NOTE ON WINDOWING AND WEIGHTED DEPTH SCALING (12 August, 2007)

The purpose of this note is to illustrate in detail how the windowed depth scaling dendrogram (*.dn2 file) is derived. The dataset used for example is ns30cen with depths of events (mainly last occurrences) for Cenozoic Foraminifera, miscellaneous shelly microfossils, dinoflagellates, and some other types of stratigraphic events such as log markers in 30 North Sea wells. Run parameters are shown in Figures 1-4. According to the RASC menu (Fig. 1) used in this experiment, only events occurring in at least 6 wells were accepted for ranking and scaling, and pairs of events co-occurring in fewer than 3 wells were not used. The resulting ranked and scaled optimum sequences are for 85 events.

According to the CASC input parameter file (Fig. 4) in this experiment, depth differences exceeding 400m (in absolute value) were not used. Relative order within ranked optimum sequence was taken as initial order, and windowing was applied with order of window (OW) set equal to 5. OW=5 implies that each pair of successive events in the ranked optimum sequence is compared to the 4 events stratigraphically above it, as well as to the 4 events stratigraphically below it. OW=1 would mean that first order depth differences are used only.

In this experiment, the last occurrence of *Spiroplectammina carinata* (Event # 18) was selected for example, tracking in detail all operations performed to first estimate its interevent distance to the last occurrence of *G*. ex.gr. *praescitula zealandica* (Event # 236) that is situated immediately below it in the ranked optimum sequence. Later, in the final windowed depth scaling dendrogram (*.dn2 file), *Spiroplectammina carinata* ends up being situated immediately above *Globigerina* ex.gr. *officinalis* (Event # 111).

In Kaminski and Gradstein (2005, Figure 18, p 50), *Spiroplectammina carinata* occurs at the top of Zone NSR8A (Late Oligocene). On the other hand, *G*. ex.gr. *praescitula zealandica* occurs within Zone NSR9A (late Early to early Middle Miocene). It will be seen that the weighted average depth difference for these two events is negative. Because of this, *G*. ex.gr. *praescitula zealandica* is moved upward in the windowed depth scaling dendrogram where it ends up among other events assigned to Zone NSR9A in Kaminski and Gradstein (2005). NSR (North Sea RASC) zones were based on a RASC scaled optimum sequence represented in dendrogram form.

A major source of uncertainty is that relatively many event pairs involving Events # 18 and 236 and their neighbors in the ranked optimum sequence co-occur in relatively few wells. In general, the frequency distribution of depth differences has thin long tails extending as much as 1000m in both the stratigraphically upward and downward directions (Agterberg, Gradstein and Liu, 2007; also see *.df1 file with histogram for all depth differences in RASCW output). This means that average depth differences calculated for small samples can become either much too large or too small if the sample happens to contain one or more depth differences that (in absolute value) are much

