APPENDIX B

FLIGHT DIRECTION CIRCULAR STATISTICS METHODS

Circular Statistical Analysis of Avian Directional Data for the New Jersey Offshore (Ship) Surveys: Methods

OVERVIEW

The objective of the following first-order and second-order circular statistical analyses is to determine the nature of the statistical distribution of avian flight directions and their variability with respect to species, taxonomic group, month, and season. The hypothesis tests are conducted to determine whether the flight directions exhibit circular uniformity (*i.e.*, a random distribution of direction) or exhibits a mean angle (*i.e.*, favoring a particular direction), both overall as well as for a given species, group, month, and season. For example, it may be feasible for one avian species to exhibit circular uniformity, whereas another species may exhibit a mean flight direction that may change from month to month or from season to season, reflecting, for example, seasonal flight migration patterns. This analysis is potentially important for subsequent calculations of avian mortality strikes, since the collision rate of birds with the wind turbine blades depends on the relative directional orientation between the birds and blades.

DESCRIPTIVE STATISTICS

Angular (directional) data consist of a sample of "n" measurements of angle (with values ranging from 0-360^o). Angular measurements (a_i for data point i, where i=1 to n) are often grouped into "J" bins, with each bin "j" defining a (usually equally-spaced) range of angles (*e.g.*, 0-10^o, 10-20^o, *etc.*), such that all J bins encompass the entire angular range 0-360^o. Each bin "j" is characterized by a midpoint angle b_j (*e.g.*, 5^o, 15^o, *etc.*). The number of measured angles (data points) a_i falling into each bin "j" is termed the angular number frequency, occurrence frequency, or number of angular measurements (f_j). Summing f_j across all J bins yields the total sample size n (*i.e.*, SUM (f_j) = n, where the summation is conducted for j=1 to J, where J = total number of bins, which encompass the entire angular range 0-360^o).

Angular data can thus be individual or grouped:

1) Individual data consist of individual angular values a_i (for measurement "i"), for i=1 to n, where n = sample size or total number of angular measurements.
2) Grouped data consist of an occurrence frequency f_j (*a.k.a.* number frequency or number of angular measurements) of individual angular measurements a_i occupying bin "j", which is defined by an angular range and a midpoint angle b_j. A total of J bins are defined whose collective ranges cover the entire 0-360^o range. Each of the a_i individual measurements (for i=1 to n) are distributed into the J bins, such that SUM (f_j) = n.

Individual and grouped angular datasets (*i.e.*, circular distributions) can be characterized based on uniformity and modality. Uniform distribution exhibit a relatively even (or random) distribution of angular measurements over the entire circular range (0-360^o), whereas a non-uniform (or clumped or aggregated) distribution tends to have a greater relative number of angular measurements concentrated around one or more particular angles or ranges of angles.

Three general types of modal distributions include unimodal, bimodal, and multimodal:

1) Unimodal distributions are characterized by a maximum occurrence frequency at one angular range (*i.e.*, the occurrence frequency for one angular bin is greater than the number frequencies for all other angular bins).
2) Bimodal distributions have a maximum occurrence frequency at 2 angular ranges (*i.e.*, 2 bins exhibit the same maximum occurrence frequency). An example is a diametrically bimodal distribution, which exhibits a maximum occurrence frequency at 2 angles spaced exactly 180^o apart (*i.e.*, on opposite sides of the circle).
3) Multi-modal distributions have several angular bins exhibiting the same maximum occurrence frequency.

The angle(s) at which the maximum occurrence frequency occurs is termed the modal angle(s). In addition to the modal angle, circular distributions are also characterized by a mean angle and median angle.

