

Wind Rose Bias Correction

SCOTT APPLEQUIST

NOAA/National Climatic Data Center, Asheville, North Carolina

(Manuscript received 9 September 2011, in final form 16 December 2011)

ABSTRACT

Wind rose summaries, which provide a basis for understanding and evaluating the climatological behavior of local wind, have a directional bias if a conventional method is used in their generation. Three techniques used to remove this bias are described and are compared for theoretical and observed wind distributions. All three techniques successfully remove the bias, with the simplest of the three performing as well as the more-complex techniques.

1. Introduction

Wind roses are histograms that depict the joint relative frequency of wind speed and direction. These diagrams are used in runway orientation, air-quality modeling, and wind-energy planning, to name only a few uses.

A well-known methodological bias exists in the creation of wind roses from directional wind data. In brief, the practice of reporting wind direction at discrete 10° increments and analyzing them in a 360° range divided into 16 equally spaced, 22.5° -wide bins results in four bins each represented by three measurements and the remaining 12 each represented by only two. This bias is often easy to identify visually, because undue prominence is seen in the frequencies of the four primary directions (north, south, east, and west). For a uniform wind distribution, this practice introduces a positive bias of 33% at these directions and a negative bias of 11% at the other 12, as illustrated in Fig. 1. A similar problem arises from the 360° range divided into eight bins with a bias of $\pm 11\%$ in alternate bins. This bias is distinct from the observational bias (Ratner 1950) in which observers give preference to cardinal directions over those between the cardinal points.

Three correction schemes to the popular 16-bin model are described here: a global rescaling, bin boundary adjustments, and proportionate binning applied to individual wind values. All address the bias and yield nearly equivalent results. For simplicity, in the work

presented here the winds are not stratified by wind speed; the analysis would be the same if they were.

2. Bias-correction techniques

The simplest correction is to globally rescale the count in each wind direction bin by the number of directions it represents relative to the average (Droppo and Napier 2008; Lea and Helvey 1971). Thus, the primary four directions are each reduced by a factor of 0.75 and the remaining 12 directions are multiplied by 1.125. As a result of this rescaling, the sum of all bins will, in most cases, be different from the original total. To keep the overall number of events constant, all bins are further scaled by the ratio of the old total to the new total.

A second correction proposed by Lea and Helvey (1971) involves an adjustment to the count in each bin that is based on where bin boundaries fall in relation to the winds represented by that bin. For example, winds counted in the east bin, defined as being from 78.75° to 101.25° , came from directions 80° , 90° , and 100° . If one assumes that winds from those three directions are evenly distributed between 75° and 105° , the number of winds counted in this bin should be reduced by the excesses of 3.75° occurring at each end, which actually fall in the ranges of neighboring bins. The count in this bin is reduced by $(30^\circ - 2 \times 3.75^\circ)/30^\circ = 0.75$. The neighboring east-southeast bin, defined as being from 101.25° to 123.75° , originally counts winds from 110° and 120° . This bin receives $3.75^\circ/30^\circ = 0.125$ of the values in the east bin and, again assuming an even distribution of winds from 105° to 125° , must itself be reduced by the excess of $1.25^\circ/20^\circ = 0.0625$, which is passed on to the

Corresponding author address: Scott Applequist, NOAA/National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801.
E-mail: scott.applequist@noaa.gov

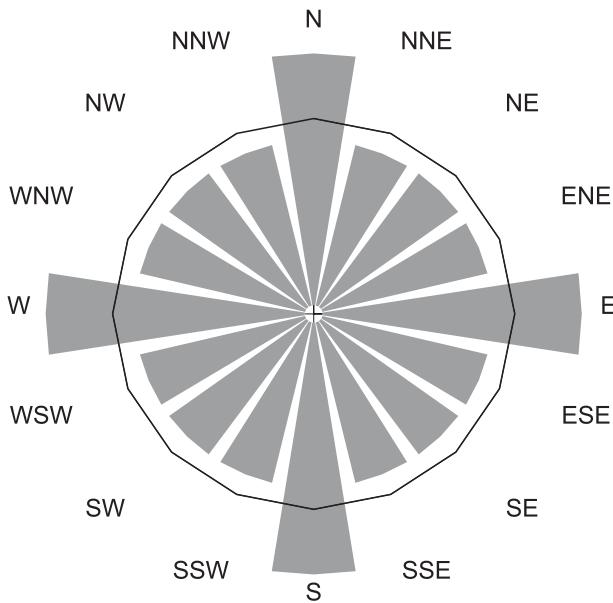


FIG. 1. Analysis of a uniform-wind case. The apparent frequencies of wind from each direction without bias correction are plotted as columns radiating from the center. The black solid line indicates the extent of the columns when the bias is removed.

