게 된다.
위해도의 계산은 우선 E_i 를 결정하는 것으로 부터 시작된다. 방사성폐기물 처분의 경우에 특별히 이 E_i 를 ‘시나리오’라고 부르게 되는데, 이 시나리오 세트가 결정이 되고 나면 이 시나리오 각각에 대하여 발생 확률 pE_i 이 부여되고 그에 따른 결과 $\mathbf{cE}_i$ 가 차례로 평가되어진다. 그러나 실제로 이 pE_i 를 결정하는 일은 상당히 복잡하고 반복적인 작업이 요구된다. 뿐만 아니라 보통 이들을 평가하는 과정에서 상당한 양의 변수값이 요구되어 이들이 부정확한 값으로 개재되는데, 어떤 경우에는 이 부정확한 값이 역으로 E_i 의 정의 자체에 영향을 줄 수 있게 되기도 한다. 가령 예를 들어 어떤 시나리오에 따른 결과 (Consequence)가 크게 나타나는 경우보다는 그 결과가 작게 나타나는 경우에 대해서는 시나리오 범위를 보다 크게 잡게 되어 pE_i 에 영향을 주게 된다는 의미이다.
이제 아래 (3) 식과 같이 벡터 형태를 갖는 입력 자료 x_j 를 생각해보자.

$$\mathbf{x} = \left[x_1, x_2, \Lambda\ , x_{nV}\right] \tag{3}$$

여기에서 x_j 는 결과해석에 필요한 부정확한 입력변수의 값이고, nV 는 총 입력변수의 수이다.
그러면 (1)식은 여러 입력변수값 x_j 를 갖는 경우에 대해 아래 (4)식으로 표현 된다.

$$R(\mathbf{x}) = \{(E_i(\mathbf{x}), pE_i(\mathbf{x}), \mathbf{cE}_i(\mathbf{x})), i = 1,\Lambda\ \ nE(\mathbf{x})\} \tag{4}$$

이렇게 되면 부정확한 입력변수 벡터 x 가 다양하게 분포를 따라 변할 때 R(x)의 결과도 어떤 분포를 갖는 형태로 나타나게 되고, 따라서 $\mathbf{cE}_i$ 에 단 한 개의 CCDF 만이 나타나는 것이 아니라 x 가 가질 수 있는 모든 값 (여기서는 nV 개)에 따라 CCDF 의 군으로서 결과가 얻어지게 된다.
그림 3 은 이렇게 시나리오에 대한 불확실성이 전파되는 과정과 파라미터의 인식적 불확실성을 표현하기 위한 분포에 따라 발생하는 결과의 전파과정을 도해한 것이다.
위해도 평가를 위한 이 입력자료 x 가 불확실하다면 평가 결과도 역시 불확실하게 될 수 밖에 없다. 따라서 부정확하게 알 수 밖에 없는 이러한 변수가 가질 수 있는 값의 범위나 그 통계적인 분포를 얻기 위한 과정, 즉 입력자료에 대한 특성화(characterization)가 요구되어 진다. 그렇게 위해도 평가 결과에 관련된 입력변수 x 의 불확실성에 대한 특성화가 이루어지면 이 결과에 따라 몬테칼로 등의 방법에 의해 위해도 결과에 대한 특성화도 비로소 이루어질 수 있게 되는 것이다. 그렇다면 이와 같이 몬테칼로방법에 이용될 수 있도록 하기 위해 입력자료의 특성화는 어떠한 방법을 통해 이루어질 수 있을까?
이러한 질문에 대한 즉답은 어렵지만, 가장 바람직한 경우로는 통계적인 절차를 이용하는 것이다. 예를 들면, 표본의 수가 작을 때 이용할 수 있는 t-분포 등을 이용하여 x 가 가지는 범위와 분포를 결정할 수도 있다. 그렇지만 이런 경우는 자료가 적절한 방법으로 수집이 되어야 하고 이 자료에 대한 예비정보를 미리 알고 있는 경우에만 가능하게 된다.

