

This discussion paper is/has been under review for the journal Solid Earth (SE).  
Please refer to the corresponding final paper in SE if available.

# Modelling complex geological angular data with the Projected Normal distribution and mixtures of von Mises distributions

R. M. Lark<sup>1</sup>, D. Clifford<sup>2</sup>, and C. N. Waters<sup>1</sup>

<sup>1</sup>British Geological Survey, Keyworth, Nottingham, UK

<sup>2</sup>CSIRO Computational Informatics, Ecoscience Precinct, G.P.O. Box 2583, Brisbane, Qld 4001, Australia

Received: 15 November 2013 – Accepted: 28 November 2013 – Published: 9 December 2013

Correspondence to: R. M. Lark (mlark@bgs.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

Angular data are commonly encountered in the earth sciences and statistical descriptions and inferences about such data are necessary in structural geology. In this paper we compare two statistical distributions appropriate for complex angular data sets: the mixture of von Mises and the projected normal distribution. We show how the number of components in a mixture of von Mises distribution may be chosen, and how one may choose between the projected normal distribution and mixture of von Mises for a particular data set. We illustrate these methods with some structural geological data, showing how the fitted models can complement geological interpretation and permit statistical inference. One of our data sets suggests a special case of the projected normal distribution which we discuss briefly.

## 1 Introduction

Angular data are commonly encountered in geology. An angular variable may be a direction, such as the direction of dip of a fault or bedding plane, or a palaeomagnetic vector. Alternatively an angular variable may be an orientation (which could be expressed by either of two opposite directions) such as the orientation of vertical faults or bedding planes or the orientation of primary palaeocurrent lineations where the direction of flow is unknown. Directional variables are distributed on the unit circle, and so have the particular property that the upper bound,  $2\pi$  radians, and the lower bound, zero, are equivalent. They therefore cannot be treated as though they were distributed over some subset of the real numbers, and require special treatment for statistical analysis (Mardia and Jupp, 2000).

The von Mises distribution is commonly used for the statistical analysis of angular data. The distribution has two parameters, a mean direction and a concentration parameter,  $\kappa$ . The latter parameter measures the dispersion of the variable about the mean. At its minimum,  $\kappa = 0$ , the von Mises distribution is the uniform distribution over

SED

5, 2181–2202, 2013

## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the unit circle. As  $\kappa$  increases so the von Mises distribution increasingly resembles the Gaussian with variance  $1/\kappa$ .

The von Mises distribution has been widely used in earth science, and software for analysis of angular data including methods based on the von Mises distribution have been made available (e.g., Jones, 2006). For example, Coblenz and Richardson (1995) examined global data on maximum horizontal compressive strength, and examined local evidence for coherence of stress direction by the Rayleigh test which is equivalent to a comparison of the von Mises distribution with  $\kappa > 0$  against a uniform alternative. Witts et al. (2012) used a similar procedure to identify trends in palaeocurrent data from the dip and dip azimuth of sand bars in a Cenozoic sedimentary succession in Indonesia. Sen and Mamtani (2006) used the  $\kappa$  parameter of the von Mises distribution to characterize the preferential orientation of biotite in thin sections of granite which was related to variations in regional strain. O'Brien et al. (2012) used von Mises distributions to characterize the orientations of fault plane solutions – with confidence intervals – before, during and after seismic swarms.

The von Mises distribution is symmetrical and unimodal. Angular data in earth sciences may often have a more complex distribution than this. There may be continuous variation in the orientation of particular features; for example the preferential direction of structures in sedimentary deposits in palaeochannels are likely to respond to channel orientation and flow direction which will vary at metre to kilometre scales. Distributions of angular data may therefore be asymmetric and multimodal.

One way to model such complex variation is to treat it as a mixture of von Mises distributions (MVM). The MVM model that we considered, with  $g$  components has  $3g - 1$  parameters:  $g$  mean directions and values of  $\kappa$  and the  $g - 1$  independent proportions of the components. By including sufficient components it is possible to model a distribution of angles with multiple modes and asymmetry.