MEAN ANGLE

Angular data take the form of individual angle measurements (a_i) over a sample size n (*i.e.*, i=1 to n). Mean angle (a_{bar}) for a sample of size n is calculated from the summation of the sine and cosine vectors of the individual angle measurements a_i. For each of the n measurements, given a_i (for i=1 to n), $\sin(a_i)$ and $\cos(a_i)$ are calculated and summed over n. The values of X and Y are then calculated as: $X = SUM[\cos(a_i)]/n$ and $Y = SUM[\sin(a_i)]/n$, where the summation is from i=1 to n. The length vector r is then calculated as $r = (X^2 + Y^2)^{0.5}$. Knowing X, Y, and r, the cosine and sine components of the mean angle are given by $\cos(a_{bar}) = X/r$ and $\sin(a_{bar}) = Y/r$.

Solving $\cos(a_{bar})$ by itself for a_{bar} yields 2 solutions for a_{bar}. Likewise, solving $\sin(a_{bar})$ by itself for a_{bar} yields 2 solutions for a_{bar}; however, solving both equations simultaneously yields a unique solution for a_{bar}, according to the following geometric and trigonometric guidelines:

1) If $\cos(a_{bar}) > 0$ and $\sin(a_{bar}) > 0$, then $0^o < a_{bar} < 90^o$.
2) If $\cos(a_{bar}) < 0$ and $\sin(a_{bar}) > 0$, then $90^o < a_{bar} < 180^o$.
3) If $\cos(a_{bar}) < 0$ and $\sin(a_{bar}) < 0$, then $180^o < a_{bar} < 270^o$.
4) If $\cos(a_{bar}) > 0$ and $\sin(a_{bar}) < 0$, then $270^o < a_{bar} < 360^o$.

Grouped angular data take the form of occurrence frequencies (f_i) of the given angle measurement (a_i), recorded over a sample size n (*i.e.*, for i=1 to n). The recorded angular value (a_i) is typically the midpoint of an angular measurement interval (or bin), and f_i is the occurrence frequency (*i.e.*, number of angular measurements) within the angular interval defined by the midpoint angle a_i. For example, for 2 bins defined by the angular ranges 0-10^o and 10-20^o, the midpoint angles a_i are 5^o and 15^o, respectively.

In the case of angular data grouped into "N" bins (or a dataset consisting of an occurrence frequency f_i associated with a distinct angular value a_i), an angle a_i occurring with a frequency f_i is equivalent to repeating the value of a_i, f_i times and recording each individual occurrence of a_i in an expanded table. In this expanded table, the expanded sample size $n = SUM(f_i)$, where the summation is from i=1 to N, where N = number of DISTINCT angular values, and n = total number of angular measurements (sample size). Working with the original data table (consisting of distinct angular values a_i and associated occurrence frequencies f_i), the values of $\sin(a_i)$, $f_i*\sin(a_i)$, $\cos(a_i)$, and $f_i*\cos(a_i)$ are calculated for each measurement "i". The values of $f_i*\sin(a_i)$ and $f_i*\cos(a_i)$ are summed over the number of distinct angular value N. Values of X and Y are then calculated as: $X = SUM[f_i*\cos(a_i)]/n$ and $Y = SUM[f_i*\sin(a_i)]/n$, where the summation is from i=1 to N, and $n = SUM(f_i)$. The length vector $r = (X^2 + Y^2)^{0.5}$ is then calculated. Knowing X, Y, and r, the cosine and sine components of the mean angle are given by $\cos(a_{bar}) = X/r$ and $\sin(a_{bar}) = Y/r$. Simultaneous solution of these 2 equations yields a unique solution for a_{bar}.

However, calculation of r from grouped data (*i.e.*, data in the form of occurrence frequencies f_i associated with distinct angular values a_i) results in negative bias, generating an underestimate of r. The following correction for this bias is applicable when the distribution is unimodal and does not deviate significantly from symmetry. The r value is multiplied by a correction factor "c" to yield the corrected r value r_c: $r_c = c*r$, where $c = (d*pi/360^o)/[\sin(d/2)]$, where the data are grouped into intervals of "d" degrees each.

MEASURES OF ANGULAR DISPERSION

Classical measures of dispersion (or spread) of the data include range, standard deviation, standard error, variance, and confidence intervals; and these measures (as well as the mean, median, and mode) are analogously applicable to circular data as well as linear data. Specific circular measures of angular dispersion include range, circular variance, angular variance, variance measure, angular deviation, and circular standard deviation.

