Modeling a Multi-modal Distribution of Wind Direction Data in Kudat, Malaysia

Nurulkamal Masseran
School of Mathematical Sciences, Faculty of Science and Technology
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
kamalmsn@ukm.edu.my

Abstract

Wind direction is the direction from which the wind is blowing. It is expressed in terms of degrees measured clockwise from geographical direction. The knowledge of the wind direction can be used to obtain information about the wind energy potential, dispersion of particulate matter in the air, the effects of engineering structures on the building, maritime study, and etc. This study provides a suitable model for the wind direction that indicates multi-modal distributional properties. A case study involves with a data from Kudat, Malaysia has been analysed. The statistical models known as a Finite Mixture of von Mises Fisher (mvMF) and Circular Distribution based on Nonnegative Trigonometric Sums (NNTS) has been fitted to the data. Then, the suitability of mvMF and NNTS models were judged based on a graphical representation and goodness-of-fit statistics. The results found that the mvMF model with $H \geq 4$ components is sufficient to provide a best model.

Keywords: Circular statistics, Dominant direction, Environmental statistics, Multi-modal distribution, Wind direction modelling.

1. Introduction

Wind direction is known as a type of circular or directional data. Thus, it has unique characteristics that are different from standard linear data sets. Such distinctive features have made directional statistics analysis substantially different from linear analysis (Mardia and Jupp, 1999; Masseran, 2015). Let θ be a random variable that measures the directional data that take values in the range 0° to 360° or 0 to 2π. An analysis of θ would depend on the selection of the starting point as the "zero-direction" and the sense of rotation. For example, in Figure 1, if the zero direction is due east, corresponding to anti-clockwise rotation, the data will take the value of 60°, whereas if the zero direction is due north, corresponding to clockwise rotation, the data will take the value of 30°. However, the "beginning" are always coincides with the "end", i.e., 0°-360°, and the measurement is also periodic, with θ being the same as $\theta + p \times 2\pi$ for any integer *p*. The starting point and rotation from this point, regardless of whether it is clockwise or anticlockwise, are taken as positive values. Observations using these two dimensions are also called circular or directional data (Jammalamadaka and SenGupta, 2001).

In addition, directional data that take values of $\theta = 0°$ to $360°$ or $\theta = 0$ to 2π are commonly termed polar coordinate data with magnitude = 1, namely, (1, θ). On the other hand, the directional data can be transformed into rectangular coordinate form, (*X*, *Y*), through $x = \cos\theta$ and $y = \sin\theta$ for every θ. Figure 2 shows the directional data in terms of both polar coordinates (1, θ) and rectangular coordinate (*X*, *Y*). There are many other unique features of directional data; for further reference, see (Mardia and Jupp, 1999; Jammalamadaka and SenGupta, 2001; Fisher, 1993).

In practice, the wind direction is an important feature that should be considered in building wind turbines and in structural and environmental design analysis. For that purpose, the statistical models always provide good information to describe about the behaviours of wind direction. In fact, the finite mixtures of von Mises distributions is among the most commonly used and was found to provide good results for the purpose of modelling the circular data, particularly for the wind direction. For examples, Masseran et al. (2013) showed that the finite mixture of the von Mises distribution with H number of components was the best distribution to describe the wind direction distributions in Malaysia. Carta et al. (2008) have showed that a the finite mixture of von Mises is a very flexible model for wind direction studies particularly for the wind direction regimes in zones with several models or prevailing wind directions. In fact, the same result have been showed by Azmani et al. (2009) in modelling the sensor data for the cases of a recursive change point estimate of the wind speed and direction. Apart from that, Heckenbergerova et al. (2015) showed that the finite mixture of von Mises distribution is able to provide an optimized solution for wind direction parameters corresponds to the Particle Swarm Optimization method. Thus, since the von Mises distribution is a flexible model for addressing wind directional data with several modes, this study attempts to describe and compare the suitability of von Mises model with the circular model based on nonnegative trigonometric sums in order to determine the best statistical model for wind direction data in Kudat.

2. Study area, data and descriptive statistics

Sabah is a state of Malaysia, located in the northern section of the island of Borneo (6.8833° N, 116.8333° E). It is the second largest state in the country after Sarawak, which it borders to the southwest. Sabah is relatively wet (annual precipitation exceeding 200 mm) due to the tail effect of typhoons, which frequently traverse the Philippine islands across the South China Sea. It is worth mentioning that from April to November each year, when typhoons frequently develop over the west Pacific and move westward across the Philippines, the south-westerly winds over the northwest coast of Sabah may reach speeds of 10.30 m/s or more (Masseran, 2013a).

The data used in this study were obtained from the Malaysian Meteorological Department. In this study, hourly wind direction data from 1 January 2009, to 30 November 2009 were used. Wind direction data are circular because they are recorded in terms of degrees, from 0° to 360°. However, for modelling, data transformation into radian units can be performed easily. Apart from that, the missing data has been estimated by using the method of single imputation (Masseran, 2013b).

