

오염토양의 효율적 조사기법과 측정불확도의 평가방법

이종천

국립환경연구원

Cost-effective Investigation on Contaminated Land and Assessment of Measurement Uncertainty

Jong-Chun Lee

National Institute of Environmental Research, Incheon 404-170, Korea

The concentration and distribution of contaminants obtained from a contaminated land investigation or an environmental geochemistry survey constitutes the basis of a decision-making process on environmental policies or of scientific researches. As the quality of data determines the reliability of the result, the investigation plan should be adjusted according to the purpose of the investigation. In general, the effort to improve the data quality had been focused mainly on the QA/QC procedures in laboratories. The rapid progress of analytical instrument has also contributed toward improving the analytical precision to a satisfiable degree. Nevertheless, in many cases, it is not the analytical precision that needs improvement for the better precision of overall measurement process: it is rather during the sampling process in the field that is responsible for the poor precision. To assess the data quality on a measured value, ISO recommends to provide information on "measurement uncertainty" along with the measured value. The measurement uncertainty in an environmental measurement context can be explained as the statistical number that expresses the degree of the uncertainty stemming from the sampling and analytical procedures. There is a cost involved in order to improve the precision of sampling and analytical methods so as to decrease the degree of measurement uncertainty. The economical point of compromise in an investigation planning can be achieved when the allowable degree of uncertainty has been set beforehand. The investigation can then be planned accordingly not to exceed the uncertainty limit. Furthermore, if the measurement uncertainty estimated from the preliminary investigation can be separated into sampling and analytical uncertainties, it can be used as a criterion where the resources for the investigation should be allotted cost-effectively to reinforce the weakest link of the whole investigation process. This paper aims to present a method of estimating the level of measurement uncertainty of a measured contamination concentration at a site used as an example and to show how the estimated uncertainty can be applied to serve the particular purpose of an investigation.

Key words : contaminated land investigation, measurement uncertainty, fitness-for-purpose criteria, data quality, cost-effective investigation

토양오염도 조사 혹은 환경지질학적 조사를 통해 측정된 오염물질의 농도나 분포에 대한 자료는 환경정책에서의 의사결정이나 과학적 연구의 기초를 제공한다. 즉 이러한 측정자료의 질적 수준에 따라 결과에 대한 신뢰도가 영향을 받게 되므로 이를 감안하여 조사 목적에 알맞은 조사기법을 채택하여야 한다. 자료의 질적 수준을 향상시키려는 노력은 대부분의 경우 실험실에서의 정도관리를 통해 이루어져왔고 이와 더불어 급속한 기기분석기술의 발전으로 인하여 분석시료에 대한 만족할 만한 분석정밀도를 가져오게 하였다. 그러나 많은 경우 환경조사로 얻어지는 자료의 질적 수준은 실험실에서의 분석과정에서 결정되기 보다는 현장에서의 시료채취과정에서 더 큰 영향을 받게 된다. 측정된 자료의 질적 수준을 가늠하는 척도로서 국제표준화기구에서는 측정불확도를 측정값과 함께 제시하도록 권고하고 있다. 이를 환경측정에 적용하면 야외시료채취과정과 실험실내에서의 분석과정에서 기인하는 불확실성을 통계적 수치로 나타낼 수 있게 된다. 측정불확도를 감소시키기 위해 시료채취방법과 분석방법의 정밀도를 높이는 데는 경제적 비용이 증가할 수밖에 없다. 경제적인 조사가 되기 위해서는 의사결정이나 연구과정에서 허용 가능한 불확도 수준을 결정해

*Corresponding author: jcleee@me.go.kr

야 하고 이를 기준으로 조사방법을 설계할 수 있다. 오염도양에 대한 예비조사에 있어서 시료채취 및 분석단계에서 기인하는 불확도를 평가할 수 있다면 그 중 더 큰 불확도를 발생시키는 단계를 우선으로 정밀도가 향상되도록 조사방법을 재설계 할 수 있다. 환경측정값에 대한 측정불확도 수준의 평가는 불필요한 경제적 자원의 낭비를 막을 수 있는 효율적인 조사방법을 설계하는데 합리적인 기준으로 응용될 수 있을 것이다. 본 논문에서는 측정불확도를 산출하는 방법을 한 오염조사지역에서 측정된 자료의 예를 들어 설명하였고, 이렇게 구해진 측정불확도는 기술한 목적에 따라 응용될 수 있음을 보여주었다.

주요어 : 오염도양조사, 측정불확도, 조사목적에 따른 기준, 자료의 질, 비용효율적 조사방법

1. 서 론

지질학적, 농학적 관점에서 주를 이루던 도양에 대한 조사는 이제 환경적 측면을 고려한 국토의 효율적 이용을 위하여 점차 그 중요성이 높아지고 있다. 환경부에서도 전국 도양에 대한 오염추세를 파악하고 오염우려지역에 대한 오염실태를 조사하여 도양오염을 예방하고 오염도양을 정화하는 등 도양보전대책을 수립·추진하기 위하여 전국에 1,500지점에 대한 도양측정망을 운영하고 2,000지점이 넘는 오염우려지역에 대한 실태조사를 수행하고 있다(환경부, 2003).

지화학적 탐사뿐만 아니라 오염이 우려되는 부지에 대한 환경계획에는 조사대상에 대한 정확한 측정결과가 요구된다. 정확하지 못한 측정결과에 대한 불확실성은 이를 바탕으로 한 의사결정이나 연구결과의 신뢰도에 큰 영향을 미치게 된다. 예를 들어 측정값에 대한 불확실성은 위해성평가나 환경영향평가, 또는 환경가치계산과정(Environmental Valuation)에서 파급, 증폭되어 그 결과가 왜곡되거나 판단오류를 초래하기도 한다.

측정결과에 대한 불확실성은 시료채취와 분석과정에서 측정대상의 불균질성, 또는 방법상의 오류나 낮은 정밀도로 인하여 발생하거나 도입된다. 그렇기 때문에 측정값의 불확실성을 감소시켜 자료의 질(data quality)을 향상시키기 위해서는 야외시료의 양과 수를 늘리고 QA/QC를 통한 분석과정의 관리, 또한 최신의 정밀분석방법을 이용하는 등 노력이 뒷받침되어야 한다.

이렇게 조사대상을 정확하게 인식하고 판단하기 위한 노력에는 많은 인적, 시간적, 경제적인 자원이 필요하다. 이와 더불어 불확실한 측정값을 근거로 잘못된 의사결정이나 연구결과가 도출될 경우 초래될 수 있는 경제적 손실은 조사에 소요되는 또 다른 경제적 요소이다. 이러한 까닭에 경제적으로 효율적인 조사를 위해서는 측정값의 불확실성의 정도를 줄이기 위해 소요되는 비용과 불확실한 측정값을 참조함으로써 발생되

는 의사결정의 오류에 의한 손실이 균형을 이루는 조사설계가 필요하다.

