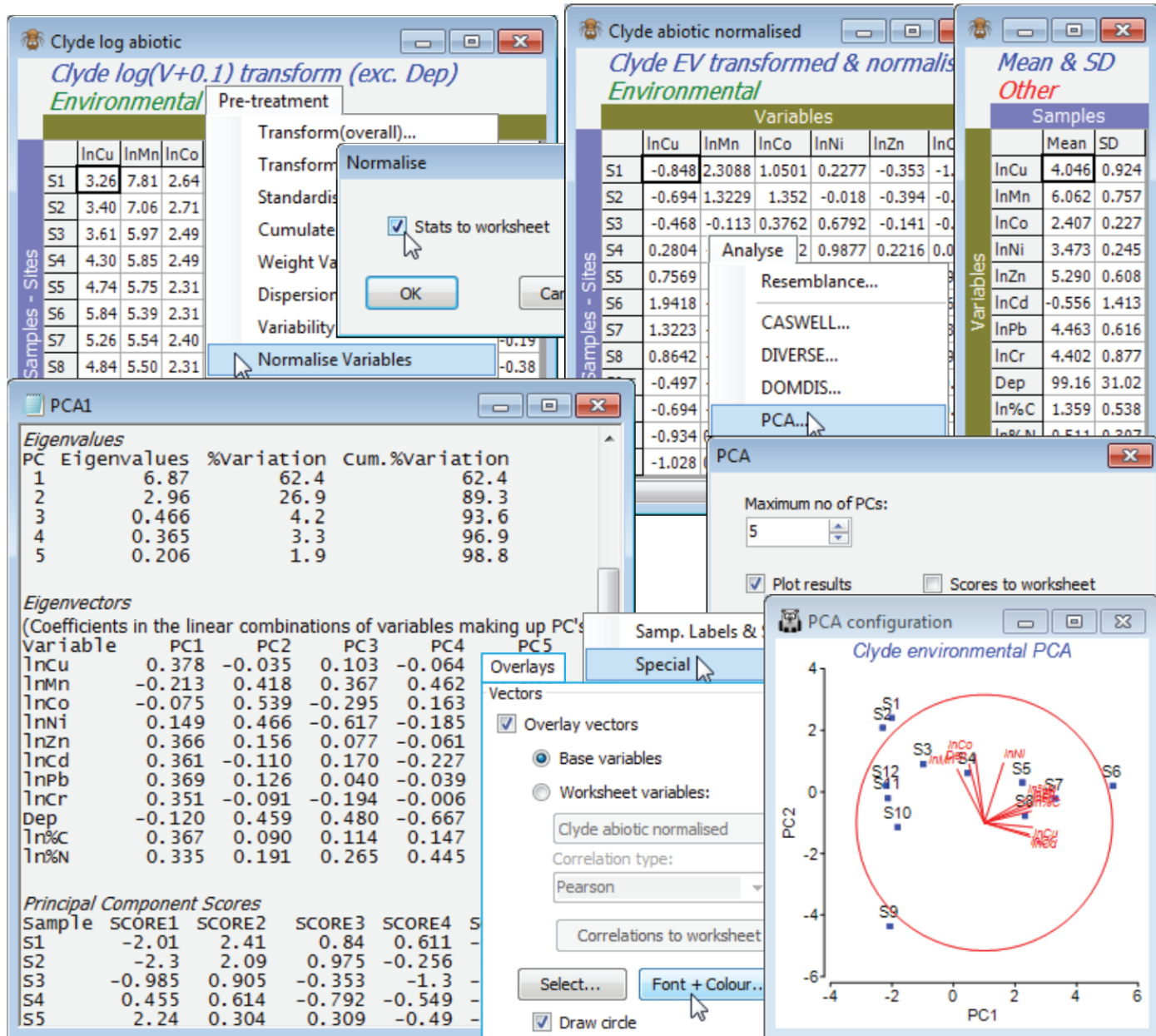# Principal Components Analysis

PCA is an ordination method in which samples, regarded as points in the high-dimensional variable space (11-d here) are projected onto a best-fitting plane, or other low-dimensional solution – the user can specify how many principal components (new axes) are required, and the routine offers 2 d and 3-d plots of any combination of these PC's. The purpose of the new axes is to capture as much of the variability in the original space as possible, and the extent to which the first few PC's allow an accurate representation of the true relationship between the samples in the original high-d space is summarised by the *% variance explained* (a percentage from *eigenvalues*). The PC's are simply a rotation of the original axes and thus a linear combination of the input variables (the coefficients are termed *eigenvectors*); PRIMER allows for superimposition of these vectors on the 2-d PCA plot. The co-ordinates of the samples on the PC axes are called the *principal component scores*, and these are output to the results, along with the %variance explained by each axis and the linear coefficients defining each PC. Chapter 4 of CiMC has a little more detail.

