

Compos Analysis Version 5.1 User's Guide

User's Guide version 5.1.0, June 2004 to Compos Analysis 5.1 standard and 5.1 plus, 08 June 2004.

Smith Ecology Microsoft® Excel tool for Compositional Analysis

Author: Dr. P. G. Smith, Smith Ecology Ltd.

1 Bettws Cottage, Bettws, Abergavenny, Monmouthshire, NP7 7LG, Wales, UK

E-mail: sales@smithecolony.com

Internet: <http://www.smithecolony.com/>

Copyright © 2004, Smith Ecology Ltd., 1 Bettws Cottage, Bettws, Abergavenny, Monmouthshire, NP7 7LG, Wales, UK

References to the Compos Analysis software:

Smith, P. G. 2003. Compos Analysis, version 5.1 standard [software]. Smith Ecology Ltd., Abergavenny, UK.

Smith, P. G. 2003. Compos Analysis, version 5.1 plus [software]. Smith Ecology Ltd., Abergavenny, UK.

Reference to published paper introducing the software:

Smith, P. G. 2004. Automated log-ratio analysis of compositional data: software suited to analysis of habitat preference from radiotracking data. *Bat Research News*, **45**: 16.

This user guide is available on the World Wide Web and can be cited as follows:

Smith, P. G. 2004. *Compos Analysis version 5.1 user's guide*. Version 5.1.0. Smith Ecology Ltd., 1, Bettws Cottage, Bettws, Abergavenny, NP7 7LG, UK. i + 18 pp. [WWW document]. URL <http://www.smithecolony.com/software.htm>

CONTENTS

QUICK START	1
INTRODUCTION.....	2
SYSTEM REQUIREMENTS.....	2
BACKGROUND.....	3
REVISION HISTORY.....	3
COMPOS ANALYSIS STANDARD	3
COMPOS ANALYSIS PLUS.....	3
INSTALLATION	3
TROUBLESHOOTING DURING INSTALLATION	4
PC.....	4
APPLE MAC.....	5
HOW TO USE THE COMPOSITIONAL ANALYSIS TOOL	6
FORMAT OF DATA IN “USED” AND “AVAIL” WORKSHEETS	6
ZEROS IN “USED” OR “AVAIL” WORKSHEETS	6
CAUTION! – TREATMENT OF ARTIFICIALLY SMALL VALUES IN “USED” OR “AVAIL” WORKSHEETS	7
STARTING THE ANALYSIS	7
SPECIFYING THE NUMBER OF ITERATIONS YOU REQUIRE	7
SPECIFYING A REPLACEMENT VALUE FOR ZEROS IN “USED” VALUES	8
PROGRESS INDICATORS	8
DATA THAT IS OUTPUT	8
WORKSHEETS HOLDING EXAMPLE DATA FROM AEBISCHER ET AL. (1993)	9
SAVING YOUR RESULTS	9
SAMPLE SIZES.....	9
REFERENCES.....	9
ACKNOWLEDGEMENTS.....	10
APPENDIX 1: WORKED EXAMPLES	11

User's Guide for Compos Analysis 5.1 standard¹ and 5.1 plus (08 June 2004): a Smith Ecology Microsoft® Excel tool for Compositional Analysis.

QUICK START

A. Copy the files to a folder and check that the program works.

1. Copy the Excel workbook file "Compositional Analysis 5_1 std.xls" or "Compositional Analysis 5_1 plus.xls" and the Excel Add-In file "CompAnal 5_1 std.xla" or "CompAnal 5_1 plus.xla" into a folder on your PC or Mac. [*Check that the copies have the same filenames as the originals as just stated (i.e. ensure that filenames have not been cropped during copying). Note that if your computer is set to hide file extensions you will not see the ".xls" or ".xla" parts of the filenames. In this case do not add ".xls" or ".xla" to the names when these extensions are already there, but hidden from view.*]
2. Start Excel and open the workbook "Compositional Analysis 5_1 std (or plus)". (See item 4 under installation on page 3 below if you get a message about macros.)
3. Click once on the tools dropdown menu and click once on "Compos Analysis" at the bottom of the menu.
4. You should see a message box appear headed Compos Analysis and stating that this is version 5.1. (If you get any other response see "trouble shooting during installation" on page 4.) Click "Cancel" and then "OK" on the next "Terminating" message box to exit from the program. [*If you click "OK" instead of cancel on the first message box you will get a message box stating that you have "Insufficient data". Clicking "OK" on this message box (the only option here) causes the program to exit.*]
5. This confirms that everything is set up correctly and the program works. If you got any errors see trouble shooting below.

B. Try running the program with some example data.

1. When the "Compositional Analysis 5_1 std (or plus)" workbook first opens the "TestResults" worksheet is selected. Now select the "Ph" worksheet. This contains some data that relate to examples from a pheasant study described by Aebischer *et al.* (1993).
2. Select and copy cells A17 to F31.
3. Select the "Used" worksheet and paste the selected data here starting from cell A1.
4. Select the "Ph" worksheet. Scroll to cells A33 to F47 and select and copy this block of data.
5. Select the "Avail" worksheet and paste the selected data here starting from cell A1.
6. Click once on the tools dropdown menu and click once on "Compos Analysis" or "Compos Analysis +" at the bottom of the menu.
7. You will see the "Compos Analysis" headed message box stating that this is version 5.1. Click "OK".
8. You will see a message box headed "Compos Analysis: Input", stating "Cases = 13, Variables = 5" and inviting you to enter the number of iterations you require for randomization or to accept the default of "1000". Click "OK".

¹ Compos Analysis standard is set to handle up to 200 cases (e.g. animals) for up to 25 variables (e.g. habitat categories), while Compos Analysis plus is set to handle up to 2000 cases for up to 30 variables. A version with increased limits can be supplied if required.