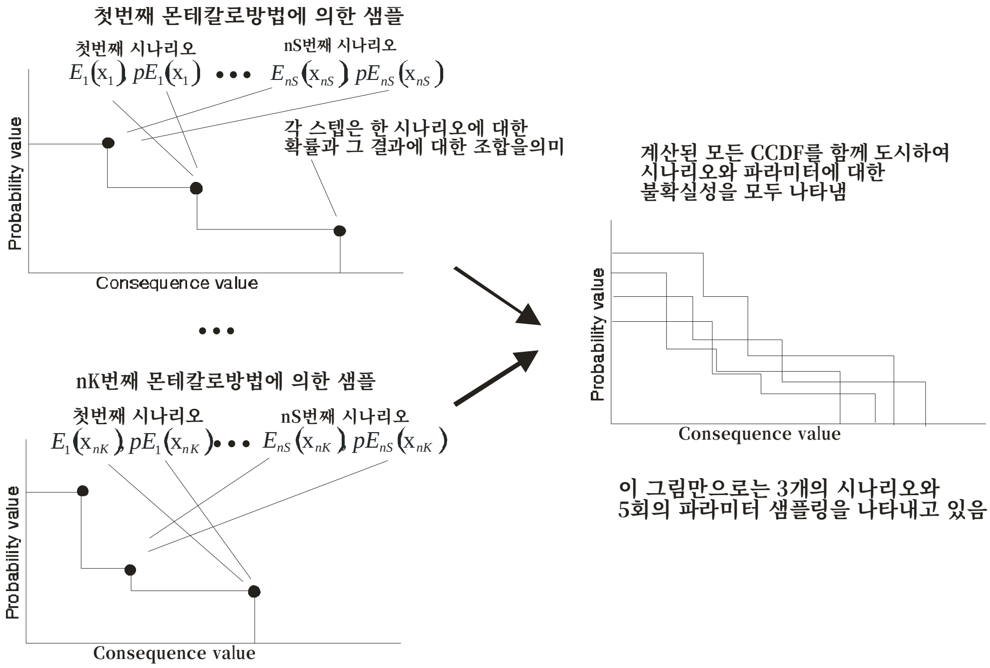


그림 3. PA 를 통하여 전파되는 시나리오의 추계적 불확실성과 그 결과에 대한 CCDF

불행히도 성능평가에 요구되는 대부분의 파라미터들은, 우선 추정하려는 변수값에 대한 시간 스케일이 실제로 사용하려는 시간 스케일에 비해 너무나 짧고 (평가하는 시간에 대해 상대적으로 엄청나게 긴 시간동안 처분장이 존재한다는 사실을 생각해 보자.), 뿐만 아니라 처분장의 전체 물리적 공간에 비해 추정하려는 물리적 스케일이 상대적으로 작기 때문에 이에 대한 정확한 사전 정보를 얻는 일은 거의 불가능하게 된다 (Helton, 1993a; Helton, 1993b). 따라서 직접적인 방법으로 입력자료의 통계적 성질을 특성화하려는 것은 사실상 불가능해 보인다. 그래서 보다 실질적이고 가능한 방법을 찾을 수밖에 없는데, 전문가 집단에 의한 검토 작업이 그 한 방법이 될 수 있다. 샌디아연구소에서 '전문가에 의한 규명 (Expert Elicitation)' 이라고 명명한 이 방법은 그 원리자체는 간단하다. (Hora and Iman, 1989)
한 특정한 변수 x_j 에 대한 함수 F 를 아래와 같은 식을 정의해보자.

$$prob(x < x_j \le x + \Delta x) = F(x + \Delta x) - F(x) \qquad (5)$$

이렇게 되면 $F(x+\Delta x) - F(x)$ 는 곧 x_j 에 대한 값이 x 와 $x + \Delta x$ 사이에 존재할 확률을 의미하게 된다.
이제 x_j 값에 대한 quantile 을 추정해가는 방법을 생각해 보자. 먼저 x_j 가 가질 수 있는 최소값 및 최대값을 상정하면 이 값이 각각 0.0 및 1.0 quantile 로 되어 그림 4 에 나타낸대로 두 양단의 점의 좌표로 정해진다.