For the Clyde log abiotic data sheet used above, which resulted from a log($0.1+x$) transform of all the environmental variables except water depth (*Dep*), take **Pre-treatment>Normalise Variables**, sending the mean and standard deviation for each of these (transformed) variables to a worksheet, and renaming the resulting data matrix Clyde abiotic normalised. On this sheet, run **Analyse>PCA**, choosing the (default) option of displaying only the first 5 PC axes, and resulting in two outputs: a detailed results window with three sections (*Eigenvalues*, *Eigenvectors* and *Scores*), and a PCA ordination with a superimposed vector plot (blue lines, text and circle). The vector overlay can be turned off (or changed in colour) for improved clarity, by unchecking **Graph>Special>Overlays**>(Vectors✓Overlay vectors), using **Font + Colour** to change colour from the default blue, or the circle (indicating a maximal vector) removed by unchecking the (✓Draw circle) box.

Clyde log abiotic

Clyde log(V+0.1) transform (exc. Dep)
Environmental | Pre-treatment

Transform(overall)...
Transform
Standardis
Cumulate
Weight Va
Dispersion
Variability
Normalise Variables

Samples - Sites	lnCu	lnMn	lnCo
S1	3.26	7.81	2.64
S2	3.40	7.06	2.71
S3	3.61	5.97	2.49
S4	4.30	5.85	2.49
S5	4.74	5.75	2.31
S6	5.84	5.39	2.31
S7	5.26	5.54	2.40
S8	4.84	5.50	2.31

Normalise

☑ Stats to worksheet

OK   Can

Clyde abiotic normalised

Clyde EV transformed & normalis
Environmental

Samples - Sites	Variables					
	lnCu	lnMn	lnCo	lnNi	lnZn	lnC
S1	-0.848	2.3088	1.0501	0.2277	-0.353	-1.
S2	-0.694	1.3229	1.352	-0.018	-0.394	-0.
S3	-0.468	-0.113	0.3762	0.6792	-0.141	-0.
S4	0.2804		2	0.9877	0.2216	0.0
S5	0.7569					
S6	1.9418					
S7	1.3223					
S8	0.8642					
	-0.497					
	-0.694					
	-0.934					
	-1.028					

Analyse
Resemblance...
CASWELL...
DIVERSE...
DOMDIS...
PCA...

Mean & SD

Other

Variables	Samples	
	Mean	SD
lnCu	4.046	0.924
lnMn	6.062	0.757
lnCo	2.407	0.227
lnNi	3.473	0.245
lnZn	5.290	0.608
lnCd	-0.556	1.413
lnPb	4.463	0.616
lnCr	4.402	0.877
Dep	99.16	31.02
ln%C	1.359	0.538
ln%N	0.511	0.307

PCA

Maximum no of PCs:
5

☑ Plot results   ☐ Scores to worksheet

PCA1

Eigenvalues
PC	Eigenvalues	%Variation	Cum.%Variation
1	6.87	62.4	62.4
2	2.96	26.9	89.3
3	0.466	4.2	93.6
4	0.365	3.3	96.9
5	0.206	1.9	98.8

Eigenvectors
(Coefficients in the linear combinations of variables making up PC's
Variable	PC1	PC2	PC3	PC4	PC5
lnCu	0.378	-0.035	0.103	-0.064	
lnMn	-0.213	0.418	0.367	0.462	
lnCo	-0.075	0.539	-0.295	0.163	
lnNi	0.149	0.466	-0.617	-0.185	
lnZn	0.366	0.156	0.077	-0.061	
lnCd	0.361	-0.110	0.170	-0.227	
lnPb	0.369	0.126	0.040	-0.039	
lnCr	0.351	-0.091	-0.194	-0.006	
Dep	-0.120	0.459	0.480	-0.667	
ln%C	0.367	0.090	0.114	0.147	
ln%N	0.335	0.191	0.265	0.445	

Principal Component Scores
Sample	SCORE1	SCORE2	SCORE3	SCORE4	S
S1	-2.01	2.41	0.84	0.611	
S2	-2.3	2.09	0.975	-0.256	
S3	-0.985	0.905	-0.353	-1.3	
S4	0.455	0.614	-0.792	-0.549	
S5	2.24	0.304	0.309	-0.49	

Samp. Labels &
Overlays | Special

Vectors
☑ Overlay vectors
◉ Base variables
○ Worksheet variables:
Clyde abiotic normalised
Correlation type:
Pearson
Correlations to worksheet
Select...   Font + Colour..
☑ Draw circle

PCA configuration

Clyde environmental PCA