반면에, 측정값에 대한 불확실성을 줄이지 않더라도 불확실성의 원인과 규모를 평가할 수 있다면 대상 부지에 대한 올바른 평가와 객관적이고 타당한 의사결정(혹은 방어적 의사결정, defensible decision-making)을 위한 토대가 될 수 있다. 측정값에 대한 불확실성을 고려하더라도 오염기준이 초과되는 경우나 더욱 확실한 측정이 필요한 경우 추가적인 정밀조사가 필요하게 된다. 이때, 앞서 평가된 불확실성의 원인과 규모를 참고하면 정밀조사를 비용효율적으로 설계하는데 응용할 수 있다. 즉 제한된 비용을 시료채취과정과 분석과정에 효율적으로 분배하여 측정값에 대한 불확실성을 최소화 하는데 이용할 수 있을 것이다. 이는 기존의 토양오염조사나 지화학적 조사와 같은 환경측정에서 측정값의 정확성 내지는 측정과정의 정밀도를 향상시키기 위하여 실험실내에서 분석과정에만 주의를 기울여 온 것과 비교하여 시료채취과정에 더욱 큰 중요성을 부과할 수 있는 근거가 될 수 있다.

다른 환경매체와는 달리 토양오염의 경우 오염이 고르지 못하게 분포(heterogeneous distribution)되는 경우가 많다. 이러한 이질성은 오염물질의 이화학적 특성, 오염의 원인과 지질학적, 도양학적 특성이 결합함으로써 다양한 결과를 보이게 된다. 이러한 부지특성과 오염특성을 고려하지 않은 시료채취방법은 분석과정과 비교되지 못할 정도로 측정결과에 불확실성을 높일 수 있다.

본 논문에서는 환경측정 및 지화학적 탐사에서 측정값에 대한 불확실성을 통계적인 방법을 이용해 정량화하여 측정불확도(measurement uncertainty)로 표현하는 방법을 알아본다. 이 측정불확도를 응용하여 환경조사결과 측정값에 근거하여 확률적인 의사결정을 내리게 할 수 있음으로써 의사결정의 오류를 감소시킬 수 있는 방법을 알아본다. 또한 측정불확도의 원인과 그 저감방안, 그리고 비용효율적인 조사기법이 되기 위

해 측정불확도를 이용하는 방법에 대해서도 살펴본다.

1.1. 측정불확도의 정의

측정값에 대한 불확실성은 국제표준화기구(ISO)에 의해 측정불확도(measurement uncertainty, 이에 대한 용어 번역은 한국표준과학연구소의 우리말 번역을 참고하였다)로 그 크기가 표시되고 “측정결과에 관련하여 측정량을 합리적으로 추정한 값의 분산특성을 나타내는 파라미터” 라고 같이 정의된다(ISO, 1993).

이는 다시 “측정값주변에서 참값이 (높은 확률로) 분포하고 있을 범위를 나타내는 변수(Thompson, 1995)” 라고 풀어 설명될 수 있다.

측정불확도를 산출하기 위해서는 먼저 대상지역 측정값들에 대한 총 분산(total variance)이 모두 세 개의 요소로 구성되어 있음을 알아야 한다. 먼저 측정대상물질이 부지 내에서 분포변화를 보임으로써 나타나는 지화학적 분산(geochemical variance, $s_{geochem}^2$)으로서, 이는 측정지점간의 실제적인 농도차이에 의해 나타난다. 이와 더불어 시료채취과정에서 비롯되는 sampling variance(s_{samp}^2), 분석과정에서 비롯되는 analytical variance(s_{anal}^2)인데(Ramsey, 1998), 이에 대한 설명은 뒤에 이어지거나 수치적인 연관성만을 살펴보면 다음과 같다.

$$s_{total}^2 = s_{geochem}^2 + s_{samp}^2 + s_{anal}^2 \quad (1)$$

측정불확도(u)는 위의 시료채취과정과 분석과정에서의 분산을 이용하여 아래와 같이 계산된다.

$$u = s_{meas} = \sqrt{(s_{samp}^2 + s_{anal}^2)} \quad (2)$$

- s_{meas} : 측정과정에서의 측정표준편차, 이하 “meas” for measurement
- s_{samp} : 시료채취과정에서의 시료채취표준편차, 이하 “samp” for sampling
- s_{anal} : 분석과정에서의 실험표준편차, 이하 “anal” for analysis

U로 표시되는 확장불확도는 측정량의 합리적인 추정값이 이루는 분포의 대부분을 포함할 것으로 기대되는 측정결과 주위의 어떤 구간을 정의하는 양으로서

(ISO, 1993), 보통 95% 신뢰도에 해당하게끔 포함인자(coverage factor)인 $k=2$ 를 적용하여 아래 식(3)과 같이 나타낸다.

$$U = ku = 2s_{meas} \quad (3)$$

1.2. 측정불확도의 원인과 종류

측정과정을 크게 야외에서의 일차시료채취과정과 실험실에서의 분석과정으로 구분한다면 측정불확도의 원인은도 각각의 과정으로 구분하여 생각할 수 있다. 또한 각 과정에서 측정불확도의 원인을 나누어 보면 환경측정대상이 가지고 있는 내재적 이질성(inherent variability)과, 측정과정에서의 오차(error)가 서로 영향을 미치게 됨으로써 발생한다(Table 1).

다시 말해, 측정대상이란 일차시료채취과정에서는 시료채취지점에서의 토양이 되겠고, 실험실에서는 일차시료를 분석하기 위해 나누어 놓은 분석시료(sub-sample)가 그 대상이 될 수 있다. 같은 지점에서 동일한 채취 방법에 따라 채취된 시료라 하더라도 서로 다른 성분을 가질 수 있다. 그 차이는 보통 채취지점에 대한 지점확인오류(locating error, 즉 시료채취 후 동일 지점을 다시 확인할 때 생기는 오차를 말하며, 시료채취지점을 표시하는데 이용한 지도의 축척 등과 관련이 있음)에 영향을 받는데(Ramsey, 1997), 이 오차가 없더라도 토양의 국지적 이질성(small-scale heterogeneity)이 큰 경우에는 그 차이가 증가하게 된다.

조사 대상물질의 분포변화는 그 공간적 범위에 따라 지역적 이질성(large-scale heterogeneity)과 국지적 이질성(small-scale heterogeneity)으로 나눌 수 있는데 전자는 (1) 식에서 $s_{geochem}^2$ 에 해당되고, 후자는 s_{samp}^2 으로 설명되어진다. 또한, 야외시료를 실험실에서 분석을 위해 필요한 양 만큼을 취하는 경우, 시료를 구성하는 입자상 이질성(particulate-scale heterogeneity)이 분석과정에서 기인하는 분산(s_{anal}^2)에 기여를 한다.