Before a detailed analysis is conducted, it is important to evaluate the descriptive statistics to obtain some preliminary information about the data. As mentioned above, directional data have many features that differ from those of standard linear data sets. For example, the arithmetic mean, which is commonly used for linear data, cannot be used as a measure of the centre of the directional data. The sample variance s^2, which depends on

the sample mean, also suffers from the same problem. Thus, we need an alternative measure of centre and dispersion when dealing with directional data (Jammalamadaka and SenGupta, 2001). Let $\theta_1, \theta_2 \ldots, \theta_n$ be a set of directional data; the mean direction can then be calculated as

$$\bar{\theta} = \begin{cases} \tan^{-1}\left(S/C\right) & , \text{ if } C > 0, S \geq 0 \\ \pi/2, & , \text{ if } C = 0, S > 0 \\ \tan^{-1}\left(S/C\right) + \pi & , \text{ if } C < 0 \\ \tan^{-1}\left(S/C\right) + 2\pi, & \text{ if } C \geq 0, S < 0 \\ \text{undefined} & , \text{ if } C = 0, S = 0 \end{cases} \qquad (1)$$

where, $C = \sum_{i=1}^{n} \cos\theta_i$ and $\mathrm{S} = \sum_{i=1}^{n} \sin\theta_i$ accomplish the polar-to-rectangular transformation. On the other hand, the measure for the dispersion of the directional data is commonly derived from the circular variance, which is given as

$$V = 1 - \frac{1}{n}\sqrt{C^2 + S^2} \qquad (2)$$

A small value of the circular variance indicates the data have a large concentration around the mean direction.

The percentile measure for directional data is same as that for linear data: it is a measure of the value of a variable below a certain percent of observations. For example, the (100*p*)-th percentile is often called the quantile of order *p*. Let $y_1 \leq y_2 \leq \ldots \leq y_n$ be an order statistic for *n* observations; y_r is then the quantile of order $p = \frac{r}{(n+1)}$ as well as the $\frac{100r}{(n+1)}$ percentile. Thus, the *p*-th percentile of the data is also a quantile of order *p* for the data. Using these descriptive measurements, Table 1 shows the descriptive statistics for the wind directional data in Kudat. Based on the descriptive statistics in Table 1, the circular mean of the wind direction is approximately 218.21°. However, the circular variance is 0.978, which implies that the data were not well concentrated around their mean direction. Thus, we suspect that the data are either approximately uniformly distributed or have a several-directional mean. The values of 25th, 50th, and 75th percentile are 80°, 220°, and 240°, respectively.

3. Statistical models for the multi-modal distribution of circular data

In this study, 2 statistical model known as Finite Mixture of von Mises Fisher and Circular distribution based on nonnegative trigonometric sums has been fitted to the data of wind direction in Kudat that indicated properties of multi-modal distribution. In order

to determine the best fitted model, the suitability of both of these models will be compare by using the Akaike's Information Criterion (AIC) method.

3.1 Finite Mixture of von Mises-Fisher distributions (mvMF)

In some applications, the observed wind direction data cannot be represented by a unimodal distribution. To overcome this problem, a finite von Mises-Fisher mixture distribution (mvMF), which is comprised of a sum of *H* von Mises probability distributions, has been proposed. Let θ be a random variable representing the wind direction in radians, and let $\mathbf{x}' = [\cos\theta_i, \sin\theta_i]'$ be a circular data point in rectangular coordinates. Then, the mixture of von Mises-Fisher distributions is given by

$$f(\mathbf{x};\boldsymbol{\mu}_h,\kappa_h) = \sum_{h=1}^{H} \omega_h f(\mathbf{x};\boldsymbol{\mu}_h,\kappa_h) = \sum_{h=1}^{H} \omega_h c_d(\kappa_h) e^{(\kappa_h \mathbf{x}'\boldsymbol{\mu}_h)} \qquad (3)$$

where $f(\mathbf{x};\boldsymbol{\mu},\kappa) = c_d(\kappa) e^{(\kappa \mathbf{x}'\boldsymbol{\mu})}$ is a single von Mises Fisher model, d is a dimension of random vector $\mathbf{x}$ (in our cases d=2), $\boldsymbol{\mu}_h, \kappa_h$ are the parameter mean direction and concentration parameter, respectively, for *h=1, 2,... H* components of the von Mises distribution, while ω_h is a mixing parameter of nonnegative quantities that sum to one, given by

$$0 \le \omega_h \le 1 \quad \text{and} \quad \sum_{h=1}^{H} \omega_h = 1 \quad \text{for} \quad (h = 1, 2\ldots, H) \qquad (4)$$