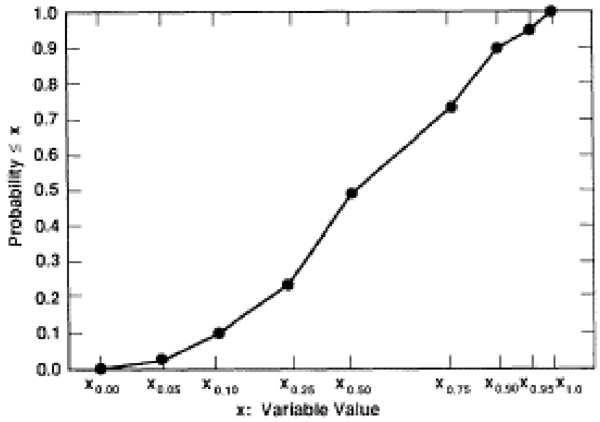


그림 4. 부정확한 입력자료에 대한 분포 함수의 예로서 0.00, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 그리고 1.00 의 quantile 을 추정하여 얻은 것

그런 다음에는 보통 중앙값 (median)을 추정하게 되는데, 이 점은 x_j 값에 대해 이 값을 넘을 확률과 그렇지 않을 확률을 균등하게 양분하게 된다. 그런 다음 0.25 및 0.75 quantile 을 차례로 추정하고, 대략 분포의 형태가 갖춰져 분포가 특성화가 될 수 있을 때까지 계속 0.05, 0.10, 0.90 이나 0.95 등의 quantile 을 차례로 추가 해 나가면 하나의 분포함수가 얻어지게 된다. 그러나 이러한 과정을 신뢰할 수 있도록 하기 위해서는 반드시 전문가에 의한 추정이 요구되는 것이다.

또 다른 대안이라면 불확실한 변수 x 에 대해 기대값이나 표준편차를 이용하여 정규분포나 로그정규분포 또는 베타분포 등의 분포를 parametric 하게 얻어내는 방안을 생각해볼 수도 있지만, 처분장에 관련된 어떤 파라미터가 하필이면 특정한 어떤 분포를 이룰 것이라는 가정이 필요하게 되는데 이에 대한 근거가 물론 존재하지도 않지만, 설사 존재한다 하더라도 그러한 분포에 대해 선택된 특정한 통계 파라미터 세트 (가령 로그정규분포의 평균치와 표준편차)에 대한 근거를 또 정당화하기는 쉽지 않으므로 이러한 방법은 바람직하지 않게 된다.

그래서 전문가에 의한 규명이 분포를 얻는 사실상 유일한 방법이 될텐데, 이러한 전문가에 의한 규명에 의해 x 에 들어 있는 각각의 변수 x_j, $j = 1,\Lambda\ nV$ 는 아래 (6)식과 같이 D_j 가 얻어지면, x 에 대한 불확실성이 $R(\mathbf{x})$ 로 전파되는 것을 알아낼 수 있게 되는 것이다. 일단 이렇게 입력자료가 추출되면 (generation), 평가를 위한 모델에 이 입력자료가 적용되어 계산이 진행되며(propagation), 불확실성 분석이 진행된다.

$$D_1, D_2, \Lambda\ , D_{nV} \tag{6}$$

그러면 이 분포를 가지고 몬테칼로 방법을 사용하여 x 에 대한 불확실성으로 부터 $R(\mathbf{x})$에 대한 불확실성을 얻는 예를 들어 보자.

다음과 같은 x 값에 대한 nS 개의 샘플을 생각해보자.

$$\mathbf{x}_s = \lfloor x_{s1}, x_{s2}, \Lambda\ , x_{s,nV}, \rfloor s = 1, \Lambda\ , nS \tag{7}$$

이제 각각의 샘플에 대해 위해도 분석을 수행하게 되면 그 위해도 결과는 샘플 s 에 대하여 다음 식으로 나타나게 된다.

$$R(\mathbf{x}_s) = \{[E_i(\mathbf{x}_s), pE_i(\mathbf{x}_s), \mathbf{cE}_i(\mathbf{x}_s)], i = 1, \ldots, nE_i(\mathbf{x}_s)\}, s = 1, \ldots, nS \quad (8)$$

그러니까 이 식은 입력자료의 한 세트 (즉, $\mathbf{x}_s$)에 대해 위해도 분석을 수행한 완전한 한 결과로서 $R(\mathbf{x}_s)$를 제공해주는 식으로 된다. 아래 그림 5 는 이렇게 부정확한 입력자료의 각 세트를 입력자료로 하여 얻은 위해도 결과를 CCDF 군으로 도시한 예이다.

