APPENDIX D

REPORT OF THE SUBGROUP ON STATISTICS

(La Jolla, USA, 14 to 18 July 1997)

REPORT OF THE SUBGROUP ON STATISTICS
(La Jolla, USA, 14 to 18 July 1997)

INTRODUCTION

1.1 The 1997 meeting of the Subgroup on Statistics was held from 14 to 18 July 1997. The meeting was convened by Dr G. Watters (USA) and held at the Southwest Fisheries Science Center in La Jolla, USA.

1.2 A provisional agenda was introduced and discussed. It was agreed that an additional item, 'Synoptic Survey Design', be added to the agenda. The agenda (Attachment A) was adopted without further modification.

1.3 The list of participants is included as Attachment B, and the list of documents submitted to the meeting is included as Attachment C.

1.4 The report was prepared by Drs I. Boyd and J. Croxall (UK), B. Manly (New Zealand), W. de la Mare (Australia), A. Murray (UK), D. Ramm (Secretariat) and G. Watters (USA).

REVIEW OF UPDATED TIME SERIES OF CEMP INDICES

2.1 Dr Ramm introducted WG-EMM-97/25 which comprises the complete tabulation of all data submitted to CEMP (section 2), a selection of figures illustrating these data (section 3) and presentations relating to the identification of anomalies following the methods proposed by the subgroup last year (section 1).

2.2 Dr Ramm and the Secretariat were thanked for the considerable work involved in producing such a comprehensive set of documents.

2.3 In reviewing the compilation of incides the subgroup noted a small number of errors which were corrected in WG-EMM-97/25 Rev. 1.

2.4 The subgroup also made some specific comments:

(i) in the illustration of data collected under Method A1B (section 3, A1B, Figures 1 to 5) the different years should be more clearly demarcated; and,

(ii) for several of the standard methods adequate data were now available to evaluate whether the recommended sampling regimes and sample sizes are appropriate. Members with such data were encouraged to undertake evaluations and report the results to WG-EMM.

FURTHER REVIEW OF IDENTIFICATION OF ANOMALIES IN CEMP INDICES

2.5 The subgroup recognised two particular issues with the identification of anomalies:

(i) identifying anomalies in data from non-normal distributions; and

(ii) some observations that are 'anomalies' from the biological point of view may not be statistically significant.

2.6 The paper by Drs Manly and MacKenzie (WG-EMM-Stats-97/6) was reviewed. The authors discussed the properties of a method for detecting anomalous years in CEMP indices, and extended the idea to situations where data contain a linear trend and autocorrelation, and where data are drawn from a constant distribution other than a normal distribution. In the case of non-normally distributed data, a Box-Cox transformation was applied prior to analysis. The method requires further investigation, but seems generally quite suitable for detecting single extreme values rather than, for example, a permanent change in the mean of a data series.

2.7 The paper by Dr de la Mare (WG-EMM-Stats-97/7) was also reviewed. This includes a proposal for combining CEMP variables to produce a smaller number of summary indices. It also notes that the currently used procedure for detecting anomalies lacks power when there are several extreme values, and that a permanent change in the mean and/or standard deviation in a series is better detected by calculating standardised residuals using the mean and standard deviation from a selected baseline derived from the series. From this point of view the detection of anomalies would include the following steps:

(i) define the classes of behaviour in a series to be detected (a change in the mean, a change in the variance, trend, etc.);

(ii) select a normalising transformation if necessary;

(iii) select a baseline derived from the series;

(iv) examine the statistical properties of the procedure taking into account possible serial correlation, missing values, etc.; and

(v) examine the power of the procedure to detect the phenomena of interest.

2.8 The need to take into account the uses for indices was discussed. It was noted that they are essentially meant to measure various aspects of the food available to predators, with integration over various spatial and temporal scales (Table 1). This emphasises the need to understand the relationship between indices through multivariate analyses, particularly if they are to be combined to produce summary indices of various kinds.

2.9 The use of the word 'anomalies' may be confusing because often what may need to be detected are extreme values that may be part of the natural variation in the system. To some extent these extreme values may just be the result of highly non-linear responses of the predators to environmental conditions. It is recommended that an alternative term be used such as VOGON (Value Outside the Generally Observed Norm). Here 'norm' is defined to be the conditions that are satisfactory for the predator populations.

2.10 Some illustrative calculations were carried out in order to demonstrate the potential value of multivariate analysis. For this purpose the data shown in Table 2 from Bird Island were used. A principal component analysis on the correlation matrix for the indices for the years 1990 to 1997 produced the output shown in Attachment D. It was found that the first component accounts for 53.0% of the variation in the data, while the second and third components account for 19.9% and 12.3%, respectively. Thus between them the first two components account for 72.9% of the variation, while the first three components account for 85.3% of the variation. Applying the analysis to transformed data gives very similar results.

2.11 The first component is essentially an average of the fur seal cow foraging duration (with a negative sign so that the least negative values represent good conditions), gentoo breeding success, macaroni fledging weight, the proportion of krill in the macaroni diet, the proportion of krill in the gentoo diet, the average of the last weighed fur seal pup mass for females, and the average of the last weighed fur seal mass for males. This component can be interpreted as the *overall biological state*. Component 2 mainly reflects the estimated fur seal pup growth rates for males and females, which may be biased because of high mortality in poor years. For this reason high values are not necessarily associated with good conditions. This can be named *fur seal pup growth*. Component 3 is mainly the *macaroni breeding success*. This may reflect the fact that these penguins are able to adapt their diet in poor years so that again it is not a good measure of overall biological conditions.

2.12 The subgroup considers that the results of this principal component analysis are helpful in clarifying the relationship between the various individual indices and the conditions in the different years and recommends that similar analyses are conducted for other sites and variables.

2.13 An initial exploration of the simple combination index suggested in WG-EMM-Stats-97/7 was prepared using CEMP dependent species data from fur seals and macaroni and gentoo penguins at South Georgia. The parameters selected for this illustration can be combined because they refer to similar temporal and spatial scales. The parameters included are listed in Table 2.

2.14 The simple index involves transforming and standardising the various parameters along the lines adopted by WG-EMM in 1996. Each parameter is transformed to have roughly a standard normal distribution. The parameter values are then added together and re-standardised using the estimated standard deviation for the sum using the covariance (correlation) matrix. The values are standardised also with respect to sign, for example, positive values indicating better than average conditions for the predator. For this reason, the sign of the transformed fur seal foraging trip duration was reversed. The simple index can be calculated for all years where some data exist.