An alternative model for more complex distributions of angular data is the projected normal (PN) distribution. If  $\mathbf{y}$  is a realization of a bivariate random variable,  $\mathbf{Y}$ , on the plane  $\mathbb{R}^2$ , and  $\Pr\{\mathbf{Y} = \mathbf{0}\} = 0$ , then its radial projection  $\|\mathbf{y}\|^{-1}\mathbf{y}$  is a random vector on the

Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



unit circle which can be converted to a vector of random angles relative to some direction treated as 0. In the PN distribution  $\mathbf{y}$  is a realization of a bivariate normal variable  $\mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  on the plane. This distribution is discussed by Mardia and Jupp (2000) and a clear and succinct summary is provided by Wang and Gelfand (2013). If  $\boldsymbol{\mu} = \mathbf{0}$  and  $\boldsymbol{\Sigma} \propto \mathbf{I}$  then the PN distribution is equivalent to a uniform distribution on the circle. Allowing  $\boldsymbol{\mu} \neq \mathbf{0}$  gives rise to a non-uniform but unimodal and symmetrical distribution on the circle, and further generalization so that  $\boldsymbol{\Sigma}$  is any valid covariance matrix gives rise to a flexible distribution on the circle which can be asymmetrical and bimodal. Wang and Gelfand (2013) explore the PN distribution in a Bayesian setting, including regression models in which the parameter  $\boldsymbol{\mu}$  is modelled as a linear function of covariates.

The MVM and PN distributions are flexible models for complex distributions of angular data. We are not aware of any examples of their use in structural geology, although they would clearly be suitable in circumstances where simpler symmetrical and unimodal distributions would not be appropriate. Some practical questions remain for their application. First, how many components of the MVM model are justified for a particular data set? Second, given that MVM for  $g > 2$  has more parameters than the PN model, how can one decide when the more complex model is justified?

In this paper we demonstrate the use the mixture of von Mises and the PN distribution as models for three sets of angular data from structural geology. We address the question of how to select the number of components,  $g$ , of the MVM for a particular data set, and the decision whether the MVM or PN model is preferable in a particular case. We use maximum likelihood to estimate parameters of the PN distribution, and show how to test a hypothesis of the uniformity of the PN model over two data sets.

## 2 Methods

### 2.1 Selection of a mixture of von Mises distributions

The mixture of von Mises distribution, as used here, comprises  $g$  different von Mises distributions, each with an associated mixture weight  $\alpha_i, i = 1, 2, \dots, g$  which is the probability that an event is drawn from the  $i$ th distribution. The parameters of the MVM distribution, for specified  $g$ , can be estimated by maximum likelihood estimation (MLE). The MVM distribution is identifiable (Holzmann et al., 2004) but a numerical approximation to the likelihood function is needed because the MLE of the  $\kappa$  parameter includes a ratio of Bessel functions. In this study we used the `movMF` procedure from the package of that name developed by Hornik and Grün (2013) for the R platform (R development core team, 2013). This finds the MLE by an Expectation Maximization algorithm (e.g., Banerjee et al., 2005).

We are interested here in how many components to specify in the MVM model for a data set. The general problem is whether the improvement in the likelihood that we obtain by fitting  $g + 1$  rather than  $g$  components with an additional three parameters is justified. In general two nested models, where the simpler (null) model is a particular case of the more general with parameters fixed to some value, may be compared on their log likelihood ratio,  $\ell_A - \ell_N$ , where  $\ell_A$  and  $\ell_N$  are the maximized log likelihoods obtained for the more complex and the null models respectively. It is necessary, for inference, to know the distribution of the ratio when the null model holds. In regular cases, asymptotically,  $L = 2(\ell_A - \ell_N)$  is distributed as  $\chi^2$  with degrees of freedom equal to the number of additional parameters in the more complex model. However, mixture models are not regular because the null model is at the boundary of the parameter space for the alternative. The distribution of  $L$  can be counterintuitive at the boundary of the parameter space, see for example, Clifford (2006). A solution to this problem has been proposed for the comparison of the two-component MVM distribution with a single von Mises distribution, in the case where the  $\kappa$  parameter is common to both components of the MVM (Fu et al., 2008). This does not address the general problem

SED

5, 2181–2202, 2013

Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of comparing MVM models for increasing values of  $g$ . For this reason we computed the distribution of  $l_{g,g+1} = \ell_{g+1} - \ell_g$  for any comparison between a MVM with  $g$  components (log likelihood  $\ell_g$ ) and one with  $g + 1$  components (log likelihood  $\ell_{g+1}$ ) by Monte Carlo simulation. This approach has been used previously to identify the number of components in a mixture model, e.g. Aitkin et al. (1981). The distribution of the log likelihood ratio  $l_{g,g+1}$  under some model with  $g$  components may depend on the parameters of the components and how well-separated they are on the circle. For this reason we accounted for uncertainty in the estimated parameters of the null model by a bootstrapping step. A bootstrap sample from the data was drawn and a MVM model with  $g$  components was fitted, a single parametric bootstrap sample from the fitted model was then drawn and the log likelihood ratio for MVM models with  $g$  and  $g + 1$  components was calculated. This step was repeated to generate the full Monte Carlo sample of the log likelihood ratio. We are not aware of a comparable combination of a bootstrap sample from the data with a parametric bootstrap, but note that it is comparable, although not identical, with the “double bootstrapping” procedure of Mclachlan and Peel (1997). The full procedure is described below.