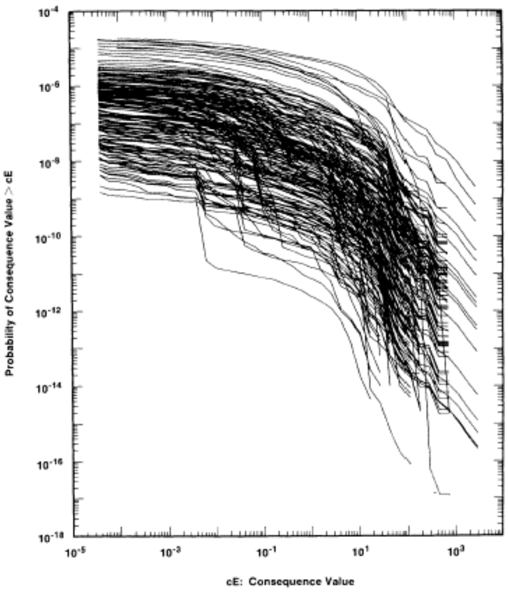


그림 5. 부정확한 입력자료를 샘플링하여 얻은 CCDF 의 예.

이 그림은 (7)식으로 주어진 불확실한 입력자료가 위해도 평가를 통해 어떻게 반영되어 전파해 나가는가 하는 것을 시각적으로 잘 보여주고 있다. 즉, CCDF 각각의 곡선 하나씩은 시나리오에 관한 추계적 불확실성 (stochastic uncertainty)만을 보여주고 있고, 이러한 각각의 CCDF 가 모두 *nS* 개의 샘플링 갯수만큼 군을 이루고 있는 것은 파라미터가 갖는 인식적 불확실성 (subjective uncertainty)을 나타내는 것이 되는 것이다.
복잡하게 얽힌 이 선들을 정리하여 보다 단순한 결과로 보여주는 방법의 한 예는 CCDF 군을 얻어낸 다음 각각의 결과로 부터 평균치와 10 번째, 50 번째 (그러니까 중앙값), 그리고 90 번째 percentile 을 얻은 후 이 percentile 값을 연결하여 그림 6 과 같은 보다 간결한 결과를 얻을 수 있다. 실제로 이 그림에 나타난 평균값을 EPA 핵종유출제한 규제치와 비교하는 방법이 WIPP 에 대한 성능평가에 사용되었다.

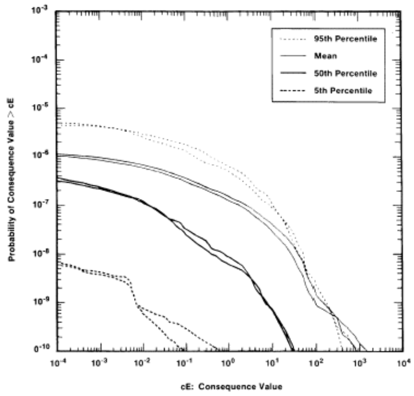


그림 6. 평균값과 percentile 을 이용하여 CCDF 를 요약한 그림.

그런가 하면 일정한 시나리오에 대하여 파라미터의 분포를 집약하여 그 불확실성만을 고려하여 CCDF 를 그리는 방법이 있을 수 있는데, 아래 그림 7 이 그 한 예이다. (Lee et al., 2002)

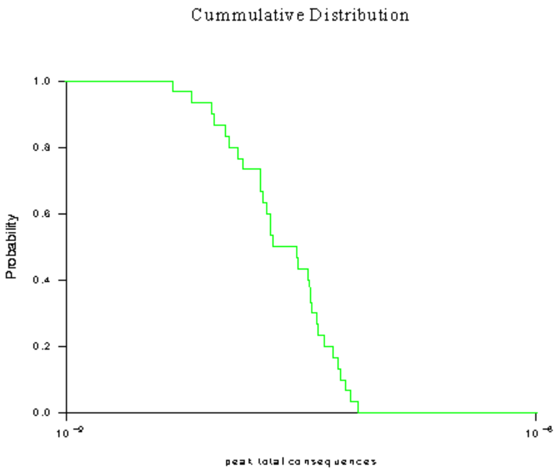


그림 7. 처분 시스템내 핵종 이동거리 분포 (Triangular distribution)에 따른 최고 총 피폭 선량(Peak total dose)결과의 CCDF

