An Examination of the Kuiper Test

Esther Birch

Department of Mathematical Sciences
Montana State University

May 3, 2018

A writing project submitted in partial fulfillment
of the requirements for the degree

Master of Science in Statistics

APPROVAL

of a writing project submitted by

Esther Birch

This writing project has been read by the writing project advisor and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Statistics Faculty.

Date	John J. Borkowski Writing Project Advisor
Date	Mark C. Greenwood Writing Projects Coordinator

Contents

Abstract

Circular data, or directional/angular data, has applications in a wide variety of fields, including biology, meteorology/oceanography, and social sciences.

In this paper, I examine Kuiper's Test for Uniformity, primarily focusing on its coding issues in both the `circular` and `circStats` packages in R, the weak documentation surrounding this test, both in the R packages and in N.I. Fisher's `Statistical Analysis of Circular Data` textbook, as well as potential inferential consequences that may result from these erroneous sources.

1 Introduction

1.1 Overview of Circular Data

Circular data describes data that are measured around a circle or a compass. Commonly referred to as directional or angular data, circular data can be recorded in several different units, depending on the application. In fields such as biology or geology, measurements are often recorded in degrees, such as when recording the vanishing direction of homing pigeons upon release (Figure 1) or the orientations of linear geologic features, such as feldspar laths (Figure 2). Circular data can also be measured in angles, radians, polar coordinates, or in units of time, such as hours, days, or months.

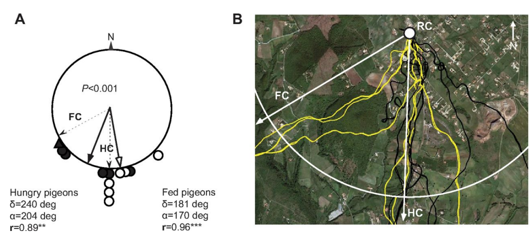

Figure 1: Vanishing directions of homing pigeons using GPS trackers [5]

1.2 Visualizations

Figure 2: Feldspar laths in igneous rock [1]

Another interesting feature of circular data is that measurements are also distinguished as "axial" or "vectorial." Vectors are considered to be lines with direction, such as in the homing pigeon example, while axial lines are "undirected." As is the case with feldspar laths, geologic fractures, and other axial data, there is little to no reason to differentiate one end of the line from the other [6].

Circular data can be plotted several different ways. One of the most common plots for this type of data is a "rose plot." Attributed to Florence Nightingale, who first wrote about these "coxcomb" plots in 1858 [6], the rose plot is very similar to a linear histogram. However, rather than the height of the sectors, or bins, representing the frequency in that sector, rose plots scale the height of the segments by taking the square root of the relative frequency. Thus, the area of that sector is proportional to the frequency of observations in that bin. As with linear histograms, rose plots are not unique; the data can be plotted with different bins or sector widths. Rose plots are frequently used for meteorological data (referred to as wind roses), measuring both the direction and frequency of winds; these plots often also add colored segments to include information about wind speeds (Figure 3).

Dotplots are also used with circular data, often in conjunction with rose plots and kernel density estimates (Figure 4). Circular dotplots are simply linear dotplots of the data, wrapped around the circumference of a circle. Kernel density estimates, which N.I. Fisher describes as a "sort of moving average" [6], describe the sum of the contributions of observations at a given point on the circle (θ), where each observation has a weight contribution of $1/n$, where n is the number of observations in the dataset.

