Computers & Geosciences 49 (2012) 112–120

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

MARD—A moving average rose diagram application for the geosciences

Mark A. Munro*, Thomas G. Blenkinsop

School of Earth & Environmental Sciences, James Cook University, Townsville, Queensland 4811, Australia

ARTICLE INFO

Article history:
Received 29 March 2012
Received in revised form
2 July 2012
Accepted 3 July 2012
Available online 24 July 2012

Keywords:
Rose diagram
Moving average
Circular statistics
Vector mean
MATLAB®
Microsoft® Excel

ABSTRACT

MARD 1.0 is a computer program for generating smoothed rose diagrams by using a moving average, which is designed for use across the wide range of disciplines encompassed within the Earth Sciences. Available in MATLAB®, Microsoft® Excel and GNU Octave formats, the program is fully compatible with both Microsoft® Windows and Macintosh operating systems. Each version has been implemented in a user-friendly way that requires no prior experience in programming with the software. MARD conducts a moving average smoothing, a form of signal processing low-pass filter, upon the raw circular data according to a set of pre-defined conditions selected by the user. This form of signal processing filter smoothes the angular dataset, emphasising significant circular trends whilst reducing background noise. Customisable parameters include whether the data is uni- or bi-directional, the angular range (or aperture) over which the data is averaged, and whether an unweighted or weighted moving average is to be applied. In addition to the uni- and bi-directional options, the MATLAB® and Octave versions also possess a function for plotting 2-dimensional dips/pitches in a single, lower, hemisphere. The rose diagrams from each version are exportable as one of a selection of common graphical formats. Frequently employed statistical measures that determine the vector mean, mean resultant (or length), circular standard deviation and circular variance are also included. MARD's scope is demonstrated via its application to a variety of datasets within the Earth Sciences.

1. Introduction

Moving averages, sometimes referred to as running averages, are a form of signal processing filter which produce averages of subsets of data series from within the master data series in 1-dimension. Data filtering of this kind serves a broad variety of applications within science and engineering, being primarily utilised to accentuate trends in data that are otherwise generally less apparent, or to recover meaningful signal components by removing the high-frequency noise. Other uses include 2-dimensional enhancements in the resolution of graphics, and providing a form of interpolation that generates intermediate values. Within the Earth Sciences, such analyses are commonly applied in various forms for image processing of aerial photographs, maps and satellite images (Diniz da Costa and Starkey, 2001; Jordan, 2007). Moving averages may be applied to rose diagrams that depict azimuthal data. At present, moving average rose diagrams are not commonly used within the Earth and Environmental Sciences; however, they are a very effective means of visualisation (e.g. Aerden, 2003,2004; Aerden and Sayab, 2008). The most obvious benefit of this representation is that the commonly segmented, or 'blocky' appearance of rose plots, due to aggregation of data in bins, is avoided.

A range of good quality software is available for plotting geoscience data on both Mac and Windows platforms. Within these programs, rose diagrams are a common function. Examples include *GEOrient*© (Holcombe, 1994), *Spheristat*™, *RJS graph*, *Stereonet*©, *OSXStereonet*©, *Grapher*™, *KaleidaGraph*©, *Rose.C* (Kutty and Ghosh, 1992) and *EZ-Rose* (Baas, 2000). However, only one of these, *Spheristat*™, presents a 'smoothing' function for rose and polar diagrams that is similar to a moving average filter. This program is relatively expensive to purchase and is only available for Windows.

The primary aim of this paper is to promote the use and benefits of moving average filters within rose diagrams via the provision of freely distributable, user-friendly programs which do not require any prior knowledge of high-level programming languages. Three formats are presented for users to readily produce a moving average rose diagrams on either Windows or Macintosh platforms. MARD is available (1) as a script written for MATLAB®, (2) a script for GNU Octave and (3), a macro implemented for use in conjunction with Microsoft Excel. The benefits of MARD are illustrated via its application to a diverse range of datasets within the Earth and Environmental Sciences.

2. Moving average rose diagrams: overview and evaluation

2.1. Moving average rose diagrams: a definition

When presenting rose diagrams, the most common convention within the Earth Sciences is to represent azimuthal data utilising 10° bins, the frequencies of which represent the sum of all of the

* Corresponding author. Tel.: +61 7 4781 6909.
E-mail address: mark.munro@my.jcu.edu.au (M.A. Munro).

0098-3004/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2012.07.012

azimuths within that bin. Binning is used in order to emphasise major trends in data, but inherently loses a certain level of detail because the distribution of data within the bins is unknown. Moreover, the arbitrary selection of bin width and boundaries may have a significant impact upon the result.

Moving average rose diagrams, on the other hand, evaluate the frequency of each azimuth within the context of those in immediate proximity to it. The frequency of each azimuth, and those of azimuths within a pre-defined range (the aperture, or moving window) either side of it are systematically summed and averaged. The resultant average for each azimuth is then assigned to the central value position and plotted on the final rose diagram. The broader the aperture, the greater the smoothing effect.