일정한 시나리오 한 개에 대하여 단순히 입력 파라미터의 기대치를 취하여 결정론적으로 결과를 낼 수 있으면 가장 바람직할 수 있지만 어떤 파라미터이든 그 불확실성을 고유하게 갖는다고 하였다. 그렇다면 이 파라미터의 분포를 찾아내어 그 분포에서 샘플링을 하여 결과를 한 그래프에 가령 시간에 대한 함수로서 몇 개라도 그려내면 되겠지만, 이 보다는 이 그림처럼 단일값으로 확률적인 관점으로 표현하는 것이 간단한 결과를 줄 수 있어 바람직할 수 있다.
X 축의 특정한 선량값을 갖거나 그 보다 큰 값을 가질 확률을 y 축에 나타냄으로서 결과를 확률론적인 관점에서 이렇게 간략하게 보여줄 수 있는 것이다.
가령 핵종 이동 거리의 분포를 모두 반영한 이 그래프에서 볼 때이동 거리가 어떤 값을 갖든 어느 경우에도 최고 총 피폭 선량이 10^{-8} Sv/yr 를 넘는 일은 없지만 어느 경우에나 10^{-9} Sv/yr 는 반드시 넘는다는 것을 의미하게 된다.
한편 민감도 분석이란 주지하듯이 계산이나 관찰을 통해 가장 영향을 주는 인자를 찾기 위한 방법이다. 누적 분포 함수 CDF 나 핵종의 유출률을 결정하는데 영향을 주는 불확실한 파라미터에 대해 그 중요도에 대한 상대적인 랭크를 얻어내는 것이라는 의미이다.

불확실성 분석이 평가 모델에 의한 결과의 예측에 사용된, 그러나 그 값을 정확히 알지 못하는 입력자료의 사용에 따른 정량적인 결과의 분포에 대한 불확실성을 분석하는 것일때, 민감도분석은 이러한 불확실성에 기여하는 각각의 입력변수의 기여도를 분석하는 작업이 되고, 따라서 민감도분석이라는 표현보다는 불확실성/민감도 분석이라는 보다 통합적인 표현이 더 타당한 표현일 수 있게 된다.
민감도분석을 위해서는 직관적인 결과를 주는 간단한 방법이 보편적으로 쓰이는 한편으로 보다 수학적으로 진보된 방법이 쓰이기도 하는데, 전자가 어떠한 시점에서 어떤 변수나 가정에 대한 정량적인 변화를 기본적인 값에 대해 그저 상대적이며 직관적으로 관찰해보는 방식인데 반해, 후자는 보다 formal 한 수학적인 방법들로 이들 방법의 공통적인 특징으로 비록 그 과정이 직관적인 방법에 비해서는 좀 복잡할 수는 있어도 계산의 수행에 있어서는 보다 적은 노력으로 보다 많은 민감도에 관한 정보를 얻을 수 있다는 장점이 있다는 것이다 (Helton and Davis, 2000).
WIPP PA 에서는 몬테칼로 분석 방법을 이용하여 불확실성 분석 및 민감도 분석을 수행하였는데, 확률론적인 방법으로 얻은 입력자료를 통한 반복적인 모델 평가에 바탕을 둔 이 방법은 모델을 통한 예측에 관련된 불확실성과 이 불확실성에 관련된 입력변수 자체에 대한 불확실성의 평가를 동시에 수행할 수 있게 한다.
몬테칼로방법은 몇 단계 정도로 그 절차를 요약해 볼 수 있는데, 아래 (9)식으로 표현되는 모델에 대해 생각해보자.

$$\mathbf{x}_j = [x_{i1}, x_{i2}, \Lambda, x_{in}], \quad i = 1,2,\Lambda, m \tag{9}$$

이 때 n 은 입력자료의 갯수를, m 은 표본의 크기 의미한다.