측정과정에서의 오차(measurement error)는 우연오차(random error)와 계통오차(systematic error)로 나눌 수 있다. 이는 차례로 측정방법의 정밀도(precision)와 편차(bias)수준을 결정한다. 여기서 시료채취과정에서의 오차는 토양의 국지적 이질성의 영향, 측정장비

Table 1. Classification of sources of measurement uncertainty during a measurement procedure.

Measurement	Sampling	Analysis
Inherent variability	Small-scale heterogeneity at the sampling location	Particulate heterogeneity of a sub-sample
Error	Locating error (sampling bias and precision)	Sub-sampling error (analytical bias and precision)

로부터의 오염, 잘못된 보관상태, 시료의 기질 조성(sample matrix) 등이 될 수 있고, 분석기기 보정의 정확도, 기기의 오염, 시약의 순도, 반응림, 영점조정 등이 분석에서의 오차의 예에 해당 된다(Eurachem, 2000). 그러나 실수로 설명되는 오차는 다른 오차와 쉽게 구분되며 따라서 논의에서 제외되어야 한다(Miller and Miller, 1993).

사실상 국제표준화기구(ISO, 1993)는 계통 오차(systematic error)는 측정과정에서 검출되어야 하고 따라서 이에 의한 영향이 측정값에 도입되는 것이 근본적으로 제거되어야함을 강조하고 있다. 그러나 사실상 판명되지 못한 계통 오차는 측정과정에서 제거되지 못하는 경우가 있을 수 있어 결국 측정불확도에 영향을 미치게 되는게 일반적이다.

1.3. 측정불확도의 저감방안

이와 같이 시료채취와 분석단계에서 각각 우연오차(random error)와 계통오차(systematic error)에서 기인하는 측정불확도를 저감시킬 수 있는 방법을 살펴보면 다음과 같다.

먼저 측정불확도의 우연적 부분(random component of measurement uncertainty)은 결국 분석방법의 정밀도(precision)에 의해 결정된다. 그러므로 보다 정밀한 분석방법을 사용하면 이 오차를 감소시킬 수 있다. 다른 방법으로는 시료의 개수를 늘리고 반복분석을 함으로써 오차를 감소시킬 수도 있다(e.g. $s = \frac{\sigma}{\sqrt{n}}$).

토양시료와 같이 분석시료가 입자상으로 되어 있을 경우 입자상 이질성(particulate-scale heterogeneity)에 의해 시료를 분취하는 단계에서 오차가 생길 수 있다. 이 과정에서 생기는 오차의 종류와 그 대책을 Pitard (1993)가 전개해 놓았는데, 이는 Gy's sampling theory(Gy, 1979)를 기초로 하고 있다.

$$s^2 = Cd^3/M \quad (4)$$

C=시료 상수(광물조성인수, m, 자유인수, l, 모양인수, f, 입자분포인수, g, 등으로 구성)

d =분석대상물질을 함유하고 있는 시료입자의 지름
M=시료의 질량

이 식은 샘플링에서 생길 수 있는 오차(여기선 시료채취과정에서의 분산, s^2 로 표현)가 시료의 입자상 성분과 시료의 양에 따라 달라짐을 나타내고 있다. 따라서 이를 참고하여 일차시료를 분취할 때 오차를 줄임으로써 측정불확도를 감소시킬 수 있다.

측정불확도의 계통적 오차부분(systematic component of measurement uncertainty)은 분석과정에서 인증시료(certified reference material)와 reagent blank를 이용하여 통계처리함으로써 편차(bias)에 의한 영향을 배제시킬 수 있다. 그러나 시료와 인증시료간의 기질(matrix)의 차이, 즉 성분 및 입자크기의 차이에 의해 정확한 편차(bias)를 도출하기 어려운 경우가 있는데 이로 인해 결국에는 측정불확도를 증가시키는 요인이 되기도 한다(Ramsey and Argyraki, 1997).

일차시료 채취에서도 채취방법에 따라 특정입자의 손실이나 이전 시료에 의한 장비의 오염 등 환경적 요인에 의해 편차(bias)가 생길 수 있다. Ramsey 등(1995)은 분석시의 인증시료를 이용한 편차 제거와 유사한 방법으로, 많은 횟수의 시료채취에 의해 그 값이 알려진 지역인 reference sampling target을 이용, 채취방법상의 편차를 도출하는 방법을 제안한 바 있다. 그러나 일반적으로 부지특성의 차이는 분석편차 보정과정에서 문제가 되는 시료성분 차이보다 더 크므로 이 방법을 현실적으로 적용하기에는 어려운 점이 있다.

1.4. 측정불확도수준의 평가방법

측정값에 대한 측정불확도를 평가하기 위한 방법은 크게 상향식 합산방식(bottom-up approach)과 하향식 계산방식(top-down approach)이 있다.

상향식 합산방식은 측정 과정에서의 세부단계별 불확도수준을 추출하여 결합시키는 방법으로써 국제표준화기구에서 채택하고 있는 방법이다(ISO, 1993). 이 방법은 주로 분석방법을 표준화시킬 때 이용되나 불확도에 기여하는 모든 세부단계를 알아내기에는 매우 어렵기 때문에 일반적 이용에 제한이 따른다(AMC, 1995).

이에 대한 대안으로는 하향식 계산방식(top-down approach)이 있는데, 각 측정의 세부단계를 분리시키기 보다는 집합적 과정으로 나누어 불확도 수준을 결정하는 방법이다. 오염부지 조사를 예로 보면 시료채취단계와 분석단계로 나눌 수 있고, 각 단계별 불확도 수준을 평가함으로써 총 측정불확도를 결정할 수 있으며 그 방법은 다음과 같다.

일차적 야외시료채취와 실험실내 분석과정에서 기인하는 측정불확도를 알아내는 기존의 방법으로는 통계분석기법인 분산분석(ANOVA)이 있는데, 이를 이용하여 반복 샘플링(replicate sampling)에 대한 반복 분석(replicate analysis)된 값을 토대로 시료채취 단계와 분석 단계의 정밀도(precision)를 추출할 수 있다(Ramsey, 1997). 이 방법은 원래 지하확탐사에서의 정도관리

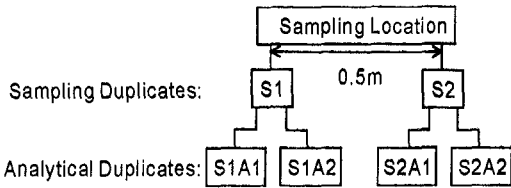


Fig. 1. Balanced experimental design for the uncertainty calculation by ANOVA showing sampling duplicate distance of 0.5 m from original sampling location.