southeast bin. Adjustments are made to the remaining 14 directions using similar logic.

The previous two methods are applicable to wind direction histograms that have already been constructed from the observations. If individual observations are available, a third technique proposed by Droppo and Napier (2008) may be applied. In this method each wind event is directly binned by a proportionate amount for the bin or bins into which it falls. For example, a wind direction of 100° is assumed to represent a uniform distribution of winds between 95° and 105° . Given the boundary between the east and east-southeast bins at 101.25° , these bins are incremented by 0.625 and 0.375, respectively. A wind direction of 180° , representing 175° to 185° , falls entirely in the south bin and simply increments its count by 1. Because it handles each wind observation individually and assigns weights proportionately, this technique removes the methodological bias at its source and provides truly unbiased wind roses.

As expected, all three methods result in perfect correction of the uniformly distributed wind case. Next we examine the performance of all three techniques on two artificial cases of increasing complexity and on actual observations.

3. Application of bias correction

In both artificial cases a pseudorandom-number generator was used to create 3600 uniformly distributed

wind directions between 0° and 360° . These were then recorded as values to the nearest 10° , from 10° to 360° inclusive. In the first case the distribution is skewed as follows: 30% of the values were replaced with normally distributed wind directions with a mean of 225° and a standard deviation of 20° and 10% of the values were replaced with normally distributed winds with a mean direction of 45° and standard deviation of 20° . All replacement values are rounded to the nearest 10° . As shown in Fig. 2, all three methods yield nearly identical results and remove the obvious bias of the original. In the second case, normally distributed values with means of 110° and 340° and a standard deviation of 25° replaced 40% and 20%, respectively, of the uniform distribution. Figure 2 again shows strong agreement among the bias corrections.

Considering the differences among the histograms between the original and bias-corrected distributions, one must first recognize that they are computed over the entire population. Any differences are due to method and not to random-sampling considerations. Inferences, however, can be made regarding the similarity of two histograms from the test between two binned datasets described by Press et al. (1992).

In this test, one accepts or rejects the null hypothesis that two samples of binned data, such as the values plotted in Fig. 2, were drawn from the same underlying population. The χ^2 statistic is greatest when the relative differences between values in a bin are most pronounced:

$$\chi^2 = \sum_{i=1}^{16} \frac{(R_i - S_i)^2}{(R_i + S_i)}.$$

In this formula, the S_i and R_i terms are the number of occurrences of an event in each bin of the two distributions, R and S , being compared. In the current example with 16 bins, when $\chi^2 > 25$ one would reject, with 95% confidence, the null hypothesis.

Given the histogram values computed here, one can ask, "Would one conclude that these describe samples taken from the same population?" When $\chi^2 > 25$, one would reject this hypothesis. Computing χ^2 among the four distributions then suggests whether the methods used are in agreement with one another. To put the χ^2 values in perspective, in the left panel of Fig. 2, $\chi^2 = 50$ between the biased distribution and that with the global rescaling correction. In the same figure, $\chi^2 = 4.2$ between the global rescaling and the bin boundary adjustment.