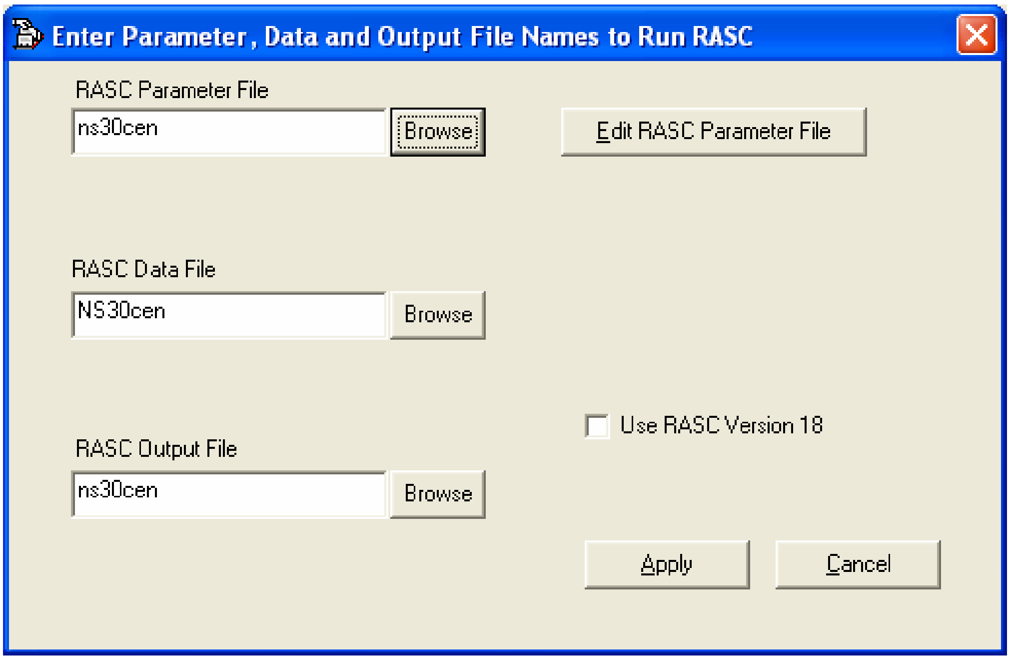

Figure 1. RASC Menu used for experiment

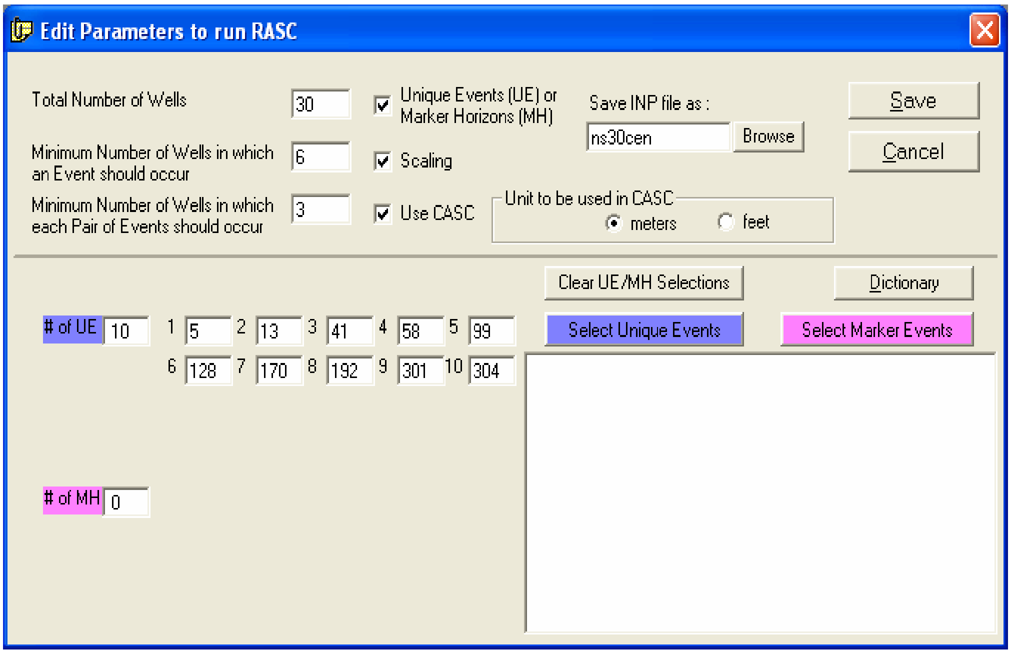

Figure 2. RASC Parameter File menu used for experiment

greater than normal. An advantage of windowing is that more than a single average depth difference is used for estimating the interevent distance. Windowing generally greatly improves quality of estimates.

Nevertheless, in order to obtain a good weighted depth scaling dendrogram, the user normally should conduct several experiments inspecting graphs such as depth difference