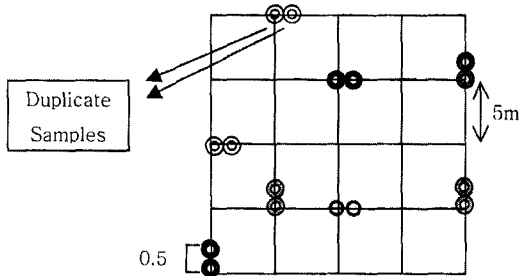


Fig. 2. 8 sampling locations randomly chosen from the regularly-spaced grid from which each duplicate sample will be taken from the distance of 0.5 m, reflecting locating error and/or small-scale heterogeneity of the soil.

(quality control)를 위해 개발되어 발전해 왔다(Miesch, 1967; Garrett and Goss, 1979). 시료채취단계와 분석단계에서 각각 반복을 통해 얻어지는 자료는 대칭적 구조(balanced design, Fig. 1)를 가지고, 이를 ANOVA를 이용하여 각 단계에서의 정확도 수준을 평가할 수 있다. 또 이 둘을 결합해 측정값에 대한 총 불확도 수준을 계산할 수 있다.

Fig. 1은 야외시료 채취단계에서 임의의 채취지점으로 부터 중복시료(duplicate sample)를 채취할 때 예상되는 지점확인오류(locating error)와 부지 특성상의 국지적 이질성(small-scale heterogeneity)을 고려한 간격(여기선 0.5 m), 그리고 각 중복시료에 대한 중복분석(duplicate analysis)의 체계가 대칭적 구조로써 정렬되는 모습을 나타낸다.

Fig. 2는 한 부지에서 격자를 이용한 시료채취방법에 의한 측정불확도수준을 알아내기 위해서 무작위로 선정된 8 지점에 대해 각각 중복시료가 채취되는 것을 보여준다. 각 시료에 대한 중복분석을 통해 얻어진 데이터는 Table 2와 같이 정렬되어 분산분석(ANOVA)을 거쳐 시료채취단계와 분석단계 각각의 분산치가 구해진다.

오염지역의 측정값들은 대개의 경우 평균적인 오염 분포와는 매우 동떨어진 차이를 보이는 값들(outliers)을 포함한다. 이를 그대로 사용할 경우 정규분포를 기반으로 한 통계방법(parametric statistics)에 큰 영향을

Table 2. Example of the balanced desinged data ready for RANOVA (Lee, 2002).

Sample	As(ug · g ⁻¹)			
	S1A1	S1A2	S2A1	S2A2
C-1-A4	112.699	110.037	88.965	97.564
C-1-B2	91.320	89.428	84.230	85.522
C-1-B5	167.575	174.829	163.598	170.420
C-1-C2	114.455	120.427	127.687	126.938
C-1-D1	35.523	36.291	21.452	18.785
C-1-D4	166.816	156.812	146.004	143.689
C-1-D5	130.564	135.232	174.711	192.017
C-1-E5	143.566	142.917	155.193	147.770

미치게 된다. 이러한 값으로부터의 영향(sensitivity)을 감소시키기 위해 log-transformation을 이용할 수도 있으나 robust statistics의 일종인 robust ANOVA가 효과적으로 사용된다(AMC, 1989). Table 3은 Table 2의 데이터로 분산분석을 실시했을 때 outliers의 영향을 감소시킨 robust ANOVA와 일반적인 ANOVA(classical ANOVA)와의 차이를 보여주고 있다.

분산분석의 결과 식(1)과 같이 분리된 각 단계에서의 분산을 식(2)와 (3)을 이용하여 다음과 같이 측정불확도를 계산할 수 있다.

$$u = s_{meas} = \sqrt{(s_{samp}^2 + s_{anal}^2)} = \sqrt{(11.212)^2 + (4.4283)^2} = 30.822 \text{ ug} \cdot \text{g}^{-1}$$

$$U_{meas} = ku = 2s_{meas} = 2 \cdot 30.822 = 61.644 \text{ ug} \cdot \text{g}^{-1}$$

이런 과정을 통해 구해진 확장측정불확도(U_{meas})를 절대불확도(absolute uncertainty, 식(5))라 하고 측정값에 대비해 다음과 같이 설명된다.

c(measured concentration) ± U_{meas}: 실제 오염물질의 농도가 95%의 확률로 존재하는 범위

$$\text{where } U_{meas} = 2s_{meas} = 2 \sqrt{(s_{samp}^2 + s_{anal}^2)} \quad (5)$$

U_{meas}: expanded measurement uncertainty

s_{meas}: measurement precision expressed in unit of standard deviation

s²_{samp}: variance of sampling method estimated from the duplicate sampling

s²_{anal}: variance of analytical method estimated from the duplicate analysis

그러므로 Fig. 2의 조사지역에서 측정된 평균농도는 Table 3에서 계산된 바 125.89 ug · g⁻¹이며 이에 대한 측정불확도 수준은 95% 신뢰도 수준에서 ±61.644 ug · g⁻¹임을 알 수 있다.

이때 U_{meas}은 동일한 측정방법이 사용된 조사지역

Table 3. Result of ANOVA with the data in Table 2. Uncertainties were estimated based on the result.

CLASSICAL ANOVA RESULTS:			
Mean = 121.03			
Standard Deviation (Total) = 48.426			
	Geochemical	Sampling	Analysis
Sums of Squares	62178	3624.9	355.81
Standard Deviation	45.906	14.678	4.7157
Percentage Variance	89.865	9.1869	0.94830
ROBUST ANOVA RESULTS:			
Mean=125.89			
Standard Deviation (Total)=42.956			
	Geochemical	Sampling	Analysis
Standard Deviation	41.230	11.212	4.4283
Percentage Variance	92.125	6.8119	1.0627
Relative Uncertainty(%) (% at 95% confidence)	-	17.811	7.0351