2.15 The mean values and covariance matrix needed for the standardisation of the data series were calculated using the data for the period 1989 to 1997; the years when data were available for all the parameters. Prior to standardisation, the data were transformed using the currently accepted transforms for each parameter. This period has been used to provide the baseline mean and covariance matrix for the calculation of the index back to the beginning of the data series in 1977. The subgroup did not examine whether this particular period would form a suitable baseline; the results presented here are for illustrative purposes only. The resulting correlation matrix is shown in Table 3.

2.16 Figure 1 shows the simple index using all the available data. It clearly indicates the two known poor years in 1977 and 1984. The index also suggests poor years in 1987, 1988 and 1994, although the last does not appear as poor as the assessment arrived at by WG-CEMP in 1994. Because the fur seal pup growth parameters were not given a high loading in the first principal component from the principal components analysis (paragraph 2.11), the index was re-calculated without using these data. Excluding these data from the index (shown with a dashed line) results in a slight further depression of the point for 1994, but otherwise there are no changes of any substantial consequence. In light of the fact that 1994 was an extremely poor year for fur seals, the insensitivity of the index to the fur seal pup growth suggests that this parameter is not effectively indexing fur seal reproductive success. It was suggested that these parameters may require further refinement, e.g. by using growth rate of total pup biomass instead of individual pup growth rates.

2.17 Figure 2 shows the simple index calculated without fur seal pup growth rates (dashed line) compared with the simple index based on the breeding success of the two penguin species only (the only parameters represented in all years). The comparison shows that, at least in this instance, the index is not particularly sensitive to the absence of some of the parameters.

2.18 The subgroup considered that the results were encouraging and recommended that further studies should be undertaken to develop some form of combined simple indices at the appropriate regional and temporal scales. The subgroup also noted that the simple index may be more robust for identifying VOGONs than the separate parameter indices because the distribution of a sum of random variables approaches a normal distribution even when the random variables themselves are not normally distributed.

2.19 The subgroup noted previous concerns that the VOGON detection method does not always identify VOGONs when these events are known to be biologically significant (SC-CAMLR-XV, Annex 4, paragraph 4.72). The subgroup agreed that in instances where the distribution of an index (or its transformation) was not approximately normal, the 0.05 α-level might be too stringent to detect biologically significant VOGONs. It was also suggested that it may be useful to develop a procedure for identifying a VOGON in cases where a high proportion of the indices are close to, but not exceeding, their critical levels in the same year.

2.20 To provide two examples of where the 0.05 α-level could be too stringent, the subgroup estimated what α-level would be required to detect all of the biologically significant VOGONs in the Bird Island time series of gentoo penguin (Index A6a) and black-browed albatross (Index B1) breeding successes. Dr Croxall identified the biologically significant VOGONs in each time series.

2.21 For each example, the calculations were made in four steps:

- the index was transformed with the log-odds transformation;
- the least extreme, biologically significant VOGON was identified;
- a critical value (Z_C) for detecting the least extreme VOGON was calculated from

$$Z_c = \frac{\bar{x} - LEV}{s}$$

where $\bar{x}$ and s are the mean and standard deviation of the transformed index, and LEV is the value of the least extreme VOGON; and

- the α-level corresponding to Z_C was identified by simulating 1 000 20-year time series of standardised normal deviates, counting the number of instances where the absolute value of simulated deviate was ≥ Z_C, and dividing this count by 20 000.

2.22 The results of the example calculations are provided in Table 4. An α = 0.22 would be required to detect all of the biologically significant VOGONs in the gentoo time series, and an α = 0.69 would be required for the albatross time series. An α = 0.05 would be too stringent in both cases.

2.23 Given the results of the example calculations, the subgroup agreed that the appropriate α-level for identifying VOGONs should be selected on an index-by-index basis after careful consideration of whether each index (or its transformation) is normally distributed. When the index (or its transformation) is not normal, α-levels between 0.2 and 0.3 may be appropriate.

CRITICAL EVALUATION OF THE ASSUMPTIONS AND PARAMETER VALUES OF THE AGNEW AND PHEGAN (1995) MODEL OF REALISED OVERLAP

3.1 Last year WG-EMM requested that the Subgroup on Statistics evaluate the assumptions and parameter values in the fine-scale model of the overlap between penguin foraging demands and the krill fishery in the South Shetland Islands and Antarctic Peninsula (Agnew and Phegan, 1995) (SC-CAMLR-XV, Annex 4, paragraph 6.80). This model calculates penguin foraging demand and is intended for the purposes of calculating an index of foraging–fishery overlap during the critical period December to March. Data from Subarea 48.1 on penguin foraging characteristics, energetic demands, and population numbers, and monthly krill catches by fine-scale grid are used as inputs to the model.

3.2 To assist in this process the Secretariat had requested (SC CIRC 97/2) data and analysis providing estimates of:

(i) monthly composition of diet (of penguins and fur seals);
(ii) maximum and mean/modal foraging distance;
(iii) mean foraging bearings; and
(iv) fine-scale data on foraging distributions.

3.3 Such data have been provided for gentoo and macaroni penguins and Antarctic fur seals for Bird Island South Georgia (Subarea 48.3) in WG-EMM-Stats-97/5. Data for chinstrap penguins at Seal Island had been submitted to the Secretariat for consideration by WG-EMM but were not available at the subgroup meeting. It was regretted that similar data have not yet been provided for other sites, particularly those in Subarea 48.1 where several extensive studies of diet and foraging have been carried out.

3.4 In reviewing the model the following main topics were considered:

(i) foraging distance;
(ii) foraging bearing;

(iii) predator consumption rates;
(iv) population counts; and
(v) model structure.

3.5 The model assumes that penguin foraging distances are normally distributed about a mean distance from the colonies. The values used in the model were: chinstrap penguin mean foraging distance of 20 km with a standard deviation of 8 km ~N(20,8); Adélie penguin ~N(38,15); gentoo penguin ~N(10,4); and macaroni penguin ~N(28,11). The maximum foraging distance was set to the mean + 2 standard deviations.

3.6 The model assumes that penguin foraging bearings are uniformly distributed about a line perpendicular to the coast on which the colony lies. Data on foraging bearings from colonies in Subarea 48.1 are limited to Seal Island. The values used in the model ranged generally 40° either side of a line perpendicular to the coast.