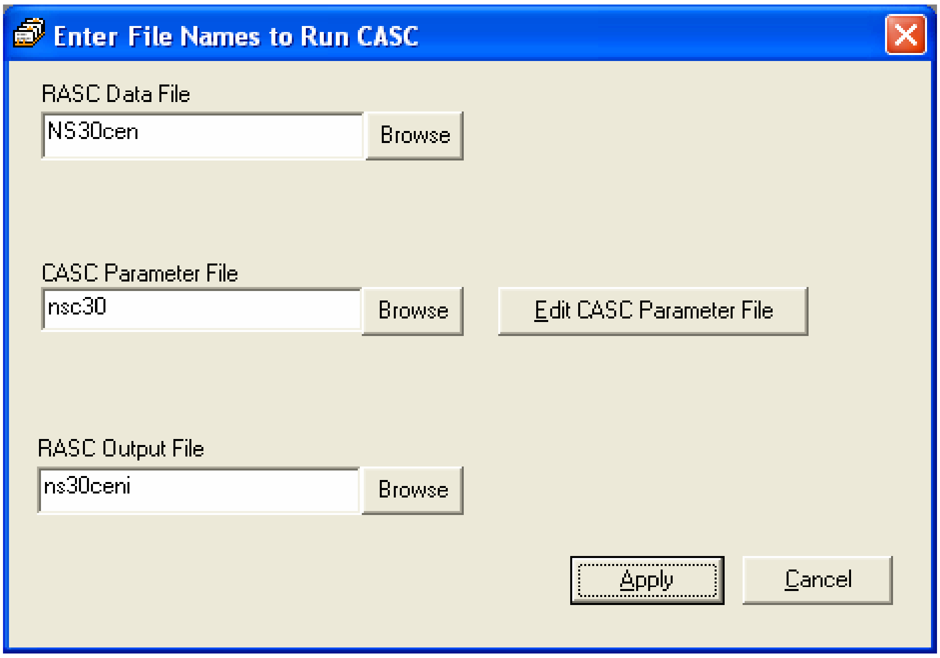

Figure 3. CASC menu used for experiment

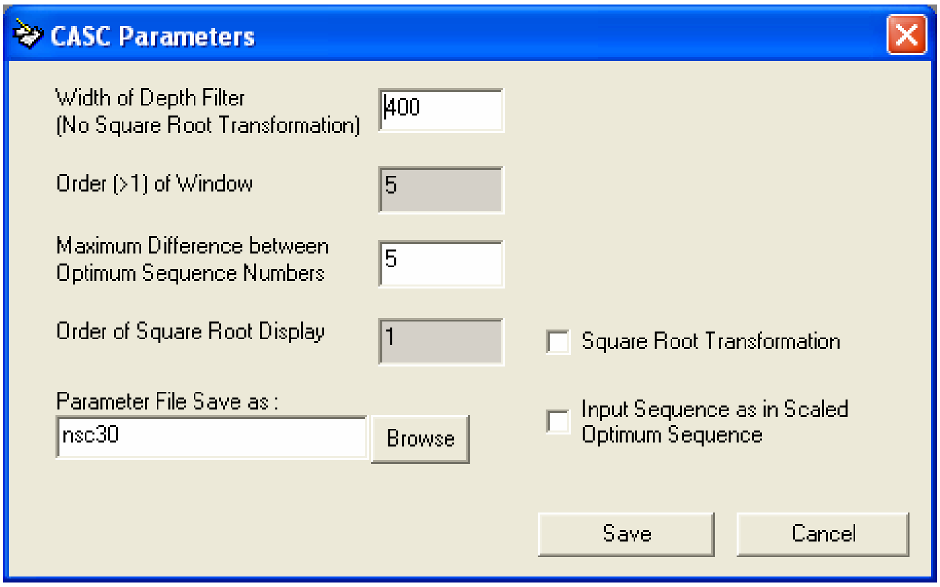

Figure 4. CASC Parameter File menu used for experiment

histograms and changing CASC depth scaling input parameters to order to ascertain that a stable solution is obtained that is acceptable from a stratigraphic point of view.

According to the depth scaling computer algorithm incorporated in the CASC portion of RASCW Version 20, sample size for co-occurring events should be at least 3. This condition was only satisfied for 2 of the 9 possible estimates (for window order = 5) of the interevent distance between Event # 18 and Event # 236 due to paucity of data in this part of the optimum sequence.

Table 1 shows sample sizes (# of Wells) for direct interevent distance estimation involving the first 40 events in the ranked optimum sequence. Note that there are as many as five 0 values in the third column (IE Dist 1) because sample size is less than 3.

Table 1. Comparison of direct interevent distances (IE Dist 1) and windowed weighted interevent distances (IE Dist 2) for first 40 events in ranked optimum sequence (ROS) in ns30cen depth scaling run used for example.

