

Facies Models Revisited: An Examination of Quantitative Methods

Charles W. Harper, Jr.
 School of Geology and Geophysics
 The University of Oklahoma
 Norman, OK 73019

Summary

In the widely-read *Geoscience Canada Facies Models* volume, Walker (1979) advocates a method for constructing facies models based on transition frequencies (see also Harper, 1984). I here discuss modifications to Walker's method, pitfalls both to Walker's and related methods, and suggest improvements, especially the use of median tetrads (defined below) and half-normal plots to recognize significant facies transitions "masked" by other approaches.

Walker's Method

Given several litho- and/or biofacies observed in successions, for each possible facies transition *i* to *j* (excluding self-transitions *i* to *i*):

a) tabulate the number of observed instances of the transition, and convert it to a relative frequency [Walker uses the relative frequency with respect to row totals, i.e., number of observed instances of *i* overlain by *j* divided by number of instances of *i* overlain by some facies other than *i* (so that frequency row totals equal 1)];

b) calculate the probability of transition from *i* to *j* assuming the null hypothesis that such transitions are random, depending only on the relative abundance of facies in the successions sampled;

c) subtract (b) from (a), and erect a facies relationship diagram emphasizing those differences which are large.

Improvements

Let us consider modifications to Walker's procedure by analyzing the large transition count matrix compiled by Johnson (1984) for the 3000 m-thick nonmarine Bellingham

Bay Member of the Chuckanut Formation of Northwest Washington (Fig. 1). Johnson's data provides an opportunity to analyze a large database (10 facies types, 2243 facies transitions counted).

1) *Observed frequencies.* Each element of the matrix in Figure 1 gives the number of occurrences of the facies listed on the left overlain by the facies listed on the top, e.g., there were 39 observations of facies C overlain by facies Fm. Walker would convert these to frequencies relative to row totals. However, it is simpler to work with actual transition counts.

2) *Estimating transition probabilities for a random sequence.* In estimating such probabilities, Walker (1979, p. 3) would use the actual numbers of instances of each facies type recorded in measured sections (rather than the row and column totals shown in Fig. 1). His procedure amounts to placing the actual facies counts as expected row (and column) totals for the purposes of computing transition probabilities (and expected cell counts) for the random sequence. Walker's procedure presumes that recorded counts of facies serve to estimate "the absolute abundance of the various facies" in the environment under study. Yet, such counts probably also reflect vagaries of outcrop pattern. For instance, facies are counted only if they occur at a facies transition; facies which tend to be covered more than others would tend to be double counted (where such facies are covered in the middle, they would tend to appear both at the top of one continuous outcrop interval yielding transition counts

and at the base of another). More fundamentally, Figure 1 gives counts of facies transitions, not facies; any model explaining it should refer to frequencies of facies transitions not frequencies of facies. My own predilection is to use *actual* row and column totals (Fig. 1) to estimate expected transition probabilities.

3) *Preserving row (and column) totals.* Walker (1979, p. 3) considers one row at a time. For each row, he estimates the probability of a transition from facies *i* to facies *j* by $n_{ij}/(N-n_i)$ where n_{ij} and n_i are the number of occurrences of facies *i* and *j* and *N* is the grand total of facies occurrences. In effect, he estimates the *conditional* probability of transition for *i* to *j* given that a transition from *i* to some other facies has occurred. While Walker's method preserves row totals, it distorts column totals (Read, 1969, p. 204; Turk, 1982; Carr, 1982, p. 907; Powers and Esterling, 1982). For example, column A and B totals obtained from Walker's Figure 2c by multiplying predicted frequencies by observed row totals (on his Fig. 2a) are 5 and 9, not 15 and 12. Also, Walker's method could just as reasonably be applied to expected column totals; in other words, he could estimate the conditional probability of transition from facies *i* to *j* given that a transition to *j* from some other facies has occurred ($= n_{ij}/(N-n_j)$). Yet, this estimate differs from his original ($= n_{ij}/(N-n_i)$); one might be judged significant yet the other not.