3.7 The foraging distance and bearing data used in the model were certainly appropriate for the Seal Island area. The group noted the paucity of available data to extend the model to include other regions within Subarea 48.1, and recommended that extrapolation to regions with no data should be made with caution.

3.8 The distribution of foraging distances is unlikely to be normal. *A priori* some kind of exponential distribution might be expected; available evidence from at-sea observations shows the pattern of distribution to be skewed. For foraging bearing there is no *a priori* reason, nor any observational evidence, to suggest that any assumption other than a uniform distribution is warranted. The distribution of both parameters should be re-examined in the light of new data, and literature on animal movements.

3.9 The model uses mean values for predator consumption rate which were the best estimates available from studies up to around 1984. There are quite extensive additional data on at-sea metabolic rate and energy requirements of penguins now available (see e.g. WG-EMM-96/19 and SC-CAMLR-XV, Annex 4, paragraph 6.41) which could improve the estimates used in the model.

3.10 The penguin population counts used in the model were derived from a long-term dataset on penguins counts, and were the best available in 1992. An updated dataset is now available (SC-CAMLR-XV/BG/29).

3.11 The subgroup examined the four steps involved in the model:

(i) estimating the total number of penguins from all colonies foraging within the area;
(ii) calculating the number of these expected to forage within each 10 x 10 n miles^2;
(iii) calculating the total consumption of krill by penguins; and
(iv) calculating the foraging–fishery overlap (FFO) index.

The subgroup agreed that the basic spatial modelling approach used was appropriate. However, it was not clear whether the temporal aspects of penguin foraging had been adequately captured in the model, and the subgroup agreed that this aspect should be developed further. The subgroup also found that the FFO index was not a direct measure of overlap, but rather was related to the total amount of krill removed from the foraging area during the critical period. The FFO index is the product [total krill consumption by penguins]∗[total krill catch in the fishery] with units of $(\text{mass})^2$.

3.12 The subgroup proposed that a new standardised index be developed based on niche overlap theory (SC-CAMLR-XV, Annex 4, Appendix H), such as Schroeder's index

$$I_t = 1 - 0.5\sum\left|p_{i,t} - q_{i,t}\right|$$

where $p_{i,t}$ is the proportion of krill consumed by a predator(s) in grid square *i* during time period *t* and $q_{i,t}$ is the proportion of krill consumed by the fishery in grid square *i* during time period *t*. This type of index would range from $I_t = 0$, no spatial overlap between predator consumption and fishery consumption during period *t* to $I_t = 1$, complete overlap between predator consumption and fishery consumption during period *t*. At present, $p_{i,t}$ can be calculated along the lines of the structure in Agnew and Phegan (1995).

3.13 It was recommended that this new index should be applied first to Subarea 48.1, initially using the existing data from Seal Island. This should be undertaken by the Secretariat so that results can be presented to the meeting of the Scientific Committee in October.

3.14 The subgroup recommended that the next tasks relating to studies of realised overlap should include:

(i) examination of the sensitivity of the index *I* to the various assumptions made about penguin foraging effort and prey consumption;

(ii) incorporation of appropriate data on foraging effort and distribution from sites in Subarea 48.1 in addition to Seal Island. These data should be submitted as soon as possible using the forms prepared by the Secretariat (SC CIRC 97/2) as a guide but, where appropriate, providing data and analyses in ways analogous to those in WG-EMM-Stats-97/5; and

(iii) applying the model to Subarea 48.3. It was noted that the fishery currently operates there in winter providing little interaction with krill-dependent predators during the December to March critical period. Useful analyses, however, might still be made by using data from earlier years when the krill fishery operated in summer.

3.15 Future desirable developments would be to examine the overlap between penguin foraging demands and the krill fishery during other potentially critical periods. Of particular importance is the post-fledging period when large numbers of chicks begin foraging independently and adults are feeding intensively in preparation for their annual molt. Recent studies are also indicating that critical periods may exist during the winter. There are little or no empirical data for most of these periods. In terms of winter studies, the priority species for concurrent investigation of the distribution of predator foraging and the krill fishery are fur seal, macaroni penguin and chinstrap penguin.

DEVELOPMENT OF INDICES OF AT-SEA BEHAVIOUR AND METHODS OF DERIVING THEM VIA ANALYSIS OF SAMPLE DATASETS

4.1 Previous discussions of WG-EMM had identified a need for a coordinated approach to the analysis of data about the at-sea behaviour of diving predators such as penguins and fur seals. The main reason for this is to allow monitoring of the behaviour of diving predators at finer spatial and temporal scales than have been available using current CEMP indices. A further objective would be to provide input to the realised overlap index (paragraph 3.12). This will also utilise several existing datasets. Methods for measuring at-sea behaviour, and for the deployment of instruments used for measuring at-sea behaviour, have already been adopted (WG-EMM-96).

4.2 The subgroup was tasked with:

(i) reviewing appropriate temporal and spatial scales for developing indices of at-sea behaviour (SC-CAMLR-XV, Annex 4, paragraphs 3.61 to 3.65 and 7.58);

(ii) considering sample datasets and analyses (SC-CAMLR-XV, Annex 4, paragraphs 4.44 and 7.58);

(iii) developing indices and methods for the calculation via analysis of the sample datasets (SC-CAMLR-XV, paragraph 5.38(i)); and

(iv) providing advice on the most appropriate indices for inclusion in the CEMP database (SC-CAMLR-XV, Annex 4, paragraphs 4.44 and 7.58).

4.3 The subgroup examined several sample datasets from Antarctic fur seals. From a bivariate dataset involving time and depth (sampled at intervals from 5 to 15 seconds) it is possible to derive several subsidiary parameters such as dive depth, dive duration and the interval spent at the surface between dives. In turn, these can provide information about dive frequency, proportion of dives made at different times of day, and bouts of diving. Past studies have shown that these have the potential to provide information about variability in at-sea behaviour between years that reflects variation in food availability.

4.4 There is little consensus in the literature as to how comparisons of at-sea behaviour between individuals and across years should be made. As a general principle, the subgroup recommended that comparisons should be based on procedures that correctly take into account the variability in the data. In particular, attention was drawn to spectral analysis as a potentially useful approach. This would have the advantage of incorporating all of the data into a single analytical approach while minimising the need to make assumptions about how individual units of behaviour, such as dives or bouts of dives, should be defined.