9. You will see a further “Compos Analysis: Input” message box. This one informs you that “Some ‘used’ values are zero.” and invites you to enter a value to be used in place of zero or to accept the recommended default value of “0.01”. Click “OK”.
10. The “TestResults” worksheet will be automatically selected and the analysis will proceed to completion. During this time progress will be indicated on the Status bar below the worksheet tabs and on the worksheet. The date and time the analysis started will be shown on row 18 of the worksheet until the worksheet is rewritten with the final results.
11. Main results are output to the “TestResults” worksheet. Intermediate results are output to the worksheets “WilksValues”, “ObsRawSSCP”, “ObsMeanCorrectedSSCP”, “ObsLRDs”, “AvailRatios” and “UsedRatios”. To save the results, either copy “TestResults” and any other worksheets you wish to save to a different workbook or save the current workbook using a different name. [N.B. In the latter case you may find multiple entries of “Compos Analysis” in the Tools Menu if you have more than one copy of the original workbook open simultaneously.]
12. You will observe that the results agree with those of the first pheasant example presented by Aebischer *et al.* (1993). The results from this example are presented as Example 1 below. When you run “Compos Analysis” again, all results from the previous analysis will be cleared once you select “OK” at the first message box (i.e. the one where the version number is given). If you realise you meant to save data from the previous run first then select “Cancel” at this message box to exit from the macro without clearing any data from the previous analysis.

INTRODUCTION

This tool for Compositional Analysis (Aitchison 1986) with randomization testing was written by Peter G. Smith of Smith Ecology Ltd., 1, Bettws Cottage, Bettws, Abergavenny, Monmouthshire, NP7 7LG, Wales, UK. Telephone: +44 (0) 1873 890055. E-mail peter@smithecolony.com. The tool was first used, in an earlier version, by Smith (2000). The copyright in this software is held by Smith Ecology Ltd. The software must not be copied or distributed without the permission of Smith Ecology Ltd.

This compositional analysis tool requires Microsoft Excel. The tool comprises an application specific Excel Add-In (the file “CompAnal 5_1 std.xla” or “CompAnal 5_1 plus.xla”) and an Excel workbook (the file “Compositional Analysis 5_1 std.xls” or “Compositional Analysis 5_1 plus.xls”). The Add-In is accessed as a library item by code in the Excel workbook to perform the analysis on data you copy into specific worksheets of the workbook.

Compositional Analysis is a technique that uses MANOVA to analyse two sets of data in which variables are represented as proportions. It is used to determine the statistical significance of differences and the rank order of differences between the variables.

SYSTEM REQUIREMENTS

This Add-In tool supports users of (on a PC) Excel Versions 95 (i.e. 7.0), 97, 2000 and 2002 on Windows 95/98 & XP and (on an Apple Mac) Excel Version 10 on OS X or any compatible versions of Excel.

The first of the two pheasant examples presented by Aebischer *et al.* (1993) is likely to take about:

3 seconds on a 2.6 GHz Pentium PC with 1 GB of RAM,

1 minute 13 seconds on a 100 MHz Pentium PC with 16 MB of RAM &

1 minute 3 seconds on an 800 MHz G4 (Power PC) Apple Mac with 512 MB of RAM.

BACKGROUND

This Add-In tool for Microsoft Excel implements fully the method for Compositional Analysis as described by Aebischer *et al.* (1993). The significance of Wilk's lambda and of t-values is determined by randomization tests as recommended by Aebischer *et al.* (1993) to overcome problems arising when the distribution of log-ratio differences is not multivariate normal. To avoid the need to always drop variables (e.g. habitat categories) or cases (e.g. animals) when availability for some of the cases is zero, the method recommended by Aebischer *et al.* (1993) in Appendix 2, option 3 is implemented, i.e. replacement of missing values in a particular residual log-ratio by the mean of all nonmissing values for that log-ratio, and then computation of a mean lambda by weighting each denominator-dependent value of lambda by the number of nonmissing values involved in its calculation, and determination of the level of significance by randomization.

REVISION HISTORY

Compos Analysis standard

Version 5.1

Contact details were updated.

Compos Analysis plus

Version 5.1

This is the first release of a version with higher limits to the number of cases and variables handled compared with the standard product.

INSTALLATION

1. Installation of this software implies your acceptance of the terms of the end-user licence agreement which is held in a ".pdf" file (in Adobe Acrobat portable document format) accompanying this help file. (You will need Adobe Acrobat Reader, which is available free from the Adobe web site to read the file.)
2. Make backup copies of the four files supplied. The files are "CompAn help 5_1_0.pdf" (holding the text of the present User's Guide in Adobe PDF format), "Compan5_1* Licence.pdf" (which holds your end-user licence), "Compositional Analysis 5_1 std.xls" or "Compositional Analysis 5_1 plus.xls" (a Microsoft Excel workbook file) and "CompAnal 5_1 std.xla" or "CompAnal 5_1 plus.xla" (the file holding the software program that performs the compositional analysis).
3. Copy all the above files to a folder on your PC or Apple Mac, e.g. you may choose to use a folder named "CA", or "Compositional Analysis". Especially when copying to a Mac, ensure that the full file name of the Add-In is retained in the copy and, if not, rename it with the full name "CompAnal 5_1 std" or "CompAnal 5_1 plus". The file type of .xla may or may not be displayed on your computer depending on whether or not it is set to hide file types. If it is there, but hidden, do not add it again in the visible part of the name. Check the name of the workbook file too though its name is not critical for the software to work.
4. Open the supplied Excel workbook file in Excel. Depending on your system settings and on your version of Excel, on this first time that you open the workbook you may encounter an option to either disable or enable macros. You should select the option to "Enable Macros". Note that the macro is not digitally signed and in Excel 2002 if your security level is set to high the macro will be disabled as you open the workbook. If this is the case,

change the security level to Medium by pointing to **Macro** on the **Tools** menu, click **Security**, click the **Security Level** tab and set the level to **Medium**. Then close the Compositional Analysis workbook, reopen it and click **Enable Macros** when asked upon loading. If you want Excel to automatically disable unsigned macros in the future, remember to change the security level back to High (though you would have to change back to Medium again to use “Compos Analysis”). Save the workbook in the form appropriate to your version of Excel. The compositional analysis tool is now available by clicking the “Compos Analysis” item on the Tools menu.