Because of the discontinuity in angular value between 0^o and 360^o (designating the same point on a circle), the range cannot simply defined as the difference between the lowest and highest angular value. Instead, the range in a circular distribution of angular data values is defined as the smallest arc of the circle that contains all of the angular data in the distribution.

In addition to range, other circular measures of angular dispersion include:

1) Circular variance: $S^2 = 1\text{-}r$, ranging from 0 to 1.
2) Angular variance is $s^2 = 2*(1-r) = 2*S^2$, ranging from 0 to 2.
3) Variance measure: $s_o^2 = -2*\ln(r)$, ranging from 0 to infinity.
4) Angular deviation: $s = (180/pi)*[2*(1-r)]^{0.5} = (180/pi)*(s^2)^{0.5}$, ranging from 0^o to 81.03^o.
5) Circular standard deviation: $s_o = (180/pi)*[-2*\ln(r)]^{0.5} = (180/pi)*(so^2)^{0.5}$, ranging from 0 to infinity.
6) Circular mean deviation: CMD = $(SUM|a_i - a_{bar}|)/n$, where a_i = angular value of measurement "i", a_{bar} = mean angle, and the summation is from i=1 to n.

CONFIDENCE INTERVAL FOR POPULATION MEAN

The lower (L_1) and upper (L_2) confidence limits for the mean angle (a_{bar}) are given by $L_1 = a_{bar} - d$ and $L_2 = a_{bar} + d$, where:

$d = \arccos[(1/R)*\{[2*n*(2*R^2 - n*X^2)]/(4*n - X^2)\}^{0.5}]$ for r<=0.9, $R>(n*X^2/2)^{0.5}$, and $n>X^2/4$
$d = \arccos\{(1/R)*[n^2 - (n^2 - R^2)*\exp(X^2/n)]^{0.5}\}$ for r>=0.9

where R = n*r.

SECOND ORDER ANALYSIS: THE MEAN OF MEAN ANGLES

If the previously-described procedure of calculating the mean angle a_{bar} (along with length vector r, X, and Y) from a group of measured angles a_i (and their associated occurrence frequencies f_i) is repeated for several groups of measured angular data, then the result is a set of mean angles. For example, for each of “k” samples of circular (angular) data, values of a_{bar}, r, X, and Y are calculated via the above procedure. It is now desired to calculate the grand mean of these “k” mean angles. For each of the “k” samples, given a_{bar} and r, values of $X = r*\cos(a_{bar})$ and $Y = r*\sin(a_{bar})$ are calculated. These values of $r*\cos(a_{bar})$ and $r*\sin(a_{bar})$ are then summed over the “k” samples. The values of X_{bar} and Y_{bar} are then calculated as: $X_{bar} = SUM[r*\cos(a_{bar})]/k$ and $Y_{bar} = SUM[r*\sin(a_{bar})]/k$, where the summation is from i=1 to k. Then the length vector $r = (X_{bar}^2 + Y_{bar}^2)^{0.5}$ is calculated. Knowing X, Y, and r, the cosine and sine components of the grand mean angle are given by $\cos(a_{bar}) = X_{bar}/r$ and $\sin(a_{bar}) = Y_{bar}/r$. Simultaneous solution of these 2 equations yields a unique solution for a_{bar}.