- 먼저 각 x_j 에 대해 값의 범위와 분포를 전에 기술한 방법, 즉 전문가에 의한 추정으로 결정한다.
- 그 다음에 이 범위와 분포로 부터 샘플을 LHS 등의 방법으로 추출한다.
- 이렇게 얻은 입력자료 표본 각각에 대하여 모델을 평가한 결과는 아래 식과 같이 된다.

$$y_i = f(x_{i1}, x_{i2}, \Lambda, x_{in}) = f(\mathbf{x}_i), \quad i = 1,2,\Lambda, m \tag{10}$$

결국 모델을 평가한다는 것은 입력자료, 즉 $\mathbf{x}_i$ 에 대하여 그 분석결과, 즉 y_i 에 대한 전사 (mapping)를 만들어내는 작업을 의미한다.

- 이제 (10)식을 기초로 하여 불확실성 분석을 하게 되는데, y 에 관한 불확실성을 나타내는 가장 간단한 방법은 평균치와 분산을 구하는 것으로, 무작위추출이나 LHS 에 의한 방법으로 (9)식과 같이 표본을 얻어낸 후에 y 에 대한 기대치와 분산을 다음 식과 같이 구할 수도 있다.

$$E(y) = \sum_{i=1}^{m} y_i / m \tag{11}$$

$$V(y) = \sum_{i=1}^{m} [y_i - E(y)]^2 / (m-1) \tag{12}$$

그러나 실제로 성능평가의 경우 출력변수값의 분포는 대개 한 쪽으로 쏠려 있어서 (highly skewed) 그러한 평균치나 분산의 변동이 심해 그 의미를 평가하기 어려울 수 있는데, 이 의미는 샘플링을 할 때마다 고른 결과를 주지않고 변동이 심하다는 의미이다. 그리고 이런 경우라면 평균과 분산만으로는 심하게 쏠린 분포에 대해 옳바르게 통계적 특성화를 할 수 없게

되면서, 불확실성에 관한 많은 정보를 단순히 한 개의 값으로 압축해버려 아까운 정보를 모두 잃게 된다.
한편 출력변수의 변화를 나타낼 때도 추정된 분포를 이용할 수 있는데 이 방법은 불확실성에 대한 정보로서 단순히 기대치와 분산을 뽑아내는 것 보다 그 불확실성을 특성화하는데는 더 유리하다. (13)식으로 부터 추정된 분포를 얻어내게 되는데 아래 그림 8 은 이렇게 추정된 분포함수의 예를 보여주고 있다.

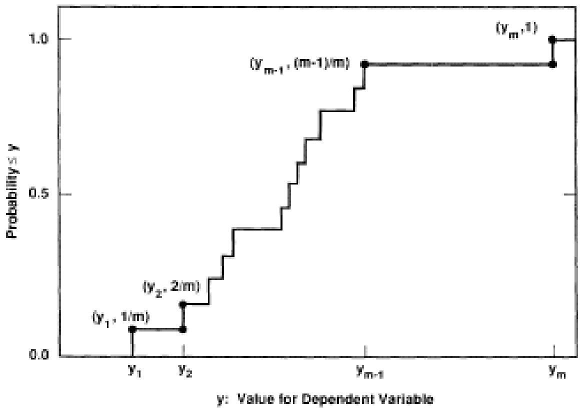


그림 8. 추정된 분포함수의 예

x 축은 출력변수의 값들을 표시하고 y 축은 누적확률 (cumulative probability)을 표시하여 x축의 값과 같거나 그 보다 작은 값을 가질 확률을 나타내게 된다. 물론 이 외에도 박스플롯 등으로 결과에 대한 분포함수를 나타내는 방법이 쓰이기도 한다.