4.5 A second approach, which also overcomes many of the assumptions with defining dives and bouts of dives, is to examine the cumulative time spent submerged during a foraging trip in relation to cumulative time spent at sea. The slope of this relationship could provide a single parameter that integrates most of the variability in at-sea behaviour within a single index.

4.6 Comparing at-sea behaviour across years is complicated by a potentially high degree of variability between individuals and because many of the parameters that are commonly used to measure at-sea behaviour often have highly-skewed distributions. Some may also show a degree of bimodality.

4.7 The subgroup recommended that the use of a randomisation test should be investigated to examine interannual variability in the indices. Dr Manly suggested that this could involve the following procedure:

(i) assume that the data consists of records for individual foraging trips and that these are from different animals;

(ii) for each pair of foraging trips measure the difference between them (e.g. a Kolmogorov-Smirnov measure of the difference between the index distribution). This gives a predator difference matrix for which $a(i, j)$, the element in row i and column j, is the difference for predators i and j;

(iii) generate a second matrix in which the elements are sample similarities as often recommended for the multi-response permutation procedure (Mielke et al., 1976). Thus the element $b(i, j)$ in row i and column j contains 0 from two cases in different years and $1/(n\text{-}1)$ for two cases in a year with a sample of size n;

(iv) test whether the correlation between $a(i, j)$ and $b(i, j)$ is significantly negative, by comparison with the distribution found by randomly permuting the sample labels for one of the matrices, i.e. do a Mantel (1967) matrix permutation test as described by Manly (1997); and

(v) the test can be done with any statistic measuring the difference between the behaviour of two predators.

4.8 The large size of the datasets and the need for detailed consideration of how these analytical techniques can be applied to measurements of at-sea behaviour meant that it was impractical for the subgroup to investigate these methods during the meeting. Drs Boyd and Murray agreed to undertake an example analysis to assess this method using multi-year data from Antarctic fur seals and to report the results to a future meeting of WG-EMM.

4.9 Scales of variability in at-sea behaviour may be defined most satisfactorily using spectral analysis. An example of such an analysis carried out by Dr Boyd showed several peaks in the spectrum that corresponded to the different scales of behaviour, namely, the dive, dive bouts and diel variability. Dr Murray suggested that alternatives to the assumptions of sine wave forms associated with Fourier transformations may provide an alternative spectrum with additional information. Drs Boyd and Murray also agreed to investigate this intersessionally.

4.10 The subgroup also considered the utility of including locational data from satellite tags as a variable describing at-sea behaviour. The precision of locational data is sufficient for input to the predator–fisheries realised overlap index (paragraph 3.12). However, at this stage, the precision of satellite locations is insufficient to allow assessments to be made of variability in foraging locations at the smallest spatial scales addressed by time–depth data.

4.11 The subgroup concluded that it was still too early to make firm recommendations about which indices of at-sea behaviour should be included within the CEMP database. Further consideration should be given to this subject once the various methods discussed by the subgroup had been tested.

METHODS FOR COPING WITH MISSING VALUES
IN MULTIPLE DATASETS

5.1 Dr Murray presented his paper WG-EMM-Stats-97/8. The paper outlines three stages in analysis of incomplete datasets:

(i) understanding the mechanisms generating the missing values (were they random or not?);

(ii) deciding on the appropriate analysis of the data in order to support the required inferences (e.g. trend estimation, identification of unusual values); and

(iii) choosing and implementing an appropriate method of missing data imputation and subsequent data analysis.

The classes of missing value mechanisms and the broad categories of imputation methods were reviewed. For a value to be considered as 'missing at random' the probability of it being missing should be independent of the observed and missing values. Analysis of an example dataset of Chinstrap penguin colony counts from Signy Island was presented to illustrate four methods of imputation.

5.2 A method of evaluating the effect of imputing missing values on the analysis would be to take a complete dataset and try various patterns (random and non-random) and extents of data deletion. Imputed values could then be compared with the original values and analyses of completed datasets compared with the analysis of the full dataset. This would give a measure of the success of the imputation procedures. Many studies of this kind have been reported in the literature and for at least some the finding has been that, although individual values may not match the original data closely, statistics such as means may be close to the original values. For illustrative purposes, an exercise of this kind may be useful for an example CEMP dataset.

5.3 WG-EMM-Stats-97/8 drew attention to the importance of understanding the mechanisms leading to missing data and called for a discussion of these in the context of CEMP series. A number of possible reasons for missing data in CEMP indices were identified.

(i) <u>Data were not collected either because there was no intention to collect or because logistic considerations</u> such as lack of means of access or equipment failure prevented collection. Such data could be considered to be missing completely at random.

(ii) <u>Data were not collected because of adverse environmental conditions</u>, such as sea-ice preventing access to a site or bad weather making completion of field work impossible. Depending on the nature of the variable in question, such reasons might not be regarded as random. For instance, for some biological parameters such as arrival time, the presence of sea ice might have an important influence so that the same reason leading to the data being missing might also affect the value. Such data could not be regarded as missing at random.

(iii) <u>Data were not collected due to biological circumstances</u>, for example the animals in question died during the course of the season (e.g. death of chicks before fledging as occurs in some years). This seems unlikely to occur at random and may, in itself, be an important biological indicator of the ecosystem status in that year.

(iv) Data were not recorded although they are known to exceed a given threshold (e.g. where data exceed storage capacity of the recording instrument). This is called censoring and is common in observations of time duration where the event, such as return from a foraging trip, is not observed before the end of the period available for observation. The reasons might be either biological in the case of extended or incomplete foraging trips in poor seasons or non-biological in the case of equipment failure or exceeding instrument data storage capacity. The former could certainly not be regarded as random although the latter might in some circumstances be so regarded. Standard statistical methods are available for estimating parameters of distributions (such as means) where observations for some units in the sample are censored. It was felt that it would be worth reviewing the standard method for foraging trip duration of fur seals (method C1) to see whether adoption of this analysis methodology would allow more complete datasets of this index to be produced.

(v) Data were not reported where in fact they were actually null values, for example certain prey items were absent from stomach contents. Such values should be identified and replaced with zeroes in the data base.