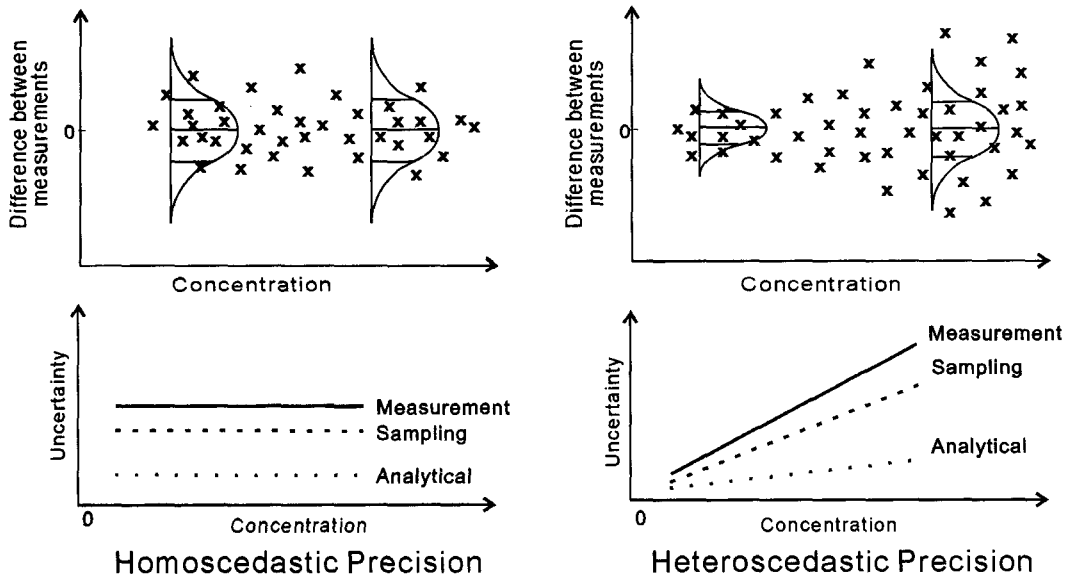


Fig. 3. Schematic representation of relationship of uncertainties as a function of concentration. Uncertainty estimation using RANOVA assumes that the analytical and sampling uncertainties are both independent of concentration or homoscedastic (left hand side). However, the heteroscedastic precision (right hand side) can be verified from the linear precision modelling with concentration. The measurement uncertainty can be derived from combining the sampling and analytical precision models after converting them into the form of variance (Lee and Ramsey, 2001).

에서의 평균값 부근 농도($c \approx \bar{x}$)에 적용될 수 있으나, 일반적인 경우 분석과정의 정밀도는 분석대상의 농도에 비례하므로(Thompson, 1995) 평균값에 대한 측정 불확도의 비율(relative uncertainty, U%)을 각 측정값에 적용하여 불확도 수준을 나타낼 수도 있다(식6).

$$c \pm U\% \cdot c \tag{6}$$

c : measured concentration

$$U\% = 200s_{\text{meas}}/c = 200 \cdot 30.822/125.89 = 48.97\%$$

이로써 조사지점중 $100 \text{ ug} \cdot \text{g}^{-1}$ 의 측정값을 갖는 지점의 측정불확도는 $48.9 \text{ ug} \cdot \text{g}^{-1}$ 이며, 이는 다시 말해 측정값 주위에 참값이 95%의 확률로 존재하는 구간이 $100 \pm 48.9 \text{ ug} \cdot \text{g}^{-1}$ 임을 설명해 준다.

그러나 U%를 적용하기 위해서는 시료채취과정에서의 불확도 수준도 채취지점의 농도에 따라 증가한다는 가정이 필요하고, 또 그렇다 하더라도 조사지역 전체의 평균값에 비해 너무 높거나 낮은 농도를 보이는 측정값에 적용하기에는 문제점이 있어왔다. 이를

해결하기 위해서 농도변화에 따른 시료채취와 분석시의 불확도 변화양상을 각각의 선형정밀도모형(linear precision model)으로 나타냄으로써 측정대상물질의 농도변화에 따른 측정불확도 수준을 보다 정확히 산출하는 방법이 개발되었다(Lee and Ramsey, 2001). 그 결과 부지특성과 오염특성에 따라 측정불확도가 농도에 따라서 통계적으로 유의한 변화를 보이는 경우와 그렇지 않은 경우로 분리할 수 있음을 알 수 있었다(Fig. 3).

1.5. 오염조사결과의 적정 측정불확도수준

그렇다면 허용가능한 측정불확도 수준은 무엇일까? 지화학탐사의 경우 Garret(1969)는 시료채취와 분석과정, 즉 측정과정에서의 기인하는 분산(measurement variance = $s^2_{meas} = s^2_{samp} + s^2_{anal}$)은 탐사지역의 측정값에 대한 총 분산(total variance = $s^2_{total} = s^2_{geochem} + s^2_{meas}$, 식1)의 25%수준이 넘어서는 안 된다고 그 기준을 제시했다. 이를 초과하는 탐사방법의 경우, 탐사에서 알아내고자 하는 분포변화(geochemical pattern)를 추적하기가 힘들게 되기 때문이다. Ramsey 등은 (1992) 오염부지 조사의 경우, 그 기준으로써 20%를 제시하고 이를 초과하는 조사방법에 의한 측정값으로는 오염중심지점(hot-spot)을 배경 오염도(background concentration)에서 분리해 내기가 어려워지기 때문이라 했다. 이와 마찬가지로 분석과정에서의 분산(s^2_{anal})이 측정과정에서의 분산(s^2_{meas})의 20%보다 작을 경우, 분석의 정밀도 수준이 시료채취과정에서의 정밀도 수준보다 필요이상으로 높음을 알려주는 기준으로서 이를 비용효율적인 조사방법을 설계할 때 참고하도록 권고하고 있다.

이러한 측정과정에서의 적정 불확도 수준은 절대적인 것으로 볼 수 없다. 환경조사나 탐사가 특별한 목적을 가지고 수행되어야 하는 경우에는, 그 목적에 알맞은 적합한(fit-for-purpose) 불확도 수준이 결정되어야

할 것이다.

1.6. 의사결정에 있어서 측정불확도의 활용

대개의 환경조사에 있어서 측정결과는 환경기준(environmental threshold)에 대비되어지고 이에 따라 의사결정(오염부지 판명에 따른 토양복원 명령, 수질개선명령, 대기방지시설 설치 등)이 이루어진다(deterministic classification). 이 때 의사결정의 오류는 측정결과의 측정불확도 수준과 관련 있으며 이를 낮추어 줌으로써 decision quality를 높일 수 있게 된다. 그러나 측정결과의 불확도를 감소시키는 것이 의사결정의 오류를 항상 줄여주는 것은 아니다. 이는 측정값과 불확도수준, 그리고 이와 대비되는 기준값(threshold)과의 관계 속에서 설명될 수 있다(Fig. 4).

여기서 가장 왼쪽의 경우, 측정값이 기준치 미만이고 측정불확도를 감안해도 ‘오염되지 않음(uncontaminated)’으로 확실하게 판명된다. 그러나 두 번째 경우에는 측정값이 기준치 미만이지만 측정불확도 수준을 감안할 때 측정지점의 오염도가 사실상 기준치를 넘었을 확률도 배제하기 힘든 경우이다(possibly contaminated). 반대로 세 번째 경우에는, 측정값이 기준치를 넘어 오염됨을 지시하나 측정불확도 수준이 이를 완전히 뒷받침 해 주지 못함을 알 수 있다(probably contaminated). 따라서 두 번째와 세 번째 경우에는 측정값의 불확도수준이 감소됨으로써 의사결정의 오류가 방지될 수 있으나 나머지 경우에는 불확도 수준의 감소가 의사결정에 영향을 주지 않음을 알 수 있다. 이러한 경우에서 알 수 있듯이, 신뢰성 있는 의사결정을 위해서는 정밀조사를 통해 불확도 수준을 감소시켜야 하나, 기준치와 대비되는 측정 불확도 수준을 감안함으로써 비용효율적인 조사방법을 설계하거나 확률적인 부지에 대한 분류(probabilistic classification)에 의거하여 이에 알맞은 대안을 찾는 데 이용할 수 있을 것이다.