Turk (1979; 1982), Powers and Esterling (1982, p. 915-917) and Carr (1982) call attention to an alternative way of estimating expected cell counts: assume a model of

	C	Fm	F1	Sm	Sr	Sx	S1	Se	Gx	Gm	ROW TOTAL
C	-	39	23	2	21	9	2	0	0	1	97
Fm	53	-	99	55	252	38	10	2	10	10	529
F1	15	93	-	13	200	38	4	4	0	4	371
Sm	2	64	18	-	27	10	2	2	0	2	127
Sr	20	284	195	53	-	77	5	5	0	4	643
Sx	5	21	31	14	131	-	11	34	32	20	299
S1	0	10	1	0	5	14	-	1	1	4	36
Se	0	4	3	2	3	31	6	-	0	2	51
Gx	0	1	2	0	1	27	3	2	-	2	38
Gm	0	6	2	7	3	27	3	1	3	-	52
COL TOTAL	95	522	374	146	643	271	46	51	46	49	2243

Figure 1 Facies transition counts for nonmarine Bellingham Bay Member of the Eocene Chuckanut Formation of Northwest Washington. (Data

taken from Johnson, 1984, Table III, p. 378). For description of sedimentary facies see Johnson (1984, Table 2 and p. 367-377)

quasi-independence and use iterative proportional fitting to both row and column totals. The technique is easily applied using the widely available SAS package (SAS Inst., 1982, p. 540-542). In our context, quasi-independence means that for all facies transitions $i \rightarrow j$ where i does not equal j , the probability of transition from facies i to j is equal to a_i times b_j , where

a_i = the probability of a transition from facies i (row effect)

b_j = the probability of a transition to facies j (column effect)

A transition from a facies i to itself may be regarded either as undefined or as unobserved. Figure 2 shows predicted numbers of facies transitions for a random sequence obtained using iterative proportional fitting to row and column totals. (Compare with Fig. 1 showing actual counts observed.)

One may test the matrix as a whole for non-randomness before proceeding further (Harper, 1984). See Powers and Esterling, 1982, p. 916, and Turk (1979) for appropriate chi-square test comparing Figures 1 and 2. The values in Figures 1 and 2 yield a chi-square value of 1126 (for 71 degrees of freedom), which is significant at the .0001 level.

4) *Recognition of Outliers.* Walker's next step would be to subtract the values in Figure 2 from corresponding values in Figure 1 (Walker, 1979, subtracted observed from expected frequencies) and use any anomalously large differences to construct a facies transition diagram. If we are to use such residuals, it would be a good idea to transform them so that they are approximately distributed as normal. Each residual might then be tested for significance. Powers and Esterling (1982, p. 922) and Turk (1979, p. 990) suggest the transformation:

$(\text{Observed value}_j - \text{Expected value}_j) / (\text{Expected value}_j)^2$. These adjusted residuals are approximately distributed as unit normal. Haberman (1973, p. 206) proposes an almost identical transformation but divides the above expression by a difficult to calculate variance term. Actually, the above residuals will be approximately distributed as unit normal only if variances of the residuals are all close to 1 (Haberman, 1973, p. 206).

Figure 3 shows the residuals comparing the observed with expected frequencies (Figs. 1 and 2).

Application of Walker's original (row-wise) procedure (1979) supplemented with binomial tests of significance (Harper, 1984) singles out the facies transitions shown in Figure 4 as significant at the .01 level. These virtually are the same transitions that Johnson regarded as significant using Walker's procedure, but without formal tests (Johnson, 1984, Fig. 14). The residuals shown in Figure 3 suggest signifi-

cant facies transitions as shown in Figure 4 (residuals > 2.32 regarded as significant at .01 level; Powers and Esterling, 1982, p. 922). The virtual identity of the residuals identified by the two methods (Fig. 4) and Johnson (his Fig. 14) suggests that, in general, Walker's method gives the same results as iterative proportional fitting and analysis of residuals. But could all these results be wrong? Analysis of raw residuals – adjusted or not – is susceptible to a serious, if not fatal, flaw: if more than one significant outlier exists, these outliers are likely to go *undetected* by the method (Hawkins, 1980, p. 51-73; Braudu and Hawkins, 1982, p. 104; Pearson and Sekar, 1936). This potential *pitfall* is discussed in the next section.