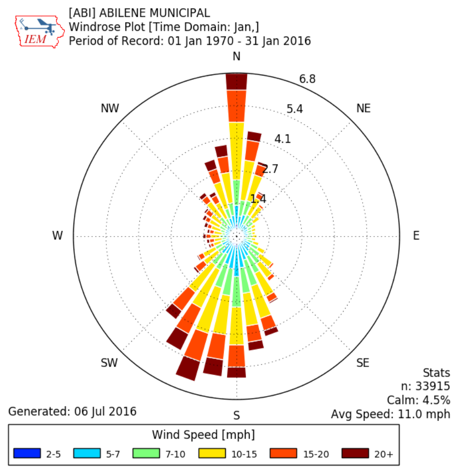

Figure 3: Wind rose for Abilene, TX, for January 1970 to January 2016 [17]

1.3 Summary Measures

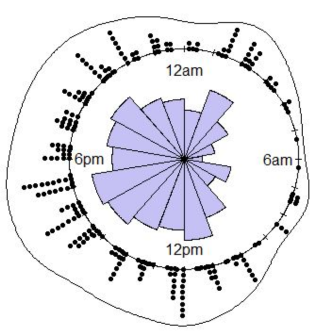

Figure 4: Rose plot of emergency room arrival times with combined with a dotplot and a kernel density estimate. Data acquired from Fisher, 1993 [6]

Another challenge unique to circular data is the calculation of summary measures such as the sample mean and variance. Taking the mean of circular data is not as simple as finding that of linear data. Rather, summary statistics for circular data are calculated using polar coordinates and trigonometric moments. Fisher describes the issue with trying to take the arithmetic mean of circular data in his 1993 textbook, *Statistical Analysis of Circular Data*, by giving the example of three data points, 359°, 1°, and 3°, whose arithmetic mean would be 121°. However, the mean direction ($\bar{\theta}$), or vector resultant, is 1°, calculated as:

$$\bar{\theta} = \frac{\sum_{i=1}^{n} cos\,\theta_i}{\sqrt{\left(\sum_{i=1}^{n} cos\,\theta_i\right)^2 + \left(\sum_{i=1}^{n} sin\,\theta_i\right)^2}}$$

The sample variance of circular data (V) is likewise found using trigonometric moments:

$$V = 1 - \frac{\sqrt{\left(\sum_{i=1}^{n} cos\,\theta_i\right)^2 + \left(\sum_{i=1}^{n} sin\,\theta_i\right)^2}}{n}$$

It thus follows that other summary measures, such as the sample median direction, sample skewness, and sample kurtosis are also calculated similarly via trigonometric moments.

1.4 Circular Probability Distributions

There are several distributions that are specifically suited or adapted for use with circular data, including the von Mises, wrapped normal, and uniform distributions. The von Mises distribution is a two-parameter, symmetric distribution, analogous to the linear normal distribution [11], and is the most commonly used unimodal distribution for directional data [6]. It closely approximates the wrapped normal distribution, which is a normal distribution "wrapped" around a unit circle. However, as Fisher notes, the von Mises distribution is mathematically simpler and easier to use, lending to its popularity over the wrapped normal. Indeed, when the concentration parameter $\kappa = 0$ (the circular equivalent of the variance parameter), the von Mises distribution is uniform [9].

The circular uniform distribution is as straightforward as its linear counterpart; all values around a circle are equally likely to occur. This distribution is most commonly used as a null model for testing against a variety of alternative models, including uni- and multimodal models [6].

2 Goodness-of-Fit Tests

When working with circular data, it is often extremely difficult to discern by eye alone if the data are uniformly distributed (i.e. if the data are randomly distributed). Goodness-of-fit tests are often employed to attempt to determine whether the directions of the data recorded are isotropic or from another non-uniform distribution.

There are several single-sample goodness-of-fit tests available for circular data. Among the most commonly used are the Rayleigh test, Rao's spacing test, Watson's U^2 test, and the Kuiper test. With the exception of the Rayleigh test, all are omnibus tests used when the assumption of a unimodal underlying distribution cannot be confidently made. Additionally, these tests are rotation-invariant. That is, the "starting" point of the data will not affect the outcome of the test.

The Rayleigh test is used specifically to test whether data are from a unimodal distribution, such as the von Mises distribution. The Rayleigh test is the uniformly most powerful invariant test [9] when the data are assumed to be unimodal, performing at its highest power when the data are indeed from a von Mises distribution. The basic premise of the test is that the length of the mean direction vector will indicate whether one-sidedness exists in the data; if the mean direction vector is small, there is little evidence for unimodality. The Rayleigh test also has an alternate form which tests the length of the mean direction vector of the sample data against a hypothesized value, but this form is much less commonly used [14]. In a 2017 article for *Biology Letters*, Graeme D. Ruxton notes that either form is "sufficiently high performing over such a broad range of shapes of unimodal distribution that no alternative test is in common usage [14]."