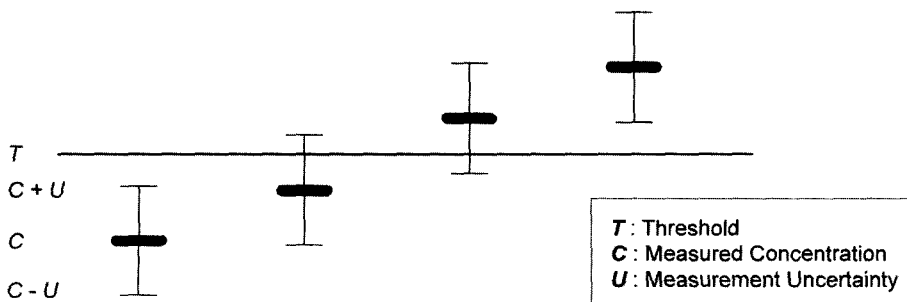


Fig. 4. Probabilistic classification of contaminated land suggesting subdivisions of classification according to uncertainty around the measured concentration. Adapted from Ramsey and Argyraki (1997).

1.7. 측정불확도수준을 응용한 효율적 조사기법의 설계

효율적인 정밀조사를 설계하기 위해 측정불확도가 어떻게 실제조사에서 응용되는가에 대한 예를 들어 보면 다음과 같다.

토양오염이 의심되는 지역의 오염 여부를 판단하기 위한 조사에는 먼저 대상지역의 오염경위, 오염원 및 부지이용기록 등에 대한 자료를 수집 분석하고, 이를 근거로 개황조사(preliminary investigation)의 범위를 결정함으로써 뒤따르는 정밀조사의 효율성과 정밀도를 높일 수 있게 된다. 개황조사에서는 수집된 자료를 분석하여 Fig. 5와 같이 대상지역을 단위 구역별(plot)로 구분하여 구역별 오염수준을 알아내고 이를 근거로 토양복원이 예상되는 구역에 대한 정밀조사의 필요 유무를 결정할 수 있다.

이러한 개황조사에서 각 구역에서 측정된 평균오염도에 대한 불확도수준을 평가할 수 있다면 각 구역에 대한 정밀조사방법을 설계할 때 참고할 수 있다. 다시 말해 시료채취방법이나 분석방법을 구역별 오염특성에 따라 설계함으로써 조사비용의 효율적인 분배를 도모할 수 있을 것이다.

개황조사가 따로 계획되어있지 않은 오염도 조사라 할지라도, 이 방법을 이용해 오염도 조사에 따른 구역별 측정불확도수준을 도출함으로써 앞 절에서 언급된 바와 같이 확률을 고려한 합리적인 의사결정을 내리는데 많은 도움이 될 것이다. 즉, 구역별로 조사된 측정 불확도 수준이 너무 높아 합리적인 의사결정을 내리는데 부적합 할 경우 해당 구역에 대한 조사방법을 정밀도를 높이는 방향으로 재설계하여 측정값의 불확도를 감소시켜야 한다. 이때 측정불확도를 시료채취과정과 분석과정에서 발생하는 불확도로 크게 구분하여 평가해 보면 측정값의 신뢰도에 가장 큰 영향을 주는 부분을 알아낼 수 있다. “사슬의 강도는 가장 약한 고리의 강도와 같다(A chain is as strong as its weakest

link)”란 서양의 속담에 있듯이 효율적인 조사가 이루어지기 위해서는 조사과정 중 가장 취약한 부분을 알아내어 보강하는 것이 필요한 것이다.

개황조사에서 평가된 측정불확도 수준에 근거하여 비용 경제적이고 효율적인 조사를 설계하는 방법은 Lee and Ramsey(2003)에 자세히 설명되어 있으나 몇 가지 예를 간략하게 정리를 해 보면 다음과 같다.

먼저 각 구역의 오염여부를 판단하기 위한 평균측정값에 대한 불확도 수준이 타당한 의사결정에 요구되는 불확도수준으로 보장되어야 하는 경우, 개황조사에서 평가된 측정불확도 수준을 기준으로 정밀조사시 필요한 측정횟수를 산출해 낼 수 있다.

$$n_{\text{meas}} = s^2_{\text{total}} / s^2_{\text{target}} \tag{7}$$

n_{meas} : 적정불확도수준을 충족시키는 측정횟수

s^2_{total} : 개황조사에서 평가된 측정불확도수준

s^2_{target} : 목표불확도수준

하지만 이 경우에 있어서 측정값에 대한 목표불확도 수준을 결정하기 위해서는 보통 전문가의 경험에 의한 주관적인 판단이 필요하나 효율적인 조사를 위한 객관적인 근거가 결여될 수 있다.

Cochran's sampling theory(Cochran, 1977)를 응용하면 측정횟수에 따른 비용과 평균측정값의 오차에 의한 손실의 합이 최소가 되는 필요측정수를 산출할 수 있다. 즉 한 측정(시료채취와 분석)에 소요되는 비용(C_{meas})과 개황조사에서 산출된 측정값의 총 분산($s^2_{\text{total}} = s^2_{\text{geochem}} + s^2_{\text{samp}} + s^2_{\text{anal}}$)을 알 때 비용경제적인 조사방법이 식(8)과 같이 도출된다.

$$n'_{\text{meas}} = \sqrt{\lambda \cdot (S^2_{\text{total}} / C_{\text{meas}})} \tag{8}$$

n'_{meas} : 최적측정횟수

λ : 대칭손실계산식에서의 상수(Cochran, 1977)

s^2_{total} : 대상지역에서 측정된 총 분산

($s^2_{\text{geochem}} + s^2_{\text{samp}} + s^2_{\text{anal}}$)

C_{meas} : 한 측정당 소요비용

그러나 이 계산에서는 임의의 대칭적 손실계산식(symmetric loss function)이 사용되고 경험적 변환상수(λ)를 결정해야 함으로써 주관적인 해석의 한계를 벗어나지 못할 수 있는 제약이 있다.