5. Now confirm the availability of the compositional analysis tool by clicking on the “Compos Analysis” item on the Tools menu. This should produce a message stating the version of the Compositional Analysis software. You will see this message every time you use the tool. The appearance of this message confirms that the tool is loaded correctly. Click OK. A message will then appear stating that Cases = 0, Variables = 0 and that the macro will terminate. This is normal and is because you have not yet set up any data for analysis as described below. Click OK and the macro will end. This completes the installation and your confirmation that the tool is working.

TROUBLESHOOTING DURING INSTALLATION

PC

If, upon selecting the “Compos Analysis” item on the Tools menu you get a **Microsoft Visual Basic** message - **Compile error: Sub or Function not defined**, or **Compile error: Can't find project or library**, or **Can't find project or library**, click **OK** (first click **Help**, then **OK** if you want instructions available on screen rather than following notes below). If using version 2000 or 2002 of Excel select **Reset** from the **Run** menu. Then, in any version of Excel, as you are now in the Visual Basic Editor rather than the workbook, select **References...** from the **Tools** menu and continue with locating “CompAnal 5_1 std” or “CompAnal 5_1 plus” as described in the appropriate paragraph below or in Excel Help. These instructions are likely to apply to Excel 97 too.

Excel version 7.0 (i.e. Excel 95)

If the compositional analysis tool is not available as described above and you are using version 7.0 of Excel, within the supplied Excel workbook, click the tab for the module named “Module1”, then select “**References...**” from the **Tools** menu. If CompAnal 5_1 is present in the list of references, but is given as “MISSING: CompAnal_5_1.xla” then click on the box to the left of the entry to remove the tick and click OK. Then select “**References...**” from the **Tools** menu again. Now the invalid CompAnal 5_1 entry should have gone from the list of references and you should then add the correct entry as follows. If CompAnal 5_1 is not present in the “Available References” list, click the **Browse** button. This opens a “Browse” window. In the “Files of Type” box, select “Microsoft Excel Files (*.xl*; *.xls; *.xla; ...)”, then navigate to the folder to which you copied it, select it and click the **OK** button. You should see the option appear ready ticked (therefore not given as “MISSING”) in the “Available References” list (possibly at the bottom initially) with the correct path given in the description box below. If it is not ticked then click on it such that a tick appears to the left of the CompAnal 5_1 reference. Click the **OK** button. Return to the “TestResults” worksheet and save the Excel workbook to preserve the change. “Compos Analysis” on the Tools menu should now work correctly. If not, then go back to **References** as described above and check that no CompAnal 5_1 entry given as “MISSING” remains in the list of references. Sometimes more than one attempt is required to remove such references and if not removed the program will continue to look for the program at the invalid location.

Excel 2000 or 2002

If the compositional analysis tool is not available as described above and you are using version 2000 or 2002 of Excel, unless already in the **Visual Basic Editor**, then within the supplied Excel workbook select **Macro** from the **Tools** menu and select **Visual Basic Editor**. Then select **References...** from the **Tools** menu of the Visual Basic Editor. If CompAnal 5_1 is present in the “Available References” list, but is given as “MISSING: CompAnal_5_1.xla” then click on the box to the left of the entry to remove the tick and click OK. Then select “**References...**” from the **Tools** menu again. Now the invalid CompAnal 5_1 entry should have gone from the “Available References” list and you should then add the correct entry as follows. If CompAnal 5_1 is not present in the “Available References” list, click the **Browse** button. This opens an “Add Reference” window. In the “Files of Type” box, select “Microsoft Excel Files (*.xls;*.xla)”, then navigate to the folder to which you copied it, select it and click the **Open** button. You should see the option appear ready ticked (therefore not given as “MISSING”) in the “Available References” list (possibly at the bottom initially) with the correct location given in the box below. If it is not ticked then click on it such that a tick appears to the left of the CompAnal 5_1 reference. Click the **OK** button. Select **Microsoft Excel** from the **View** menu, or **Close and return to Microsoft Excel** from the **File** menu, to return to the workbook. Save the Excel workbook to preserve the change. “Compos Analysis” on the Tools menu should now work correctly. If not, then go back to **References** as described above and check that no CompAnal 5_1 entry given as “MISSING” remains in the “Available References” list. Sometimes more than one attempt is required to remove such references and if not removed the program will continue to look for the program at the invalid location.

Apple Mac

If the compositional analysis tool is not available as described above and you are using Excel X on an Apple Mac, first ensure the Mac is a PowerPC G4 or compatible and that you have installed Mac OS X version 10.1 or later (though Compos Analysis has also been shown to run on a blue & white G3 with Mac OS 9.2.2 and Office 2001 with Service Release 1 for Office installed). Check that the latest upgrades to OS X and Excel X or Office X are installed. You can find these on the internet or contact your supplier. You may have had an error message such as “error loading dll”. If so you need to download the “VBA converter for Office v. X” or the “VBA converter for Excel v. X” from the internet and install it. Go to <http://www.microsoft.com/mac/download> to find this. “Compos Analysis” on the Tools menu should now work correctly.

HOW TO USE THE COMPOSITIONAL ANALYSIS TOOL

First open the supplied Compositional Analysis Excel workbook file within Microsoft Excel. The analysis will use the “Compos Analysis” macro supplied in the file CompAnal 5_1.xla. This is run from within the Compositional Analysis Excel workbook either by clicking the “Compos Analysis” item on the Tools menu, or by selecting the “Compos_Analysis” macro available via the Tools/Macro menu. Compos Analysis standard is set to handle up to 200 cases (e.g. animals) for up to 25 variables (e.g. habitat categories). Compos Analysis plus is set to handle up to 2000 cases for up to 30 variables. A version with increased limits can be supplied if required.

Format of data in “Used” and “Avail” worksheets

On running, the macro accesses “used” and “available” data from the “Used” and “Avail” worksheets within the Compositional Analysis Excel workbook. The data in these worksheets should be set up in a standard format (e.g. see worked examples below), i.e.: Cell A1 - Project name; Cell A2 - Source of data, e.g. study area (SA), minimum convex polygon (MCP), or radio location fix (Fix); Cells in row 2 from B2 - names of variables (e.g. habitat categories) [It is best to limit names to, say, 12 lower case characters for brevity and for clarity in output of final results.]; Cells in column A from A3 down - names of cases, e.g. identity of each animal.