CONFIDENCE LIMITS FOR THE SECOND-ORDER MEAN ANGLE

The precision of the above-calculated second order mean angle (*i.e.*, grand mean angle) can be quantified by the calculation of confidence limits. The following equations are used to calculate b_1 and b_2, which are instrumental in calculating the lower and upper confidence limits:

$A = (k-1)/[SUM(x^2)]$
$B = -[(k-1)*SUM(x*y)]/[SUM(x^2) * SUM(y^2)]$
$C = (k-1)/[SUM(y^2)]$
$D = 2*(k-1)*\{1 - [SUM(x*y)^2]/[SUM(x^2) * SUM(y^2)]\} * F(a(1),2,k-2)/[k*(k-2)]$
$H = A*C - B^2$
$G = A*X_{bar}^2 + 2*B*X_{bar}*Y_{bar} + C*Y_{bar}^2 - D$
$U = H*X_{bar}^2 - C*D$
$V = (D*G*H)$
$W = H*X_{bar}*Y_{bar} + B*D$
$b_1 = (W+V)/U$
$b_2 = (W-V)/U$

where $SUM(x^2) = SUM(X^2) - [SUM(X)]^2/k$, $SUM(y^2) = SUM(Y^2) - [SUM(Y)]^2/k$, and $SUM(x*y) = SUM(X*Y) - [SUM(X*Y)]/k$.

The values of b_1 and b_2 each yields one of the confidence limits (*i.e.*, either the lower limit or the upper limit). An M value is calculated for each of the b_i values as follows: $M = (1+b_i^2)^{0.5}$. Then sine = b_i/M and cosine = 1/M are calculated. A unique angle is obtained from the calculated "sine" and "cosine" value. Then, one of the confidence limits is either this unique angular value or the value of this angle plus or minus 180^o (whichever is greater than 0^o and less than 360^o), whichever is closer to the mean angle.

TESTS FOR CIRCULAR UNIFORMITY

It is often desired to determine the nature of the distribution of the n angles (a_i) around the circle of 0-360^o. Are the directional data (either individual or grouped into angular bins) distributed relatively uniformly (*i.e.*, exhibit circular uniformity) around the circle (from 0^o to 360^o), or do the data exhibit directional bias by favoring one or more particular angles or ranges of angles? This question can be answered via statistical tests for circular uniformity, where the null (H_o) and alternative (H_a) hypotheses are stated as:

H_o: The sample came from a population with a uniform circular distribution (no directional bias).
H_a: The sample came from a population exhibiting bias in one or more directions (*i.e.*, a non-uniform circular distribution).

Among the available parametric tests for circular uniformity, the Rayleigh test and modified Rayleigh test (or V-test) for significance of the mean angle assume a unimodal distribution, whereas the Hodges-Ajne, and Batschelet Omnibus tests for significance of the median angle do not have such a restrictive assumption (*i.e.*, can be applied to unimodal, bimodal, and multi-modal distributions).

RAYLEIGH TEST AND V-TEST FOR SIGNIFICANCE OF THE MEAN ANGLE (UNIMODAL DISTRIBUTIONS)

The parametric Rayleigh test is used to test whether the given sampled population is uniformly distributed around a circle (*i.e.*, possesses circular uniformity, having no mean direction), and is valid only for unimodal populations following a von Mises (VM) angular normal distribution. Non-unimodal (*e.g.*, axially bimodal) datasets must first be transformed into unimodal data before Rayleigh's test can be applied to test for circular uniformity. Axially bimodal data, where the 2 angles with the highest occurrence frequencies are spaced exactly 180^o apart, can be transformed into unimodal data by simply doubling all of the angles. "Rao's spacing test" can be used when the circular data are neither unimodal nor axially bimodal (Batschelet, 1981: 66-69; Rao, 1976; Russell and Levitin, 1994). Populations not following the von Mises distribution must undergo the following procedure before Rayleigh's test can be applied:

1) Convert the angular data from angular scale to a linear scale.
2) Conduct statistic tests to determine if the data are normal (*e.g.*, Kolmogorov-Smirnov test) and homoscedastic (possess homogeneity of variances) (*e.g.*, Bartlett test).
3) Apply an appropriate data transformation (*e.g.*, logarithmic, arcsin, square root) to the resultant linear data if the data are found to be non-normal or non-homoscedastic.
4) Convert the transformed linear data back to the angular scale.