측정값에 대한 측정불확도 수준은 조사비용과 밀접한 관계를 가지고 있으나 이를 무시할 경우 발생할 수 있는 의사결정의 오류에 따른 경제적 손실과도 이어질 수 있다. 이를 동시에 참고한 조사방법 설계에 대한

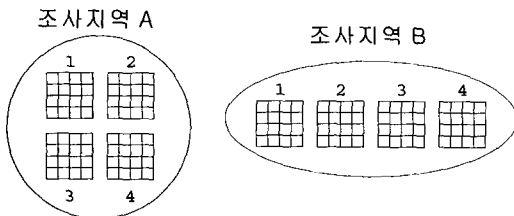


Fig. 5. Division of the target area into plots prior to investigation based on information such as the history of the land use, geology and possible contamination route, etc.

개요는 뒤에 이어진다.

1.8. 조사방법의 비용효율적 최적불확도 수준

정밀조사 결과에 요구되는 객관적인 최적의 불확도 수준이 결정되면 그 불확도 수준이 기대되도록 시료채취방법과 분석방법을 설계하여야 한다. 여기서 최적의 불확도 수준을 결정하기 위하여 경제적 제한요소를 고려하는 방법이 있다. 즉, 측정결과와 불확도가 감소하려면 정밀한 시료채취방법이나 분석방법을 고안하여 적용하거나 야외시료 및 분석시료의 수와 양을 늘리는 등의 노력이 필수적이다. 그러나 이에 소요되는 비용은 기하급수적으로 증가하게 되므로 조사목적이나 조사 후 부지사용에 따른 경제성과 같은 경제적 제한요소를 고려해 이에 알맞은 정밀도 수준을 결정하여 조사계획에 반영하여야 한다.

이러한 경제적 측면을 고려한 오염조사방안에서 흔히 간과되는 부분은 측정값에 근거한 의사결정이 잘못될 경우, 즉 측정된 값의 불확도(measurement uncertainty)에 의해 오염기준에 따른 구분에 오류(Fig. 4 참조)가 생김으로써 발생할 수 있는 경제적 손실이다. 이러한 오류의 예로서, 사실상 기준치 이하의 오염도를 가진 부지에 대한 오염판정(False Positive), 반대로 오염상황을 파악하지 못함으로써 내려지는 잘못된 의사결정(False Negative) 등이 있을 수 있다. 이로 인한 손실에는 불필요한 복원비용, 검출하지 못한 오염에 의한 건강상의 위해, 또 이에 대한 보상 또는 법률적 비용 등이 있을 수 있는데, 이는 측정불확도가 증가함에 따라 이러한 손실이 야기될 수 있는 확률 또한 증가하게 된다.

따라서 최적의 조사방법설계를 위해서는 환경조사에 소요되는 직접적인 비용 뿐만 아니라 조사결과를 바탕으로 내려진 의사결정의 오류에 의한 경제적 손실까지도 예상함으로써 최종 조사비용이 최소화 되는 조사방법의 정밀도 수준을 도출할 필요가 있다. 이러한 과정은 측정값에 대한 측정불확도수준을 확률적 손실계산식(probabilistic loss function, 식(9))을 통해 조사비용과 의사결정오류에 의한 손실의 합이 최소가 되는 최적의 불확도값을 산출하고 이를 기준으로 조사방법의 정밀도에 적용시킴으로써 완성할 수 있다(Lee and Ramsey, 2002).

Fig. 6는 측정에 소요되는 비용(Measurement Cost, 식(9)의 오른쪽 term)과 의사결정의 오류에 의한 손실(Loss from Misclassification, 식(9)의 왼쪽 term)이 조사방법의 불확도 수준에 따라 달라지는 것을 보여준다.

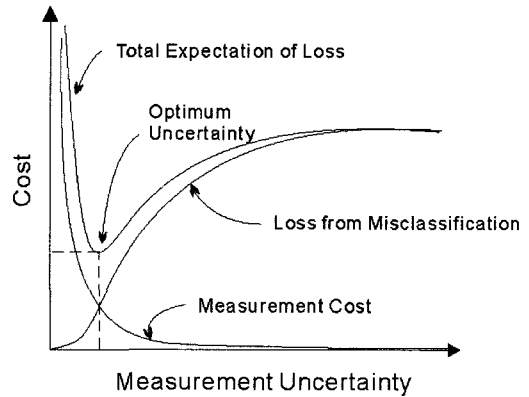


Fig. 6. Illustration for the loss function on the misclassification error and measurement cost (Equation 9). Minimization of this function enables to locate the optimum level of measurement uncertainty for the optimization of contaminated land investigation .

이들을 더함으로써 총예상소요비용(E(L), Total Expectation of Loss)이 산출됨을 볼 수 있다. 이 두 식은 모두 측정불확도(s_{meas})에 의한 함수로 표현되며 이 총예상소요비용을 최소로 만들어 주는 측정불확도가 정밀조사의 설계에 반영됨으로써 효율적이고 비용경제적인 조사가 이루어 질 수 있게 된다.

$$E(L) = \left\{ C_c \left[1 - \Phi \left(\frac{\epsilon}{s_{meas}} \right) \right] + \frac{D}{s_{meas}^2} \right\} \quad (9)$$

E(L) : 총예상소요비용(Total Expectation of Loss)

C_c : consequence cost(의사결정의 오류로 인해 발생할 수 있는 최대손실)

Φ : standard normal cumulative distribution function

ϵ : Threshold concentration - measured concentration

D : optimized cost for unit measurement variance(Thompson and Fearn, 1996)

s_{meas} : 측정불확도

측정이란 야외시료채취와 분석과정으로 나뉘어 지므로, 다음은 이렇게 구해진 측정불확도 수준을 각 단계별 불확도 수준으로 할당해야 한다. 같은 측정불확도 수준이라도 가장 경제적인 방법으로 시료채취단계의 불확도 수준과 분석과정에서의 불확도수준으로 할당될 수 있다. 그 최적의 비율은 시료채취단계와 분석단가를 고려하는 Thompson and Fearn(1996)의 “cost rule”을 응용하여 도출할 수 있다. 이를 이용하여 최소의 비용으로 시료채

취과정과 분석과정에서 기대되는 불확도 수준이 각각 나누어지게 되는데 그 결과는 식(10), 식(11)과 같다.

$$(s_{\text{samp}}^2)' = s_{\text{samp}}^2 \left(\frac{\sqrt{A}}{\sqrt{A} + \sqrt{B}} \right) \quad (10)$$

$$(s_{\text{anal}}^2)' = s_{\text{anal}}^2 \left(\frac{\sqrt{B}}{\sqrt{A} + \sqrt{B}} \right) \quad (11)$$

$(s_{\text{samp}}^2)'$: 시료채취방법의 최적 불확도수준
(분산으로 표현된 시료채취방법의 정밀도 수준)

$(s_{\text{anal}}^2)'$: 분석방법의 최적 불확도수준
(분산으로 표현된 분석방법의 정밀도 수준)

A: sampling cost per unit sampling variance

B: analytical cost per unit analytical variance

이렇게 구해진 시료채취방법의 최적 불확도수준과 분석방법의 최적 불확도 수준을 개황조사에서 평가된 각각의 불확도 수준과 비교하여 차후 정밀조사를 수행할 때 기존의 채취방법이나 분석방법에 대한 정밀도를 개선한다든지 아니면 반대로 지나치게 정밀한 조사방법을 덜 정밀한 방법을 사용함으로써 비용을 절감시킬 수 있게 된다. 결과적으로 부지특성과 오염특성을 고려한 부지별 최적 조사방법을 설계할 수 있는 것이다.