We fitted MVM distributions with  $g$  and  $g + 1$  components to our  $n$  data, using the `movMF` procedure. We computed the log of the ratio of likelihoods for the two distributions, which we denote by  $\hat{l}_{g,g+1}$ . To obtain a distribution for this ratio under the null model we undertook the following steps.

1. A bootstrap sample was drawn from the data and the parameters of the MVM model with  $g$  components were estimated with the `movMF` procedure.
2. The `rmovMF` procedure from the `movMF` library was used to generate a random sample from the MVM distribution ( $g$  components) with the parameters estimated from the bootstrap sample.
3. We then fitted the MVM distribution with  $g$  and with  $g + 1$  components to the simulated values, and computed  $\hat{l}_{g,g+1}$ .
4. Steps 1–3 above were repeated 1000 times in total.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The proportion of values of the log likelihood ratio in the resulting sample larger than  $\hat{l}_{g,g+1}$  was computed,  $\hat{P}_g$ . If  $\hat{P}_g \geq 0.05$  then the null model ( $g$  components) was accepted. This procedure was followed, starting with  $g = 1$  and testing an alternative distribution with two components. If the null model was rejected then a distribution with three components was compared to the distribution with two and so on until the null model for some  $g$  was accepted.

## 2.2 Maximum likelihood estimation of the PN distribution and comparison with the selected MVM distribution

A PN distribution of angles can be specified by the parameters of the bivariate normal distribution whose radial projection corresponds to points on the unit circle. To give a unique parameterization for any distribution on the circle we followed Wang and Gelfand (2013) by fixing the variance of one of the normal variables to 1.0. There are four parameters to be estimated: the two elements of the mean vector,  $\boldsymbol{\mu}$ , the correlation coefficient,  $\rho$  which takes values in the interval  $(-1, 1)$  and the standard deviation of the second normal variable,  $\tau$  which takes values  $\tau > 0$ . These were estimated by maximum likelihood, using the PN density given by Wang and Gelfand (2013):

$$f(\theta|\boldsymbol{\mu}, \rho, \tau) = C(\boldsymbol{\eta})^{-1} \left[ \phi_2\{\mu_1, \mu_2|\mathbf{0}, \boldsymbol{\Sigma}\} + aD(\boldsymbol{\eta})\Phi_1\{D(\boldsymbol{\eta})\}\phi_1\left\{a(\mu_1 \sin \theta - \mu_2 \cos \theta)/\sqrt{C(\boldsymbol{\eta})}\right\} \right], \quad (1)$$

where  $\theta$  is an angle, obtained from the radial projection of the bivariate normal distribution by the arctan\* function defined by Wang and Gelfand (2013),  $\boldsymbol{\eta}$  is the vector of values  $[\theta, \mu_1, \mu_2, \rho, \tau]^T$ ,  $\phi_2(x_1, x_2|\boldsymbol{\mu}, \boldsymbol{\Sigma})$  is the bivariate normal probability density function (pdf) with specified parameters,  $\phi_1(x)$  is the standard normal pdf and  $\Phi_1(x)$  is the

standard normal distribution function,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \tau^2 & \tau\rho \\ \tau\rho & 1 \end{bmatrix},$$

$$a = \left( \tau \sqrt{1 - \rho^2} \right)^{-1},$$

$$C(\boldsymbol{\eta}) = a^2 \left( \cos^2 \theta - \rho\tau \sin 2\theta + \tau^2 \sin^2 \theta \right)$$

$$D(\boldsymbol{\eta}) = \frac{a^2 \{ \mu_1 (\cos \theta - \rho\tau \sin \theta) + \mu_2 \tau (\tau \sin \theta - \rho \cos \theta) \}}{C(\theta)^{-\frac{1}{2}}}$$

The maximum likelihood estimates  $\hat{\boldsymbol{\mu}}$ ,  $\hat{\rho}$  and  $\hat{\tau}$  maximize the likelihood function  $\mathcal{L}(\boldsymbol{\mu}, \rho, \tau | \theta) = f(\theta | \boldsymbol{\mu}, \rho, \tau)$ , given the data in  $\theta$ . The likelihood was maximized with the `optim` procedure in R.

For  $g \geq 2$  the MVM distribution has more parameters than the PN. Having fitted both to a data set, the question remains whether the greater complexity of the MVM is justified by its goodness of fit. We addressed this by computing the Akaike Information Criterion (AIC) (Akaike, 1973) for each of the alternative models:

$$A = -2\ell + 2N_p, \tag{2}$$

where  $\ell$  is the maximized (full) log likelihood for a model with  $N_p$  parameters. The model with the smallest  $A$  is selected, so effectively the selection is based on the likelihood with a penalty for the model complexity as measured by  $N_p$ .