$c_d(\kappa_h)$ is a normalising constant given by $c_d(\kappa_h) = \dfrac{\kappa_h^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(\kappa_h)}$, where $I_r(.)$ represents the modified Bessel function of the first kind and order *r*. The MLEs for the mvMF are very difficult to derive in a standard way. However, Banerjee et al. (2005) provided a solution of the parameter estimates for the mvMF distribution based on the expectation maximisation (EM) approach. Let $\boldsymbol{\alpha}_h = (\boldsymbol{\mu}_h, \kappa_h)$ denote the parameters of the von Mises-Fisher distribution, $\boldsymbol{\alpha}_h = (\boldsymbol{\mu}_h, \kappa_h)$, for $1 < h < H$. Then, the mvMF distribution can be written as

$$f(\mathbf{x};\Theta) = \sum_{h=1}^{H} \omega_h f_h(\mathbf{x} \mid \boldsymbol{\alpha}_h) \qquad (5)$$

where $\Theta = (\omega_1, \omega_1, \ldots, \omega_H, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_H)$. According to Banerjee et al. (2005), to generate a random sample from this mixture distribution, the *h*-th von Mises distribution is randomly chosen with probability ω_h. Let, $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}'$ be a data set of *n* independent sample points following Equation (17), and let $Z = \{z_1, z_2, \ldots, z_n\}'$ be the corresponding set of hidden random variables that indicate a particular von Mises distribution from which a sample is generated. In particular, $z_i = h$ if $\mathbf{x}_i$ is generated from $f_h(\mathbf{x} \mid \boldsymbol{\alpha}_h)$. Thus, the log-likelihood can be written as

$$\ln f(X, Z \mid \Theta) = \sum_{i=1}^{H} \ln\left(\omega_{z_i} f_{z_i}(\mathbf{x}_i \mid \boldsymbol{\alpha}_{z_i})\right) \qquad (6)$$

Assume that the posterior distribution, $p(h|\mathbf{x}_i,\Theta)$, of the hidden variables $Z|(X,\Theta)$ are known. Then, the expectation of the log-likelihood over the given posterior distribution *p* is given by

$$E_p\left[\ln P(X,Z|\Theta)\right]=\sum_{i=1}^{n}E_p\left[\ln\left(\omega_{z_i}f_{z_i}\left(\mathbf{x}_i|\boldsymbol{\alpha}_{z_i}\right)\right)\right] \quad (7)$$

$$=\sum_{i=1}^{n}\sum_{h=1}^{H}\ln\left(\omega_h f_h\left(\mathbf{x}_i|\boldsymbol{\alpha}_h\right)\right)p(h|\mathbf{x}_i,\Theta)$$

$$=\sum_{h=1}^{H}\sum_{i=1}^{n}(\ln\omega_h)p(h|\mathbf{x}_i,\Theta)+\sum_{h=1}^{H}\sum_{i=1}^{n}\ln\left(f_h\left(\mathbf{x}_i|\boldsymbol{\alpha}_h\right)\right)p(h|\mathbf{x}_i,\Theta)$$

Next, the parameter Θ is re-estimated to maximise the expectation function. To maximise the expectation function with respect to ω_h, the Lagrangian multiplier λ corresponding to the constraint $\sum_{h=1}^{H}\omega_h=1$ is used, and by taking the partial derivatives with respect to each ω_h from the Lagrangian, the following is obtained:

$$\sum_{i=1}^{n}p(h|\mathbf{x}_i,\Theta)=-\lambda\omega_h \quad (8)$$

Next, by summing both sides of Equation (20) over all *h*, Banerjee at al. (2005) found that $\lambda=-n$; thus, the parameter estimate for ω_h is given by

$$\hat{\omega}_h=\frac{1}{n}\sum_{i=1}^{n}p(h|\mathbf{x}_i,\boldsymbol{\Theta}) \quad (9)$$

The parameter estimates for $\boldsymbol{\alpha}_h=(\boldsymbol{\mu}_h,\kappa_h)$ can be derived under the constraints $\|\boldsymbol{\mu}\|=1$ and $\omega_h\geq 0$ for *h=1, 2, ... H.* Let λ_h be the Lagrange multiplier corresponding to the constraint; if $\kappa=0$, then $f(\mathbf{x}|\boldsymbol{\alpha})$ is the uniform distribution on the sphere, and if $\kappa>0$, then the multiplier for the inequality constraint has to be zero. Thus, the Lagrangian is given as

$$L\left(\{\boldsymbol{\mu}_h,\kappa_h,\lambda_h\}_{h=1}^{H}\right)=\sum_{h=1}^{H}\sum_{i=1}^{n}\ln\left(f_h\left(\mathbf{x}_i|\boldsymbol{\alpha}_h\right)\right)p(h|\mathbf{x}_i,\Theta)+\sum_{h=1}^{H}\lambda_h\left(1-\boldsymbol{\mu}_h'\boldsymbol{\mu}_h\right)$$

$$=\sum_{h=1}^{H}\left[\sum_{i=1}^{n}\left(\ln c_d\left(\kappa_h\right)\right)p(h|\mathbf{x}_i,\Theta)+\sum_{i=1}^{n}\kappa_h\boldsymbol{\mu}_h'\mathbf{x}_i\,p(h|\mathbf{x}_i,\Theta)+\lambda_h\left(1-\boldsymbol{\mu}_h'\boldsymbol{\mu}_h\right)\right] \quad (10)$$