When the data are not assumed to be unimodal, however, there are several omnibus tests that perform well. As explained in *Circular Statistics in R*, none of the three most commonly used (Rao's spacing test, Watson's U^2, or Kuiper's test) are inherently better performing than another under all circumstances, and none have very high power [13]. However, when the departures from normality observed in the data are not unimodal, all three tests have higher power than the Rayleigh test [13].

Rao's spacing test examines the arc length between two adjacent points in the data. Under the assumption of randomness, points should be approximately evenly spaced around the circle, so the average distance between two neighboring points should be of length $360°/n$. Large deviations from this value are considered to be evidence of non-uniformity [4]. The only assumption required for this test is that the data are not grouped, or collected in measured intervals.

Watson's U^2 test is a Cramér-von Mises type test, which tests whether the data are from a hypothesized distribution by comparing the empirical distribution of the data against the hypothesized cumulative distribution. If the mean square deviation of the data are greater than a critical value [15], the null hypothesis of uniformity is rejected. Watson's U^2 test is extremely similar to the Kuiper test, in that both examine deviations from uniformity; the difference is that Watson's test uses the squared deviations, while the Kuiper test does not.

3 Kuiper's Test

Kuiper's test was developed in 1960 by Nicholaas Kuiper as the result of a problem posed by a fellow professor, who asked him how to not only define and estimate the degree of orientation of a hypothetical group of birds, but how to test the null hypothesis that said birds have no preference of direction [7].

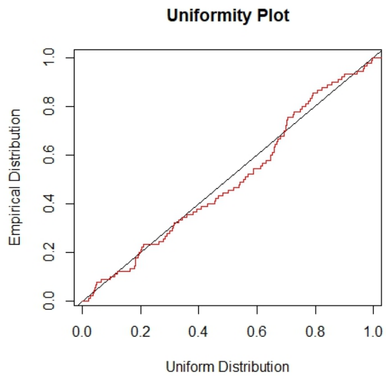

Figure 5: Example of simulated uniform data compared to the uniform CDF

Kuiper, whose primary work was in differential geometry [12], adapted the Kolmogorov-Smirnov test so it could be applied to points around a circle as opposed to points on a line. The Kuiper test takes a random sample of n angular values and tests whether the points are randomly dispersed around a circle by examining the deviations between the empirical distribution (EDF) and the cumulative distribution function (CDF) of a uniform distribution (Figure 5), much like the Watson's U^2 test.

The Kuiper test statistic V_n is calculated by finding the maximum vertical deviation of the EDF above the CDF (D^+) and the maximum vertical deviation of the EDF below the CDF (D^-). These two values are summed together to get the test statistic V_n:

$$D^+ = \max_{1 \leq i \leq n} \left(\frac{i}{n} - \theta_i \right)$$

$$D^- = \max_{1 \leq i \leq n} \left(\theta_i - \frac{i-1}{n} \right)$$

$$V_n = D^+ + D^-$$

Larger values indicate greater deviations from uniformity, leading to the conclusion that the data are not randomly dispersed. Kuiper's original paper for *Biometrika* provided critical regions for the test statistic, but in 1965, M. A. Stephens calculated the asymptotic upper tail probabilities for

$$V_n{}' = n^{\frac{1}{2}} V_n$$

In 1970, he further modified the test statistic V_n to its current form

$$V = V_n \left(n^{\frac{1}{2}} + 0.155 + 0.24 n^{-\frac{1}{2}} \right)$$

When $n \geq 8$, this form can be used to find a critical value which in turn will give a range of upper tail probabilities (Table 1). Further work has since been done using Monte Carlo methods to derive exact p-values for this modified test statistic [8].