The text entered in cell A1 of the “Used” worksheet will appear at the head of the “TestResults” worksheet. This should be the project title and should ideally not exceed a page width, though the maximum is usually 255 characters. Text indicating what is being compared with what is automatically appended to the project title at the head of the “TestResults” worksheet. This is derived from the A2 cells of the “Used” and “Avail” worksheets. The text ‘(used)’ and ‘(available)’ are appended to the text derived from each of these worksheets respectively, so that it is clear how the source data was set up. Hence, if the project title was given simply as ‘Pheasants’, the “TestResults” header text might read: ‘Pheasants - MCP (used) vs. Study area (available)’, assuming A2 of “Used” says ‘MCP’ and A2 of “Avail” says ‘Study area’.

The data within the “Avail” and “Used” worksheets are percentage data, with each row nominally adding to 100% if you have included all original variables. However, after converting your data to percentages, you can drop variables not required before copying to the “Used” and “Avail” worksheets without the need to recalculate percentages, as ratios of habitat pairs would not be affected [as in the examples given in Aebischer *et al.* (1993) and presented below]. Hence cases may total less than 100% if you have done this.

Zeros in “Used” or “Avail” worksheets

Where any “available” data are zero or missing, the value(s) must be set to zero in the “Avail” worksheet. Where zero is set in the “Avail” worksheet Wilk's lambda is calculated as a weighted mean as recommended in Aebischer *et al.* (1993). However, where any “used” data are zero the figure(s) may either be set to zero in the “Used” worksheet (recommended), or a small percentage can be substituted, e.g. 0.01, to be an order of magnitude less than existing non-zero values in either available or utilized data compositions [see Aebischer *et al.* (1993)]. If “used” zeros are left as zeros, the macro will automatically determine an appropriate value to substitute for these “used” zeros in the calculations and will prompt the user to accept this as a default or to enter their own figure. The program will not replace the zeros on the “Used” worksheet itself but it will state the substitution value used in the report of results and will report the number of “used” zeros that are present and which correspond with nonzero “avail” values and hence were replaced in the calculations. Note that any “used” value that corresponds with a zero on the “avail” worksheet will be treated as missing in the analysis.

Ensure that **ONLY** the relevant data as demonstrated in the examples below is present in the “Avail” and “Used” worksheets so that the macro runs correctly. **DO NOT ADD OTHER DATA IN EXTRA ROWS OR COLUMNS.** Always include the appropriate number of variable (e.g. habitat category) labels in row 2. Case labels are optional and may be entered in column A (e.g. P1 to P13 in column A in the first worked example below, or 2 to 13 in column A in the second example below). Do not add any extra rows or columns of data that are not part of the required compositional analysis. Do not change the names of the worksheets supplied in the Compositional Analysis Excel workbook. TIP: Before entering data in the “Used” and “Avail” worksheets you can clear each worksheet’s contents completely by, within each worksheet in turn, clicking the box at the top left corner of the worksheet (i.e. at the intersection of the column letters and row numbers), right clicking your mouse to get a menu, and left clicking on the “Clear Contents” item in the menu so obtained.

Caution! – Treatment of artificially small values in “Used” or “Avail” worksheets

Sometimes you may derive either “used” or “avail” data from a geographical information system (GIS) by, for example, calculating the total area of each habitat within a Minimum Convex Polygon or other outline that bounds the area deemed to be “used” or “available”. If the boundary line just clips the edge of a habitat that is otherwise not represented within the defined area the total area of that habitat included may be extremely small – perhaps far smaller than the smallest unit recognised during data collection. If retained, such an artificially small value may further influence the data as the replacement value for any zeros in the “used” data is determined by reference to the smallest “used” or “avail” value present and will be an order of magnitude less than this. It is recommended that values that fall below the smallest unit recognised during data collection are either rounded up to the smallest recognised unit or replaced by zero as appropriate before commencing compositional analysis. This will avoid the potential for artificially small values to influence the analysis.

Starting the analysis

Once data is set up in the “Avail” and “Used” worksheets, click the “Compos Analysis” item in the Tools drop down menu. A message will appear stating the version number of the software. Click OK. The worksheets “TestResults”, “WilksValues”, “ObsRawSSCP” (observed raw sums of squares and cross-products matrix), “ObsMeanCorrectedSSCP” (observed mean-corrected sums of squares and cross-products matrix), “ObsLRDs” (observed log-ratio differences), “AvailRatios” and “UsedRatios”, holding output from any previous analysis will be cleared automatically to make way for the new results.

If you meant to save the results of the previous run click “Cancel” instead of “OK” at the first message box and the macro will exit without clearing any results from the previous analysis. If you wish to retain results from your previous analysis you should make a copy of what you require into other worksheet(s) or save them in a different workbook or text file before running the macro again with new data. If data has not been set up, or has not been set up correctly in the “Avail” and “Used” worksheets an error message will appear explaining the problem and the macro will end. The seven results worksheets named above must be available for the macro to output its results for the current analysis.

Specifying the number of iterations you require

Assuming that the data in the “Avail” and “Used” worksheets appears to be OK, a dialogue box will appear asking you to set the number of iterations you require for randomization or to accept the default of 1000. With 1000 iterations (made up of the original data plus 999 permutations of the data) the smallest obtainable level of probability is $P \leq .001$. The default is given as 1000 because Manly (1997) recommends that a minimum of 1000 iterations is used

for tests at a significance level of .05. (TIP: If you are simply checking that the macro is working, then change the value to a small number, such as 10, so that the macro completes in a short time.)