By taking the partial derivative with respect to $\{\boldsymbol{\mu}_h,\kappa_h,\lambda_h\}_{h=1}^{H}$ from Equation (22), and setting it equal to zero, for each *h*, Banerjee et al. (2005) obtained

$$\boldsymbol{\mu}_h=\frac{\kappa_h}{2\lambda_h}\sum_{i=1}^{n}\mathbf{x}_i\,p(h|\mathbf{x}_i,\Theta) \quad (11)$$

$$\boldsymbol{\mu}_h'\boldsymbol{\mu}_h=1 \quad (12)$$

$$\frac{c_d^{'}(\kappa_h)}{c_d(\kappa_h)}\sum_{i=1}^{n} p(h\,|\,\mathbf{x}_i,\Theta) = -\boldsymbol{\mu}_h^{'}\sum_{i=1}^{n}\mathbf{x}_i\, p(h\,|\,\mathbf{x}_i,\Theta) \qquad (13)$$

Using Equations (23) and (24), they found that

$$\lambda_h = \frac{\kappa_h}{2}\left\|\mathbf{x}_i\, p(h\,|\,\mathbf{x}_i,\Theta)\right\| \qquad (14)$$

and

$$\boldsymbol{\mu}_h = \frac{\sum_{i=1}^{n}\mathbf{x}_i\, p(h\,|\,\mathbf{x}_i,\Theta)}{\left\|\sum_{i=1}^{n}\mathbf{x}_i\, p(h\,|\,\mathbf{x}_i,\Theta)\right\|} \qquad (15)$$

Next, substituting Equation (27) into Equation (25) provides the parameter estimates for κ_h as given by

$$\frac{c_d^{'}(\kappa_h)}{c_d(\kappa_h)} = \frac{\left\|\sum_{i=1}^{n}\mathbf{x}_i\, p(h\,|\,\mathbf{x}_i,\Theta)\right\|}{\sum_{i=1}^{n} p(h\,|\,\mathbf{x}_i,\Theta)} \qquad (16)$$

where $p(h\,|\,\mathbf{x}_i,\Theta) = \dfrac{\omega_h f_h(\mathbf{x}_i\,|\,\Theta)}{\sum_{l=1}^{k}\omega_l f_l(\mathbf{x}\,|\,\Theta)}$. Readers desiring a detailed discussion of the parameter estimates for the von Mises-Fisher distribution should consult (Banerjee et al., 2005; Hornik and Grun, 2014; Mooney et al., 2003).

3.2 Circular Distribution based on Nonnegative Trigonometric Sums (NNTS)

The NNTS model has been found to be flexible enough to model directional data sets exhibiting multimodality or skewness. The nonnegative trigonometric sum series for a circular variable θ has been expressed by Fejer (1915) as the squared modulus of a sum of complex numbers, which can be written as

$$\left\|\sum_{k=0}^{M} c_k e^{ik\theta}\right\|^2, \quad \text{for } k = 0, 1, 2, \ldots, M \qquad (17)$$

where $\theta \in 2\pi$, $i = \sqrt{-1}$, and c_k is a complex parameter. Using this series, Fernandez-Duran [15] proposed a new family of distributions for circular random variables, given as

$$f(x_j; M, \underline{\theta}) = \left\|\sum_{k=0}^{M}\theta_k e^{ikx_i}\right\|^2 = \sum_{k=0}^{M}\sum_{l=0}^{M}\theta_k\bar{\theta}_l e^{i(k-l)x_i} \qquad (18)$$

where the parameters (a_k, b_k) are expressed in terms of the complex parameter $a_k - ib_k = 2\sum_{v=0}^{n-k} c_{v+k}\bar{c}_v$. An additional constraint, $\sum_{k=0}^{n}\|c_k\|^2 = \dfrac{1}{2\pi} = a_0$, is imposed to make the

trigonometric sum integrate to 1. Thus, there are 2**M* free parameters, where the parameter c_0 must be real and positive.

Next, to ensure that the integration for the probability density function is equal to one, $\int_\theta f(\theta; M, \boldsymbol{\alpha})\, d\theta = 1$, a constraint on the parameter α is needed, namely,

$$\left\| \sum_{k=0}^{M} \alpha_k \right\|^2 = \frac{1}{2\pi} \qquad (19)$$

If $\alpha_{c0} = 0$ and $\alpha_{rk} > 0$, then α_k is a non-negative real number. It then follows that Equation (3) can be written in the quadratic form

$$f(\theta_j; M, \boldsymbol{\alpha}) = \boldsymbol{\alpha}^T \mathbf{e}_j \mathbf{e}_j{}^T \boldsymbol{\alpha} \qquad (20)$$

where $\mathbf{e}_j = \left(1, e^{-i\theta_j}, e^{-2i\theta_j}, e^{-3i\theta_j}, \ldots, e^{-Mi\theta_j}\right)'$ and $\boldsymbol{\alpha}^T$ is a transposition and conjugate for vector $\boldsymbol{\alpha}$. Then, the likelihood function can be written as

$$L(M, \boldsymbol{\alpha} \mid \theta_1, \theta_2, \ldots \theta_n) = \prod_{k=1}^{n} \boldsymbol{\alpha}^T \mathbf{e}_j \mathbf{e}_j{}^T \boldsymbol{\alpha} \qquad (21)$$

and the log-likelihood function is derived as

$$\ln\left[L(M, \boldsymbol{\alpha} \mid \theta_1, \theta_2, \ldots \theta_n)\right] = \sum_{k=1}^{n} \boldsymbol{\alpha}^T \mathbf{e}_j \mathbf{e}_j{}^T \boldsymbol{\alpha} \qquad (22)$$