	Percentage Point 100α			
α	0.15	0.10	0.05	0.01
V	1.537	1.620	1.757	2.001

Table 1: Upper tail probabilities for different critical values of the Kuiper test [15]

4 Issues Implementing the Kuiper Test

4.1 The Kuiper Test in R

There are two packages in the statistical software R which can be used to calculate the Kuiper test statistic: the `kuiper.test` function in the `circular` package [3] and the `kuiper` function in the `CircStats` package [2], both authored by Claudio Agostinelli and Ulrich Lund. Both provide the same results, and share the exact same coding.

4.2 The Feldspar Problem

It should be strongly noted that while circular data is used in many fields, there are comparatively little resources regarding its analysis. At the time of this writing, there are only seven textbooks devoted to the topic. Of these, three are highly cited: *Statistics of Directional Data* (Mardia, 1972), *Circular Statistics in Biology* (Batschelet, 1981), and *Statistical Analysis of Circular Data* (Fisher, 1993). Additionally, M. A. Stephens' 1970 article *Use of the Kolomogorov-Smirnov, Cramer-Von Mises and Related Statistics Without Extensive Tables* is also very well known and cited for its modification of the Kuiper test (among other goodness-of-fit tests for circular data).

However, the Fisher textbook, while arguably the most approachable of the three, has several issues which prove problematic for the reader including inconsistent notation and incorrectly described data sets. This problem is then carried over into the `circular` data packages in R, leading to potentially incorrect conclusions and erroneous inferences.

These issues were discovered when this author tried to replicate the Kuiper test as detailed in the Fisher textbook both by hand and in R. To illustrate the Kuiper test calculations, Fisher uses a sample dataset of "measurements of the long-axis orientations of 60 feldspar laths in basalt," randomly selected from a larger dataset of 164 laths. It is important to note here that the data are recorded in degrees, and while they are axial data, only one endpoint was recorded for each observation; this is evident as the data range in values from 1° to 176°. When plotted, the data only encompass one half of the circle (Figure 6).

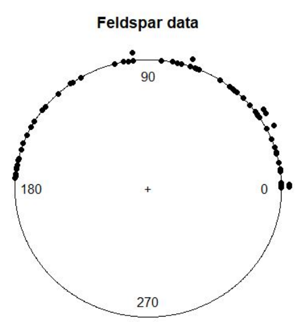

Figure 6: Plot of the feldspar data set [6]

The Fisher text, while giving the correct instructions for calculating the test statistic, neglects to clearly explain that the notation θ_i represents the measurements in radians, not degrees. Indeed, θ_i is used interchangeably in this text for both units of measurement, often with little indication of which is currently being used.

The reader is instructed to first order the data, then convert the data to angles by dividing each value by 2π before calculating the test statistic. However, if the reader attempts to apply this to data that are in degrees, such as the feldspar data, "by hand" calculations result in a drastically inflated test statistic ($V = 350.1185$ in this case). There is a supplementary note for this example that indicates that the axes should be converted to vectors by "doubling them and reducing modulo 360°" [6] before calculating the test statistic. However, it is unclear where this step falls in the given algorithm. If the data are converted first from axes to vectors, then to angles, $V = 6.7046$, which does not match Fisher's test statistic of 1.586. This is also the case if the order of conversion is switched, where the data are first converted to angles, then the doubling and modulo 360° is applied; the results remain the same. Additionally, if the reader uses modulo operations in R, rather than dividing by 360, the data remain unchanged from degrees, resulting in the original by-hand test statistic of $V = 350.1185$. This is because the `modulo` operation in R returns the remainder of the modulo operation; if the dividend is less than the divisor, the dividend is simply returned.

Since this author was unable to replicate Fisher's results by hand, statistical software (R) was then implemented to ensure against calculation errors. When the feldspar data is used with the `kuiper.test` function in the `circular` package, the function returns a test statistic of 1.3257, which again does not agree with the Fisher text. A warning message is also output, to indicate to the reader that the data were coerced to a circular object (Figure 7). When the data are manually coerced with the specifications given in the warning, using the `circular` function within the `circular` package, the `kuiper.test` function returns the same results, albeit without the warning message.