Specifying a replacement value for zeros in “Used” values

Once you respond to the iterations dialogue box the macro will check for zeros among “used” values and, should there be any present, a new dialogue box will appear. This will show a recommended value (determined to be at least an order of magnitude less than the smallest non-zero value present in the “Used” and “Avail” worksheets) to automatically substitute for any zeros in the “used” values (though any “used” value that corresponds with a zero in “avail” will be treated as missing in the analysis). [See also comments under the heading “Caution – Treatment of artificially small values in “Used” or “Avail” worksheets” above.] This default may be accepted or an alternative value may be entered. Click OK to continue. Zeros will not actually be replaced in the “Used” worksheet, though the chosen value will be used as a substitute in the analysis and reported in the results.

Progress indicators

The analysis will now run to completion. The date and time the analysis started will be displayed in cell A18 of the “TestResults” worksheet. An increasing count will show on the status bar and on the screen (in cell B21 of the “TestResults” worksheet) indicating the number of iterations completed as the macro runs. Results will be output to the seven previously cleared worksheets as the macro runs.

Data that is output

The key test results are output to the “TestResults” worksheet. Wilk's lambda values from randomization are sorted into ascending order and the smallest 5% are output to the “WilksValues” worksheet. *** is placed in column B against the value calculated from the observed data, assuming this is within the lowest 5% of values determined by randomization. Matrices of raw and mean-corrected sums of squares of cross-products of log-ratio differences of observed data are output to the “ObsRawSSCP” and “ObsMeanCorrectedSSCP” worksheets respectively for each variable as denominator in turn. (Wilk's lambda is derived as a ratio of the determinants of these matrices.) The observed log-ratio differences (for one set of variables as denominators only) are output to the “ObsLRDs” worksheet. Ratios of each pairing of “available” variables are output to the “AvailRatios” worksheet, while the equivalent ratios of “used” variables are output to the “UsedRatios” worksheet. Clearly log-ratios can be derived easily from these if required. (In the case of large numbers of variables – typically 16 or more - output of some of the intermediate results will be limited according to the number of columns available in the worksheets.)

The sorting of lambda values for output can take some time after the randomization iterations are complete. A message “PLEASE WAIT - sorting in progress” will display during this time and a counter will increment showing the percentage of the sort that is complete. In the “TestResults” worksheet, the simplified matrix of ranks is produced using the significance levels determined from randomization testing, but with the significance levels from t-tests on the OBSERVED data shown in parentheses where this differs.

Variables are presented in rank order above the simplified rank matrix. Variables are separated with > symbols, with those to the left being of higher rank (preference) than those to the right. Where there is a significant difference at $P < .05$ between two consecutively ranked variables, a >>> symbol is used. Where variables are of equal rank they are separated by the = symbol.

Worksheets holding example data from Aebischer et al. (1993)

Four further worksheets are supplied. These are named “DskPh”, “Ph”, “DskSq” and “Sq”. “DskPh” and “DskSq” hold the percentage data for Pheasant and Squirrel studies respectively as presented by Aebischer *et al* (1993). “Ph” and “Sq” are derived from this test data, but are arranged ready to copy appropriate blocks to the “Used” and “Avail” worksheets for test runs. To carry out analyses on the examples given in Aebischer *et al* (1993), simply copy and paste appropriate data from “Ph” or “Sq” to the “Avail” and “Used” worksheets. The results output to the TestResults worksheet for each of these example analyses is presented below. [Note that the simplified ranking matrix results for the second of the two squirrel examples presented by Aebischer *et al.* (1993) in their table 5b, i.e. the Fix v. MCP comparison, is in error. Hence the results differ from those obtained by running the present compositional analysis tool. The error in the paper was due to mislabelling of variables in the original analysis of this example by Aebischer *et al.* (1993). The error does not apply to any of the overall results (lambda or randomization).] A beep will sound on completion of the macro.

Saving your results

Always copy and save results that are required into another file before carrying out another analysis. The macro can be stopped during an analysis at any time by pressing the Escape key.

SAMPLE SIZES

As noted by Aebischer *et al.* (1993), with reference to Siegel (1956), if values of a log-ratio difference all have the same sign, 6 is the minimum sample size required to show a significant difference from zero at $P < .05$ by randomization. Aebischer *et al.* (1993) advise that for a comparison of utilized with available habitats, 6 radio-tagged animals constitute an absolute minimum. They recommend sample sizes above 10, and preferably above 30, to represent a population adequately.

Missing values (zeros) in “Avail” have a direct effect on sample size. The effective sample size of each category of the log-ratio differences is 1 more than the number of degrees of freedom reported in the univariate test results. If the sample size of any one of these categories is less than 6 a warning message is included in the TestResults output on lines 19 and 20.

The program will not accept data with fewer cases than the number of variables as a valid Wilk’s Lambda value could not then be determined.

REFERENCES

- Aebischer, N. J., P. A. Robertson & R. E. Kenward 1993. Compositional analysis of habitat use from animal radio-tracking data. *Ecology*, **74**(5): 1313-1325.
- Aitchison, J. 1986. The statistical analysis of compositional data. Chapman and Hall, London, England, 416 pp.
- Manly, B. F. J. 1997. *Randomization, bootstrap and Monte Carlo methods in biology*. Second edition. Texts in Statistical Science. Chapman and Hall, London.
- Siegel, S. 1956. *Nonparametric statistics for the behavioural sciences*. McGraw-Hill, New York, New York, USA.
- Smith, P. G. 2000. *Habitat preference, range use and roosting ecology of Natterer’s bats (Myotis nattereri) in a grassland-woodland landscape*. PhD thesis, University of Aberdeen.
- Smith, P. G. 2004. Automated log-ratio analysis of compositional data: software suited to analysis of habitat preference from radiotracking data. *Bat Research News*, **45**(1): 16.

ACKNOWLEDGEMENTS

The development of this program has especially benefited from the helpful comments and suggestions of Dr. N.J. Aebischer of the Game Conservancy, Fordingbridge, UK and also from the feedback of various users.

APPENDIX 1: WORKED EXAMPLES

WORKED EXAMPLE 1 (Pheasants) [taken from an example presented by Aebischer *et al.* (1993)]

The zeros in “Used” should be replaced with small values >0 , such as 0.01, see above. This will be done automatically by the program if you leave them as zeros (see explanation above).