ROS #	# of Wells	IE Dist 1	IE Dist 2	Event #	Event Name
1	15	15.4	15.4	77	Elphidium spp.
2	7	28.571	28.571	228	Cassidulina teretis
3	7	34.286	34.286	338	Bulimina marginata
4	8	7.5	7.5	1	Neogloboquadrina pachyderma
5	6	12.5	-14.614	31	Cibicidoides scaldisensis
6	7	27.857	25.128	171	Buccella frigida
7	5	152	111.572	270	Cibicidoides grossa
8	3	60	-14.718	339	Globigerina bulloides
9	5	50	22.968	4	Globorotalia inflata
10	3	96.667	3.367	181	Cibicidoides dutemplei
11	4	-32.5	14.381	316	Trifarina fluens
12	5	-22	8.729	159	Cibicidoides grossa LCO
13	12	81.917	73.492	23	Sigmoilopsis schlumbergeri
14	5	-5.2	-10.55	269	Neogloboquadrina atlantica
15	2	0	-38.003	266	Globorotalia puncticulata
16	1	0	121.726	330	Cibicidoides pachyderma
17	4	-39.75	-0.105	219	Martinotiella cylindrica
18	3	66.667	78.884	130	Siphonina advena
19	3	-72	-73.6	285	Caucasina elongata
20	5	72.8	99.508	91	Diatoms/radiolarians LCO
21	3	9.667	-22.452	282	Uvigerina ex.gr. semiornata
22	6	-10.5	4.42	123	Globigerinoides trilobus
23	5	-53	-31.823	207	NS Log G
24	1	0	-34.841	125	Neogloboquadrina continuosa
25	3	236.667	90.159	71	Epistomina elegans
26	3	-32.333	38.463	15	Globigerina praebulloides
27	1	0	-56.47	18	Spiroplectammina carinata
28	11	58.091	42.184	236	G. ex.gr. praescitula zealandica
29	5	46.8	13.611	17	Asterigerina gurichi
30	2	0	17.777	138	Tenuitella angustiumbilicata
31	3	56.667	22.005	111	Globigerina ex.gr. officinalis
32	4	-3.75	-19.263	413	Silicious biofacies
33	5	28.8	80.803	20	Gyroidina girardana
34	3	93.333	53.226	24	Turrilina alsatica
35	7	37.429	18.14	142	Gyroidina soldanii mamilligera
36	23	49.261	51.886	25	Coarse agglutinated spp.
37	17	47.353	41.21	97	Cyclammina placenta
38	11	68.636	53.361	182	Spirosigmoilinella compressa
39	6	25.5	24.858	262	Karrerulina horrida
40	5	72.6	0.465	259	Ammodiscus latus

Although the same minimum sample size criterion (>2 co-occurring pairs) was used for weighted interevent distance estimation (IE Dist 2), there are no 0 values in column 4 of Table 1. The values in this column also tend to be smaller than those in the third column. This is because they are subject to less uncertainty.

The weighted estimate of the distance between *Spiroplectammina carinata* (18) and *G*. ex.gr. *praescitula zealandica* (236) is a relatively large negative value (= -56.47). It is based on interevent distances between last occurrences of *Spiroplectammina carinata* (18) and *G*. ex.gr. *praescitula zealandica* (236) in connection with last occurrences of *Globigerina praebulloides* (15) and *Asterigerina gurichi* (17) immediately above 18 and below 236, respectively, in the ranked optimum sequence. Wells with co-occurrences of two or more of these four events are listed in Table 2.

Table 2. Depths (m) of events with Ranked Optimum Sequence (ROS) numbers 26 to 29. Wells with only one of the four events are not listed.

ROS #	26	27	28	29
Event #	15	18	236	17
Well #				
3	1100			1090
6	1130		1160	
8	1030			1140
12			990	1260
14	560	816		541
16	865	612		554
17		722		685
18	899		899	908
19	920		920	920
22	1514		1539	1539
23			1400	1400
24			609	609
25	1383		1383	1356
27	2200		1554	1554
28	1527		1685	1740
29	1834	1734	1654	1834
30	1493		1493	1645

An example (see Table 3) of unweighted average depth difference calculation involving *Spiroplectammina carinata* and *Asterigerina gurichi* (27→29) is as follows: Depth difference for Well #14: 541-816 = -275; for Well #16: 554-612 = -58; for Well #17: 685-722 = -37; and for Well #29: 1834-1734 = 100. Consequently, Total Depth Difference = -270; and Average Depth Difference = -67.5m

Table 3. Average Depth Difference (ADD) and Sample Size (SS); ADDs for SS=1 or SS=2 were not used; depth differences greater than 400m in (absolute value) were not used.

