A NEW UNBALANCED NESTED ANOVA MODEL IN GEOLOGY - A DOWN-TO-EARTH DESIGN

Thomas I. Goss, Systems Approach Consultants Limited
Robert G. Garrett, Geological Survey of Canada

1. Introduction

This paper describes a new unbalanced 4-stage analysis of variance (ANOVA) model developed as part of Canada's 1977 National Geochemical Reconnaissance (NGR) Program. Implementation to several center-lake sediment surveys covering over 55,000 square miles (see Figure 1) has been completed during the summer of 1977. A generalized computer program that calculates point and interval estimates for the variance components and approximate F-tests for the individual components differing from zero will be highlighted. Outstanding statistical issues inherent in the design will be briefly addressed in the last section.

Multi-stage nested sampling designs have been used in large scale regional geochemical surveys to determine the relative magnitudes of the data variability across the region, on some local scales, and in the laboratory. Appropriate interpretation of the variation of the geological and geochemical data (e.g., see Miesch [23]) leads to the production of reliable (i.e., statistically reproducible) geochemical maps in addition to the development of more efficient sampling plans for future surveys. The literature on such geochemical studies has been extensive since the landmark paper of Krumbein and Slack [19], which employed the classical ANOVA method using balanced data (i.e., equal sub-class numbers) for simultaneously estimating the components of variance at the nested sampling stages. Similar ANOVA methods were applied to regional and exploration geochemical problems during the mid-1960's to early 1970's, as exemplified by Miesch [22], Howarth and Lowenstein [18], Garrett [11,12], Michie [21], and Bolviken and Sinding-Larsen [6]. A major limitation of the completely balanced layout was the large concentration of samples (hence degrees of freedom) at the lowest sampling stage that sometimes led to large disparities in the precision of the variance component estimates, as well as making the survey prohibitively expensive. Figure 2 illustrates a 4-stage balanced and unbalanced nested design where N refers to the total number of samples to be collected and analysed, L is the number of levels (stages) in the design, and k is the number of replication groups at level one. The UB/B ratio of Figure 3 clearly shows the sample efficiency of unbalanced designs, particularly where greater than 4 levels are used.

During the 1970's, Miesch and his co-workers in the Branch of Regional Geochemistry of the U.S. Geological Survey have done much to develop the use of unbalanced nested ANOVA models in geochemical studies (e.g., see Miesch [24], Ebens and McNeal [9], Tidball and Severson [32]). The latter two articles illustrate their application of the favourable staggered (unbalanced) design (e.g., see Bainbridge [4], Leone et al. [20], Anderson [1]), in which the degrees of freedom for the ANOVA variance component estimation are spread quite evenly over all sampling levels.

2. Statistical Model and Computational Method

Since the first Canadian lake sediment regional geochemical reconnaissance survey in 1974 (see Hornbrook and Garrett [17] for a detailed description of the survey methodology), the major sampling plan consideration was adequate coverage of a geographic or geologic unit to ensure with high probability the detection of mineral exploration targets when they existed, whereupon further mineral exploration activity would hopefully lead to the discovery of significant mineral occurrences (e.g., see Garrett [13] for recent developments and references on this aspect). The secondary (but still very important) aspect of these surveys was the evaluation of sampling and analytical variation (in a quality control sense), leading to the production of stable regional geochemical maps that highlighted major regional geochemical trends. See Garrett and Goss [14] for further details on this topic.

The above considerations saw the evolution of an unbalanced nested sampling design for the 1977 lake sediment regional geochemical reconnaissance surveys where the proportion of the variation at the regional (i.e., between grid cells), 5 sq. mile grid cell, lake and analytical levels could be determined; furthermore, all 5 sq. mile cells containing suitable lakes were at least sampled once to satisfy the mineral exploration constraint, while every randomly chosen 16th cell (i.e., about 6% replication) involved sub-sampling within the cell (to estimate cell variability), within a lake (to estimate lake variability), and for analytical purposes to satisfy the quality control constraints. Figure 4 illustrates the typical field sampling grid. See Garrett and Goss [14] for details on the survey design. The resulting sample scheme, denoted a 4-stage inverted nested design (Bainbridge [4]) due to the large concentration of degrees of freedom at the top sampling level, is displayed for one replication block of 16 cells in Figure 5. The data structure Type I was the designed layout but Type II occurred in places where sample material drawn from a lake duplicate was insufficient to generate the analytical duplicate.