5) *Recognition of Outliers – The Problem of Masking.* Any procedure designed to search for anomalous values, i.e., outliers, is susceptible to two types of errors – masking and swamping (Bradu and Hawkins, 1982). Masking occurs if the procedure declares fewer outliers than actually exist (i.e., so-called "type-two" errors occur). Swamping occurs if the procedure declares more outliers than actually exist (i.e., "type-one" errors occur). Analysis of raw residuals as outlined above, while perhaps relatively free of swamping, runs a high risk of *masking* significant residuals, if more than one exist. In particular, all but the largest residual may well be masked.

As an example, suppose that transition counts for ten facies have expected values identical to those given in Figure 2 above, except for elements at (3,2), (5,1), (5,2), and (5,6). Set the expected values of $n_{3,2}$, $n_{5,1}$, $n_{5,2}$ and $n_{5,6}$ to 350, 100, 2500 and 308, respectively, (as opposed to 106, 34,

235 and 105 in Fig. 2). Any artificial dataset produced under such constraints results from a perfectly quasi-independent (random) model with four glaring outliers. One such dataset is given in the appendix. If we apply iterative proportional fitting to the dataset (as was done in Fig. 2), and calculate residuals (as in Fig. 3), three of the four outliers produce residuals which are zero! (All but that for $n_{5,2} = 2500$ are 0).

At first glance it might appear that a multi-step procedure identifying outliers one at a time could circumvent the problem of masking. For example, Carr (1982, p. 908) would identify significant extreme cells as follows: perform a chi-square test of the matrix as a whole to test its fit to a random model (in his case, a log-linear model of quasi-independence). Carr's model is generated using iterative proportional fitting as outlined above. If the quasi-independence model provides an inadequate fit to the observed data: a) select the cell that causes the greatest reduction in the chi-square statistic; b) recalculate the chi-square statistic with the cell eliminated; c) continue steps (a) and (b) until the recalculated chi-square value is not significant. However, such multi-step procedures also run a high risk of masking (Hawkins, 1980, p. 51-52, 63-67). Carr's procedure applied to the above artificial dataset does locate 3 outliers. However, if outliers "351" and "2500" are changed to "500" and "1900", respectively, then the procedure misses three of the four defined outliers (including "1900" – the most deviant)! Rather, it deletes one spurious "deviant" cell after another (mostly ones with negative residuals, 7 in all) leaving real outliers intact. I have applied Carr's procedure applied to the Bellingham Member data using a .05

		Fm	F1	Sm	Sr	Sx	S1	Se	Gx	Gm	ROW TOTAL
C	-	24.2	15.5	5.4	34.3	10.8	1.6	1.8	1.6	1.8	97
Fm	24.0	-	107.6	37.3	237.7	74.8	11.4	12.7	11.3	12.2	529
F1	15.1	105.8	-	23.5	149.6	47.1	7.1	8.0	7.1	7.7	371
Sm	4.6	32.4	20.7	-	45.8	14.4	2.2	2.4	2.2	2.3	127
Sr	33.5	235.4	150.6	52.2	-	104.7	15.9	17.7	15.9	17.0	643
Sx	11.6	80.8	51.7	17.9	114.2	-	5.5	6.1	5.4	5.8	299
S1	1.3	8.6	5.6	2.0	12.5	3.9	-	.7	.6	.6	36
Se	1.8	12.5	8.0	2.8	17.7	5.6	.8	-	.8	.9	51
Gx	1.3	9.3	6.0	2.1	13.2	4.1	.6	.7	-	.7	38
Gm	1.8	12.8	8.2	2.8	18.0	5.7	.9	1.0	.9	-	52
COL TOTAL	95	522	374	146	643	271	46	51	46	49	2243

Figure 2 Iterative Proportional Fit to Bellingham Member data of Figure 1

stopping level of significance; this produces very different results from those shown in Figure 4. While eight of the transitions shown in Figure 4 are singled out as deviant (Sr to Fm; Sr to Fl; Fl to Sr; Sl to Sx; Sx to Se; Sx to Gx; Sx to Gm; and Sl to Gm) by the method, the remaining 14 are not regarded as significant (doubtless due to masking). For more discussion of Carr's method, see the appendix.