Event Pair	ADD	SS	Note:
27-28	-80	1	Not used because SS < 3
27-29	-67.5	4	
28-29	58.09	11	
26-28	4.125	8	1554-2200 = -646 not used
26-27	-32.33	3	

Indirect distances for 27→28 are (a) difference between 27→29 and 28→29, and (b) difference between 26→28 and 26→27. In general, if one or both differences used for indirect distance estimation would be for a sample with SS < 3, the estimate is not used. The sample sizes (4, 11, 8 and 3) involving events 26 and 29 all satisfy this cut-off criterion and both indirect distance estimates are used. However, as mentioned before, although OW=5, none of the 6 indirect estimates involving events 23, 24, 25, 30, 31 and 32 were used because they did not meet the cut-off criterion. Consequently, setting FW=2 would have yielded the same answer as FW=5, but only for 27→28 and perhaps a few other pairs of this type that is probably either unique or rare.

For the example, the direct distance for 27→28 is 0 because SS =1 is less than 3. The two unweighted indirect distance estimates would be -67.5 – 58.09 = -125.591, and 4.125 – (-32.33) = 36.458. However, precision of ADDs depends on sample size: the distance estimates for 28→29 (= 58.09) and 26→28 (= 4.125) should be given more weight than those for 27→29 (= -67.5) and 26→27 (= -32.33) because they are based on larger samples.

Weights are calculated as follows: Suppose that n_0 represents sample size of average depth difference of the first order (successive events) written as x_0; e.g. in Table 3: n_0 = 1, x_0 = -80. Suppose further that n_1 and n_2 represent sample sizes of average depth differences of the second order written as x_1 and x_2, respectively; e.g. in Table 3: n_1 = 4, x_1 = -67.5 and n_2 =11, x_2 =58.09.

In Agterberg, Gradstein and Liu (2007, Figs. 8A and 11) it is shown that standard deviations of square root transformed depth differences of the first five orders satisfy approximately normal frequency distributions. The slopes of best-fitting lines in the Q-Q plots for square root transformed data are approximately equal. This implies that the variance of depth differences of any low order is nearly constant and can be written as σ^2. Consequently, the variance of x_0 is σ^2/n_0, and the variance of any difference (x_1 - x_2) becomes:

$$\sigma^2(x_1 - x_2) = \sigma^2/n_1 + \sigma^2/n_2 = \frac{n_1 + n_2}{n_1 n_2}\sigma^2$$

Weights of average depth differences can be defined as $w_0 = n_0/\sigma^2$, $w_1 = n_1/\sigma^2$, and $w_2 = n_2/\sigma^2$, respectively. Also, the weight of w_{12} of any estimate (x_1 - x_2) is:

$$w_{12} = \frac{n_1 n_2}{n_1 + n_2} \cdot \sigma^{-2}$$

Table 4 shows relative weights (after setting σ^{-2} =1) for selected values of n_1 and n_2. For $n_2 = n_1 = n_0$, the weight w_{12} of an indirect estimate is half of w_0 representing the weight of the direct estimate. If $n_2 > n_1 = n_0$, w_{12} is less than w_0 but greater than $0.5 \cdot w_0$.