### 2.3 Modelling variations in the mean vector of the PN distribution

Wang and Gelfand (2013) showed that the PN distribution for angular observations can be fitted in the form of a regression model in which the elements of the mean vector  $\boldsymbol{\mu}$  are expressed as linear functions of predictors. The parameters  $\tau$  and  $\rho$  remain fixed,

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



but considerable flexibility in the form of the angular distribution is still possible. In this paper we consider the case where the predictor is a categorical variable with  $P$  levels (e.g.  $P$  stratigraphic units). The modelled values of the mean vector for  $n$  observations are

$$\boldsymbol{\mu} = \mathbf{MX}, \quad (3)$$

where  $\mathbf{X}$  is a  $P \times n$  design matrix with element  $\mathbf{X}\{i, j, \}$  set to 1 if the  $i$ th observation is in the  $j$ th level of the categorical predictor and set to 0 otherwise and

$$\mathbf{M} = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,P} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,P} \end{bmatrix}$$

where  $\mu_{1,j}$  is the modelled value of  $\mu_1$  for an observation corresponding to the  $j$ th level of the categorical predictor. The modelled values of  $\mu_2$  are defined similarly from the elements of the second row of  $\mathbf{M}$ . The elements of  $\mathbf{M}$  may be estimated by maximum likelihood by substituting Eq. (3) for  $\boldsymbol{\mu}$  in the likelihood function.

## 3 Case studies

### 3.1 West Cumbria dip directions

#### 3.1.1 The data

The first two data sets consist of the observations of dip direction recorded for two units in the British Geological Survey's map sheets at 1 : 50 000 in West Cumbria, north-west England. The units are the Sherwood Sandstone Group (Triassic sandstone) and the Windermere Supergroup (Ordovician–Silurian mudstone, sandstone and limestone). A total of 90 observations were available for the Sherwood Sandstone Group and 572 for the Windermere Supergroup. The data are shown in Fig. 1a and b by rose diagrams which show the relative frequency of observations in bins of width  $\pi/10$  radians.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 3.1.2 MVM distribution and comparison with PN

Table 1 shows the results from comparison of MVM distributions with increasing numbers of components. In the case of the Sherwood Sandstone Group there was evidence to select the MVM distribution with two components over a single von Mises distribution, but not to reject the model with two components in favour of one with three. In the case of the Windermere Supergroup a mixture model with five components was selected.

Table 2 shows the results from the comparison of the selected MVM model with a PN model for both the Sherwood Sandstone Group and the Windermere Supergroup dip directions, and Fig. 2a and b shows the probability densities for these two distributions wrapped around the circle. In both cases the AIC was smaller for the more complex MVM distribution.

In the case of the Sherwood Sandstone Group the clear difference between the PN distribution and the MVM is that in the latter there is a stronger contrast between the tightly distributed subset of dip directions towards the south-west (mean direction is 4.07 radians or 233 degrees) and a second subset with a mean direction close to north (6.03 radians or 346 degrees) and a wider dispersion. Although the PN distribution is bimodal, the contrast between the two modes is less pronounced, and the MVM distribution better captures this heterogeneity in the data. Geologically, the south-west dips are found in the west of the study area in a relatively small area where there is a relatively simple consistent structure and good exposure of the geology. The roughly northern dips are found mainly in northern Cumbria, and their dispersion may reflect the fact both that they are spread out along the crop of the Sherwood Sandstone Group encompassing greater structural variability, and that they are subject to greater observation error because the dips are smaller.

The greater complexity of the MVM model for the Windermere Supergroup may partly reflect the fact that there are more data available to support a complex model, but the fitted model also makes geological sense. In the study area the Windermere

SED

5, 2181–2202, 2013

Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Angular data**

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Supergroup is subject to cylindrical folding which would be expected to give rise to dips of approximately equal frequency in a north-west and south-east direction. Both fitted distributions as illustrated in Fig. 2b show modes in these directions, but the dominant mode is in a south-east direction. The structure of the Windermere Supergroup in the study areas is complicated by a major fault trending north-east–south-west. This may introduce different dominant dip directions locally, which may explain the rather more complex form of the selected MVM distribution. In particular, some of the folding in the vicinity of the fault is overturned, so that both limbs of the fold dip to the south-east. This accounts for the asymmetry between the two dominant modes of both fitted distributions.