6) *Recognition of Outliers – A Supplemental Procedure.* Bradu and Hawkins (1982) provide a way to identify multiple outliers in a single step while avoiding the problem of masking: median tetrads (defined below) are used instead of residuals. Half-normal plots help to locate deviant outliers. The method requires at least five different facies types in the data matrix.

Suppose we have shown that our data matrix as a whole does not fit a quasi-independent (random) model (in section 3 above). Then, as Powers and Esterling (1982, p. 922-23) point out, we should use a model which includes *both* signal and noise. A log-linear model (Carr, 1982, p. 907-908) with an interaction term does just this. For two facies, *i* and *j*, suppose quasi-independence holds. Then the expected number of transitions from *i* to *j*, E_{ij} , equals $E_{ij} = Na_{ij}$ (1) where *N* = total number of facies transitions observed (see discussion of quasi-independence above). Thus:

$\log(E_{ij}) = \log(N) + \log(a_{ij}) + \log(b_{ij})$ (2) where $\log(N)$ is a constant. Suppose, however, quasi-independence does not hold, but *i* and *j* interact (i.e., outliers exist). We can model this interaction with a term g_{ij} so that

$\log(E_{ij}) = \log(N) + \log(a_{ij}) + \log(b_{ij}) + g_{ij} = Y_{ij}$ say. (3) (4)

[Carr (1982, p. 908) misleadingly refers to this model as the model of quasi-independence; this is true only if the interaction terms are all zero (eqn. 2 above)]. Now, suppose that there is a subset *T* of aberrant cells such that

$g_{ij} = 0$ for (i,j) not in *T* (5)
 $g_{ij} \neq 0$ for (i,j) in *T*. (6)

The subset *T* contains the outliers.

Following Bradu and Hawkins (1982), we may define a tetrad $T_{ij;eg}$ as:
 $T_{ij;eg} = Y_{ij} + Y_{eg} - Y_{ig} - Y_{ej}$. (7)
 Tetrads for which $i = e$ or $j = g$ are identically zero and should be ignored.

Similarly, in our case, any tetrad involving a diagonal element is undefined and may be ignored (Bradu and Hawkins, 1982, p. 107).

If all four cells of the tetrad are non-deviant, then the expected value of $T_{ij;eg}$ is zero, as is the expected value of g_{ij} . If, however, cell (i,j) is an outlier, whereas

cells (e,j) , (i,g) and (e,g) are not, then $T_{ij;eg}$ is an unbiased estimator of g_{ij} . (Bradu and Hawkins call such a tetrad a "clean" tetrad. A "contaminated" tetrad is one where (e,j) and/or (i,g) and/or (e,g) are outliers.)

Bradu and Hawkins (1982, p. 105) recommend the following procedure:

- (1) For each cell calculate $Q_2(i,j)$ the median tetrad (median not mean; Noether, 1971, p. 7).
- (2) Rank cells in decreasing order of $|Q_2(i,j)|$, the absolute value of the median tetrad.
- (3) Generate half-normal plot and use it to locate outliers (see details below).

To compute the median tetrad $Q_2(1,2)$ for, let us say, the matrix shown in Figure 1, calculate the tetrads $T_{ij;eg}$ for $(e,g) = (4,3), (5,3), (6,3), (7,3), (8,3), (9,3), (10,3), (3,4), (5,4) \dots (9,10)$. (In other words, for all non-