It is especially interesting to note here that the `circular` package contains

```
kuiper.test(feldspar)

Warning in as.circular(x): an object is coerced to the class 'circular' using
default value for the following components:
  type: 'angles'
  units: 'radians'
  template: 'none'
  modulo: 'asis'
  zero: 0
  rotation: 'counter'
conversion.circularxradians0counter2pi


       Kuiper's Test of Uniformity

Test Statistic:  1.3257
P-value > 0.15
```

Figure 7: Output from the `kuiper.test` function

all the datasets used in Fisher's text, presented as both the raw data, and the data in "circular" form. The feldspar data is referred to as `fisherB5` in its original form, and `fisherB5c` in the circular form. The circular form has been coerced with the specifications `type = directions` and `Units = degrees`. However, when the `fisherB5c` data are entered into the `kuiper.test` function, a test statistic of $V = 4.0761$ is returned (still failing to match that given in the Fisher text).

Within the `kuiper.test` function, there are no available options that can be changed, however since the function requires a circular object, this author attempted to alter the options available within the `circular` function, which raw data is passed through by the `kuiper.test` function itself:

```
circular(x, type = c("angles", "directions"),
  units = c("radians", "degrees", "hours"),
  template = c("none", "geographics", "clock12", "clock24"),
  modulo = c("asis", "2pi", "pi"),
  zero = 0, rotation = c("counter", "clock"), names)
```

The options `zero` and `rotation` apply to plotting commands, and do not alter the test output, which makes sense, as both are related to rotation, which the Kuiper test is invariant to. The `type` of the data likewise had no effect, as the `circular` documentation indicates it is an unused command. In fact, the only option available which *did* change the test statistic was the `modulo` option, which indicates how the data should be reduced, and that depended on the `units` selected (Table 2). Interestingly, when the `circular`

Table 2: comparing the Kuiper test statistic using different `circular` options.

Units	Modulo	V
degrees	asis	4.0761
	2π	4.0761
	π	4.0761
angles	asis	1.3257
	2π	1.3257
	π	4.2006
hours	asis	1.8508
	2π	1.8508
	π	4.2965

function is applied to the data, with `units` changed to "radians" and modulo changed to "asis," and the object is viewed again, the data appear unchanged; documentation implies the `circular` function only sets the properties, but does not actually convert the data.

Since the issue of differing test statistics appears not to lie in the `circular` conversion, the source code for the `kuiper.test` function itself was examined to verify the function was indeed doing what it should:

```
KuiperTestRad <- function(x, alpha) {
  if (!any(c(0, 0.01, 0.025, 0.05, 0.1, 0.15)==alpha))
    stop("'alpha' must be one of the following values: 0, 0.01, 0.025, 0.05,
0.1, 0.15")
  x <- sort(x %% (2 * pi))/(2 * pi)
  n <- length(x)
  i <- 1:n
  D.P <- max(i/n - x)
  D.M <- max(x - (i - 1)/n)
  V <- (D.P + D.M) * (sqrt(n) + 0.155 + 0.24/sqrt(n))
  return(V)
}
```

The root of the erroneous output lies in the line assigning the object `x`. Here, the function first applies the `modulo` operation to the data, then divides by 2π. However, as stated previously, the `modulo` function in R does not return the value of $x/2\pi$; it returns the remainder of x modulo 2π if $x \geq 2\pi$ or simply x if $x < 2\pi$. When the data are then divided by 2π, the effect is that the data are now scaled between 0 and 1, although not correctly.

The scaling of the data to be between 0 and 1 is the crux of all the errors this author encountered, both in the by-hand calculations and when using the `circular` package in R. What is painfully unclear to the casual reader of the Fisher text, as well as other supporting literature (including the Stephens [16] article referenced by the `circular` package), is that for the Kuiper test to be correctly calculated, the data must be scaled to be between 0 and 1. This means that if the data are measured in a unit other than radians, by the algorithm alone, the `kuiper.test` function in R will not calculate the correct value, nor will the reader be able to replicate worked examples.