5.4 The subgroup agreed that it was important to assess the CEMP series to determine the reasons for the missing data before proceeding to formal analysis. Such an assessment should be done as soon as possible. The originators of the data should be encouraged to supply the necessary information and it was felt that such a request could be phrased in the form of a multiple choice along the lines in paragraph 5.3.

5.5 There are two levels at which missing data may arise in the CEMP series. The first is at the level of the samples which go to make up the calculated value which is submitted; the second at the level of the calculated CEMP indices.

5.6 It is important to discover if any missing value techniques have been applied to sample data in the calculation of values which have been already submitted to CCAMLR. In certain cases, for example a colony count is missing from a set of colony counts at a site, missing value imputation could be used to calculate a site value. The subgroup recommended that where such cases can be identified the raw data should be submitted so that appropriate statistical techniques can be examined and applied.

5.7 Missing values in time series incorporated into the CEMP database should only be imputed in the course of analyses for particular purposes. The methods used should take into account the reasons for the missing data supplied by the originators of the data and the intent of the analysis. Such imputed data should not be stored in the CCAMLR database. The imputed values should not be used as if they are real data. They serve solely to allow the analysis of values which do exist and, indeed, different values may be imputed in the context of different analyses. It is important to ensure that the imputation methods which are used serve to allow the use of all observed data without adding artificial effects to the data. That is, the imputed values should be as far as possible 'neutral' in their effect on estimates of means, correlations, trends, etc.

5.8 Imputation should be as realistic as possible with consideration being given to the appropriate biological, spatial and temporal factors in deciding which data to use in multivariate imputation techniques. For example, imputation might be 'cross-sectional' based on using values for the same variable or related variable(s) at different colonies or sites in the same year, or 'longitudinal' using values from adjacent years, or a combination of both.

SYNOPTIC SURVEY DESIGN

6.1 The subgroup reiterated the view that the primary objective of the synoptic survey is to provide an estimate of krill biomass and its variability for use in the krill yield model. Other objectives (e.g. to study the spatial structure of krill aggregations) are secondary. The subgroup noted that there are two key issues with regard to the design of the synoptic survey: stratification, and random versus systematic placement of transect lines.

6.2 The subgroup agreed with WG-EMM's previous opinion (SC-CAMLR-XV, Annex 4, paragraph 3.75(v)) that the survey should be stratified according to large-scale spatial differences in krill density. The subgroup noted that there are many historical datasets (e.g. FIBEX, AMLR, LTER) that can be used to estimate how sampling effort should be allocated between strata.

6.3 The subgroup initiated the discussion on transect placement by noting that random placement should facilitate both design-based (e.g. Jolly and Hampton estimators) and model-based (e.g. geostatistics) estimates of variance in krill biomass. Systematic transect placement requires model-based variance estimation. Model-based variance estimators can be more efficient than design-based estimators, but such estimators are conditional on the adequacy of the model. A simulation study is needed to compare the relative efficiencies of random and systematic transect placement in a synoptic survey for krill. Such a study is the only quantitative way of comparing the two survey designs.

6.4 The subgroup agreed that a simulation study should receive high priority; it would be best if the work could be completed within about one year. A small panel of interested parties should be convened as soon as possible to define some realistic goals and boundaries for the simulation study. The subgroup did note that the simulation should, at a minimum, consider the following points:

(i) the cost (e.g. in ship-hours) of alternative designs (including the cost of various degrees of randomisation);

(ii) the biases introduced by the diel vertical migrations of krill; and,

(iii) the effects of the spatial coherence of the krill distributions being different in different directions.

It might also be valuable to consider whether there is a point at which the marginal utility of reducing the variance becomes small. This could be studied by considering when the results of the krill yield model become more sensitive to variability in krill recruitment rather than to uncertainty in krill biomass.

6.5 Drs Manly and Murray stated that they would be willing to develop the simulation study in collaboration with a colleague from New Zealand who specialises in geostatistics. Drs Manly and Murray also noted that they would be grateful for input from other interested parties, especially those with historical krill survey datasets. Dr de la Mare undertook, in conjunction with the Secretariat, to examine the marginal utility of reducing the variance in biomass estimates.

6.6 In the absence of a simulation study, the subgroup agreed that randomly-spaced parallel transects would be a conservative design because both design- and model-based variance estimators could be used to analyse the data.

ADVICE TO WG-EMM

7.1 The subgroup summarised its recommendations.

Agenda Item 2

7.2 The term VOGON (Value Outside the Generally Observed Norm) should be used in place of anomaly (paragraph 2.9).

7.3 Principal components analysis should be carried out for appropriate sites and indices (paragraph 2.12).

7.4 The fur seal pup growth index (C2b) may not be an effective measure of reproductive success and should be examined for further refinement (paragraph 2.16).

7.5 Further studies should be undertaken to develop combinations of CEMP indices at appropriate regional and temporal scales that may be more robust for identifying VOGONs than individual indices (paragraph 2.18).

7.6 Consideration should be given to the development of a procedure for identifying situations where a high proportion of indices give near VOGONs (paragraph 2.19).

7.7 Appropriate α-levels for identifying VOGONs should be done on an index-by-index basis, with levels higher than 0.05 being considered for non-normal data (paragraph 2.23).

Agenda Item 3

7.8 Modify the Agnew and Phegan (1995) model to improve temporal aspects (paragraph 3.11).

7.9 A new index of niche overlap, such as Schroeder's Index, should be applied to Subarea 48.1 (paragraph 3.12).

7.10 Further work on the study of realised overlap, including sensitivity analyses, incorporation of new data from Subarea 48.1, and application to Subarea 48.3 should be undertaken (paragraph 3.14).

7.11 Future developments of a realised overlap index should examine penguin–fishery interactions during other potentially critical periods (paragraph 3.15).

7.12 Additional data should be submitted so that the work outlined above can progress (paragraph 3.3).

Agenda Item 4

7.13 Methods of comparing at-sea behaviour indices between sites and across years should be developed with randomisation tests (paragraphs 4.7 and 4.8).

7.14 Indices that summarise at-sea behaviour, including the use of satellite data (paragraph 4.10), should be developed and the properties of these indices should be investigated (paragraph 4.9).

7.15 Items in paragraphs 7.13 and 7.14 need to be dealt with before a decision can be made about which indices can be incorporated into the CEMP database.