	C	Fm	Fl	Sm	Sr	Sx	Sl	Se	Gx	Gm
C	-	3.0	-1.9	-1.5	-2.3	-0.5	.3	-1.4	-1.3	-0.6
Fm	5.9	-	-0.8	2.9	.9	-4.3	-0.4	-3.0	-0.4	-0.6
Fl	-0.0	-1.2	-	-2.2	4.1	-1.3	-1.2	-1.4	-2.7	-1.3
Sm	-1.2	5.6	-0.6	-	-2.8	-1.2	-0.1	-0.3	-1.5	-0.2
Sr	-2.3	3.2	3.6	.1	-	-2.7	-2.7	-3.0	-4.0	-3.2
Sx	-1.9	-6.7	-2.9	-0.9	1.6	-	2.4	11.3	11.4	5.9
Sl	-1.1	.4	-2.0	-1.4	-2.1	5.0	-	.4	.5	4.2
Se	-1.3	-2.4	-1.8	-0.5	-3.5	10.8	5.6	-	-0.9	1.1
Gx	-1.2	-2.7	-1.6	-1.4	-3.4	11.2	3.0	1.6	-	1.6
Gm	-1.3	-1.9	-2.1	2.5	-3.5	8.9	2.3	.0	2.3	-

Figure 3 Residuals of Bellingham Member data
 $Residual_{ij} = (Obs_{ij} - Exp_{ij}) / (Exp_{ij})^{1/2}$

Obs_{ij} = observed transition count (from Fig. 1)
 Exp_{ij} = expected count (from Fig. 2)

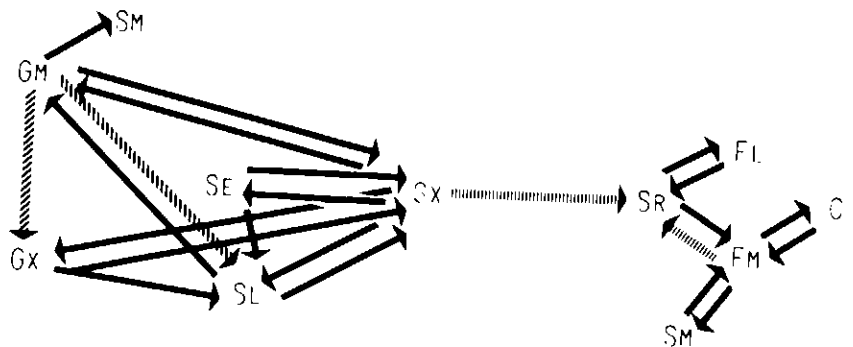


Figure 4 Facies relationship diagram for Bellingham Member. Solid arrow = facies transition indicated as significant both by (A) the procedure of Walker (1979) and Harper (1984), as well as (B) the analysis of residuals in Figure 3. Dashed arrow = transition indicated as signifi-

cant only by (A). Diagonally dashed arrow = transition indicated as significant by (B). Level of significance = .01 (.05 for transition Sx to Sl using (A)). (Compare with Johnson, 1984, Fig. 14, p. 377.)