“Used” Worksheet

	A	B	C	D	E	F
1	Pheasants					
2	MCP	scrub	broad-leaf	conifer	grass	
3	P1	19.6	1.97	0	0	
4	P2	20.58	14.7	0	0.4	
5	P3	8.32	5.73	0	58.9	
6	P4	8.78	23.86	0	12.24	
7	P5	4.41	30.29	10.32	54.98	
8	P6	5.43	30.52	10.11	53.94	
9	P7	4.34	31.24	11.02	53.39	
10	P8	8.66	38.71	0	52.61	
11	P9	9.64	53.16	0	37.17	
12	P10	11.72	8.82	0	79.44	
13	P11	14.36	32.56	0	53.05	
14	P12	9.7	10.95	0	79.33	
15	P13	5.53	20.08	4.11	70.28	

“Avail” Worksheet

	A	B	C	D	E	F
1	Pheasants					
2	Study area	scrub	broad-leaf	conifer	grass	
3	P1	3.22	9.23	0.75	52.83	
4	P2	3.22	9.23	0.75	52.83	
5	P3	3.22	9.23	0.75	52.83	
6	P4	3.22	9.23	0.75	52.83	
7	P5	3.22	9.23	0.75	52.83	
8	P6	3.22	9.23	0.75	52.83	
9	P7	3.22	9.23	0.75	52.83	
10	P8	3.22	9.23	0.75	52.83	
11	P9	3.22	9.23	0.75	52.83	
12	P10	3.22	9.23	0.75	52.83	
13	P11	3.22	9.23	0.75	52.83	
14	P12	3.22	9.23	0.75	52.83	
15	P13	3.22	9.23	0.75	52.83	

Results generated by Worked Example 1 (Pheasants) as they appear in the TestResults worksheet.

NB: this analysis was run on a 2.6 GHz Pentium PC with 1 GB of RAM.

Pheasants - MCP (used) vs. Study area (available).

Multivariate (Wilk's lambda) test Smith Ecology Compos Analysis 5.1 plus: test run created 08/06/2004 at 23:13:19 Time to complete (h:mm:ss): 0:00:03

Lambda = 0.0849
Chi Sq= 32.0591 d.f.= 4
P = 0.0000 P < .0001
Rand P = 0.0020

None of the 'Avail' values was zero. 19 'Used' values of 0% were replaced by 0.01.

Univariate tests (t-tests)

Numerator	scrub	scrub	scrub	scrub	broad-leaf	broad-leaf	broad-leaf	conifer	conifer	grass
Denominator	broad-leaf	conifer	grass	crop	conifer	grass	crop	grass	crop	crop
t-statistic	1.0905	3.2146	2.4989	6.0746	3.4310	2.8052	4.6793	-1.2032	1.8682	2.5256
d.f.	12	12	12	12	12	12	12	12	12	12
Mean	0.3785	3.2701	2.0965	6.5166	2.8916	1.7181	6.1381	-1.1736	3.2465	4.4200
StErr	0.3471	1.0173	0.8390	1.0728	0.8428	0.6125	1.3118	0.9754	1.7377	1.7501
Obs p value	0.2969	0.0074	0.0280	0.0001	0.0050	0.0159	0.0005	0.2521	0.0863	0.0266
Obs sig level	+	+++	+++	+++	+++	+++	+++	-	+	+++
Rand p value	0.3220	0.0120	0.0020	0.0020	0.0070	0.0040	0.0070	0.2430	0.1100	0.0410
Rand sig level	*	***	***	***	***	***	***	*	*	***
nCases	13									
nVariables	5									
Iterations	1000									

Matrix of means and standard errors

	scrub	broad-leaf	conifer	grass	crop
scrub		0.378 ± 0.347	3.270 ± 1.017	2.097 ± 0.839	6.517 ± 1.073
broad-leaf	-0.378 ± 0.347		2.892 ± 0.843	1.718 ± 0.612	6.138 ± 1.312
conifer	-3.270 ± 1.017	-2.892 ± 0.843		-1.174 ± 0.975	3.246 ± 1.738
grass	-2.097 ± 0.839	-1.718 ± 0.612	1.174 ± 0.975		4.420 ± 1.750
crop	-6.517 ± 1.073	-6.138 ± 1.312	-3.246 ± 1.738	-4.420 ± 1.750	

Ranked variable sequence (most to least used): **scrub>broad-leaf>>>grass>conifer>crop**

>>> denotes a significant difference between two consecutively ranked variables

Simplified ranking matrix

	scrub	broad-leaf	conifer	grass	crop	Rank
scrub		+	+++	+++	+++	4
broad-leaf	-		+++	+++	+++	3
conifer	---	---		-	+	1
grass	---	---	+		+++	2
crop	---	---	-	---		0

WORKED EXAMPLE 2 (Pheasants) [taken from an example presented by Aebischer *et al.* (1993)]

The zeros in “Used” should be replaced with small values >0 , such as 0.01, see above. This will be done automatically by the program if you leave them as zeros (see explanation above).

“Used” Worksheet

	A	B	C	D	E
1	Pheasants				
2	Fix	scrub	broad-leaf	grass	
3	2	57	25.5	0	
4	3	47.6	14.3	4.8	
5	4	19.6	73.9	6.5	
6	5	6.7	45	8.3	
7	6	6.7	90	0	
8	7	0	53.3	0	
9	8	16.7	80	3.3	
10	9	6.7	78.3	15	
11	10	40.7	14.8	44.5	
12	11	39.3	14.3	46.4	
13	12	26.9	9.6	63.5	
14	13	13	37	22.2	

“Avail” Worksheet

	A	B	C	D	E
1	Pheasants				
2	MCP	scrub	broad-leaf	grass	
3	2	20.58	14.7	0.4	
4	3	8.32	5.73	58.9	
5	4	8.78	23.86	12.24	
6	5	4.41	30.29	54.98	
7	6	5.43	30.52	53.94	
8	7	4.34	31.24	53.39	
9	8	8.66	38.71	52.61	
10	9	9.64	53.16	37.17	
11	10	11.72	8.82	79.44	
12	11	14.36	32.56	53.05	
13	12	9.7	10.95	79.33	
14	13	5.53	20.08	70.28	

Results generated by Worked Example 2 (Pheasants) as they appear in the TestResults worksheet.