Once the directional dataset is confirmed to be unimodal and follow the VM angular normal distribution, then the parametric Rayleigh test can be applied to test for circular uniformity. The null (H_o) and alternative (H_a) hypotheses and assumptions of the Rayleigh test are as follows:

H_o: The population has a uniform circular distribution with no mean direction.
H_a: The population has a non-uniform circular distribution with a mean direction.

"Rayleigh's R", given by $R = n*r$ (where n = sample size), can be used to test how large a sample r must be to confidently indicate a non-uniform population distribution. "Rayleigh's z", given by $z = R^2/n = n*r^2 = R*r$ can be used to test the null hypothesis (H_o). Critical values of z as a function of sample size n and confidence level "a" are tabulated in published statistical tables. An approximation of the probability of Rayleigh's R is given by $P = \exp\{[(1 + 4*n + 4*(n^2 - R^2)]^{0.5} - (1+2*n)\}$. If Rayleigh's z is greater than the critical z value, then H_o is rejected. Otherwise, H_o is accepted.

The non-parametric V-test (Greenwood and Durand, 1955; Durand and Greenwood, 1958) is a modification of the Rayleigh test such that a specific mean direction or expected mean angle (*e.g.*, 90^o) is specified when testing the null hypothesis (H_o). The H_o and H_a hypotheses are given by:

H_o: The population has a uniform circular distribution with no mean direction, OR has a non-uniform circular distribution with a mean direction that is different from the specified mean direction.
H_a: The population has a non-uniform circular distribution with a specified mean direction.

Because a mean direction is specified, the V test is more powerful than the Rayleigh test (Batschelet, 1972; 1981: 60). For the Rayleigh test, rejection of H_o indicates only that the population has a mean direction, without indicating the specific mean direction. In contrast, for the V test, rejection of H_o provides additional information on the specific mean direction; however, if the data has an actual mean direction that is different from the specified (tested) mean direction, the test may accept H_o for 2 reasons:

1) Either the population has a uniform circular distribution, OR,
2) The population has a non-uniform circular distribution with a mean direction that is different from the specified mean direction.

The test statistic for the V test is given by $V = R*\cos(a_{bar} - u_o)$, where u_o is the predicted mean angle. The significance of V is determined from the "u" statistic: $u = V*(2/n)^{0.5}$. Critical values of u for different values of sample size n and confidence level "a" are tabulated in published statistical tables. As sample size n increases, the "u" statistic approaches a 1-tailed normal deviate $Z_a(1)$. If the data are grouped, then R is calculated from r_c rather than from r. If the u statistic is greater than the critical u value, then H_o is rejected. Otherwise, H_o is accepted.

PARAMETRIC 1-SAMPLE TEST FOR MEAN ANGLE

The parametric 1-sample test for mean angle (analogous to the 1-sample t-test for data on a linear scale) is used to test whether the population mean angle (u_a) is equal to a specified value (u_o). The H_o and H_a hypotheses are given by:

H_o: The population has a mean angle (u_a) equal to the specified mean angle (u_o).
H_a: The population's mean angle (u_a) is NOT equal to the specified mean angle (u_o).

This test calculates the (1-a) confidence interval (CI) for u_a and determines whether u_o falls within this CI. H_o is rejected if u_o is outside the CI, and is accepted if u_o falls within the CI.

HODGES-AJNE, AND BATSCHELET OMNIBUS TESTS FOR SIGNIFICANCE OF THE MEDIAN ANGLE (UNIMODAL, BIMODAL, AND MULTI-MODAL DISTRIBUTIONS)

The Hodges-Ajne and Batschelet tests are omnibus tests in that they do not require the assumption of a unimodal data distribution (unlike the Rayleigh and V-tests). That is, these omnibus tests are applicable to unimodal, bimodal, and multimodal distributions; however, for unimodal distributions, the Rayleigh and V tests are more powerful than these omnibus tests.

The null (H_o) and alternative (H_a) hypotheses for the Hodges-Ajne test are given by:

H_o: The population has a uniform circular distribution.
H_a: The population has a non-uniform circular distribution.