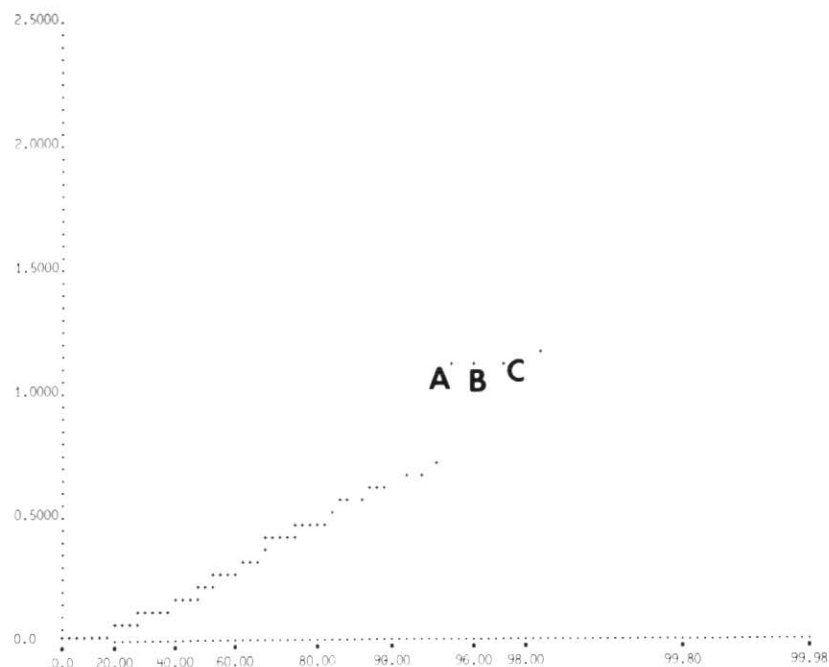


Figure 5 Half-normal plot of simulated data matrix with four outliers. Outlier A = transition from Sr to C; B = transition from Sr to Sx; C = transition from Fl to Fm; and D = transition from Sr to Fm

	C	Fm	Fl	Sm	Sr	Sx	Sl	Se	Gx	Gm	TOTAL
C	-	23	15	5	33	9	0	0	0	0	85
Fm	23	-	105	37	237	72	11	13	12	11	521
Fl	14	351	-	22	149	46	6	8	7	8	611
Sm	3	32	18	-	44	14	2	2	3	1	119
Sr	99	2500	149	51	-	307	15	18	15	16	3170
Sx	8	80	50	17	114	-	6	7	5	5	292
Sl	1	9	5	0	11	1	-	0	0	0	27
Se	3	10	9	2	18	4	0	-	0	0	46
Gx	2	9	5	1	12	2	0	0	-	0	31
Gm	0	11	7	1	18	3	0	0	0	-	40
COL											----
TOTAL	153	3025	363	136	636	458	40	48	42	41	4942

Figure 6 Simulated data matrix with four outliers (for explanation see text)

linear; suspected outliers will lie at the end of the plot where a deviation from linearity occurs. Sparks (1970) provides a short Fortran program which will generate half-normal plots (see also Munford, 1972).

Figure 5 shows a half-normal plot for the artificial dataset defined above (listed in the appendix). All four outliers show up on the plot.

A half-normal plot for the Bellingham Member data (Fig. 1) does not show outliers. Rather, it is perfectly linear. This result suggests that median tetrads and half-normal plots should be used to supplement, not replace, the methods discussed earlier in this paper.

Acknowledgements

I would like to express my thanks to R.N. Hiscott, Memorial University of Newfoundland, who called to my attention several key references.

References

Bradu, D. and D.M. Hawkins, 1982, Location of multiple outliers in two-way tables, using tetrads: *Technometrics*, v. 24, p. 103-108.

Carr, T.R., 1982, Log-linear models, Markov chains and cyclic sedimentation: *Journal of Sedimentary Petrology*, v. 52, p. 905-912.

Haberman, S.J., 1973, The analysis of residuals in cross-classified tables: *Biometrics*, v. 29, p. 205-220.

Harper, C.W., Jr., 1984, Improved Methods of Facies Sequence Analysis: *in* Facies Models, R.G. Walker, ed., Geoscience Canada Reprint Series no. 1, 2nd ed., p. 11-13.

Hawkins, D.M., 1980, Identification of Outliers: Chapman and Hall: New York, p. 1-188.

Johnson, N.L. and F.C. Leone, 1977, Statistics and experimental design in engineering and physical sciences, v. 2, 2nd ed., John Wiley and Sons, N.Y., p. 593-1082.

Johnson, S.Y., 1984, Cyclic Fluvial Sedimentation in a rapidly subsiding basin, Northwest Washington: *Sedimentary Geology* v. 38, p. 361-391.

Munford, A.G., 1972, A remark on algorithm AS 30: Half normal plotting: *Applied Statistics*, v. 21, p. 351.

Noether, G.E., 1971, Introduction to Statistics: Houghton Mifflin Co., N.Y. p. 1-253.

Pearson, E.S. and C. Sekar, 1936, The efficiency of statistical tools and a criterion for the rejection of outlying observations: *Biometrika*, v. 28, p. 308-320.

Powers, D.W. and R.G. Esterling, 1982, Improved methodology for using embedded markov chains to describe cyclical sediments: *Journal of Sedimentary Petrology*, v. 52, p. 913-23.