## 3.2 Bangladesh anticlinal axial planes and Landsat-derived lineaments

### 3.2.1 The data

These two data sets, from eastern Bangladesh, are presented by Davis and Sampson (2002). They are both orientation data rather than observations with a single direction in the interval  $[0, 2\pi]$ . The first set comprises 32 observations of the orientation of the axial plane of a series of anticlines. The second set comprises orientations of 40 major lineaments identified by interpretation of Landsat imagery. Davis and Sampson (2002) compare the orientations in the two data sets by an analysis assuming that each has an underlying von Mises distribution. In this paper we compare the MVM and PN distributions for the Landsat data set, and then evaluate evidence that the two sets of orientations are different by fitting PN distributions. The original orientation data were first presented as angular data over the range  $[0, \pi]$ , and then doubled so that they are distributed on the circle. Figure 1c and d shows the data as rose diagrams.

### 3.2.2 MVM distribution and comparison with PN

As shown in Table 1, both these variables are better-fitted by a MVM with two components than by a single von Mises distribution, but adding a third component is not justified. As shown in Table 2, the AIC was smallest for the MVM model in the case of the anticline planes, but for the PN model in the case of the Landsat-derived lineaments; in this latter case the likelihood for the PN distribution was larger than for the more complex MVM distribution. Figure 2d shows how the asymmetric and bimodal distribution seen in the rose diagram (Fig. 1d) is fitted by the two models. The MVM represents the complex variation of the data with two distinct components, the PN distribution is also asymmetric and bimodal, but note in particular that the MVM has a tail giving non-zero density between zero and about  $\pi/4$  radians where there are no observations. The Landsat lineaments data illustrate the flexibility of the PN distribution, which in this case fits the data better than the MVM distribution but with fewer parameters.

### 3.2.3 PN distributions for the combined Bangladesh data sets

We considered the two data sets on orientations as a combined data set, allowing us to examine the evidence that the two variables correspond to the same structural features. Alternative models were fitted to the combined data set, and these are detailed in Table 3 along with the maximized log likelihood. Figure 3 shows the density functions wrapped around the circle for all three models.

Model 1 is a PN distribution with single values for all parameters, i.e. all parameters pooled for both the anticlines and the Landsat lineaments. Model 2 has separate PN distributions for the two orientations, but with common values of the variance parameters  $\tau$  and  $\rho$ . This corresponds to Eq. (3) with different values of the mean vector  $\mu$  for the orientation of the anticline planes and the Landsat-derived lineaments.

A PN distribution was fitted to each data set separately, as described in the previous section, for comparison with an alternative MVM distribution. These distributions

SED

5, 2181–2202, 2013

Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



considered together may be treated as a model for the combined data set with all parameters of the PN distribution differing between the anticlines and Landsat lineaments. This is denoted Model 3 in Table 3 and Fig. 3. The log likelihood for this model is the sum of the two log likelihoods for the separate fittings. Note that the parameter  $\rho$  approaches the boundary value 1 for both the anticline planes and the Landsat lineaments. We checked that the maximum likelihood estimation was reliable by computing the profile likelihood for this parameter, and found that the likelihood increased smoothly as  $\rho$  approached 1.

This sequence of models can be regarded as nested: Models 2 and 3 are particular cases of Model 1 with certain parameters set to common values. A simpler model may be compared with a more complex one by computing the  $L$  statistic, twice the log likelihood ratio. Under the null model, the simpler one nested in the alternative, the asymptotic distribution of this statistic is  $\chi^2$  with degrees of freedom equal to the difference in the number of independent parameters. However, in this case the distribution of  $L$  is not a simple  $\chi^2$  for any comparison with Model 3, because  $\rho$  is at the boundary. In Table 3 we present the result for a comparison of Model 1 with Model 2. Note that we can reject Model 1 in favour of Model 2 with different parameters  $\mu$  for the anticlines and the Landsat lineaments, so we can conclude that there is reason to believe that the orientations of the Landsat lineaments differ from those of the anticline axial planes. This is consistent with the conclusion of Davis and Sampson (2002), but our analysis treats the complex distribution of the data more plausibly.

In Model 3,  $\rho$  approaches the boundary at  $\rho = 1$ . The density function given by Eq. (1) is undefined at this boundary. When  $\rho = 1$  and  $\tau\mu_2 \neq \mu_1$  the PN distribution, which we denote by  $PN_1$ , is continuous with support over half the circle. The density of the PN distribution in this case is

$$f_1(\theta|\mu, \tau) = \frac{|c| \times \phi_1(c \times (\tau + \cot(\theta - \theta_0)) - \mu_2)}{\sin(\theta - \theta_0)^2} \quad (4)$$

## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



for values of  $\theta$  between  $\theta_0 = \arctan(1/\tau)$  and  $\theta_0 + \pi$  on the side of the circle that faces  $\boldsymbol{\mu} = (\mu_1, \mu_2)$  and 0 on the side that faces away from  $\boldsymbol{\mu}$ . The constant  $c$  that appears in Eq. (4) is defined as  $c = \frac{\tau\mu_2 - \mu_1}{(\tau^2 + 1)}$  and is guaranteed to be non-zero in this case. When  $\rho = 1$  and  $\tau\mu_2 = \mu_1$  (i.e., when  $c = 0$ ) the PN distribution is discrete with probabilities of  $\Phi_1(-\mu_2)$  and  $1 - \Phi_1(-\mu_2)$  in directions  $\theta_0$  and  $\theta_0 + \pi$  respectively. Similar results hold at the boundary  $\rho = -1$ .

The  $\text{PN}_1$  distribution might be useful because it is more parsimonious than the general PN distribution while retaining flexibility for both unimodal and bimodal distributions on the circle. However, as seen above, the support of  $\text{PN}_1$  is restricted to an interval of width  $\pi$  and so it does not include the uniform distribution on the circle as a special case, which limits its usefulness for inference.

## 4 Conclusions

Only the Bangladesh data presented here have been analysed before, by Davis and Sampson (2002) who assumed a simple von Mises distribution. Our results suggest that this distribution is not appropriate for these data, in both cases there was evidence that a mixture of two von Mises distributions was more suitable than a single von Mises, and in the case of the Landsat-derived lineaments, the projected normal distribution was favoured over the mixture of von Mises. The projected normal distribution and a mixture of von Mises distributions can be used to model variation of angular data in the earth sciences which may be distributed in a complex multimodal and asymmetric way. It is possible to select the number of components in a mixture of von Mises distribution by a sequential testing procedure in which the log likelihood ratio is used as a test statistic to decide whether to add a component to the model, with the distribution of the statistic under the simpler model computed by a Monte Carlo simulation. One may select the number of components for a mixture of von Mises distributions and compare the result with the projected normal distribution, which is potentially more parsimonious model, by computing the Akaike Information Criterion for the two distributions.



Hornik, K., Grün, B., and the R Development Core Team: Package movMF, R package version 0.1–2, 2013. 2185

Jones, T. A.: MATLAB functions to analyze directional (azimuthal) data – I: Single-sample inference, *Comput. Geosci.*, 32, 166–175, 2006. 2183

5 Mardia, K. V. and Jupp, P. E.: *Directional Statistics*, John Wiley & Sons, Chichester, 2000. 2182, 2184

McLachlan G. J. and Peel, D.: On a resampling approach to choosing the number of components in normal mixture models, in: *Computing Science and Statistics*, Vol. 28, edited by: Billard, L., Fisher, N. I., Interface Foundation of North America, Fairfax Station, VA, 260–266, 10 1997. 2186

O'Brien, J. F., Roman, D. C., Dixon, J. P., Power, J. A., and Arnold, R.: Multiple causes for non-eruptive seismic swarms at Mt. Martin, Katmai Volcanic Cluster, Alaska (2004–2008), *J. Volcanol. Geoth. Res.*, 229, 13–22, 2012. 2183

R Development Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3–900051–07–0, available at: 15 <http://www.R-project.org/>, 2013. 2185

Sen, K. and Mamtani, M. A.: Magnetic fabric, shape preferred orientation and regional strain in granitic rocks, *J. Struct. Geol.*, 28, 1870–1882, 2006. 2183

Wang, F. and Gelfand, A. E.: Directional data analysis under the general projected normal distribution, *Statistical Methodology*, 10, 113–127, 2013. 2184, 2187, 2188

20 Witts, D., Hall, R., Nichols, G., and Morley, R.: A new depositional and provenance model for the Tanjung formation, Barito Basin, SE Kalimantan, Indonesia, *J. Asian Earth Sci.*, 56, 77–104, 2012. 2183

## SED

5, 2181–2202, 2013

### Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Table 1.** Selection of Mixture of von Mises (MVM) distribution.

Number of components		$\hat{I}_{g,g+1}$	$\hat{P}_g$
$g$ (null model)	$g + 1$ (proposed model)		
Sherwood Sandstone Group			
1	2	19.8	< 0.001
2	3	2.0	0.273
Windermere Supergroup			
1	2	196.3	< 0.001
2	3	13.2	0.001
3	4	18.2	0.024
4	5	17.1	0.031
5	6	11.6	0.064
Bangladesh anticline axial planes			
1	2	10.3	0.003
2	3	0.1	0.823
Bangladesh Landsat-derived lineaments			
1	2	5.8	0.022
2	3	2.0	0.263

## Angular data

R. M. Lark et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Table 2.** Comparison of mixture of von Mises (MVM) with  $g$  components and projected normal (PN) distributions.