To determine the maximum likelihood estimator for parameter α, Fernandez-Duran & Gregorio-Dominguez (2010) suggested the application of the Newton-like manifold algorithm using the procedures described by Absil et al. (2008). Using the Newton-like manifold algorithm, the maximum likelihood estimator for the NNTS model can be solved by

$$gradl(\boldsymbol{\alpha}) = \mathbf{0} \qquad (23)$$

where $gradl(\boldsymbol{\alpha})$ is a gradient for the log-likelihood function of the value of $\boldsymbol{\alpha}$ in Equation (6). The solution of Equation (7) is the critical point of the real log-likelihood function for the NNTS model. The Newton-like manifold algorithm method provides the solution for the maximum likelihood estimator through several steps:

i) Determine the initial value for $\boldsymbol{\alpha}_0$

ii) For $k = 0, 1, \ldots,$ solve the Newton equation, $Hessl(\boldsymbol{\alpha}_k)\boldsymbol{\eta}_k = -gradl(\boldsymbol{\alpha}_k)$ for the unknown $\boldsymbol{\eta}_k$ in the tangent space at $\boldsymbol{\alpha}_k$.

iii) Set $\boldsymbol{\alpha}_{k+1} = R_\theta(\boldsymbol{\eta}_k)$, where R_θ is a retraction from the tangent space onto the manifold at α_k

This algorithm terminates when the norm of the gradient or the norm of the difference $gradl(\boldsymbol{\alpha}_k) = -gradl(\boldsymbol{\alpha}_{k-1})$ is less than some specified error. For example, Fernandez-Duran & Gregorio- Dominguez (2010) have used the differentiation rule of real functions of a complex vector, as follows

$$gradl(\boldsymbol{\alpha}) = P_\theta \left(\frac{\partial(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^T} \right)^T = \frac{1}{2\pi} \sum_{k=1}^{n} \frac{\mathbf{e}_k}{\boldsymbol{\alpha}^T \mathbf{e}_k} - n\boldsymbol{\alpha} \qquad (24)$$

Apart from that, the NNTS model in Equation (2) can also be expressed as

$$f(\theta; M, \boldsymbol{\alpha}) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{M} \left[a_k \cos(k\theta) + b_k \sin(k\theta) \right] \qquad (9)$$

where $a_k - ib_k = 2\sum_{\nu=0}^{M-k} \alpha_{\nu+k} \bar{\alpha}_\nu$ for $k = 1, 2, \ldots, M$.

4. Results and discussion

As mentioned above, the objective of this study was to identify the most appropriate distribution for wind direction at the Kudat station to better understand the wind regime in this area. Tables 2 and 3 show the parameter estimates for the mvMF and NNTS model respectively. Figure 3 presents the fitted mvMF and NNTS for the wind direction at the Kudat station. From the figure, it is clear that the single mvMF distribution (H=1) failed to model the wind direction data at the Kudat station accurately. The same conclusion is draw from NNTS model (M=1). However, as the number of components of the mvMF and NNTS model increase, both of the models fit the data in a more precise way. As a result, the fitted mvMF and NNTS models with 2, 3, …, 8 components model the data with similar accuracies. It is quite difficult to determine the suitability precision of both models based on graphical representations only. Thus, the Akaike's Information Criterion (AIC) was used to evaluate the performance of both models.

Table 4 and Figure 4 show the AIC comparison each fitted mvMF and NNTS model. From Table 4 and Figure 4, by comparing each mvMF, it is clear that the single mvMF has the highest AIC value, implying that the single mvMF is not a good model for wind direction in Kudat. In fact, this result is similar for the NNTS model with M=1 component. However, as the number of components increase for both mvMF and NNTS, the AIC values decrease, which implies that the use of more components in the mvMF and NNTS models provides a model that more adequately fits the data. In addition, by comparing the values of AIC for both models, it is found that the AIC values for the mvMF models are lower than those for the NNTS models for all components. For example, the value of AIC and BIC for the mvMF model with H=4 components are lower than those for the NNTS model with M=1, 2, 3 4, 5, and 6 components. Therefore, the mvMF models were able to provide better results in fitting the wind directional data in Kudat compared to the NNTS models. Thus, the mvMF model is preferred to the NNTS model for fitting the wind directional data in Kudat. In fact, a suitable mathematical equation for the Kudat wind directional data that can written as a mvMF (H=4) model is given by