Agenda Item 5

7.16 Various missing value scenarios should be explored with a complete CEMP dataset (paragraph 5.2).

7.17 Information on the reasons for missing values in CEMP data should be collected, as soon as possible, along the lines suggested in paragraph 5.3 (paragraph 5.4).

7.18 Work should be undertaken to identify series and methods whereby missing sample data can be imputed in order to provide a value for a parameter which would otherwise be missing from the CEMP series (paragraph 5.6).

7.19 Work should be undertaken to explore the methodology for analyses of multivariate series with missing values so that such analyses can be performed in the future (paragraphs 5.7 and 5.8).

Agenda Item 6

7.20 A simulation study should be conducted to compare random versus systematic transect spacing for the synoptic krill survey, and a panel should be convened to define realistic goals and boundaries for the study (paragraph 6.4).

7.21 Work should be undertaken to use the krill yield model to examine the marginal utility of reducing uncertainty in the krill biomass estimate (paragraph 6.5).

7.22 Random transect spacing should be used in the synoptic survey if a simulation study is not completed (paragraph 6.6).

CLOSE OF THE MEETING

8.1 The report was adopted. In closing the meeting the Convener thanked the Southwest Fisheries Science Center and Dr R. Holt for hosting the meeting. The Convener also thanked all the meeting participants.

REFERENCES

Agnew, D.J. and G. Phegan. 1995. Development of a fine-scale model of land-based predator foraging demands in the Antarctic. *CCAMLR Science*, 2: 99–110.

Manly, B.F.J. 1997. *Randomisation, Bootstrap and Monte Carlo Methods in Biology*, 2nd Edition. Chapman and Hall, London.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209–220.

Mielkle, P.W., K. J. Berry and E.S. Johnson. 1976. Multi-response permutation procedures for *a priori* classifications. *Communications in Statistics*, A5: 1409–1424.

Table 1: Temporal scales of integration of variables monitored for predators.

2 – 10 years	1 Year	0.5 – 2 Years	About 6 Months (winter)	1 – 6 Months (summer)
Juvenile survival	Adult survival	Population size	Adult mass at arrival	Foraging trip duration Pup growth rate Weaning/fledging mass Breeding success Diet composition Meal mass

Table 2: Data from Bird Island used for illustrative purposes for multivariate analysis and the production of summary indices. The sign of the fur seal foraging duration is given a negative sign in order that the least negative values represent good conditions.

Year	C1 Fur Seal Cow Foraging Duration * (-1)	C2b Fur Seal Pup Growth Female	C2b Fur Seal Pup Growth Male	A6a Macaroni Breeding Success	A6a Gentoo Breeding Success	A7 Macaroni Fledging Weight	A7 Gentoo Fledging Weight	A8 Macaroni Proportion Krill in Diet	A8 Gentoo Proportion Krill in Diet	Fur Seal Last Weighed Mass Female	Fur Seal Last Weighed Mass Male
1977				0.476	0.598						
1978				0.250	0.006						
1979				0.473	0.294						
1980				0.602	0.577						
1981				0.527							
1982				0.509	0.048						
1983				0.491	0.506						
1984				0.092	0.285						
1985				0.477	0.428						
1986				0.504	0.418						
1987				0.361	0.427						
1988				0.364	0.468						
1989				0.608	0.457	3450	5464				
1990	-80	1.89	2.38	0.592	0.356	3237	5800	0.998	0.594	11.24	13.07
1991	-203	2.77	3.26	0.583	0.010	3112	5043	0.694	0.191	11.48	12.73
1992	-94	2.14	2.58	0.408	0.631	3507	5791	0.988	0.499	12.84	14.81
1993	-123	2.67	3.69	0.553	0.894	3318	5482	0.833	0.845	12.45	15.02
1994	-469	2.48	2.66	0.456	0.040	2913	5065	0.112	0.129	10.66	11.89
1995	-103	2.12	3.31	0.505	0.583	3025	5239	0.536	0.544	11.21	13.92
1996	-90	2.25	2.78	0.445	0.789	3179	5502	0.999	0.243	11.84	14.31
1997	-97	2.25	2.95	0.484	0.500	3300	5960	0.986	0.362	11.93	14.95

Table 3: Correlation matrix for CEMP parameters of dependent species during the breeding season at South Georgia for the years 1989 to 1997.

Variable	Fur Seal Foraging Trip Duration	Fur Seal Pup Growth Females	Fur Seal Pup Growth Males	Macaroni Breeding Success	Gentoo Breeding Success	Macaroni Fledging Weight	Gentoo Fledging Weight	Macaroni Krill Proportion in Diet	Gentoo Krill Proportion in Diet	Fur Seal Female Pup Mass at Weaning	Fur Seal Male Pup Mass at Weaning
Fur seal foraging trip duration	1.000	0.611	0.069	0.047	0.710	0.665	0.743	0.808	0.590	0.552	0.730
Fur seal pup growth females		1.000	0.672	-0.160	0.367	0.225	0.607	0.571	0.197	-0.051	0.132
Fur seal pup growth males			1.000	-0.307	-0.139	0.109	0.389	0.451	-0.347	-0.185	-0.311
Macaroni breeding success				1.000	-0.272	-0.168	-0.166	-0.073	0.310	-0.283	-0.270
Gentoo breeding success					1.000	0.549	0.590	0.506	0.717	0.607	0.835
Macaroni fledging weight						1.000	0.784	0.630	0.546	0.919	0.766
Gentoo fledging weight							1.000	0.769	0.432	0.560	0.663
Macaroni krill proportion in diet								1.000	0.213	0.454	0.491
Gentoo krill proportion in diet									1.000	0.528	0.624
Fur seal female pup mass at weaning										1.000	0.851
Fur seal male pup mass at weaning											1.000

Table 4: Determination of α-levels that are required for detecting biologically identified VOGONs.