Read, W.A., 1969, Analysis and simulation of Namurian Sediments in Central Scotland using a Markov process model: *Mathematical Geology*, v. 1, p. 199-219.

- SAS Institute Inc., 1982, SAS Users Guide: Statistics, SAS Institute Inc., Cary, N. Carolina, p. 1-584.
- Sparks, D.N., 1970, Half Normal Plotting: Applied Statistics, v. 19, p. 192-6 (see also Corrigenda, 1971, v. 20, p. 118).
- Turk, G., 1979, Transition analysis of structural sequences: Discussion and Reply: Geological Society of America bulletin, v. 90, p. 989-991.
- Turk, G., 1982, Markov Chain Analysis: Mathematical Geology v. 14, p. 539-42.
- Walker, R.G., 1979, Facies and Facies Models 1: General Introduction in Facies Models, R.G. Walker, ed., Geoscience Canada Reprint Series no. 1., 1st ed., p. 1-7.

Appendix

Figure 6 shows an artificial dataset produced by assuming that the transition counts for ten facies have expected values identical to those given in Figure 2 above, except for elements at (3,2), (5,1), (5,2), and (5,6), which have expected values of 350, 100, 2500, and 308, respectively. A loglinear model is assumed such that the log of each dataset element is distributed as unit normal; for this reason, transition counts listed may differ slightly from those given in Figure 2.

A few details regarding Carr's method follow: First, Carr (1982, p. 909) recommends the stepwise technique of Brown (1974) which selects the cell at each step whose deletion brings about the greatest reduction in chi-square. However, Brown's procedure only approximates the value of chi-square that will result from the deletion of each cell. It is much safer to calculate these resultant chi-square values exactly (with a modest increase in computer time using SAS). Brown's procedure can result in an incorrect suite of cells being selected as outliers, especially where differences between the effects of cells on chi-square are slight; indeed, it introduces errors when applied to the Bellingham data.

Secondly, Brown (1974, p. 408) suggests an alternative criterion for the stepwise identification of cells as significant: instead of selecting the cell at each step whose deletion produces the minimum chi-square, select the cell with the maximum residual (Fig. 3). This alternative applied to the Bellingham data produces the same results as shown in Figure 4(B) with the notable exception that residuals for transitions C to Fm, Sr to Fm, Sr to Fl and Fl to Sr are not regarded as significant.

Reference

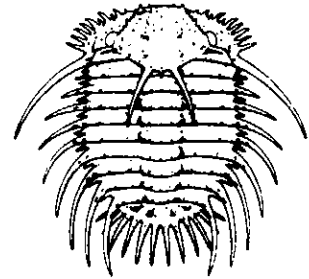
- Brown, M.B., 1974, Identification of the Sources of Significance in Two-way contingency tables Applied Statistics, v. 23, p. 405-413.

MS received, as revised, July 19, 1984

PALAEONTOGRAPHICA CANADIANA No.1



Silicified Silurian
odontopleurid trilobites from
the Mackenzie Mountains
by
Brian D.E. Chatterton
and
David G. Perry



Palaeontographica Canadiana is a series devoted to major contributions to Canadian paleontology, is dominantly, but not exclusively, systematic in content. It is sponsored jointly by the Canadian Society of Petroleum Geologists and the Geological Association of Canada.



Soft Cover, 126 pages, 21 x 28 cm (8 1/2" x 11") 36 Plates, 1983 ISSN 0821-7556

CANADIAN SOCIETY OF PETROLEUM GEOLOGISTS
GEOLOGICAL ASSOCIATION OF CANADA



Cost \$21.00 + Mailing and Handling, \$2.50 in Canada,
\$5.00 elsewhere (Canadian funds)

Send cheque or money order in Canadian dollars to:

Canadian Society of Petroleum Geologists
#505, 206 - 7th Ave. S.W., Calgary, Alberta, Canada T2P 0W7

or

Geological Association of Canada Publications,
Business and Economic Service Ltd.,
111 Peter Street, Suite 509, Toronto, Ontario, Canada M5V 2H1