However, if the data are first scaled by the total unit of measurement for the circle around which they were recorded, then converted to angles by dividing by 2π, the by-hand calculations will agree with the worked example in the Fisher text. For example, if the feldspar data is first divided by 180° (since the data only wrap around half the circle), then the values lie between 0 and 1, and the test statistic is $V = 1.586$.

Unfortunately, this correction does not hold for the `kuiper.test` function, because of the additional conversion step noted above. To offset this, the scaled data must be multiplied by 2π before being entered into the function:

```
feldspar.circ <- as.circular(feldspar, type="angles", units="radians",
template="none", modulo = "asis", zero=0, rotation="counter")

feldspar <- feldspar.circ*pi*2/180
kuiper.test(feldspar)


      Kuiper's Test of Uniformity

Test Statistic:  1.5864
0.10 < P-value < 0.15
```

This essentially doubles the values of the data points and then converts them to radians.

5 Discussion

While Kuiper's test is not as well known or as often employed as the Rayleigh test, it is still used by a variety of disciplines, which makes these findings troublesome. The critical values for Kuiper's test are very low, making this test quite sensitive to mathematical errors. Unfortunately, it is extremely easy to miss the scaling requirement, which almost invariably causes the test statistic to be far greater than it should be, leading to the rejection of randomness. This is the case with the Fisher text, which compounds the confusion by both adding the extra axes-to-vectors conversion and by not being clear with the units θ_i currently represents.

More problematic, however, is the fact that the `circular` and `CircStats` packages also fail to make the scaling requirement clear in their respective documentation. In fact, the documentation for both only states "Kuiper's test statistic is a rotation-invariant Kolmogorov-type test statistic. The critical values of a modified Kuiper's test statistic are used according to the tabulation given in Stephens (1970)" [2, 3].

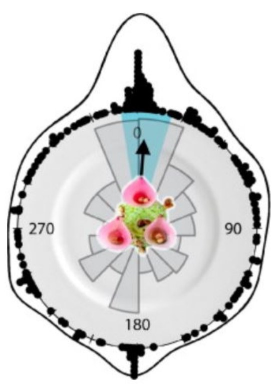

Figure 8: A circular plot depicting the preferred orientation of food [10]

It is very easy to see how the results of this function could lead to incorrect conclusions and inferences, especially for those unfamiliar with the analysis of circular data. This author has identified at three papers as of this writing which use the `kuiper.test` function in their analyses. Several others have been found which cite the `kuiper.test` in the `circular` package, but these articles were unavailable without purchase, so were not examined. Of the three available, all concluded their data were not isotropic, however, none had freely accessible data. Only one, a study of the orientation of the plating of food, published in *Food Quality and Preference* in 2015 [10], provided circular plots(Figure 8).

Based on those figures, it is reasonable to agree with the authors' conclusions on non-randomness. However, without appropriate figures, one is left to question the validity of the conclusions of other papers.

6 Conclusion

Researchers and scientists analyzing circular data need to be aware of the inconsistencies and lack of clarity that have been found in the available literature and R packages. Better documentation and explanation of the measurement units and scaling required to get correct results from the `kuiper.test` function are required before resulting conclusions made from these results can be assumed valid.

References

[1] Tibbet Hill Metavolcanic (image of feldspar laths). https://anrweb.vt.gov/PubDocs/DEC/GEO/Franklin/TibbetHillK.jpg, 2018.

[2] Agostinelli, C., and Lund, U. *R package `CircStats`: Circular Statistics, from "Topics in circular Statistics" (2001).* S-plus original by Ulric Lund and R port by Claudio Agostinelli, 2012. R package version 0.2-4.

[3] Agostinelli, C., and Lund, U. *R package `circular`: Circular Statistics (version 0.4-93).* CA: Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University, Venice, Italy. UL: Department of Statistics, California Polytechnic State University, San Luis Obispo, California, USA, 2017.