Because these histograms are generated from a circular distribution of wind directions, one must also consider whether the arbitrary locations of the origin and

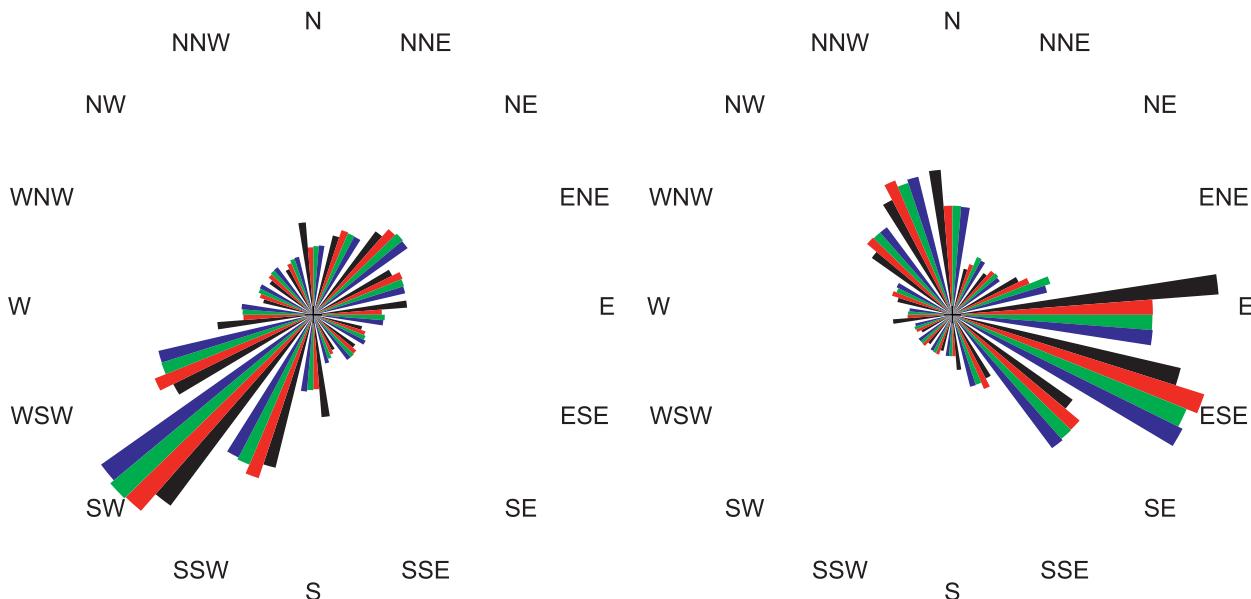


FIG. 2. Original biased and corrected wind distributions for artificial cases described in the text, showing modes at (left) 225° and 45° and (right) 110° and 320°. In each sector there are four bars: black is the original biased distribution, red is with global rescaling, green is with bin boundary adjustment, and blue is with proportionate bins. Evidence of the bias is still easily seen in the north, south, east, and west directions.

bin boundaries have an impact on the value of the test statistic (Jammalamadaka and SenGupta 2001). To address this, the boundary definitions for the 16 sectors were each shifted by 2.5° clockwise, resulting in a different bias pattern requiring recomputation of the global rescaling weights, bin boundary adjustments, and proportionate binning values as described above. This shift was performed eight times to cover all unique combinations of how the directional bins could be analyzed.

To assess in a practical manner whether the choice of bias-correction scheme makes a difference, the techniques described above were applied to a selection of 30 years (1981–2010) of wind records from the Integrated Surface Database (Smith et al. 2011). On the first and 16th day of each month, samples were gathered containing wind directions at 3-hourly intervals for 15 days centered on that date. Wind reports with speeds of less

than 3 mi h⁻¹ (1.34 m s⁻¹) were excluded. Thus each sample had a maximum of 3600 observations. These samples were taken 2 times per month at 262 locations in the United States and its territories, yielding a total of 6288. Values of χ^2 were computed among the original biased wind distributions and the three bias-correction methods using the original and eight shifted boundary configurations. The distribution of the χ^2 values is shown in Table 1.

4. Results

Several key points are seen in these comparisons:

- 1) From the first three rows in Table 1, in over 99% of the cases, the bias-corrected versions would be deemed different from the original distribution. An example of one that was not is shown in Fig. 3. Here, for data

TABLE 1. The distribution of χ^2 associated with intercomparisons between the original biased distribution and the three bias-correction methods using 6288 samples at the original and eight shifted wind-sector definitions. Columns *R* and *S* indicate the distributions compared: 1 = original biased, 2 = global rescaling, 3 = bin boundary adjustment, and 4 = proportionate bins. A high χ^2 leads to rejection of the null hypothesis that the distributions describe similar underlying populations. Probabilities of the right-tail area associated with the χ^2 are provided in parentheses underneath the χ^2 values.