The Hodges-Ajne test (Ajne, 1968) is applied by calculating, from the given sample of circular data of sample size n, the smallest number of data (m) falling within a range of 180^o. The data are plotted on a circle, a diameter line is drawn through the circle (*e.g.*, connecting 0^o and 180^o), and the number of data on each side of the line are counted. The diameter line is rotated by a certain angular resolution (*e.g.*, by 1^o, so that the new line intersects 1^o and 181^o), and the above procedure is repeated until all angle pairs (spaced 180^o apart) are covered. Thus, for each angle pair defining the diameter line, the numbers of data on each side of the line are counted (with the smaller of these 2 numbers being designated as “m”, and the other number being “n-m”). The smallest “m” value is desired and is used as the “m” test statistic in running the Hodges-Ajne test.

The probability P of an m value at least this small is given by (Hodges *et al.*, 1955): $P = 2^{(1-n)} * (n-2m)*{}_nC_m$, where ${}_nC_m = n!/[m! * (n-m)!]$. Critical values for m as a function of sample size n (up to n=50) and confidence level “a” are tabulated in published statistical tables. For larger sample sizes (n>50), P is calculated as $P = [(2*pi)^{0.5}/A]*exp[-pi^2/(8*A^2)]$, where $A = pi*(n^{0.5})/[2*(n-2*m)]$ and pi = 3.1415927. If the m statistic is less than the critical m value, then H_o is rejected. Otherwise, H_o is accepted.

The non-parametric Batschelet test (Batschelet, 1981: 64-66) is a modification of the Hodges-Ajne test such that a specific angle is specified when testing the null hypothesis (H_o). The H_o and H_a hypotheses are given by:

H_o: The population has a uniform circular distribution, OR has a non-uniform circular distribution that is concentrated around an angle that is different from the specified concentration angle.
H_a: The population has a non-uniform circular distribution that is concentrated around a specified concentration angle.

Because a concentration angle is specified, the Batschelet test is more powerful than the Hodges-Ajne test. For the Hodges-Ajne test, rejection of H_o indicates only that the population has a concentration angle, without indicating the specific angle. In contrast, for the Batschelet test, rejection of H_o provides additional information on the specific angle; however, if the data has an actual concentration angle that is different from the specified (tested) concentration angle, the test may accept H_o for 2 reasons:

1) Either the population has a uniform circular distribution, OR,
2) The population has a non-uniform circular distribution with a concentration angle that is different from the specified concentration angle.

The test statistic for the Batschelet test is given by C = n – m’, where n = sample size and m’ = number of data within 90^o of the specified angle. A 2-tailed binomial test is then conducted, with p = 0.50 and with C counts in one category and m' counts in the other.

Critical values for C as a function of sample size n, number of tails (1 or 2), and confidence level “a” are tabulated in published statistical tables. If the C statistic is less than the critical C value, then H_o is rejected. Otherwise, H_o is accepted.

TEST FOR SIGNIFICANCE OF THE MEDIAN ANGLE: BINOMIAL TEST

This non-parametric test (Zar, 1999) can be used to determine whether the population median angle is equal to a specified value. The H_o and H_a hypotheses are given by:

H_o: The population median angle is equal to the specified value.
H_a: The population median angle is NOT equal to the specified value.

The angular data are plotted on a circle, a diameter line is drawn through the specified angle, and the data on each side of this line are counted. A 2-tailed binomial test is then conducted, with p = 0.50.

TEST FOR SYMMETRY AROUND THE MEDIAN ANGLE: NON-PARAMETRIC WILCOXON PAIRED-SAMPLED (SIGNED-RANK) TEST

The non-parametric Wilcoxon paired-sample test (or signed-rank test) is used to test the symmetry of a distribution around the median, using either 1-tailed or 2-tailed tests (Zar, 1999). The H_o and H_a hypotheses are given by:

2-tailed test:

H_o: The underlying distribution is symmetrical around the median.
H_a: The underlying distribution is NOT symmetrical around the median.

1-tailed test (with "T-" as test statistic):

H_o: The underlying distribution is NOT skewed clockwise from the median.
H_a: The underlying distribution is skewed clockwise from the median.