The statistical model may be expressed following Leone et al. [20] or Anderson and Bancroft [2] as: $x_{ijkm} = \mu + a_i + b_{j(i)} + c_{k(ij)} + e_{m(ijk)}$ where x_{ijkm} represents an individual analysis; μ is the true overall mean concentration of the element of interest; a_i, b_j, c_k, and e_m represent additive contributions due to the sampled cells, lakes in cells, samples in lakes and split samples-it is assumed that a,b,c, and e are random and NID variables with zero means and constant variances σ_a^2, σ_b^2, σ_c^2, and σ_e^2 respectively. Although several techniques exist for estimating variance components from unbalanced data (e.g., see Searle [28], Rao [26]) the classical ANOVA approach was used because of its simplicity, and its successful and complementary usage in geochemical surveys at the United States Geological Survey. By pooling all sums of squares for each stage of sampling (following Ganguli's [10] procedure), the variance components are straightforwardly estimated for the inverted and multi-structured design.

Reprinted from the 1978 Statistical Computing Section
Proceedings of the American Statistical Association.

The analysis of variance formulas are presented in Table 1.

Apart from the establishment of the relative variability in the data it is desirable to construct approximate F-tests for the observed variance components of the design. For example, it would be impractical to map the variation among grid cells if there was a good chance (statistically) that no variation actually exists. Essentially, the test is accomplished by forming synthesized mean squares for the denominator of the observed F ratio and then employing Satterthwaite's formula for computing the associated degrees of freedom (Satterthwaite [27]). This procedure has been widely applied to unbalanced nested designs (e.g., see Tietjen and Moore [33], Snee [29]), and was used in our model to test whether the variance component estimates, σ_a^2 and σ_b^2, are significantly different from zero. Note that the test of the hypothesis $\sigma_c^2 = 0$ can be done directly as in balanced designs (i.e., by forming the F ratio as MSC/MSE).

The theory for constructing confidence interval estimates for variance components from unbalanced data is largely incomplete (e.g., see Searle [28], Anderson and Bancroft [2]). However, there exist several different methods for obtaining such estimates for balanced or slightly unbalanced data. Although our inverted nested design violates this latter assumption, it was felt that much could be learned by experimenting with "approximate" confidence interval estimates for the variance components. Thus, following Boardman's [5] investigation on the accuracy of nine methods for constructing an approximate confidence interval about σ_a^2 in a two level balanced nested design (where $y_{ij} = \mu + a_i + b_{j(i)}$, $a_i \sim$ NID $(0, \sigma_a^2)$, $b_j \sim$ NID $(0, \sigma_b^2)$), we have adapted the Satterthwaite and Williams-Tukey procedures for calculating confidence interval estimates about σ_a^2, σ_b^2, and σ_c^2 (see Boardman [5; p.255] for a summary of the formulas used). In addition, the exact 95% confidence interval estimate for σ_e^2 (assuming normality) was calculated as $[SSE/\chi^2_{.025}(DF(E)), SSE/\chi^2_{.975}(DF(E))]$, as illustrated in Anderson and Bancroft [2]. Incidentally, a recent improvement on the Williams-Tukey interval estimate for unbalanced data (see Thomas and Hultquist [31]) will be incorporated into the computer program at a later date.

Finally, a useful empirical variance ratio, v, derived from the estimated variance components to assist in evaluating the "effectiveness" of the sampling design was calculated by $v = s_a^2 (s_b^2 + s_c^2 + s_e^2)^{-1}$; large values of v (>>1) usually indicate no further sampling or analytical effort is required to describe the compositional differences among cells at the regional level (Miesch [24]).