<i>R</i>	<i>S</i>	<5 (0.992)	5–10 (0.992–0.820)	10–15 (0.820–0.451)	15–20 (0.451–0.172)	20–25 (0.172–0.050)	25–30 (0.050–0.012)	≥30 (<0.012)
1	2	0	0	32	86	250	608	55 616
1	3	0	0	9	61	148	464	55 910
1	4	0	0	9	58	254	727	55 544
2	3	31 647	15 675	5015	1925	839	433	1058
2	4	46 094	8692	1171	291	126	61	157
3	4	52 367	2911	612	255	141	80	226

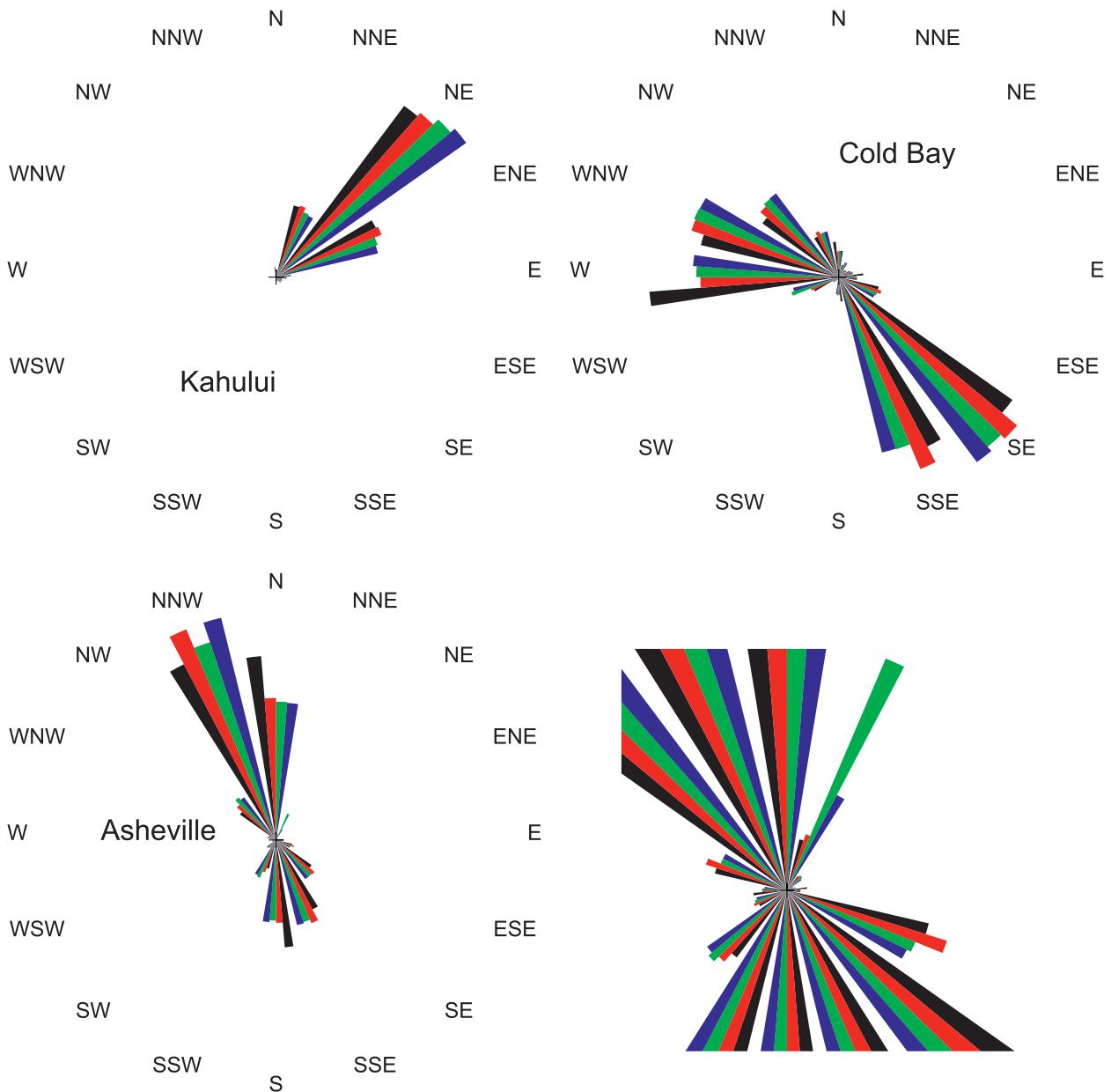


FIG. 3. Wind distributions with bias corrections as shown in Fig. 2. Observations are every 3 h for the dates indicated in years 1981–2010. Shown are (top left) Kahului for 10–25 Jul, (top right) Cold Bay for 10–25 Jul, (bottom left) Asheville for 25 Dec–8 Jan, and (bottom right) a magnification of the center of the Asheville distribution.

collected at Kahului, Hawaii, in mid-July, the winds have such a dominant northeasterly mode that any bias correction makes little difference.

- 2) Fewer than 0.4% of the cases were found to have the global rescaling and proportionate binning techniques compared with $\chi^2 > 25$. This indicates that a similar result is expected from either method. Among the original bin definitions, the worst-case comparison between the two ($\chi^2 = 20.9$; right tail area = 0.14) is

shown in Fig. 3. At Cold Bay, Alaska, coincidentally again for mid-July, we see a relatively large difference between these two methods in the west-southwest and south-southeast sectors. Here the global rescaling resulted in a decrease in the west sector largely compensated by increases in the south-southeast and southeast bins.

- 3) The bin boundary adjustment technique yielded significantly different distributions from the global

rescaling and proportionate binning methods 2.6% and 0.5% of the time, respectively. An example in which this method shows a difference appears in Fig. 3 for Asheville, North Carolina, for dates within a week of 1 January. The bin boundary adjustment moves such a large number of values from the north bin to the north-northeast bin that the contribution to the χ^2 sum from this bin alone is enough to conclude that the methods yield different results.