$$f(\mathbf{x};\boldsymbol{\mu},\kappa) = (0.3474064)\frac{11.658^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(11.658)} e^{(11.658\boldsymbol{\mu}_1^T\mathbf{x})} + (0.1439757)\frac{9.227^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(9.227)} e^{(9.227\boldsymbol{\mu}_2^T\mathbf{x})}$$

$$+ (0.1412261)\frac{10.752^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(10.752)} e^{(10.752\boldsymbol{\mu}_3^T\mathbf{x})} + (0.3673919)\frac{0.971^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(0.971)} e^{(0.971\boldsymbol{\mu}_4^T\mathbf{x})}$$

with the parameter of mean directions, $\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\\ \boldsymbol{\mu}_3\\ \boldsymbol{\mu}_4\end{bmatrix}^T = \begin{bmatrix}0.56772170 & 0.8232205\\ -0.31381192 & -0.9494852\\ -0.63765746 & -0.7703200\\ 0.02245815 & 0.9997478\end{bmatrix}^T$.

Since the mvMF has been determined to be a good model for the data, it can be used to describe some characteristics of the wind direction in Kudat. In this study, the parameter $\boldsymbol{\mu}$ for mvMF has been defined in terms of rectangular coordinates. The interpretation of the dominant direction of the wind is not suitable to be described in this way. Thus, by transforming the results into units of degrees, $0 \le \mu < 360$ may be more appropriate. Based on Equation (31), the measured parameters for the mean directions in terms of degrees are 233.37°, 256.06°, 55.52° and 82.65°. In addition, Figure 5 shows a circular density plot for the mvMF with *H*=4 components. This figure clearly shows that most of the wind was blowing from the north-northeast and the west-southwest and some from the east-southeast. The circular density plot reveals that the wind direction has two different dominant directions: from 190°-270° with mean directions of 233.37° to 256.06° with respect to the parameter concentration $\kappa = 11.658$ and $\kappa = 10.752$, while the other dominant direction are found to be in the range of 30°-90° with mean directions of 55.52° and 82.65° and also the concentration parameter $\kappa = 9.227$ and $\kappa = 0.971$. These imply that a stronger concentration about the mean direction of the wind blow comes from the South-West direction and follow by the minor dominant direction of the wind blows from the North-East direction.

Apart from that, the others direction are found to be quite uniformly distributed. Determining the dominant wind direction will contribute valuable information to planning or forecasting activities in such sectors as wind energy generation, air pollution assessment, climate change, construction, and maritime activities. For example, in wind energy evaluation, based on this information, the wind turbine can be positioned such that the production of energy is maximised.

6. Conclusions

Our study focused on determining the best statistical model for wind direction in the Kudat region. The mixtures von Mises-Fisher distribution and Circular distribution based on nonnegative trigonometric sums thereof were fit to the data. The results obtained showed that the mixtures von Mises-Fisher distribution was able to be presented as a better model than a Circular distribution based on nonnegative trigonometric sums. In

fact, it was found that mixtures von Mises-Fisher distributions with $H \geq 4$ components adequately modelled the wind direction distribution in Kudat. Apart from that, circular plots of the mvMF model clearly show that several wind directions are more dominant in Kudat, while the other directions show an approximately uniform dispersion.

Appendix

(a) Bessel Function and its Important Properties:

Bessel function, $J_v(z)$ for the order- v^{th}, is a solution for the differential equation given by

$$\frac{\partial^2 x}{\partial z^2} + \frac{1}{z}\frac{dx}{dz} + \left(1 - \frac{v^2}{z^2}\right)x = 0$$

In addition, the Bessel function, $J_v(z)$ can also be written in term of the integration form given by

$$J_v(z) = \frac{\left(z/2\right)}{\Gamma\left(1/2\right)\Gamma\left(v+1/2\right)}\int_0^{\pi} e^{iz\cos\theta}\sin^{2v}\theta d\theta \quad , \quad \left(v > -\frac{1}{2}\right)$$

While in the the expansion series for Bessel function, $J_v(z)$ can be written as

$$J_v(z) = \sum_{r=0}^{\infty}(-1)^r \frac{\left(z/2\right)^{v+2r}}{r!\Gamma\left(v+r+1\right)}$$

For any positive integer *n*, the integration form of the Bessel function can also be written as

$$J_n(z) = \frac{1}{\pi}\int_0^{\pi}\cos\left(n\theta - z\sin\theta\right)d\theta$$

In fact, $J_{-v}(z) = (-1)^v J_v(z)$, dan $\frac{d}{dz}\left\{z^v J_v(z)\right\} = z^v J_{v+1}(z), \quad v > 0.$