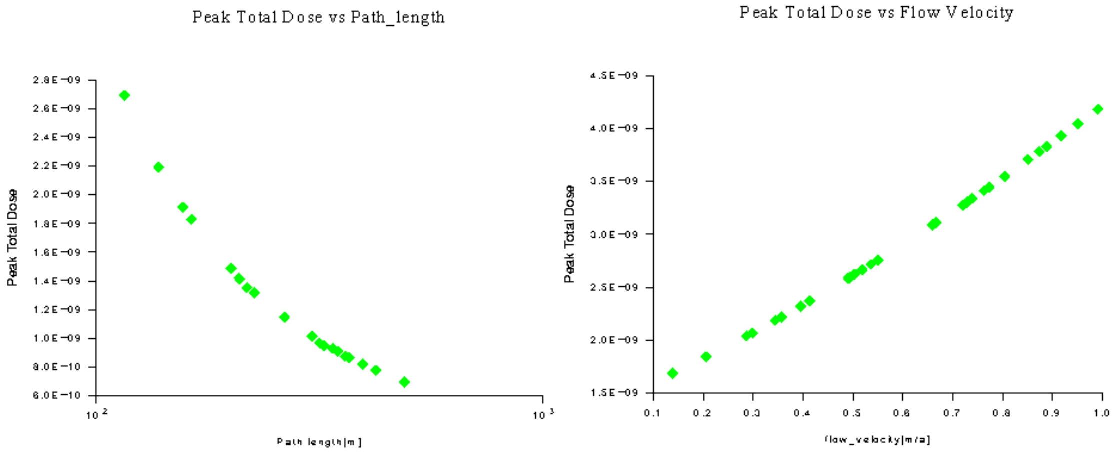
$$F(y)=\begin{cases} 0 & \text{if } y < y_1 \\ i/m & \text{if } y_i \le y < y_{i-1}, \quad i=1,2,\Lambda\ ,m-1 \\ 0 & \text{if } y_m \le y \end{cases} \qquad (13)$$

여기에서도 $y_i \le y_{i-1}$과 같이 y_i 가 순서를 가진다고 본다.
이 계단함수는 (10)식에 있는 모든 정보를 가지고 y 에 관한 불확실성을 표현하는 그래프를 만들어준다. 즉, $1-F(y)$ 로 부터 CCDF 를 얻어 낼 수 있게 된다. 이렇게 쉽게 CCDF 를 얻어낼 수 있는 점이 몬테칼로방법의 큰 특징이다.

• 이렇게 불확실성 분석이 끝나게 되면 마지막 단계로 민감도분석을 하게 된다. 이 민감도 분석은 결국 입력자료가 (10)식으로 표현되는 결과에 미치는 전사 (mapping)를 조사해보는 것인만큼 가장 단순하면서도 유용한 방법은 그저 산란그래프 (Scatterplot)를 작성해 보는 것이다. 즉, 입력변수 x_j 에 대해 출력변수 y 를 (14)식과 같은 점으로 그리는 것이다.

$$(x_{ij}, y_i), \quad i=1,2,\Lambda\ ,m \qquad (14)$$

이 점들의 분포는 x_j 와 y 의 관계를 보여줄텐데, 그 결과는 그림 9 에 예로 든 산란 그래프에서처럼 어떤 임계치 (threshold)를 갖는 형태나 또는 비선형인 방법을 보여주든지, 아니면 그저 전 영역에 걸쳐 퍼져 있는 점들의 무질서한 모습을 보여주거나 하게 된다. 이들 그래프중 x 축으로 표시된 핵종 이동거리와 y 축으로 표시된 최고 총 피폭선량의 관계를 도시한 (a)의 경우나, 지하수 유속과총 피폭선량의 관계를 도시한 (b)의 경우 모두 입력 변수와 그 선량 결과 사이에 매우 직접적인 상관 관계가 있음을 뚜렷이 보여주는 것을 알 수 있다.
한편 이 그래프의 결과중 (a)는 그림 7 과 연계하여 얻어낸 그래프이다.
이 방법은 특히 LHS 샘플링과 병행하여 좋은 결과를 줄 수 있는데 이는 이 방법의 속성상 입력변수의 어떤 범위에 한정되지 않고, 전 범위에 걸쳐 샘플링 범위가 나뉘어지기 때문이다.