- 4) The two-bin χ^2 test can only indicate when it is safe to reject the null hypothesis that two samples were drawn from the same population. The high occurrence of low χ^2 values at the left end of Table 1 indicates, however, that in the vast majority of cases the bias-correction distributions are very similar among the three methods. In other words, in most cases the differences in the distributions are nowhere near the threshold required to reject the null hypothesis. By extension, the bias corrections yield results that should be similar enough for practical purposes.

This process has been repeated using eight bins with analogous results. Because the bias correction is not as large in the eight-bin case, there were more examples in which the original and bias-corrected distributions could not be distinguished. Nonetheless, the basic pattern of agreement among the three correction methods remained the same.

5. Summary

The methodological bias in creating wind roses should be eliminated using one of the three methods described here. If individual wind reports are available, the proportionate binning method has the appeal of representing each wind report in an unbiased manner. For wind direction histograms that have already been computed, the global rescaling and bin boundary adjustment methods offer nearly identical bias correction, with the first being easier to implement and demonstrating more consistent agreement with the proportionate binning.

REFERENCES

- Droppo, J. G., and B. A. Napier, 2008: Wind direction bias in generating wind roses and conducting sector-based air dispersion modeling. *J. Air Waste Manage. Assoc.*, **58**, 913–918.
- Jammalamadaka, S. R., and A. SenGupta, 2001: *Topics in Circular Statistics*. World Scientific, 322 pp.
- Lea, D. A., and R. A. Helvey, 1971: A directional bias in wind roses due to mixed compass formats. *J. Appl. Meteor.*, **10**, 1037–1039.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Vol. 1, Cambridge University Press, 933 pp.
- Ratner, B., 1950: A method for eliminating directional bias in wind roses. *Mon. Wea. Rev.*, **78**, 185–188.
- Smith, A., N. Lott, and R. Vose, 2011: The Integrated Surface Database: Recent developments and partnerships. *Bull. Amer. Meteor. Soc.*, **92**, 704–708.