In particular,

$$\begin{aligned} J_0(z) &= \frac{1}{2\pi}\int_0^{2\pi} e^{iz}\cos\theta d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi}\cos\left(z\cos\theta\right)d\theta \\ &= \frac{2}{\pi}\int_0^{\pi/2}\cos\left(z\cos\theta\right)d\theta \end{aligned}$$

and $\frac{d}{dz}\left(zJ_1(z)\right) = zJ_0(z)$. Next, the modified Bessel function of the 1st kind with order-*v* has been descibe as the following Equations

$$
\begin{aligned}
I_v(z) &= i^{-v} J_v(iz) \\
&= \frac{1}{2\pi}\int_0^{2\pi} \cos v\theta e^{z\cos\theta} d\theta \\
&= \sum_{r=0}^{\infty} \frac{\left(z/2\right)^{v+2r}}{r!\Gamma(v+r+1)}
\end{aligned}
$$

In particular, $I_0(\rho) = \frac{1}{2\pi}\int_0^{2\pi} e^{\rho\cos\alpha} d\alpha$ has been determined as normalizing constant for the von Mises-Fisher distribution. The other properties of the modified Bessel function are explained based on the following Equations

$$I_{-v}(z) = I_v(z),$$

$$I_{v-1}(z) + I_{v+1}(z) = 2I_v'(z),$$

$$I_{v-1}(z) - I_{v+1}(z) = \frac{2v}{z} I_v(z),$$

$$\frac{d}{dz} I_0(z) = I_0'(z) = I_1(z),$$

$$
\begin{aligned}
I_0''(z) = I_1'(z) &= \frac{1}{2}\left(I_0(z) + I_2(z)\right) \\
&= I_2(z) + \frac{I_1(z)}{z} \\
&= I_0(z) - \frac{I_1(z)}{z}
\end{aligned}
$$

By referring to the above mentioned properties, it is found that $\frac{d}{d\kappa}\left(\kappa I_1(\kappa)\right) = \kappa I_0(\kappa)$, and $\frac{d}{d\kappa}\left(\frac{I_1(\kappa)}{\kappa}\right) = \frac{I_2(\kappa)}{\kappa}$. Next, the properties of monotonic increasing function for the parameter κ in the Bessel function can be explain based on the ratio of $A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$ with $A'(\kappa) = 1 - \frac{A(\kappa)}{\kappa} - A^2(\kappa)$. Apart from that, some approximations for the Bessel function that having a small or large value of κ can be solve using the following equations:

If κ is large,

$$I_p(\kappa)=\frac{e^{\kappa}}{\sqrt{2\pi\kappa}}\left\{1-\frac{\left(4p^2-1\right)}{8\kappa}+\frac{\left(4p^2-1\right)\left(4p^2-9\right)}{2\left(8\kappa\right)^2}\right\}$$
$$-\frac{\left(4p^2-1\right)\left(4p^2-9\right)\left(4p^2-25\right)}{6\left(8\kappa\right)^3}+\ldots$$

In particular $I_0(\kappa)\sim\frac{e^{\kappa}}{\sqrt{2\pi\kappa}}$, and $A(\kappa)\sim\left(1-\frac{1}{2\kappa}-\frac{1}{8\kappa^2}-\ldots\right)$.

If κ is small, $I_0(\kappa)\sim 1+\frac{\kappa^2}{4}+\frac{\kappa^4}{64}+\ldots,$ and $A(\kappa)\sim\frac{\kappa}{2}\left(1-\frac{\kappa^2}{8}+\frac{\kappa^4}{48}-\ldots\right)$.

For more detail regarding the properties of Bessel function, please refer to Jammalamadaka and SenGupta, 2001; Stephenson, 1973).

Acknowledgements

The authors are indebted to the staff of the Malaysian Meteorology Department for providing wind direction data. This research would not have been possible without sponsorship from the Ministry of Education, Malaysia (grant number: FRGS/1/2014/SG04/UKM/03/1 and GGPM-2014-056).

References

1. Mardia, K. V. and Jupp, P. E. (1999). *Directional Statistics*. John Wiley, Chichester.
2. Masseran, N. (2015). Markov chain model for the stochastic behaviors of wind-direction data. *Energy Conversion and Management*, 92; 266-274.
3. Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in circular statistics*. World Scientific Publishing, Singapore.
4. Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge University Press, Cambridge.
5. Masseran, N., Razali, A. M., Ibrahim, K., and Latif, M. T. (2013). Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. *Energy Conversion and Management*, 72: 94-102.
6. Carta, J. A., Bueno, C. and Ramirez, P. (2008). Statistical modelling of directional wind speeds using mixture of von Mises distributions: Case study. *Energy Conversion and Management*, 49: 897-907.
7. Azmani, M., Reboul, S., Choquel, J. B. and Benjelloun, M. (2009). A recursive, change point estimate of the wind speed and direction. *IEEE 7th International Conference on Computational Cybernetics*, Palma de Mallorca, Spain.
8. Heckenbergerova, J., Musilek, P. and Kromer, P. (2015). Optimization of wind direction distribution parameters using particle swarm optimization. *Advances in Intelligent System and Computing*, 334: 15-26.
9. Masseran, N., Razali, A. M., Ibrahim, K., Zaharim, A. and Sopian, K. (2013). The probability distribution model of wind speed over east Malaysia. *Research*

Journal of Applied Sciences, Engineering and Technology, 6: 1774-1779.