3. Illustration of Computer Program

A generalized computer program (called UANOVA) was developed to handle complex (e.g., multi-structured data sets) unbalanced nested designs for up to nine stages employing the ANOVA method for estimating the variance components. Initial program development began with the Postma and White [25] program. Note that the ANOVA formulas as exemplified by Table 1 assume a completely random effects model for infinite populations (see Snee [29] for further details). The results of a test set of data for uranium in the 1977 Ontario survey are presented in Table 2 solely as an example of the computer output. Logarithmic (base 10) and square root transforms are options in the program to assist the user in altering the data so as to more readily conform to the model assumptions (i.e., normality, additivity, and homogeneity of variance). The common use of the log transform in geochemical studies is discussed in Miesch [24]. Most of the printout abbreviations of Table 2 are self-explanatory, along with the following summary of terminology and ad hoc decisions:

- Unit Size = cumulative sample size at each level (LVL) or stage of the design.
- Percent of Total = relative contribution of each variance component expressed as a percent of total variation; any negative estimate of variance component is treated as zero under this column but is otherwise used in its negative form in the (unbiased) variance of the mean, $v(\bar{x})$, calculations (defined below) as long as the negative contributions do not lead to an overall negative $v(\bar{x})$; see Searle [28] for a summary of alternative suggestions for handling negative estimates.
- Synthesized Mean Square - the denominator, or error mean square, has been synthesized (e.g., see Snee [29]) for the ensuing approximate F-test only at the levels corresponding to the footnote (2) attached to the "Computed DOF" column; negative variance component estimates may be involved in the calculation of the synthesized mean square as long as the net result is not negative (in this situation, the F-test is aborted); see Snee [29], Gaylor and Hopper [15], Searle [28], and Cummings [7] for additional work and references in this area.
- Computed DOF - the accuracy of this approximation is also discussed by the above authors; in section 4, this topic is addressed from the context of our inverted nested design.
- Approx. F-Ratio = ratio of "Mean Square" to "Synthesized Mean Square" columns for the respective level of the design.
- Prob. value = probability of the observed F-ratio given the respective degrees of freedom (column 4, column 10).
- Estimated Population Mean = $\bar{x}_{....}$ as defined in Table 1.
- Estimated Variance of Mean = $v(\bar{x}_{....})$ is given as: (Leone et al. [20])

$$[(\Sigma n_{i..}^2)s_a^2 + (\Sigma n_{ij.}^2)s_b^2 + (\Sigma n_{ijk}^2)s_c^2 + n_{...}s_e^2]\, n_{...}^{-2}$$

- 95% Confidence Limits Around Mean = $\bar{x}_{....} \pm t(.025, n-1)\sqrt{v(\bar{x}_{....})}$ where t is the usual Student's t value corresponding to .025 confidence probability with n-1 degrees of freedom (as suggested by Anderson and Bancroft [2]).

4. Conclusions and Outstanding Statistical Issues

The application of multi-stage inverted nested sampling designs to regional geochemical reconnaissances such as Canada's NGR is a natural choice considering the constraints of mineral

exploration surveys (i.e., fine areal sampling) and the costs of sample collection and analysis. Although this application has extended studies into areas where little theoretical work has been undertaken, the results obtained are sensible when viewed empirically by geochemists. As approximate as the results may be they have provided objective criteria which have helped in the critical assessment of the survey data and sampling design (see Garrett and Goss [14]). Bainbridge [4] had considered such designs as potentially very efficient because the confidence interval estimate for the variance component at the highest sampling level has the largest degrees of freedom, precisely where they are most needed. Empirical studies on the sampling distributions of variance components for Bainbridge's 4-stage inverted nested design (Leone et al. [20]) indicated the design is reasonably robust, although further research is needed.

The theoretical properties of multi-stage (>3) inverted nested designs having mixed data structures is sorely lacking, particularly with respect to variance component estimation techniques, the design effects on the model's distributional assumptions, and the hypothesis testing of variance components. Anderson and Crump [3], and Goldsmith and Gaylor [16] have studied the optimum combination of sampling structures for 2 and 3-stage nested designs using the ANOVA method. This pooling procedure across multiple data structures avoids more complicated analyses such as maximum likelihood estimation of variance components at the expense of losing information about the structure effects (Anderson [1]). Nonetheless, recent developments of alternative variance component estimation procedures (e.g., see Swallow and Searle [30], Searle [28]) are seriously questioning the efficiency and desirable optimality criteria of the ANOVA approach to unbalanced data.