	Gentoo	Albatross
Years with biologically significant VOGONs	1978, 1982, 1991, 1994	1980, 1984, 1987, 1991, 1994
Years excluded from analysis – reason for exclusion	1981 – no data	1988, 1995 – adverse environmental conditions identified as main cause of breeding failure
Adjusted time series length	20 years	20 years
Year with least extreme VOGON	1982	1987
Mean of transformed index	-0.7210	-1.4650
Standard deviation of transformed index	1.8508	2.1379
Level of least extreme VOGON	-2.9874	-2.3259
Critical value required to detect least extreme VOGON	1.2245	0.4027
α-level for critical value	0.22	0.69

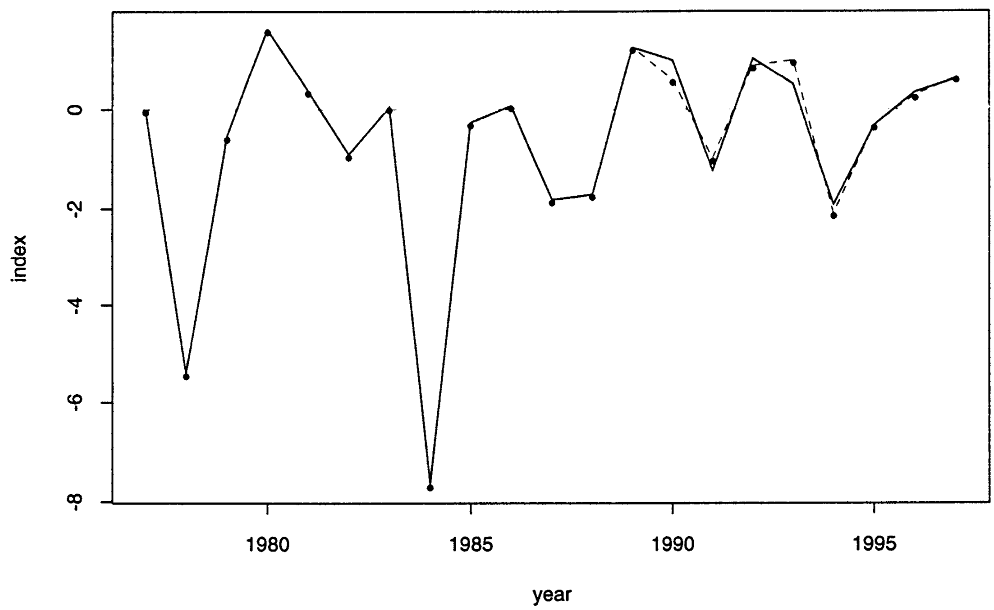

Figure 1: Illustration of the simple index for dependent species at South Georgia which combines fur seal and penguin data relevant to the breeding season. The full line is the index using all the data values, the dashed line shows the effect of deleting the fur seal pup growth data.

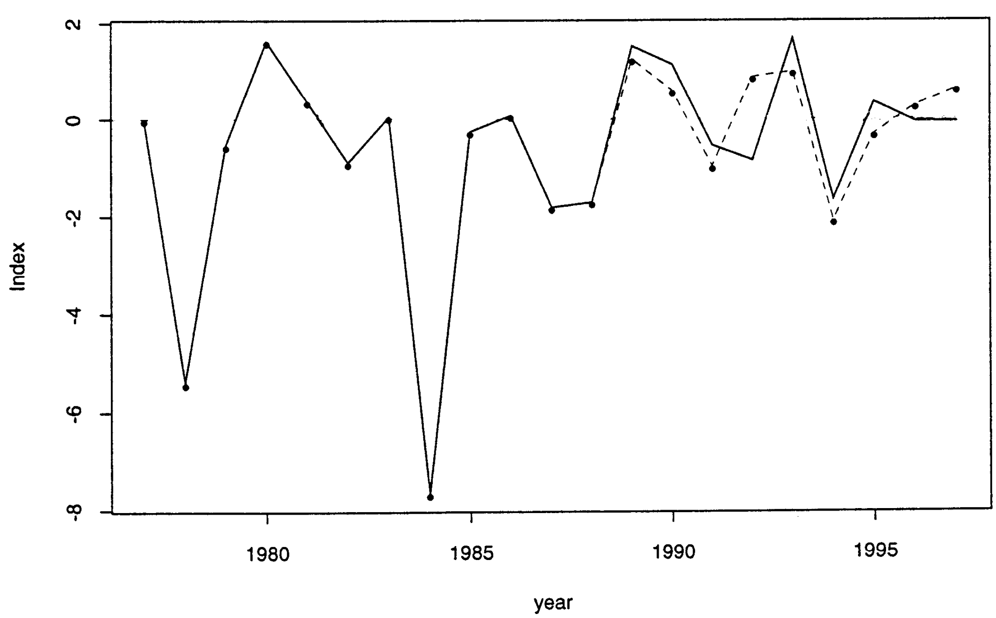

Figure 2: Illustration of the simple index for dependent species at South Georgia which combines fur seal and penguin data relevant to the breeding season. The full line is the index using only the penguin breeding success data, the dashed line shows the effect of including all the other data, apart from the fur seal pup growth data.

ATTACHMENT A

AGENDA

Subgroup on Statistics
(La Jolla, USA, 14 to 18 July 1997)

1. Introduction

   (i) Opening of the Meeting
   (ii) Organisation of the Meeting and Adoption of the Agenda

2. Further Review of Identification of Anomalies in CEMP Indices

   (i) Review updated time series of CEMP indices

   (ii) Summarise recent problems with/suggestions for identifying anomalies (various problems and suggestions can be found in SC-CAMLR-XV, Annex 4, paragraphs 4.58 to 4.61, 4.70, 4.72, 4.75 and 7.1)

   (iii) Discuss and develop methods to deal with problems/take up suggestions in identifying anomalies (SC-CAMLR-XV, paragraph 5.38(ii))

3. Critical Evaluation of the Assumptions and Parameter Values of the Agnew and Phegan (1995) Model of Realised Overlap

   (i) Review and summarise data and analyses submitted in response to SC CIRC 97/2 ('WG-EMM Subgroup on Statistics – Request for Data and Analyses')

   (ii) Evaluate assumptions and parameter values used in the Agnew and Phegan model (SC-CAMLR-XV, paragraph 5.38(iv))

   (iii) Determine whether the data submitted in response to SC CIRC 97/2 could be used to refine the Agnew and Phegan model or develop an alternative index of realised overlap

4. Development of Indices of At-sea Behaviour and Methods of Deriving them via Analysis of Sample Datasets

   (i) Review appropriate temporal and spatial scales for developing useful indices (background information on this topic is presented in SC-CAMLR-XV, Annex 4, paragraphs 3.61 to 3.65 and 7.58)

   (ii) Consider sample datasets and analyses (SC-CAMLR-XV, Annex 4, paragraphs 4.44 and 7.58)