NB: this analysis was run on a 2.6 GHz Pentium PC with 1 GB of RAM.

Pheasants - Fix (used) vs. MCP (available).

Multivariate (Wilk's lambda) test

Smith Ecology Compos Analysis 5.1 plus: test run created
08/06/2004 at 21:40:18 Time to complete (h:mm:ss):
0:00:02

Lambda = 0.3656
Chi Sq= 12.0755 d.f.= 2
P = 0.0024 P < .01
Rand P = 0.0030

None of the 'Avail' values was zero. 4 'Used' values of 0% were replaced by 0.01.

Univariate tests (t-tests)

	scrub	scrub	broad-leaf
Numerator	scrub	scrub	broad-leaf
Denominator	broad-leaf	grass	grass
t-statistic	-0.4527	4.3687	3.3401
d.f.	11	11	11
Mean	-0.2795	2.8501	3.1296
StErr	0.6174	0.6524	0.9370
Obs p value	0.6596	0.0011	0.0066
Obs sig level	-	+++	+++
Rand p value	0.8710	0.0020	0.0040
Rand sig level	*	***	***
nCases	12		
nVariables	3		
Iterations	1000		

Matrix of means and standard errors

	scrub	broad-leaf	grass
scrub		-0.280 ± 0.617	2.850 ± 0.652
broad-leaf	0.280 ± 0.617		3.130 ± 0.937
grass	-2.850 ± 0.652	-3.130 ± 0.937	

Ranked variable sequence (most to least used):

broad-leaf>scrub>>>grass

>>> denotes a significant difference between two consecutively ranked variables

Simplified ranking matrix

	scrub	broad-leaf	grass	Rank
scrub		-	+++	1
broad-leaf	+		+++	2
grass	---	---		0

WORKED EXAMPLE 3 (Squirrels) [taken from an example presented by Aebischer *et al.* (1993)]

The zeros in “Used” should be replaced with small values >0 , such as 0.01, see above. This will be done automatically by the program if you leave them as zeros (see explanation above).

“Used” Worksheet

	A	B	C	D	E	F	G
1	Squirrel						
2	MCP	young	<i>Thuja</i>	larch	mature	open	
3	1	0	0	37.85	61.53	0.6	
4	2	60.01	0	1.88	35.53	2.57	
5	3	88.82	0	0	3.56	7.61	
6	4	19.73	0	29.49	41.6	9.18	
7	5	5.78	21.84	0	71.1	1.27	
8	6	24.02	0	7.88	35.99	32.1	
9	7	37.63	0.4	4.52	16.76	40.69	
10	8	62.57	0	0	30.4	7.01	
11	9	0.1	0	12.54	74.82	12.54	
12	10	30.32	0	10.2	35.86	23.61	
13	11	0	0	25.64	62.76	11.58	
14	12	0	0	27.55	67.52	4.91	
15	13	61.75	0	0	11.17	27.05	
16	14	0	0	33.99	61.14	4.86	
17	15	8.47	0	18.52	56.05	16.94	
18	16	0	0	18.83	50.62	30.53	
19	17	52.56	0	2.56	41.52	3.35	

“Avail” Worksheet

	A	B	C	D	E	F	G
1	Squirrel						
2	Study area	young	<i>Thuja</i>	larch	mature	open	
3	1	7.35	4.77	1.83	31.7	54.35	
4	2	7.35	4.77	1.83	31.7	54.35	
5	3	7.35	4.77	1.83	31.7	54.35	
6	4	7.35	4.77	1.83	31.7	54.35	
7	5	7.35	4.77	1.83	31.7	54.35	
8	6	7.35	4.77	1.83	31.7	54.35	
9	7	7.35	4.77	1.83	31.7	54.35	
10	8	7.35	4.77	1.83	31.7	54.35	
11	9	7.35	4.77	1.83	31.7	54.35	
12	10	7.35	4.77	1.83	31.7	54.35	
13	11	7.35	4.77	1.83	31.7	54.35	
14	12	7.35	4.77	1.83	31.7	54.35	
15	13	7.35	4.77	1.83	31.7	54.35	
16	14	7.35	4.77	1.83	31.7	54.35	
17	15	7.35	4.77	1.83	31.7	54.35	
18	16	7.35	4.77	1.83	31.7	54.35	
19	17	7.35	4.77	1.83	31.7	54.35	

Results generated by Worked Example 3 (Squirrels) as they appear in the TestResults worksheet.

NB: this analysis was run on a 2.6 GHz Pentium PC with 1 GB of RAM.

Squirrel - MCP (used) vs. Study area (available).

Multivariate (Wilk's lambda) test Smith Ecology Compos Analysis 5.1 plus: test run created 08/06/2004 at 23:38:55 Time to complete (h:mm:ss): 0:00:03

Lambda = 0.0681
Chi Sq= 45.6791 d.f.= 4
P = 0.0000 P < .0001
Rand P<= 0.0010

None of the 'Avail' values was zero. 24 'Used' values of 0% were replaced by 0.01.