The approximate F-test for variance component testing described in section 2 assumes incorrectly that the ANOVA mean squares are independent and have chi-square type distributions. The robustness of these assumptions was analysed (using a very similar F-test) for selected 3-stage unbalanced nested designs having mixed sampling structures that deliberately violated the assumptions (see Cummings and Gaylor [8]). In general, the F-test appeared satisfactory for a wide range of variance component ratios for extremely unbalanced nested designs, but extensions for the testing of 4th and higher stage variance components in unbalanced nested designs remain unclear.

As a final note, it is hoped that the availability of this computer program to handle multistage unbalanced nested designs will encourage greater use and further improvements in the techniques adopted, particularly in a variety of geological and geochemical studies. The program is being submitted for publication in Computers and Geoscience, a journal of the International Association of Mathematical Geology. Immediate requests may be sent to R.G. Garrett, Geological Survey of Canada, 601 Booth Street, Ottawa, Ontario, K1A 0E8, Canada.

5. Acknowledgements

The authors would like to express their appreciation to E.M. Cameron and A.T. Miesch for their past contributions and helpful discussions in the formulation of this study. Also, much appreciation goes to Rob Stuart, Tolon Gucu and Lise Campeau for their programming services and support.

6. References

[1] Anderson, R.L., Designs and Estimates for Variance Components, in A Survey of Statistical Design and Linear Models, Ed. J.N. Srivastava, North-Holland Publishing Co. (1975), 1-29.

[2] Anderson, R.L. and Bancroft, T.A., Statistical Theory in Research, New York: McGraw-Hill (1952), 399 p.

[3] Anderson, R.L. and Crump, P.P., Comparisons of Designs and Estimation Procedures for Estimating Parameters in a Two-Stage Nested Process, Tech., 9 (1967), 499-516.

[4] Bainbridge, T.R., Staggered, Nested Designs for Estimating Variance Components, Am. Soc. Quality Control, Ann. Conf. Trans. (1963), 93-103.

[5] Boardman, T.J., Confidence Intervals for Variance Components - A Comparative Monte Carlo Study, Biometrics, 30 (1974), 251-262.

[6] Bolviken, B. and Sinding-Larsen, R., Total Error and Other Criteria in the Interpretation of Stream Sediment Data, in Geochemical Exploration, Ed. J.M. Jones, I.M.M. London (1972), 285-295.

[7] Cummings, W.B., Variance Component Testing in Unbalanced Nested Designs, North Carolina State University Instit. of Stat. Series No. 843 (1972).

[8] Cummings, W.B. and Gaylor, D.W., Variance Component Testing in Unbalanced Nested Designs, Jour. Am. Stat. Assoc., 69 (1974), 765-771.

[9] Ebens, R.J. and McNeal, J.M., Geochemistry of the Fort Union Formation, U.S. Geol. Surv. Open File Report 76-729 (1976), 94-111.

[10] Ganguli, M., A Note on Nested Sampling, Sankhya, 5 (1941), 449-452.

[11] Garrett, R.G., The Determination of Sampling and Analytical Errors in Exploration Geochemistry, Econ. Geol., 64 (1969), 568-569.

[12] Garrett, R.G., The Determination of Sampling and Analytical Errors in Exploration Geochemistry - A Reply, Econ. Geol., 68 (1973), 282-283.

[13] Garrett, R.G., Sample Density Investigations in Lake Sediment Geochemical Surveys of Canada's Uranium Reconnaissance Program, Symposium on Hydrogeochemistry and Stream Sediment Reconnaissance for Uranium in the United States, United States Dept. of Energy, Grand Junction, Colorado (1977), 173-185.

[14] Garrett, R.G. and Goss, T.I., The Evaluation of Sampling and Analytical Variation in Regional Geochemical Surveys, in Press, Geochemical Exploration Proceedings (1978).

[15] Gaylor, D.W. and Hopper, F.N., Estimating the Degrees of Freedom for Linear Combinations of Mean Squares by Satterthwaite's Formula, Tech., 11 (1969), 691-706.

[16] Goldsmith, C.H. and Gaylor, D.W., Three Stage Nested Designs for Estimating Variance Components, Tech., 12 (1970), 487-498.