(a) (b)

그림 9. 산란 그래프의 한 예 (Lee et al., 2002)

그러나 민감도 분석이 언제나 이렇게 간단하진 않다. 입력변수중에 결과에 영향을 지대하게 끼치는 변수가 있지 않는 한 언제나 이와 같이 뚜렷한 결과를 얻을 수는 없는 것이다. 그렇기때문에 여기서는 따로 논의하지는 않지만, 산란그래프 방법외에 회귀분석 (Regression analysis)이나 상관분석 (Correlation analysis) 등의 방법을 쓰기도 하는 것이다 (Helton, 1993b).

6. 맺는 말

이 논문을 통해 검토한 내용은, 결국 이러한 작업을 통해 EPA 규제에 대해 부합시킬 확률을 도출하기 위한 방법론이다.
WIPP 사이트에 대하여 그 안전성에 대한 책임을 지고 있는 DOE 가 EPA 에 신청하기 위한 CCA 를 위하여, 샌디아연구소가 수십년간 수행한 방대한 PA 관련 연구는, 그 연구결과나 과정이 우리나라에서 고려되고 있는 처분환경과는 상이한 것임에도 그 방법론에 대한 고찰은 우리나라 고준위 폐기물 처분 연구 일정과 방향에 비추어 매우 유용한 일일 것이다.
우리의 경우도 점점 단순한 결정론적 결과로 나타나는 규제기준이 마련되는 대신 통계적으로 불확실성을 정량화하는 확률론적인 규제 기준을 마련하는 방향으로 나가는 추세이고 이에 대한 규제 기관의 기준이 곧 마련될 전망이어서, 비록 이 보고를 통해 깊이있는 검토가 이뤄지지는 않았지만 향후 고준위 처분 시스템의 설계 및 인허가에 대비한 우리의 연구 방향을 생각해 볼 수 있는 계기가 필요한 것으로 사료되어 향후 보다 심도있는 관련 연구가 필요하다는 생각에서 이렇게 간략하게나마 방법론에 대한 검토를 수행하여 보았다.

우리도 서둘러 이 논문을 통해 제시된 연구방향과 유사하게 성능평가에 요구되는 파라미터의 불확실성을 수용하고 이 들 파라미터의 상대적인 중요도를 파악하기 위해 파라미터의 민감도 및 불확실성 분석관련 연구를 수행해 나가야 할 것이다.
다만 WIPP 사이트의 경우와는 다르게 부지 특성적인 처분시스템을 아직 가지지 못한 우리나라의 경우로서는 일정한 시나리오에 대해서 입력 파라미터의 확률론적 분포를 고려하는 민감도 및 불확실성 분석을 통한 방법론적 접근을 통한 확률론적 성능평가 및 파라미터 비교우위 도출 방안이 보다 현실성 있는 방안으로 제안될 수 있을 것이다.

참고문헌

"Environmental Standards for the Management and Disposal of Spent Nuclear Fuel, High-Level and Transuranic Radioactive Waste: Final Rule," 40CFR191, *Federal Register*, 50, 38066 (1985).

B.A. Howard, "Regulatory basis for the Waste Isolation Pilot Plant performance assessment," *Reliab. Eng. Sys. Saf.*, 69, 109 (2000).

J.C. Helton et al., Sensitivity Analysis Techniques and Results for Performance Assessment at the Waste Isolation Pilot Plant, SAND90-7103, UC721, Sandia National Laboratories (1991).

J.C. Helton, "Risk, uncertainty in risk, and the EPA release limits for radioactive waste disposal," *Nucl. Tech.*, 101, 18 (1993a).

J.C. Helton, " Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal," *Reliab. Eng. Sys. Saf.*, 42, 327 (1993b).

J.C. Helton and F.J. Davis, " Chapter 6. Sampling-Based Methods," p.101, Sensitivity Analysis, Wiley, (2000).

J.C. Helton et al., " Characterization of stochastic uncertainty in the 1966 performance assessment for the Waste Isolation Pilot Plant," *Reliab. Eng. Sys. Saf.*, 69, 167 (2000).

J.C. Helton, " Conceptual structure of the 1966 performance assessment for the Waste Isolation Pilot Plant," *Reliab. Eng. Sys. Saf.*, 69, 151 (2000).

S.C. Hora and R.L. Iman, " Exprt opinion in risk analysis: The NUREG-1150 Methodology," Nucl. Sci. Eng., 102, 323 (1989).

Y.M. Lee et al., being prepared for publication (2002).