	$g$	$N_p$	MVM $\ell$	AIC	$N_p$	PN $\ell$	AIC
Sherwood Sandstone Group	2	5	-137.39	284.78	4	-145.33	298.66
Windermere Supergroup	5	14	-614.88	1257.76	4	-635.38	1278.76
Bangladesh anticline axial planes	2	5	-14.86	39.72	4	-18.3	44.6
Bangladesh Landsat-derived lineaments	2	5	-34.84	79.68	4	-33.67	75.34

## Angular data

R. M. Lark et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

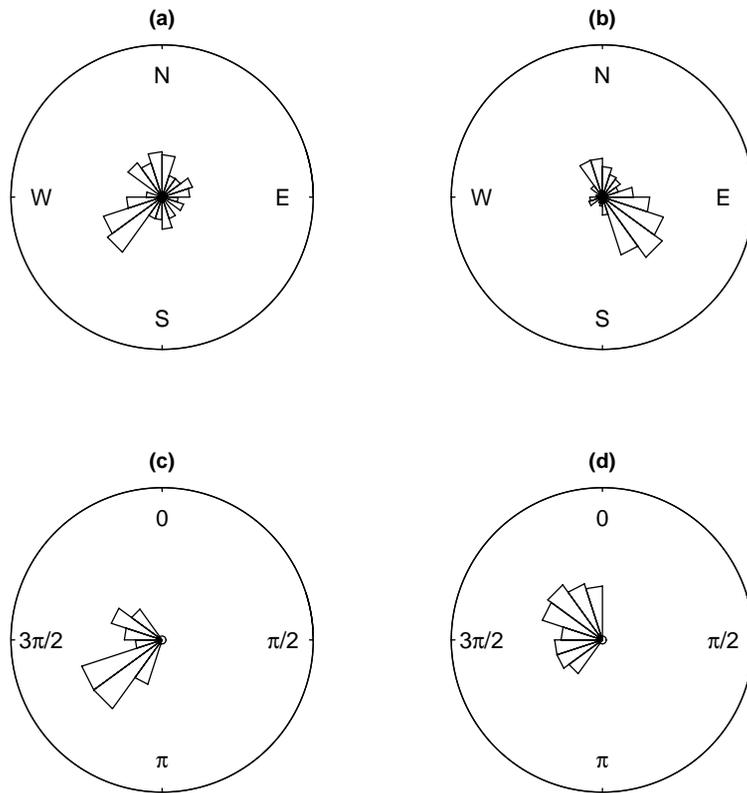
[Interactive Discussion](#)



**Table 3.** Projected normal distributions fitted to Bangladesh anticline axial plane and Landsat-derived lineament orientations (doubled directions). Subscript “A” for a parameter denotes that it pertains only to observations of anticline axial plane orientations, subscript “L” denotes a parameter of the distribution of Landsat lineaments, subscript “P” denotes a pooled parameter.

Model	Parameters								$\ell$
Model 1	$\mu_{1,P}$ 0.05		$\mu_{2,P}$ -3.49		$\tau_P$ 3.94		$\rho_P$ 0.83		-67.5
Model 2	$\mu_{1,A}$ -2.94	$\mu_{1,L}$ 2.34	$\mu_{2,A}$ -4.86	$\mu_{2,L}$ -3.39	$\tau_P$ 3.66		$\rho_P$ 0.72		-52.9
Model 3	$\mu_{1,A}$ -9.30	$\mu_{1,L}$ 2.25	$\mu_{2,A}$ -16.56	$\mu_{2,L}$ -3.5	$\tau_A$ 12.07	$\tau_L$ 3.76	$\rho_A$ 0.999*	$\rho_L$ 0.999*	-52.0
Comparison between models									
Null	Alternative	L	P	$\chi^2$ degrees of freedom					
Model 1	Model 2	29.2	$4.6 \times 10^{-7}$	2					

\* These estimates of  $\rho$  are at the boundary of the parameter space.