[17] Hornbrook, E.H.W. and Garrett, R.G., Regional Geochemical Lake Sediment Survey, East-Central Saskatchewan, Geol. Surv. Can. Paper 75-41 (1976), 20 p.

[18] Howarth, R.J. and Lowenstein, P.L., Sampling Variability of Stream Sediments in

Broad-Scale Regional Geochemical Reconnaissance, Trans. I.M.M., 80 (1971), B363-372.
[19] Krumbein, W.C. and Slack, H.A., Statistical Analysis of Low-Level Radioactivity of Pennsylvanian Black Fissile Shale in Illinois, Bull. Geol. Soc. Am., 67 (1956), 739-761.
[20] Leone, F.C., Nelson, L.S., Johnson, N.G. and Eisenstat, S., Sampling Distributions of Variance Components II. Empirical Studies of Unbalanced Nested Designs, Tech., 10 (1968), 719-737.
[21] Michie, U. McL., The Determination of Sampling and Analytical Errors in Exploration Geochemistry - Discussion, Econ. Geol., 68 (1973), 281-282.
[22] Miesch, A.T., Effects of Sampling and Analytical Error in Geochemical Prospecting, in Computers in the Mineral Industries (Pt. 1), Ed. G.A. Parks, Stanford Univ. Publ. Geol. Sci., 9 (1964), 156-170.
[23] Miesch, A.T., Theory of Error in Geochemical Data, U.S. Geol. Surv. Prof. Paper 574-A (1967), 17 p.
[24] Miesch, A.T., Geochemical Survey of Missouri-Methods of Sampling, Laboratory Analysis and Statistical Reduction of Data, U.S. Geol. Surv. Prof. Paper 954-A (1976), 39 p.
[25] Postma, B.J. and White, J.S., Computer Program: Variance Components for Unbalanced n-Level Hierarchic Designs, Jour. Qual. Tech., 7 (1975), 144-149.
[26] Rao, C.R., Estimation of Variance and Covariance Components in Linear Models, Jour. Am. Stat. Assoc., 67 (1972), 112-115.
[27] Satterthwaite, F.E., An Approximate Distribution of Estimates of Variance Components, Biometrics, 2 (1946), 110-114.
[28] Searle, S.R., Linear Models, New York: Wiley (1971), 532 p.
[29] Snee, R.D., Computation and Use of Expected Mean Squares in Analysis of Variance, Jour. Qual, Tech., 6 (1974), 128-137.
[30] Swallow, W.H. and Searle, S.R., Minimum Variance Quadratic Unbiased Estimation (MIVQUE) of Variance Components, Tech., 20 (1978), 265-272.
[31] Thomas, J.D. and Hultquist, R.A., Interval Estimation for the Unbalanced Case of the One-Way Random Effects Model, Ann. Math. Statist.,6 (1978), 582-587.

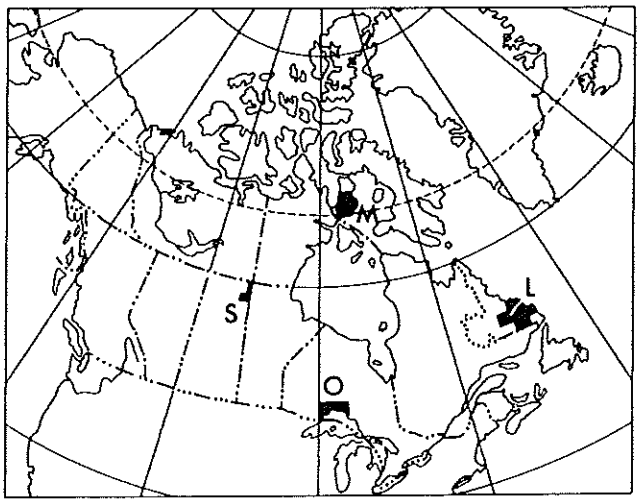

FIGURE 1: LOCATION OF SURVEY AREAS

[32] Tidball, R.R. and Severson, R.C., Chemistry of Northern Great Plains Soils, U.S. Geol. Surv. Open File Report 76-729 (1976), 57-81.
[33] Tietjen, G.L. and Moore, R.H., On Testing Significance of Components of Variance in the Unbalanced Nested Analysis of Variance, Biometrics, 24 (1968), 423-429.