Table 4. Relative weights w_{ij} for n_1 >2 and n_2 >2

n_1/n_2	3	4	5	7	10	20
3	1.50					
4	1.71	2.00				
5	1.88	2.22	2.50			
7	2.10	2.55	2.92	3.50		
10	2.31	2.86	3.33	4.12	5.00	
20	2.61	3.33	4.00	5.19	6.67	10.00

If there are other indirect distances that can be estimated, the subscripts 1 and 2 in the preceding expression can be replaced by *i* and *j*. The final Weighted Average Depth Difference (WADD) satisfies:

$$WADD = \frac{\Sigma w_{ij} \cdot \bar{x}_{ij}}{\Sigma w_{ij}}$$

where the summation sign applies to all available indirect estimates $\bar{x}_{ij}$ with weights w_{ij}. Obviously, WADD does not depend on the value of σ^2. Consequently, we can use relative weights that, for convenience, also can be written as w_{ij} by setting σ^2=1 without loss of generality. If depth differences (in absolute value) greater than a pre-defined truncation depth (= 400m in the experiment) are excluded, the variance of depth differences becomes less than σ^2. However, this variance reduction does not introduce bias into the relative weights, because depth differences in all samples then have the same reduced variance.

In general, windowing consists of averaging a larger number of ADDs than in the example of Table 3. The direct estimate x_0 is combined with estimates of the type $(x_i - x_j)$ that are stratigraphically lower and higher. For the example of Table 3, the direct estimate is not used because SS =1 is less than 3 and there are only two indirect estimates of the type $(x_i - x_j)$. These are -9.41 and 28.205 (see before) with relative weights 2.933 and 2.192, respectively. Multiplication of the two indirect estimates by their weights and adding the results gives -2.941 x (-125.591) + 2.192 x 36.458 = -288.855. Division of this total by the sum of the weights (=5.115) gives -56.47 representing our final estimate of the weighted average depth difference (WADD) between events 27 and 28.

Table 5. ROS# = Ranked Optimum Sequence Number; CumDist = Cumulative Depth Difference; WADD = Weighted Average Depth Difference; IE Dist = Interevent Distance in Depth Scaling Dendrogram

ROS #	Event #	WADD	Cum Dist	Event Name	Event #	Reordered Cum Dist	IE Dist(m)
1	77	15.4	0	Elphidium spp.	77	0	15.4
2	228	28.5714	15.4	Cassidulina teretis	228	15.4	28.5714
3	338	34.2857	43.9714	Bulimina marginata	338	43.9714	27.1717
4	1	7.5	78.2571	Neogloboquadrina pachyderma	171	71.1432	7.114
5	31	-14.6139	85.7571	Cibicidoides scaldisensis	1	78.2571	7.5
6	171	25.1279	71.1432	Buccella frigida	31	85.7571	10.514
7	270	111.5722	96.2711	Cibicidoides grossa	270	96.2711	96.8538
8	339	-14.7184	207.8433	Globigerina bulloides	4	193.1249	14.7184
9	4	22.9684	193.1249	Globorotalia inflata	339	207.8433	8.25
10	181	3.3666	216.0933	Cibicidoides dutemplei	181	216.0933	3.3666
11	316	14.3813	219.4599	Trifarina fluens	316	219.4599	14.3813
12	159	8.7292	233.8412	Cibicidoides grossa LCO	159	233.8412	8.7293
13	23	73.4921	242.5704	Sigmoilopsis schlumbergeri	23	242.5704	24.9386
14	269	-10.5504	316.0625	Neogloboquadrina atlantica	330	267.509	38.0031
15	266	-38.0031	305.5121	Globorotalia puncticulata	266	305.5121	10.5504
16	330	121.7261	267.509	Cibicidoides pachyderma	269	316.0625	73.068
17	219	-0.1045	389.2351	Martinotiella cylindrica	130	389.1306	0.1046
18	130	78.884	389.1306	Siphonina advena	219	389.2351	5.1794
19	285	-73.6	468.0146	Caucasina elongata	91	394.4146	14.8123
20	91	99.5081	394.4146	Diatoms/radiolarians LCO	71	409.2268	34.8408
21	282	-22.4523	493.9227	Uvigerina ex.gr. semiornata	125	444.0677	23.9469
22	123	4.4201	471.4704	Globigerinoides trilobus	285	468.0146	3.4558
23	207	-31.8228	475.8905	NS Log G	123	471.4704	4.4201
24	125	-34.8409	444.0677	Neogloboquadrina continuosa	207	475.8905	5.4888
25	71	90.1595	409.2268	Epistomina elegans	236	481.3793	12.5434
26	15	38.4634	499.3863	Globigerina praebulloides	282	493.9227	5.4636
27	18	-56.4704	537.8497	Spiroplectammina carinata	15	499.3863	24.1768
28	236	42.1838	481.3793	G. ex.gr. praescitula zealandica	17	523.5631	13.6108
29	17	13.6109	523.5631	Asterigerina gurichi	138	537.174	0.6757
30	138	17.7774	537.174	Tenuitella angustiumbilicata	18	537.8497	17.1017
31	111	22.0052	554.9514	Globigerina ex.gr. officinalis	111	554.9514	2.7424
32	413	-19.2628	576.9566	Silicious biofacies	20	557.6938	19.2628
33	20	80.8032	557.6938	Gyroidina girardana	413	576.9566	61.5404
34	24	53.226	638.497	Turrilina alsatica	24	638.497	53.226
35	142	18.1402	691.723	Gyroidina soldanii mamilligera	142	691.723	18.1402
36	25	51.886	709.8632	Coarse agglutinated spp.	25	709.8632	51.886
37	97	41.2097	761.7492	Cyclammina placenta	97	761.7492	41.2097
38	182	53.3608	802.9589	Spirosigmoilinella compressa	182	802.9589	53.3608
39	262	24.8575	856.3197	Karrerulina horrida	262	856.3197	24.8575
40	259	-881.1772	881.1772	Ammodiscus latus	259	881.1772	0.465