Univariate tests (t-tests)

	young	young	young	young	Thuja	Thuja	Thuja	larch	larch	mature
Numerator	Thuja	larch	mature	open	larch	mature	open	mature	open	open
t-statistic	4.2712	-0.9797	-1.3137	0.6685	-5.3675	-11.0729	-5.8376	0.1541	2.6133	5.1572
d.f.	16	16	16	16	16	16	16	16	16	16
Mean	4.2427	-1.5061	-1.4003	0.6159	-5.7488	-5.6430	-3.6268	0.1058	2.1220	2.0162
StErr	0.9933	1.5373	1.0659	0.9213	1.0710	0.5096	0.6213	0.6867	0.8120	0.3909
Obs p value	0.0006	0.3418	0.2075	0.5133	0.0001	0.0000	0.0000	0.8795	0.0188	0.0001
Obs sig level	+++	-	-	+	---	---	---	+	+++	+++
Rand p value	0.0010	0.3460	0.2140	0.5110	0.0010	0.0010	0.0010	0.8750	0.0190	0.0010
Rand sig level	***	*	*	*	***	***	***	*	***	***
nCases	17									
nVariables	5									
Iterations	1000									

Matrix of means and standard errors

	young	Thuja	larch	mature	open
young		4.243 ± 0.993	-1.506 ± 1.537	-1.400 ± 1.066	0.616 ± 0.921
Thuja	-4.243 ± 0.993		-5.749 ± 1.071	-5.643 ± 0.510	-3.627 ± 0.621
larch	1.506 ± 1.537	5.749 ± 1.071		0.106 ± 0.687	2.122 ± 0.812
mature	1.400 ± 1.066	5.643 ± 0.510	-0.106 ± 0.687		2.016 ± 0.391
open	-0.616 ± 0.921	3.627 ± 0.621	-2.122 ± 0.812	-2.016 ± 0.391	

Ranked variable sequence (most to least used): **larch>mature>young>open>>>Thuja**

>>> denotes a significant difference between two consecutively ranked variables

Simplified ranking matrix

	young	Thuja	larch	mature	open	Rank
young		+++	-	-	+	2
Thuja	---		---	---	---	0
larch	+	+++		+	+++	4
mature	+	+++	-		+++	3
open	-	+++	---	---		1

WORKED EXAMPLE 4 (Squirrels) [taken from an example presented by Aebischer *et al.* (1993)]

The zeros in “Used” should be replaced with small values >0 , such as 0.01, see above. This will be done automatically by the program if you leave them as zeros (see explanation above).

“Used” Worksheet

	A	B	C	D	E	F
1	Pheasants					
2	Fix	young	larch	mature	open	
3	1	0	29.55	70.45	0	
4	2	31.03	0	65.52	3.45	
5	3	70.59	0	5.88	23.53	
6	4	4.76	16.67	54.76	23.81	
7	5	3.57	0	89.29	3.57	
8	6	27.59	1.72	56.9	13.79	
9	7	18.52	0	70.37	11.11	
10	8	50	0	39.29	10.71	
11	9	0	12.5	62.5	25	
12	10	26.67	0	56.67	16.67	
13	11	0	20	70	10	
14	12	0	13.79	82.76	3.45	
15	13	86.67	0	10	3.33	
16	14	0	14.58	77.08	8.33	
17	15	6.9	10.34	67.24	15.52	
18	16	0	41.67	55	3.33	
19	17	32.61	4.35	47.83	15.22	

“Avail” Worksheet

	A	B	C	D	E	F
1	Pheasants					
2	MCP	young	larch	mature	open	
3	1	0	37.85	61.53	0.6	
4	2	60.01	1.88	35.53	2.57	
5	3	88.82	0	3.56	7.61	
6	4	19.73	29.49	41.6	9.18	
7	5	5.78	0	71.1	1.27	
8	6	24.02	7.88	35.99	32.1	
9	7	37.63	4.52	16.76	40.69	
10	8	62.57	0	30.4	7.01	
11	9	0.1	12.54	74.82	12.54	
12	10	30.32	10.2	35.86	23.61	
13	11	0	25.64	62.76	11.58	
14	12	0	27.55	67.52	4.91	
15	13	61.75	0	11.17	27.05	
16	14	0	33.99	61.14	4.86	
17	15	8.47	18.52	56.05	16.94	
18	16	0	18.83	50.62	30.53	
19	17	52.56	2.56	41.52	3.35	

Results generated by Worked Example 4 (Squirrels) as they appear in the TestResults worksheet.

NB: this analysis was run on a 2.6 GHz Pentium PC with 1 GB of RAM.

Squirrel - Fix (used) vs. MCP (available).

Multivariate (Wilk's lambda) test

Smith Ecology Compos Analysis 5.1 plus: test run created 08/06/2004 at 22:18:58 Time to complete (h:mm:ss): 0:00:07

Weighted mean lambda = 0.3445 0.2517 0.5533 0.2874 0.3034

(Chi Sq= 18.1148 d.f.= 3 Warning: weighted mean lambda is most unlikely to conform to standard Chi Sq distribution.)

(P = 0.0004 P < .001)

Rand P = 0.0200

9 'Avail' values were zero, so treated as missing. 5 'Used' values of 0% where corresponding 'Avail' was nonzero were replaced by 0.01.

Univariate tests (t-tests)

	young	young	young	larch	larch	mature
Numerator	young	young	young	larch	larch	mature
Denominator	larch	mature	open	mature	open	open
t-statistic	1.5386	-3.9418	-1.4408	-2.4477	-1.5103	1.6062
d.f.	7	11	11	12	12	16
Mean	1.8326	-0.8848	-0.6438	-1.9854	-1.2514	0.5839
StErr	1.1911	0.2245	0.4469	0.8111	0.8285	0.3635
Obs p value	0.1678	0.0023	0.1775	0.0307	0.1568	0.1278
Obs sig level	+	---	-	---	-	+
Rand p value	0.2030	0.0040	0.1890	0.0120	0.1610	0.1470
Rand sig level	*	***	*	***	*	*
nCases	17					
nVariables	4					
Iterations	1000					

Matrix of means and standard errors

	young	larch	mature	open
young		1.833 ± 1.191	-0.885 ± 0.224	-0.644 ± 0.447
larch	-1.833 ± 1.191		-1.985 ± 0.811	-1.251 ± 0.829
mature	0.885 ± 0.224	1.985 ± 0.811		0.584 ± 0.364
open	0.644 ± 0.447	1.251 ± 0.829	-0.584 ± 0.364	

Ranked variable sequence (most to least used):

mature>open>young>larch

>>> denotes a significant difference between two consecutively ranked variables

Simplified ranking matrix

	young	larch	mature	open	Rank
young		+	---	-	1
larch	-		---	-	0
mature	+++	+++		+	3
open	+	+	-		2