Journal of Geochemical Exploration 107 (2010) 410–422

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

journal homepage: www.elsevier.com/locate/jgeoexp

Use of measurement uncertainty in a probabilistic scheme to assess compliance of bottled water with drinking water standards

Alecos Demetriades *

Institute of Geology and Mineral Exploration, 1 Spirou Louis Street, Entrance C, Olympic Village, Acharnae, Athens, Gr-136 77, Hellas

ARTICLE INFO

Article history:
Received 19 September 2010
Accepted 2 November 2010
Available online 5 November 2010

Keywords:
Robust analysis of variance
Measurement uncertainty
Water standards
Compliance assessment
Legislation
Bottled water
Drinking water

ABSTRACT

Ground water bodies are important resources for drinking water, including bottled water, and national regulatory bodies should assess their quality continuously. For this purpose, an effective assessment system of bottled water at source should be installed. A hierarchical nested balance design for the collection of random primary duplicate water samples, and their replicate analyses, is described, and the use of robust analysis of variance to estimate measurement uncertainty. The latter is subsequently used for the development of four probabilistic categories for the classification of element concentrations in bottled water with respect to legislative standard values, i.e., (a) compliant (below Lower Threshold Limit), (b) possibly non-compliant (possibly above Standard Value), (c) probably non-compliant (probably above Standard Value), and (d) non-compliant (above Upper Threshold Limit), for the reliable assessment of compliance to European Union and national drinking water standards. Overall, the quality of European bottled water is considered good, with the exception of a few that have concentrations in Mn, B, Ba, As, Fe, Ni, Se, and Al, which are definitely above the estimated respective Upper Threshold Limit and, thus, exceed the corresponding legislative standard value defined by European Union directives. National regulatory bodies should verify these results, and install an efficient assessment system of compliance to regulatory limits using the methodology described in this paper.

1. Introduction

In applied geochemical surveys, field duplicate samples at a predetermined proportion of sampling sites are taken, and replicate analyses on splits of both duplicate samples in a balanced hierarchical nested design are performed (Garrett, 1969, 1973, 1983; Miesch, 1964, 1967, 1973, 1976). The analytical results are processed by classical analysis of variance (ANOVA) to estimate sampling, analytical and geochemical variance.

Classical ANOVA relies on the assumptions of normality of the unimodal distribution of the studied parameter and homoscedasticity of variances, i.e., homogeneity or uniformity of variances (Lee and Ramsey, 2001). If these assumptions are not met, the estimated variances become less reliable. The first assumption of normality is not met in many instances, particularly in the case of geochemical data, where analyte concentrations often display positively skewed or lognormal distributions, and are generally multimodal (Ahrens, 1954a,b; Tennant and White, 1959; Lepeltier, 1969; Sinclair, 1976, 1983, 1986). Furthermore, in order for homoscedasticity to hold, ANOVA assumes that there is no change of the variance within the concentration range, so that the estimated uncertainty could be applicable for the whole range of analyte concentration in the samples. Thus, the estimation of uncertainty by ANOVA is only applicable at analyte concentrations close to the mean value of a unimodal distribution, and does not apply in cases with a wide range of concentration and multimodal distributions, where a change in measurement precision with concentration is expected, and is strongly affected by a few outlying values.

To overcome the problems of non-normally distributed data, the use of robust statistics has been suggested. Log-transformation of the raw data and expression of the estimated uncertainty as a percentage value relative to concentration has also been proposed for the same purpose. However, these techniques resolve the problem partially. Recent research has shown that the use of a linear measurement precision modelling method may be more appropriate for the evaluation of measurement uncertainty in instances of the raw data being log-normal and heteroscedastic, i.e., the variables have different variances (Lee and Ramsey, 2001; Saari et al., 2008).

Robust ANOVA can overcome the problems of classical ANOVA and the novelty of the particular computer program used, ROBCOOP4.EXE, is that it accommodates outlying values up to 10% of the total population (Ramsey, 1998; Ramsey et al., 1992). If, however, the outliers are more than 10% of the total population, then the estimates of component variances could be erroneous. Hence, before proceeding with the calculations, it is important to estimate the percentage proportion of outlying values, and then to decide accordingly on the steps to be followed. In case the outliers are more than 10%, then the ROBAN programme (Roban, 2001) can be used instead of ROBCOOP4,

* Tel.: +30 210 2413272.
E-mail addresses: ademetriades@igme.gr, alecos.demetriades@gmail.com.

0375-6742/$ – see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.gexplo.2010.11.001

because there is an option to increase the number of outlying values to be accommodated, i.e., from mean + 1.5 × Standard Deviation (S.D.) to mean + 2 × S.D., or + 2.5 × S.D., or 3 × S.D., etc.

The RANOVA hierarchical nested balance design according to Ramsey (1998) requires a minimum of eight randomly selected sample sites to produce reasonably acceptable results, although a larger number would be preferable (Fig. 1). The minimum number of eight duplicate/replicate samples has been, however, questioned, because the number of duplicated sites will affect the quality of subsequent estimates of sampling and analytical uncertainties. Hence, Lyn et al. (2007) tested the method, and verified that it does provide estimates of uncertainty that are both acceptably accurate and cost effective.

In the EuroGeoSurveys project "*European Ground water Geochemistry: Bottled Water*" (Reimann and Birke, 2010), an extensive quality control programme was performed, which could be considered as unique in applied hydrogeochemical investigations, and is described by Birke et al. (2010a). Apart from utilising the duplicate results to calculate the practical detection limit, it was decided to use the quadruple analyses, performed on bottled water samples, which satisfy the conditions of a modified balance nested hierarchical design for the purpose of illustrating the procedure of estimating the two components of measurement uncertainty (sampling and analytical variance), and geochemical variance by the robust ANOVA (RANOVA) technique proposed by Ramsey (1998).