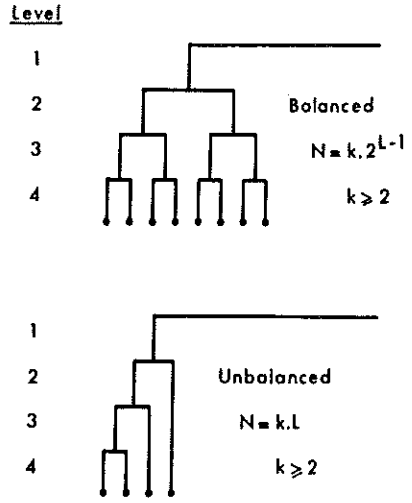

FIGURE 2: ILLUSTRATION OF BALANCED AND UNBALANCED NESTED DESIGN

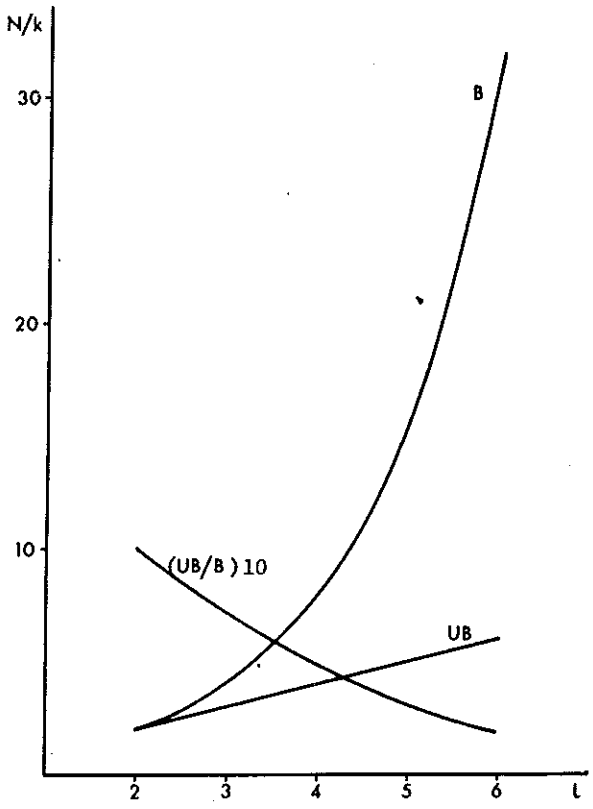

FIGURE 3: COMPARISON OF SAMPLING EFFICIENCY OF FIGURE 2 NESTED DESIGNS

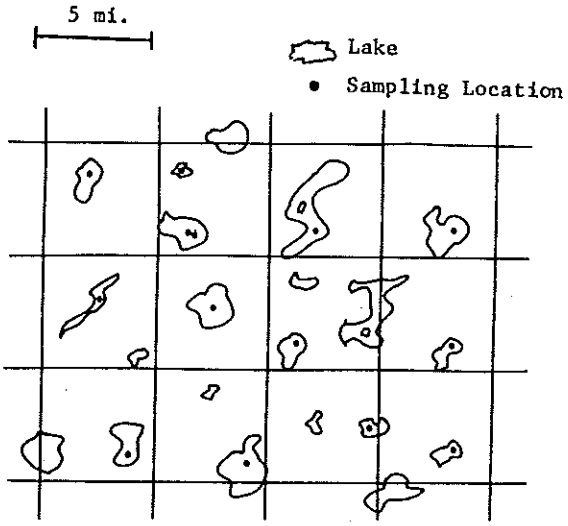

FIGURE 4: TYPICAL FIELD SAMPLING GRID

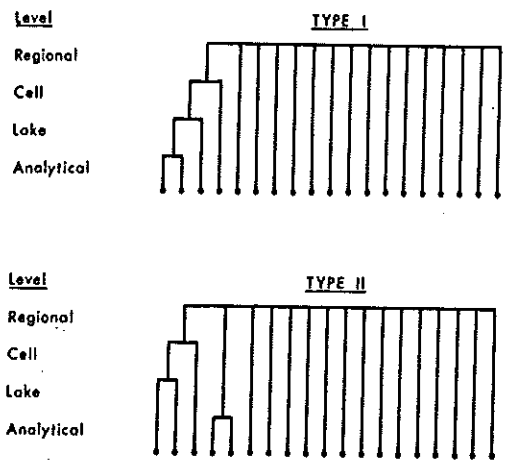

FIGURE 5: DATA STRUCTURES EMPLOYED IN THE INVERTED NESTED DESIGN

TABLE 1: ANALYSIS OF VARIANCE FOR 4-STAGE NESTED DESIGN[1]

Variation Between	Sum of Squares	Degrees of Freedom (DF)	Mean Squares	Expected Mean Squares[2]
Cells (a)	$\sum_i n_{i..}(\bar{x}_{i...}-\bar{x}_{....})^2=A$	a-1	A/DF(A)	$\sigma_e^2 + (\sum_{ijk} n_{ijk}^2 f_i)\sigma_c^2 + (\sum_{ij} n_{ij.}^2 f_i)\sigma_b^2 + (\sum_i n_{i..}^2 f_i)\sigma_a^2$
Lakes (b) in Cells	$\sum_{ij} n_{ij.}(\bar{x}_{ij..}-\bar{x}_{i...})^2=B$	$\sum_i n_i - a$	B/DF(B)	$\sigma_e^2 + (\sum_{ijk} n_{ijk}^2 f_{ij})\sigma_c^2 + (\sum_{ij} n_{ij.}^2 f_{ij})\sigma_b^2$
Samples (c) in lakes	$\sum_{ijk} n_{ijk}(\bar{x}_{ijk.}-\bar{x}_{ij..})^2=C$	$\sum_{ij} n_{ij} - \sum_i n_i$	C/DF(C)	$\sigma_e^2 + (\sum_{ijk} n_{ijk}^2 f_{ijk})\sigma_c^2$