2. 결 론

환경측정은 현장에서의 시료채취와 이에 대한 분석으로 그 과정을 크게 구분할 수 있다. 이 과정을 통해 구해진 측정값은 참값일 수 없으며 항상 불확실성을 내포하게 된다. 이러한 불확실성은 측정대상의 불균질성(heterogeneity)와 시료채취방법과 분석방법의 정밀도 차이에 의해 기인하며 모두 측정불확도(measurement uncertainty)에 영향을 미치는 요소들이다.

대개의 경우 환경조사에서의 측정값은 측정불확도를 고려하지 않은 채 이용됨으로써 의사결정에 오류를 일으키거나 재측정시 보일 수 있는 측정값의 차이는 결국 자료에 대한 불신과 이를 근거로 환경정책을 수립하는 경우 대중으로부터 지지를 얻지 못하는 결과를 초래할 수 있게 된다.

따라서 오염도 조사시 측정불확도를 평가하여 측정값과 같이 제시를 함으로써 의사결정에 있어서 측정값에 대한 객관성을 보장하게 되고, 필요하다면 측정값의 신뢰도를 높이기 위한 노력을 경주하게 하는 기준으로서의 역할을 할 수 있게 된다. 즉 오염도 조사를

위한 예비조사(preliminary investigation) 내지는 개황조사에서 시료채취과정과 분석과정의 정밀도에 의한 부지별 측정값의 불확도 수준을 평가할 수 있다면, 이를 근거로 적은 비용으로 측정값의 신뢰도를 효율적으로 높이는 방향으로 정밀조사를 설계할 수 있을 것이다.

더 나아가서는 비용경제적인 측면을 고려하는 최적의 조사방법을 설계해야 하는 경우, 조사비용에 측정값의 불확도에 의해 의사결정에 오류가 생기므로써 야기되는 경제적 손실까지도 포함시켜야 할 것이다. 즉, 조사방법의 정밀도를 높여 측정불확도수준을 낮추는 데는 많은 비용이 소요되나 이는 불확도수준을 낮추지 않음으로써 증가될 수 있는 의사결정의 오류에 의한 경제적 손실과 균형을 이루어야 한다. 이렇게 비용경제적인 측면을 고려하는 조사설계방법은 최적의 조사설계에 대한 하나의 기준이 될 수 있다. 이를 위하여 최소한의 경제적 손실이 발생할 수 있는 측정불확도수준을 손실계산식으로부터 추정할 수 있다면 기존의 주관적인 조사방법 설계가 아닌 보다 객관적이고 과학적인 조사방법 설계에 대한 기준을 제시해 줄 수 있을 것이다.

Reference

- AMC (1989) Robust statistics-how not to reject outliers-Part 1. Basic concepts. Analyst 114, p. 1693-1697.
- AMC (1995). Uncertainty of measurements: Implications of its use in analytical science. Analyst 120, p. 2303-2308.
- Cochran, W.G. (1977) Sampling techniques. 3rd ed. Wiley series in probability and mathematical statistics. Wiley, New York.
- Eurachem (2000) Quantifying uncertainty in analytical measurement. 2nd ed. S. L. R. Ellison, M. Rosslein and A. Williams, Eds. Eurachem.
- Garrett, R. (1969) The Determination of Sampling and Analytical Errors in Exploration Geochemistry. Economic Geology 64, p. 568-569.
- Gy, P. (1979) Sampling of particulate materials: Theory and practice. Elsevier, Amsterdam.
- ISO (1993) Guide to the expression of uncertainty in measurement. 1st ed. International Organisation for Standardisation, Geneva, Switzerland.
- Lee, J.C. and Ramsey, M.H. (2001) Modelling measurement uncertainty as a function of concentration: an example from a contaminated land investigation. Analyst 126, p. 1784-1791.
- Lee, J.C. (2002) Measurement uncertainty in contaminated land investigations related to analyte concentration and cost. PhD Thesis. Department of Environmental Science and Technology, Imperial College of Science, Technology and Medicine, London.
- Lee, J.C. and Ramsey, M.H. (2003) Optimization of contaminated land investigation based on different fit-

- ness-for-purpose criteria. *Korea Society of Economic and Environmental Geology* 36(3), p. 191-200.
- Miesch, A.T. (1967) *Theory of error in geochemical data*. U.S. Geological Survey Professional Paper Part A, p. 574.
- Miller, J.C. and Miller, J. N. (1993) *Statistics for analytical chemistry*. 3rd ed. Ellis Horwood and Prentice Hall, Chichester, UK.
- Pitard, F.F. (1993) *Pierre Gy's sampling theory and sampling practice: heterogeneity, sampling correctness and statistical process*. 2ed. Boca Raton London: CRC Press, London.
- Ramsey, M.H., Thompson, M. and Hale, M. (1992) Objective evaluation of precision requirements for geochemical analysis using robust analysis of variance. *J. of Geochemical Exploration* 44, p. 23-36.
- Ramsey, M.H., Argyraki, A. and Thompson, M. (1995) Estimation of sampling bias between different sampling protocols on contaminated land. *Analyst* 120, p. 1353-1356.
- Ramsey, M.H. (1997) Measurement uncertainty arising from sampling: Implications for the objectives of geoanalysis. *Analyst* 122(11), p. 1255-1260.
- Ramsey, M.H. and Argyraki, A. (1997) Estimation of measurement uncertainty from field sampling: Implications for the classification of contaminated land. *Science of the Total Environment* 198(3), p. 243-257.
- Ramsey, M.H. (1998) Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources. *Journal of Analytical Atomic Spectrometry* 13(2), p. 97-104.
- Thompson, M. (1988) Variation of precision with concentration in an analytical system. *Analyst* 113, p. 1579-1587.
- Thompson, M. and Fearn, T. (1996) What exactly is fitness for purpose in analytical measurement? *Analyst* 121(3), p. 275-278.
- 환경부 (2003) 2002년도 토양 측정망 및 실태조사 결과. 행정간행물등록번호 11-1480000-000669-01.

2003년 12월 11일 원고접수, 2004년 2월 11일 게재승인.