Ideally, the duplicate analyses should be performed on two different bottled water samples of the same brand, bottled at different times, but within an hour between the first and second bottled water sample, or in this project, the two bottled water samples of the same brand should have been purchased from different supermarkets, in order to satisfy the conditions of a balance nested hierarchical design (Fig. 1). In this case, however, the quadruple analyses were carried out on the same bottle brands, but at different times, giving, thus, the temporal variation. Although this is not the ideal case, the quadruple analyses performed, are assumed to satisfy the statistical conditions, since the purpose of this paper is to propose a procedure to assess compliance of bottled water analytical results.

In other words, the purpose of this paper is, in fact, to propose an objective method to assess compliance of bottled water analytical results with statutory standards defined by EU directives (EC, 1998; EU, 2003, 2009), national legislation (Reimann and Birke, 2010), and international organisations, such as the World Health Organisation (WHO, 2008), the Food and Agricultural Organisation (FAO, 1997) and the US Environmental Protection Agency (US EPA, 1992, 2009). The quality of bottled water and, of course that of potable water in general, is of paramount of importance to human health, and its quality should be efficiently assessed by national regulatory bodies to ensure that it meets legislative standards.

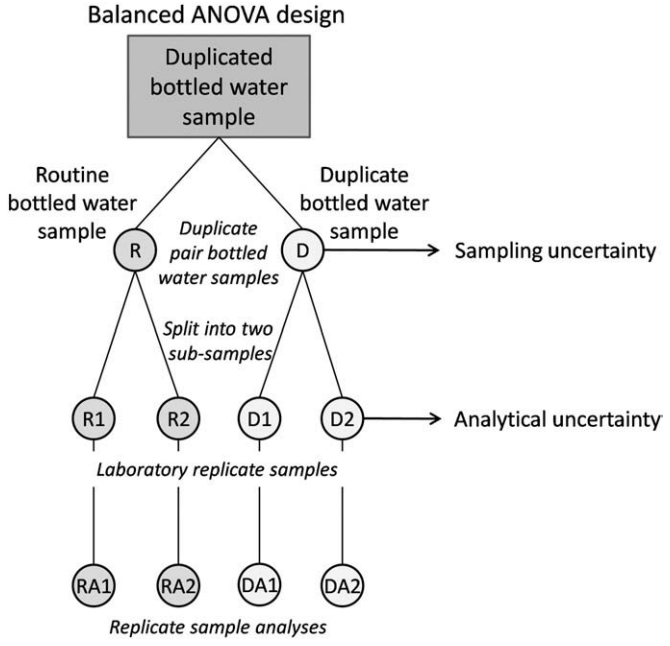

Fig. 1. Balanced hierarchical sampling and analytical design for the estimation of random components of measurement uncertainty.

The United Kingdom Drinking Water Inspectorate posed an interesting question "*What are the drinking water standards?*" (DWI, 2010), and the answer given is that "*drinking water must be 'wholesome' and this is defined in law by standards for a wide range of substances, organisms and properties of water in regulations. The standards are set to be protective of public health and the definition of wholesome reflects the importance of ensuring that water quality is acceptable to consumers*".

The quality of bottled water, and potable water in general, to be acceptable to consumers, the Regulatory body must set up an objective and effective procedure to assess compliance to legislation. The key to assessment of compliance of bottled water to legal standards depends on the applied "*decision rule*". The normal approach is to compare directly the measured concentration value of a determinand or analyte with its legislative drinking water standard, and to assess if it is above or below the latter, and to reject or accept the particular batch of bottled waters, respectively. This is the deterministic approach, which does not consider the uncertainties due to the measurement process, and, in fact, assumes erroneously that measurement uncertainty is zero.

Regulatory bodies should use, however, the probabilistic approach, because it is based on measurement uncertainty and the specification of limits are defined by considering an acceptable level of probability of making a wrong decision (Ellison and Williams, 2007). Consequently, based on the probabilistic decision rule, an "*acceptance zone*" and a "*rejection zone*" can be estimated. If the measurement of an analyte or determinand lies in the acceptance zone, the bottled water is declared compliant, and if in the rejection zone, it is declared noncompliant to that particular analyte, and further action should be taken.

This paper describes concisely the RANOVA method for the estimation of the components of measurement uncertainty (sampling and analytical variance) and geochemical variance, and the use of the former in the probabilistic assessment of compliance with legislative drinking water standards. The RANOVA results also indicate which component of sampling and analytical variance contributes most to measurement uncertainty. For a more detailed discussion of measurement uncertainty, methodology and terminology refer to Ramsey et al. (1992), Ramsey and Argyraki (1997), Ramsey (1998, 2009), Ellison et al. (2000), Lee and Ramsey (2001), Ellison and Williams (2007), Lyn et al. (2007), Ramsey and Ellison (2007), Saari et al. (2008) and the references therein.

For the estimation of quality control parameters and measurement uncertainty by robust statistics apart from the ROBCOOP4 program (Ramsey, 1998), there is also the Windows based ROBAN program, which draws a pie chart of component variances and also has the option to set the proportion of outliers to be accommodated at levels beyond the range of mean ± 1.5 standard deviation that is considered as the optimal for 10% outliers (AMC, 2009; Roban, 2001).

2. Estimation of uncertainty due to sampling and analysis

Measurement uncertainty is simply defined as '*the interval around the result of a measurement that contains the true value with high probability*' (Ramsey, 1998, p.97). A more internationally recognised definition of measurement uncertainty is '*an estimate attached to a test result which characterises the range of values within which the true value is asserted to lie*' (ISO, 1993, 3.25). The most current definition of