1-tailed test (with "T+" as test statistic):

H_o: The underlying distribution is NOT skewed counter-clockwise from the median.
H_a: The underlying distribution is skewed counter-clockwise from the median.

For each angle (X_i), the deviation of X_i from the median angle (X_{median}) is calculated: $d_i = X_i - X_{median}$. The d_i values are ranked over the number of data points (*i.e.*, sample size n), and the ranks are "signed" based on the sign of the deviation (*i.e.*, positive deviations d_i have a positive rank, and negative deviations d_i have a negative rank). All of the positive signed ranks over the sample size n are summed to obtain the "T+" value. Likewise, all of the negative signed ranks are summed, and the absolute value of this sum is calculated, to obtain the "T-" value. Critical values T_{crit} as a function of sample size n (up to n=100) and confidence level "a" for both 1-tailed and 2-tailed tests are tabulated in published statistical tables. For the 2-tailed test, H_o is rejected if either "T+" or "T-" is less than T_{crit}, and is accepted if neither "T+" nor "T-" is less than T_{crit}. For the 1-tailed test, we use either "T+" or "T" as the test statistic T_{stat}. H_o is rejected if $T_{stat} < T_{crit}$, and is accepted if $T_{stat} >= T_{crit}$.

GOODNESS-OF-FIT (GOF) TESTING FOR CIRCULAR DISTRIBUTIONS: NON-PARAMETRIC CHI-SQUARE (X^2) TEST, WATSON'S 1-SAMPLE U^2 TEST, KOLMOGOROV-SMIRNOV (KS) TEST, AND KUIPER'S TEST

Goodness of fit (GOF) tests are used to test the statistical significance of conformity of a given sample of nominal scale data (*e.g.*, counts, abundances) to a specified or expected frequency distribution. A test statistic is used to quantify the deviation of the sample from the specified theoretical distribution, and is compared to an associated tabulated critical value, the latter of which is a function of the degrees of freedom (DF) and confidence level of interest (*e.g.*, 90%, 95%, 99%). From the comparison between the test statistic and critical value, a probability is calculated along with a decision of whether to accept or reject the null hypothesis (H_o) of no difference. For GOF tests, H_o and the alternative hypothesis (H_a) can be generally stated as follows:

H_o: The sample came from a population that follows the specified theoretical (expected) frequency distribution.
H_a: The sample did NOT come from a population that follows the specified theoretical (expected) frequency distribution.

For the case of testing for uniform circular uniformity, H_o and H_a would be stated as:

H_o: The sample came from a population with a uniform circular distribution.
H_a: The sample did NOT come from a population with a uniform circular distribution.

The critical value is a threshold such that values of the test statistic in excess of the threshold represent sufficiently high deviations from the theoretical distribution to warrant a conclusion that the sample does not come from a population that follows the theoretical distribution (*i.e.*, H_o is rejected). Conversely, if the test statistic is less than the critical value (threshold), then it can be concluded that the deviation of the sample from the theoretical distribution is sufficiently small as to conclude that the sample did indeed come from a population that follows the theoretical distribution (*i.e.*, H_o is accepted, or not rejected). In summary:

1) If test statistic > critical value: H_o is rejected.
2) If test statistic < critical value: H_o is not rejected.

For the circular statistical analysis of bird directional data, the non-parametric chi-square (X^2) GOF test (Zar, 1999), Watson's 1-sample U^2 GOF test (Watson, 1961; 1962), the Kolmogorov-Smirnov (KS) GOF test (Zar, 1999), and Kuiper's GOF test are used to test the null hypothesis (H_o) of a uniform circular distribution. The H_o and H_a hypotheses are stated as:

H_o: The sample came from a population that follows a uniform circular distribution.
H_a: The sample did NOT come from a population that follows a uniform circular distribution.