**Fig. 1.** Data: direction of dip for **(a)** Sherwood Sandstone group (90 data) and **(b)** Windermere Supergroup (572 data); orientations (doubled from original range of  $[0, \pi]$ ) for **(c)** Bangladesh anticline axial planes (32 data) and **(d)** Bangladesh Landsat-derived lineaments (40 data). Note that segments of the rose diagrams are proportional to relative frequency within each data set separately, so are not comparable between data sets with respect to numbers of observations. If all data appeared within a single bin of the rose diagram then the corresponding segment would be equal in length to the radius of the circle.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

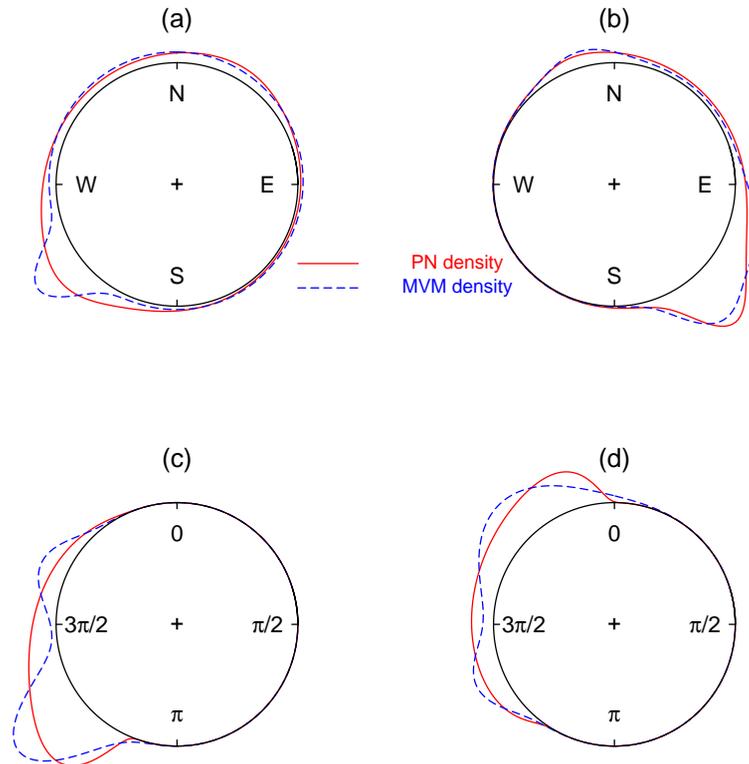
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



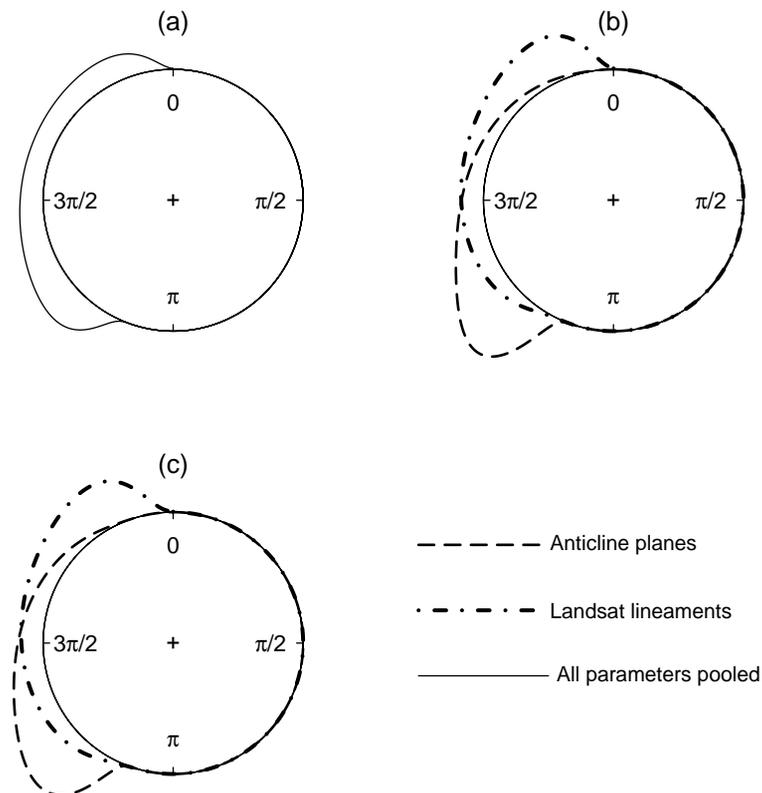


**Fig. 2.** Density of fitted Projected Normal distributions (PN) and mixture of von Mises distributions (MVM) wrapped around the circle: direction of dip for **(a)** Sherwood Sandstone group and **(b)** Windermere Supergroup; orientations (doubled from original range of  $[0, \pi]$ ) for **(c)** Bangladesh anticline axial planes and **(d)** Bangladesh Landsat-derived lineaments.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


## Angular data

R. M. Lark et al.



**Fig. 3.** Density of projected normal distributions fitted to combined Bangladesh data and wrapped around the circle. **(a)** Model 1, with all parameters pooled for the combined data set; **(b)** Model 2, with  $\mu$  modelled separately for anticline axial planes and Landsat lineaments; **(c)** Model 3, with all parameters separate for anticline axial planes and Landsat lineaments.