   (iii) Develop indices and methods for their calculation via analysis of the sample datasets (SC-CAMLR-XV, paragraph 5.38(i))

(iv) Provide advice on the most appropriate indices for inclusion in the CEMP database (SC-CAMLR-XV, Annex 4, paragraphs 4.44 and 7.58)

5. Methods for Coping with Missing Values in Multiple Datasets

(i) Examine methods for interpolating missing data in matrices of time series of CEMP indices collected from a group of predator colonies (SC-CAMLR-XV, paragraph 5.38(iii) and Annex 4, paragraph 4.63)

6. Synoptic Survey Design

7. Advice to WG-EMM

8. Close of the Meeting.

LIST OF PARTICIPANTS

Subgroup on Statistics
(La Jolla, USA, 14 to 18 July 1997)

BOYD, Ian (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom i.boyd@bas.ac.uk
CROXALL, John (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom
DE LA MARE, William (Dr)	Australian Antarctic Division Channel Highway Kingston Tasmania 7050 Australia bill_de@antdiv.gov.au
EVERSON, Inigo (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom i.everson@bas.ac.uk
HOLT, Rennie (Dr)	US AMLR Program Southwest Fisheries Science Center PO Box 271 La Jolla, Ca. 92038 USA rholt@ucsd.edu
ICHII, Taro (Mr)	National Research Institute of Far Seas Fisheries Orido 5-7-1, Shimizu Shizuoka 424 Japan ichii@enyo.affrc.go.jp
MANLY, Bryan (Dr)	University of Otago PO Box 56 Dunedin New Zealand bmanly@maths.otago.ac.nz

MILLER, Denzil (Dr)	Sea Fisheries Research Institute Private Bag X2 Roggebaai 8012 South Africa dmiller@sfri.wcape.gov.za
MURRAY, Alastair (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom a.murray@bas.ac.uk
WATTERS, George (Dr)	US AMLR Program Southwest Fisheries Science Center PO Box 271 La Jolla, CA 92038 USA gwatters@amlr.ucsd.edu Present address: Inter-American Tropical Tuna Commission 8604 La Jolla Shores Drive La Jolla CA 92038 USA gwatters@iattc.ucsd.edu

SECRETARIAT:

RAMM, David (Dr) (Data Manager) MARAZAS, Rosalie (Secretary)	CCAMLR 23 Old Wharf Hobart Tasmania 7000 Australia ccamlr@ccamlr.org

LIST OF DOCUMENTS

Subgroup on Statistics
(La Jolla, USA, 14 to 18 July 1997)

WG-EMM-Stats-97/1	PROVISIONAL AND ANNOTATED PROVISIONAL AGENDA FOR THE 1997 MEETING OF THE WG-EMM SUBGOUP ON STATISTICS
WG-EMM-Stats-97/2	LIST OF PARTICIPANTS
WG-EMM-Stats-97/3	LIST OF DOCUMENTS
WG-EMM-Stats-97/4	DEVELOPMENT OF INDICES OF AT-SEA BEHAVIOUR I.L. Boyd (UK)
WG-EMM-Stats-97/5	DIET AND FORAGING RANGE OF PENGUINS AND FUR SEALS AT SOUTH GEORGIA J.P. Croxall, I.L. Boyd, K. Reid and P.N. Trathan (UK)
WG-EMM-Stats-97/6	TESTS FOR ANOMALOUS YEARS IN THE CCAMLR INDEX SERIES (DRAFT) B.F. Manly and D. MacKenzie (New Zealand)
WG-EMM-Stats-97/7	SOME CONSIDERATIONS FOR THE FURTHER DEVELOPMENT OF STATISTICAL SUMMARIES OF CEMP INDICES W.K. de la Mare (Australia)
WG-EMM-Stats-97/8	TREATMENT OF MISSING VALUES IN CEMP DATA SETS A. Murray (UK)

OTHER DOCUMENTS

WG-EMM-97/25	CEMP INDICES 1997: SECTIONS 1 TO 3 Secretariat

RESULTS OF A PRINCIPAL COMPONENTS ANALYSIS
ON BIRD ISLAND DATA 1990–97

The variables are in the order shown in Table 2, with obvious abbreviations for names.

Bird Island data (all untransformed)

PCA axis	1	2	3	4	5	6	7
Eigenvalue	5.83	2.19	1.36	0.82	0.47	0.20	0.13
% of Total	53.02	19.92	12.32	7.46	4.27	1.78	1.22
Cumulative %	53.02	72.94	85.26	92.72	96.99	98.78	100.00
Eigenvectors (component loadings)							
SEALFD (C1)	0.36	0.02	0.27	-0.02	-0.49	0.26	-0.33
SEALPG-F (C2b)	-0.16	0.51	-0.28	0.45	0.03	-0.12	0.35
SEALPG-M (C2b)	0.02	0.65	-0.04	-0.13	-0.20	-0.25	-0.35
MACBS (A6a)	-0.06	0.29	0.73	0.26	0.04	-0.06	0.17
GENBS (A6a)	0.34	0.15	-0.16	-0.47	-0.13	0.13	0.65
MACFW (A7)	0.37	-0.05	-0.10	0.37	0.34	0.16	-0.17
GENFW (A7)	0.34	-0.29	0.10	0.10	0.17	-0.74	0.08
MACPK (A8)	0.36	-0.09	0.17	0.34	-0.34	0.09	0.33
GENPK (A8)	0.27	0.27	0.31	-0.36	0.61	0.13	-0.02
SEALWT-F	0.35	0.14	-0.31	0.28	0.19	0.31	-0.12
SEALW-M	0.38	0.14	-0.21	-0.12	-0.16	-0.38	-0.17
Principal component scores							
1990	0.22	-0.60	0.90	0.03	0.15	0.04	0.08
1991	-0.88	0.50	0.17	0.60	-0.19	0.10	-0.08
1992	0.99	-0.44	-0.50	0.16	0.24	0.18	-0.13
1993	0.71	1.07	-0.00	-0.09	0.26	-0.03	0.12
1994	-1.74	-0.29	-0.36	-0.14	0.26	-0.07	0.07
1995	-0.21	0.23	0.18	-0.61	-0.17	0.05	-0.19
1996	0.32	-0.25	-0.30	-0.10	-0.42	0.10	0.21
1997	0.59	-0.21	-0.08	0.16	-0.12	-0.37	-0.07

Plots of principal components for each year