CHI-SQUARE (X^2) GOF TEST

A typical sample of nominal scale data consists of "k" classes or categories, each of which is quantified by a frequency (or number of counts). In the current application of bird directional data, the nominal directional classes include the 8 directions (*i.e.*, k=8): N (0^o or 360^o), NE (45^o), E (90^o), SE (135^o), S (180^o), SW (225^o), W (270^o), and NW (315^o). The non-parametric chi-square (X^2) GOF test (Zar, 1999) uses the following test statistic to measure the deviation of the sample from the specified theoretical distribution: X^2 = SUM[f_i – $f_{hat,i}$] ^{2}/$f_{hat,i}$, where f_i and $f_{hat,i}$ are the actual and expected frequencies, respectively, of class "i", and the summation is conducted over the "k" classes. The f(i) values are obtained directly from the sample, and the $f_{hat,i}$ values are calculated based on the theoretical frequency distribution specified in the null hypothesis (H_o). Two checks on the calculations are the following: 1) SUM[f_i] = n; 2) SUM[$f_{hat,i}$] = n, where n = sample size. The X^2 test statistic is compared to the critical X^2 value (tabulated in published statistical tables) for k-1 degrees of freedom (*i.e.*, DF=k-1) for the desired confidence level CL (*e.g.*, 90%, 95%, 99%). The critical X^2 value correlates positively with DF and negatively with CL. If the X^2 statistic is greater than the critical X^2, then H_o is rejected. Otherwise, H_o is accepted.

WATSON'S 1-SAMPLE U^2 GOF TEST

The non-parametric Watson's 1-sample U^2 GOF test (Watson, 1961; 1962) normalizes the angular directional data from a 0-360^o range to a range of 0-1, via the transformation: $u_i = a_i/360^o$, where a_i = angular measurement (degrees) of data point "i", and u_i = normalized value (dimensionless, range=0-1). The U^2 statistic (Mardia, 1972: 182) is calculated as: U^2 = SUM u_i^2 – (SUM $u_i)^2$/n – (2/n)*SUM(i*u_i) + (n+1)*u_{bar} + n/12, where n = sample size; the summations are conducted for i=1 to n; and the mean value u_{bar} = SUM(u_i)/n. The U^2 test statistic is compared to the critical $U^2[a,n_1,n_2]$ value (tabulated in published statistical tables) for n_1=n and n_2=n for the desired confidence level "a" (*e.g.*, 90%, 95%, 99%). The critical U^2 value correlates negatively with CL and n_1 and can correlate either positively or negatively with n_2. If the U^2 statistic is greater than the critical U^2, then H_o is rejected. Otherwise, H_o is accepted.

KOLMOGOROV-SMIRNOV (KS) AND KUIPER GOF TESTS

The non-parametric Kolmogorov-Smirnov (KS) GOF test (Zar, 1999) represents an improvement over the X^2 GOF test for ordered categories of data, and is more powerful than the X^2 test for small sample sizes (n) and/or small expected frequencies; however, a drawback is that the results of the KS test depend on the order of the categories, yielding different results for different starting points on a circle. Fortunately, this shortcoming is resolved with the Kuiper GOF test, which generates results that are unrelated to the starting point. The Kuiper test is preferred over the X^2 test for ungrouped data.

For each sample “i” (for i=1 to n), cumulative observed frequency F_i is calculated as the summation of observed frequency f_i as “i” is increased from 1 to n; and cumulative relative frequency (CRF_i) is simply F_i/n. The test statistic is given by D = max[max(D_i), max(D_i’)], where D_i = abs[CRF(i) – CRF_{exp}(i))] and D_i’ = abs[CRF(i-1) – CRF_{exp}(i)]. Thus, D represents the maximum deviation between the observed frequency distribution and the expected (hypothesized) frequency distribution. The D test statistic is compared to the critical D[a,n] value (tabulated in published statistical tables) for sample size “n” and the desired confidence level “a” (*e.g.*, 90%, 95%, 99%). The critical D value correlates negatively with both n and “a”. If the D statistic is greater than the critical D, then H_o is rejected. Otherwise, H_o is accepted.