2.2. Types of moving average: weighted versus unweighted

Moving averages may be either unweighted (a.k.a. a 'simple' moving average) or weighted. Where no weighting is applied, each azimuth frequency within the aperture is counted as its full value, i.e. each is deemed equally important in determining the local average:

$$M_\alpha = \frac{1}{A}\sum_{i=\alpha-((A-1)/2)}^{n=\alpha+((A-1)/2)} F_i \tag{1}$$

where M_α is the unweighted average value to be subsequently plotted on the rose diagram for that azimuth, α is the azimuth for which the average is determined, F_i is the raw frequency at angle i, A is the aperture size.

Alternatively, when a weighted moving average is applied, the raw values of the data in azimuths outside the central one are reduced by a magnitude that depends upon their proximity to the centre. The raw value of immediately adjacent data is reduced by less than those nearer to the margins of the aperture. Weighting permits more emphasis to be placed upon those data closer to the centre of the aperture. A weighted mean is thus less than, and represents a proportion of, an equivalent unweighted mean.

In many applications, especially when dealing with large datasets, unweighted moving averages are appropriate. However, in a number of circumstances the application of a weighted average might be more beneficial. One such example is when handling small datasets. In particular, where data is clustered, an unweighted moving average may result in a 'plateau' distribution like a binned plot. The use of a weighted moving average in scenarios such as this provides a means of smoothing the data markedly whilst still preserving the local maxima.

2.3. Weighted moving averages in MARD

In MARD, weight is distributed among the azimuths in a non-linear fashion, because linear methods are intrinsically restrictive: for example, reducing each value progressively by 10% from the central position can only be achieved for 10 positions either side before becoming negative. However, the non-linear method is applicable to any aperture. The user specifies a proportion, expressed as the weighting factor, which reduces the contribution of each neighbouring raw value with increasing distance from the reference position. Thus, in calculating the moving average around an azimuth of 87° using a weighting factor of 0.9, the frequencies of data at azimuths of 86° and 88° (one position removed from the centre) are reduced to 0.9 of their full values, and the frequencies at azimuths of 85 and 89° (two positions removed) are reduced to a proportion of $0.9\times0.9=0.81$ of their full values. The moving average for that azimuth is then the average of these weighted values within the aperture in the same manner as an unweighted average, and is represented as

$$Mw_\alpha = \frac{1}{A}\sum_{i=\alpha-((A-1)/2)}^{n=\alpha+((A-1)/2)} F_i \, w^{|\alpha-((A-1)/2)|} \tag{2}$$

where Mw_α is the weighted average value to be subsequently plotted on the rose diagram for that azimuth, α is the azimuth for which the average is determined, F_i is the raw frequency at angle i, A is the aperture size, w is the value of weighting factor applied.

2.4. The 'up-scaling factor' for weighted moving averages

One effect of weighting the averages as described above is that it lowers the frequencies of the plot relative to those within an equivalent unweighted plot. The degree to which this occurs is proportional to the degree of weighting applied. Where the adjacent values are still weighted relatively strongly (e.g. a weighting factor of 0.95) then the frequencies will not be reduced to a large degree. On the other hand, if the adjacent values are allocated much lower weight than the central (e.g. 0.7) then the magnitude of frequencies will be reduced considerably more, resulting in a notable loss in magnitude of the moving average.

In order to counter-act this effect, an 'up-scaling factor' is applied to the weighted moving average frequencies. The value of this factor is inversely proportional to the value of weighting factor selected and is given as

$$\frac{1}{D} = 1 + 2\sum_{i=1}^{n=((A-1)/2)} w^i \tag{3}$$

where $1/D$ is the up-scaling factor, A is the aperture size, w is the weighting factor.

To upscale the weighted moving average, the proportion of the equivalent unweighted mean that the value of the weighted mean should represent is calculated for the selected aperture size and weighting factor applied. Each weighted frequency is then multiplied (up-scaled) by the inverse of this proportion. This restores the absolute frequencies back into a range equivalent to that of an unweighted moving average, so that they can be compared to the unweighted moving averages.

2.5. Benefits and limitations

The primary advantage of moving average rose diagrams over their binned counterparts is the removal of the coarse, blocky appearance of the former, with artificial steps at bin boundaries. The product is a plot that is more visually appealing and notably more informative. Moreover, the suppression of minor variations by averaging accentuates significant changes or trends in the data. The main objective of applying moving averages to a dataset is to smooth the plot to emphasise the significant trends present, whilst retaining its original character as much as possible. Filtering data in this way inherently provides more context for the distribution of frequencies around the compass, and within individual bins, than conventional plots. Consequently, moving averages commonly reveal a more accurate representation of the distribution of modal maxima within a dataset than are otherwise portrayed in an equivalent binned counterpart (Fig. 1).

One limitation of the moving average method by calculating a mean value (as implemented in MARD) is that it may potentially be influenced by outliers, i.e. spurious maxima. A median average method (e.g. Jordan, 2007) may ameliorate this problem, and this could be implemented in future versions of MARD. At present, however, the mean method is proposed because it is the most common and readily understood type of moving average.