The preceding estimate (-56.47) of the WADD between events 27 and 28 is negative. This implies that, whereas 28 was stratigraphically below 27 in the ranked optimum sequence used as input, it ends up above 27 after depth scaling when the window is used. Table 5 shows weighted average depth differences for the first 30 events of the ns30cen example. The estimate derived for example (-56.47) can be found in column 3 of Table 5, in row 27 for *Spiroplectammina carinata* (18).

In order to eliminate the occurrence of negative distances, the events can be re-ordered following the same procedure as used in RASC scaling. Cumulative WADD values are measured from an origin with 0 value set at the event with highest stratigraphic position (see column 4 in Table 5). These cumulative distances are reordered and new interevent distances are determined by taking successive differences between re-ordered cumulative values as shown in the last three columns of Table 5. These final interevent distances (in meters) are graphically displayed in the depth scaling dendrogram (*.dn2 file).

The choice of CASC input parameters (Fig. 4) used in this experiment is probably not the best for achieving optimum results. Revisions that can be made are:

1. Width of Depth Filter (set at 400m) can be changed. Note that this only affects results based on depths in meters (as in the experiment). If depth differences are square root transformed, all depth differences will be used.
2. Order of Window (>1) can be changed.
3. Square root transformation can be applied.
4. Order of events in optimum sequence can be changed from order in ranked optimum sequence to order in scaled optimum sequence.

Comparison of final results of the experiment (see Table 5) to NSR (North Sea RASC) zones based on the RASC scaled optimum sequence represented in dendrogram form (Kaminski and Gradstein, 2005, Figure 18, p 50) shows that Event 24 (last occurrence of *Turrilina alsatica*) is probably out of place. It is classified as Early Oligocene (NSR7B) in Kaminski and Gradstein (2005) but its neighbors in Table 5 are Late Oligocene (NSR8A). This is further illustrated in Table 6 (left hand side).

The experiment was run again after changing the menu of Figure 4 by activating both yes-no buttons on the right: Square Root Transformation (Yes instead of No) and Input Sequence as in Scaled Optimum Sequence (Yes instead of No). Results for this modification are shown on the right hand side of Table 6 (first 40 events only).

An advantage of using the square root transformation is that this normalizes low order depth difference distributions (see Agterberg, Gradstein and Liu, 2007) thus reducing abnormally strong influence exercised by the largest depth differences (both positive and negative) depth differences. However, interevent distances are not measured in meters but in √m. Of course, the square root transformation does not prevent used from constructing a dendrogram. In RASC scaling dendrograms, interevent distances are measured along a *Z*-scale that is more difficult to interpret than the √m scale.

Table 6. Comparison of dendrogram results on right hand side of Table 5 with NSR zones of Kaminski and Gradstein (2005) and results obtained by means of new experiment described in the text. Note downward move in position of Event 24 that occurred stratigraphically too high (**) on left hand side of Table 6.