[4] Batschelet, E. *Circular Statistics in Biology.* Mathematics in biology. Academic Press, 1981.

[5] Blaser, N., Dell'Omo, G., Dell'Ariccia, G., Wolfer, D. P., and Lipp, H.-P. Testing cognitive navigation in unknown territories: homing pigeons choose different targets. *The Journal of Experimental Biology 216*, 16 (aug 2013), 3123 LP – 3131.

[6] Fisher, N. I. *Statistical Analysis of Circular Data.* Cambridge University Press, 1993.

[7] Kuiper, N. H. Tests concerning random points on a circle. *Indagationes Mathematicae (Proceedings) 63* (1960), 38–47.

[8] Louter, A. S., and Koerts, J. On the Kuiper test for normality with mean and variance unknown. *Statistica Neerlandica 24*, 2, 83–87.

[9] Mardia, K. V. *Statistics of directional data*. Probability and mathematical statistics. Academic Press, London, New York, 1972.

[10] Michel, C., Woods, A. T., Neuhäuser, M., Landgraf, A., and Spence, C. Rotating plates: Online study demonstrates the importance of orientation in the plating of food. *Food Quality and Preference 44* (2015), 194–202.

[11] NCSS LLC. NCSS Statistical Software Documentation — NCSS Software Help, 2011.

[12] O'Connor, J. J., and Robertson, E. F. Kuiper biography. http://www-history.mcs.st-and.ac.uk/Biographies/Kuiper.html, 2008.

[13] Pewsey, A., Neuhauser, M., and Ruxton, G. D. *Circular Statistics in R*. Oxford University Press, 2013.

[14] Ruxton, G. D. Testing for departure from uniformity and estimating mean direction for circular data. *Biology Letters 13*, 1 (jan 2017), 20160756.

[15] Stephens, M. A. The distribution of the goodness-of-fit statistic, UN2. II. *Biometrika 51*, 3-4 (1964), 393–397.

[16] Stephens, M. A. Use of the Kolmogorov-Smirnov, Cramer-Von Mises and Related Statistics Without Extensive Tables. *Journal of the Royal Statistical Society. Series B (Methodological) 32*, 1 (1970), 115–122.

[17] US Department of Commerce, NOAA, N. W. S. Abilene Wind Rose. https://www.weather.gov/sjt/climate-abiwindroses.

Appendix - R Code

```
require(circular)

## Figure 4: Rose plot code
icu <- fisherB1
icu.c <- circular(icu, units="hours", template = "clock24",
                rotation="clock")

## dotplot
plot.circular(icu.c, stack=T, sep=0.07, bins=360, axes = F,
                shrink = 1.2, cex=.75, tol=.18)

## rose plot
rose.diag(icu.c, pch = 16, cex = 1, axes = F, shrink = 1,
                col="#c5c1f4", prop = 2.5, bins=18, upper=TRUE,
                ticks=F, units="hours", tcl = 0.05, add=T)

## labels and tick marks
axis.circular(at=circular(seq(0, 2*pi-pi/2, pi/2)),
                labels=c("6am", "12am", "6pm", "12pm"),
                tick=T, tcl.text=.18)
ticks.circular(circular(seq(0,2*pi-pi/12,pi/12)), tcl=.05)

## kernel density estimates
res25 <- density(icu.c, bw=25,
control.circular=list(units="degrees"))
lines(res25, shrink=.45)

## Figure 5: Uniformity plot code
rand.dat.deg <- sample(360, 90, replace=T)
pp.unif.plot(rand.dat.deg, col="white",
                main="Uniformity Plot")
lines.edf(rand.dat.deg/360, col=2)
(rand.dat.deg/360)%%(2*pi)
```

```
## Code for Feldspar plots
feld <- fisherB5
feld.c <- circular(feld, units="degrees")
plot.circular(feld.c, stack=T, sep=0.06, bins=360,
                cex=1, pch=16, main="\nFeldspar␣data")
```