Analyses (e) in Samples	$\sum_{ijkm}(x_{ijkm}-\bar{x}_{ijk.})^2=E$	$\sum_{ijk} n_{ijk} - \sum_{ij} n_{ij}$	E/DF(E)	σ_e^2
Total	$\sum_{ijkm}(x_{ijkm}-\bar{x}_{....})^2$	$\sum_{ijk} n_{ijk} - 1$		

[1] Notation is as follows:
summation limits for subscripts (i,j,k,m) are 1 to $(a, n_{i.}, n_{ij.}, n_{ijk})$ respectively;
the dot notation is used to indicate summation over an index:

$n_{...} = \sum_{ijk} n_{ijk}$ = total no. of analyses

$n_{i..} = \sum_{jk} n_{ijk}$ = total no. of analyses in i^{th} cell

$n_{ij.} = \sum_{k} n_{ijk}$ = total no. of analyses in j^{th} lake in i^{th} cell

$\bar{x}_{....} = \sum_{ijkm} x_{ijkm}/n_{...}$ = overall element mean (estimate of μ)

etc.

VARIANCE COMPONENTS ($\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_e^2$) ESTIMATION

- equate the "Mean Squares" and "Expected Mean Squares" columns then solve the linear equations

e.g., $s_e^2 = E/DF(E)$

$s_c^2 = (C/DF(C) - s_e^2)/(\sum_{ijk} n_{ijk}^2 f_{ijk})$

etc.

[2] $f_i = (1/n_{i..} - 1/n_{...})/DF(A)$; $f_{ij} = (1/n_{ij.} - 1/n_{i..})/DF(B)$; $f_{ijk} = (1/n_{ijk} - 1/n_{ij.})/DF(C)$
Note: these formulas assume random effects model for infinite populations, see text.