	Event #	IED (m)	Event Name	Zone	Event #	IED ($m^{0.5}$)	Zone
1	77	15.4	Elphidium spp.	NSR13	77	2.4561	NRS13
2	228	28.5714	Cassidulina teretis	NSR13	228	3.0848	NSR13
3	338	27.1717	Bulimina marginata		338	1.7601	NSR12
4	171	7.114	Buccella frigida	NSR12	171	1.4397	NSR12
5	1	7.5	Neogloboquadrina pachyderma	NSR12	31	1.074	
6	31	10.514	Cibicidoides scaldisensis		270	0.4225	
7	270	96.8538	Cibicidoides grossa	NSR12	1	7.3983	NSR12
8	4	14.7184	Globorotalia inflata	NSR12	4	0.059	NSR12
9	339	8.25	Globigerina bulloides		181	0.06	
10	181	3.3666	Cibicidoides dutemplei		316	0.6834	
11	316	14.3813	Trifarina fluens		339	0.1306	
12	159	8.7293	Cibicidoides grossa LCO		159	0.5433	
13	23	24.9386	Sigmoilopsis schlumbergeri	NSR12	23	3.0787	NSR12
14	330	38.0031	Cibicidoides pachyderma		330	2.298	
15	266	10.5504	Globorotalia puncticulata	NSR11	269	1.7406	
16	269	73.068	Neogloboquadrina atlantica		266	6.7681	NSR11
17	130	0.1046	Siphonina advena	NSR9B	130	0.2865	NSR9B
18	219	5.1794	Martinotiella cylindrica	NSR9B	219	2.0738	NSR9B
19	91	14.8123	Diatoms/radiolarians LCO	NSR9B	91	1.7821	NSR9B
20	71	34.8408	Epistomina elegans		125	1.2834	NSR9A
21	125	23.9469	Neogloboquadrina continuosa	NSR9A	71	0.8642	
22	285	3.4558	Caucasina elongata	NSR9A	282	0.1476	NSR9A
23	123	4.4201	Globigerinoides trilobus	NSR9A	285	0.4237	NSR9A
24	207	5.4888	NS Log G		236	0.0025	NSR9A
25	236	12.5434	G. ex.gr. praescitula zealandica	NSR9A	207	0.6652	
26	282	5.4636	Uvigerina ex.gr. semiornata	NSR9A	123	1.4817	NSR9A
27	15	24.1768	Globigerina praebulloides	NSR9A	17	0.7504	NSR9A
28	17	13.6108	Asterigerina gurichi	NSR9A	15	1.1588	NSR9A
29	138	0.6757	Tenuitella angustiumbilicata	NSR8A	413	0.2326	
30	18	17.1017	Spiroplectammina carinata	NSR8A	18	2.6293	NSR8A
31	111	2.7424	Globigerina ex.gr. officinalis	NSR8A	20	4.2542	NSR8A
32	20	19.2628	Gyroidina girardana	NSR8A	25	0.1701	NSR8A
33	413	61.5404	Silicious biofacies		138	1.5962	NSR8A
34	24	53.226	Turrilina alsatica	NSR7B**	142	1.3286	
35	142	18.1402	Gyroidina soldanii mamilligera		97	1.7649	NSR8A
36	25	51.886	Coarse agglutinated spp.	NSR8A	111	0.7507	NSR8A
37	97	41.2097	Cyclammina placenta	NSR8A	182	0.166	NSR8A
38	182	53.3608	Spirosigmoilinella compressa	NSR8A	259	0.4668	NSR7B
39	262	24.8575	Karrerulina horrida	NSR7B	183	2.1813	NSR7B
40	259	0.465	Ammodiscus latus	NSR7B	24	1.8059	NSR7B

References

Agterberg, F.P., Gradstein, F.M., and Liu, G., 2007, Frequency distributions of thickness of sediments bounded by Cenozoic biostratigraphic events in wells drilled offshore Norway and along the northwestern Atlantic margin: *Natural Resources Research*, Vol. 16, No. 3 (in press).

Kaminski, M.A., and Gradstein, F.M., 2005, Atlas of Paleogene cosmopolitan deep-water agglutinated Foraminifera: Grzybowski Foundation Spec. Publication No. 10, Drukarnia Narodowa, Kraków, 548 p.