TABLE 2: TEST DATA FOR AN UNBALANCED ANOVA - DESIGN STRUCTURE IS INVERTED AND CONSISTS OF 5 DATA BLOCKS FROM THE 1977 ONTARIO NGR/URP SURVEY

ANALYSIS OF VARIANCE FOR URANIUM
DATA TRANSFORM = LOG10

LVL	VARIATION BETWEEN	SUM OF SQUARES	DOF	MEAN SQUARE	ESTIMATION OF VARIANCE COMPONENTS: UNIT SIZE	VARIANCE COMPONENT	PERCENT OF TOTAL	SYNTHESIZED MEAN SQUARE	COMPUTED DOF	HYPOTHESIS TEST (VARIANCE COMPONENT = 0): APPROX. F-RATIO	PROB. VALUE	SIGNIFICANCE(1)
4	CELLS	13.052	79	.16522	80	.13531	96.54	.52247E-02	5.50(2)	31.62	.9995	***
3	LAKES IN 4	.33384E-01	5	.66768E-02	85	.37126E-02	2.65	.12316E-02	5.41(2)	5.42	.9564	*
2	SMPLS IN 3	.66132E-02	5	.13226E-02	90	.68255E-03	.49	.45809E-03	5.00	2.89	.8653	NS
1	ANALS IN 2	.22904E-02	5	.45809E-03	95	.45809E-03	.33					
TOTALS		13.095	94			.14017						

(1) NS = NOT SIGNIFICANT, F .LE. 0.95
* = F .GT. 0.95 BUT .LE. 0.99
** = F .GT. 0.99 BUT .LE. 0.999
*** = F .GT. 0.999

(2) ERROR MEAN SQUARES SYNTHESIZED AND DEGREES OF FREEDOM COMPUTED BY SATTERTHWAITE'S APPROXIMATION

EXPECTED MEAN SQUARE FORMULAE:

E(MS(4)) =	1.0000	1.0366	1.0933	1.1824
E(MS(3)) =	1.0000	1.1333	1.4667	
E(MS(2)) =	1.0000	1.2667		
E(MS(1)) =	1.0000			

SUMMARY STATISTICS (ASSUMING NORMALITY) FOR URANIUM

DATA TRANSFORM = LOG10

ESTIMATED POPULATION MEAN = .70256
ESTIMATED VARIANCE OF MEAN = .23273E-02

STANDARD ERROR OF MEAN = .48242E-01

	LOWER	POINT	UPPER	
95% CONFIDENCE LIMITS AROUND MEAN ARE:	.60654	.70256	.79859	
LIMITS IN ORIGINAL UNITS ARE:	4.0415	5.0415	6.2891	
95% CONFIDENCE LIMITS AROUND LEVEL 1 VARIANCE COMPONENT ARE:	.20689E-03	.45809E-03	.19995E-02	
APPROXIMATE CONFIDENCE LIMITS FOR HIGHER LEVELS, THESE ASSUME A BALANCED, OR ONLY SLIGHTLY UNBALANCED, DESIGN				
95% CONFIDENCE LIMITS AROUND LEVEL 2 VARIANCE COMPONENT ARE:	.22784E-03	.68255E-03	.13307E-01	BY SATTERTHWAITE'S METHOD (1)
	0.	.68255E-03	.71721E-02	BY WILLIAMS-TUKEY METHOD (2)
95% CONFIDENCE LIMITS AROUND LEVEL 3 VARIANCE COMPONENT ARE:	.14250E-02	.37126E-02	.31655E-01	BY SATTERTHWAITE'S METHOD (1)
	0.	.37126E-02	.36880E-01	BY WILLIAMS-TUKEY METHOD (2)
95% CONFIDENCE LIMITS AROUND LEVEL 4 VARIANCE COMPONENT ARE:	.10514	.13531	.18183	BY SATTERTHWAITE'S METHOD (1)
	.87773E-01	.13531	.21494	BY WILLIAMS-TUKEY METHOD (2)

EMPIRICAL VARIANCE RATIO, V, AS A MEASURE OF SAMPLING EFFICIENCY (3) = 27.88

(1) SEE SATTERTHWAITE, F.E., 1946, BIOMETRICS, V.2, N.2, PP. 110-114

(2) SEE WILLIAMS, J.S., 1962, BIOMETRIKA, V.49, N.1&2, PP. 278-281

(3) SEE MIESCH, A.T., 1976, U.S. GEOL. SURV. PROF. PAPER 954-A, PP. A6-A10 AND A30-A32