10. Masseran, N., Razali, A. M., Ibrahim, K., Zaharim, A. and Sopian, K. (2013). Application of the single imputation method to estimate missing wind speed data in Malaysia. *Research Journal of Applied Sciences, Engineering and Technology*, 6: 1780-1784.
11. Banerjee, A., Dhillon, I. S., Ghosh, J. and Sra, S. (2005). Clustering on the unit Hypershere using von Mises-Fisher Distributions. *Journal of Machine Learning Research*, 6: 1345-1382.
12. Hornik, K. and Grun, B. (2014). movMF: An R package for fitting mixtures of von Mises-Fisher Distributions. Journal of Statistical Software 58: 1- 31.
13. Mooney, J. A, Helms, P. J. and Jolliffe, I. T. (2003). Fitting mixture of von Mises distributions: a case study involving sudden infant death syndrome. *Computational Statistics & Data Analysis*, 41: 505-513.
14. Fejer, L. (1915). Uber trigonometrische polynome. *Journal fur die Riene und Angewandte Mathematik*, 146: 53-82.
15. Fernandez-Duran, J. J. (2004). Circular distributions based on nonnegative trigonometric sums. Biometrics, 60: 499-503.
16. Fernandez-Duran, J. J. and Gregorio-Dominguez, M. M. (2010). Maximum likelihood estimation of nonnegative trigonometric sum models using a Newton-like algorithm on manifolds. *Electronic Journal of Statistics*, 4: 1402-1410.
17. Absil, P. A. Mahony, R. and Sepulchre, R. (2008). *Optimization Algorithm on Matrix Manifolds*. Princeton University Press, Princeton.
18. Stephenson, G. (1973). *Mathematical methods for science students.* Pearson Education Limited, Edinburgh.

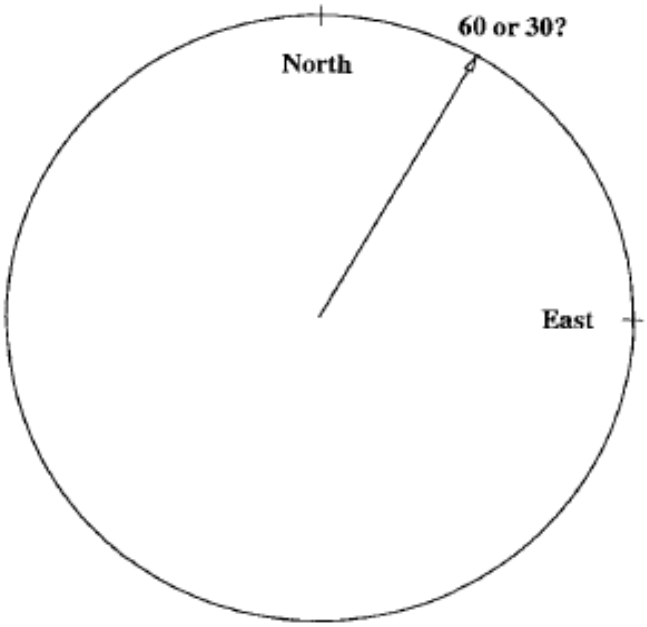

Figure 1: The observed directional data depend on choice of origin and the sense of rotation [1].

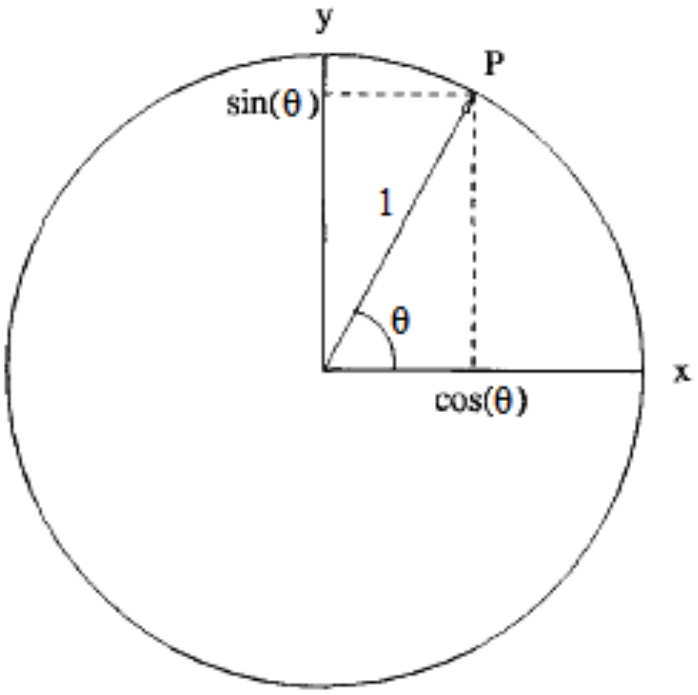

Figure 2: Relationship between polar coordinate data (1, θ) and rectangular coordinate data (*X*, *Y*) [1].

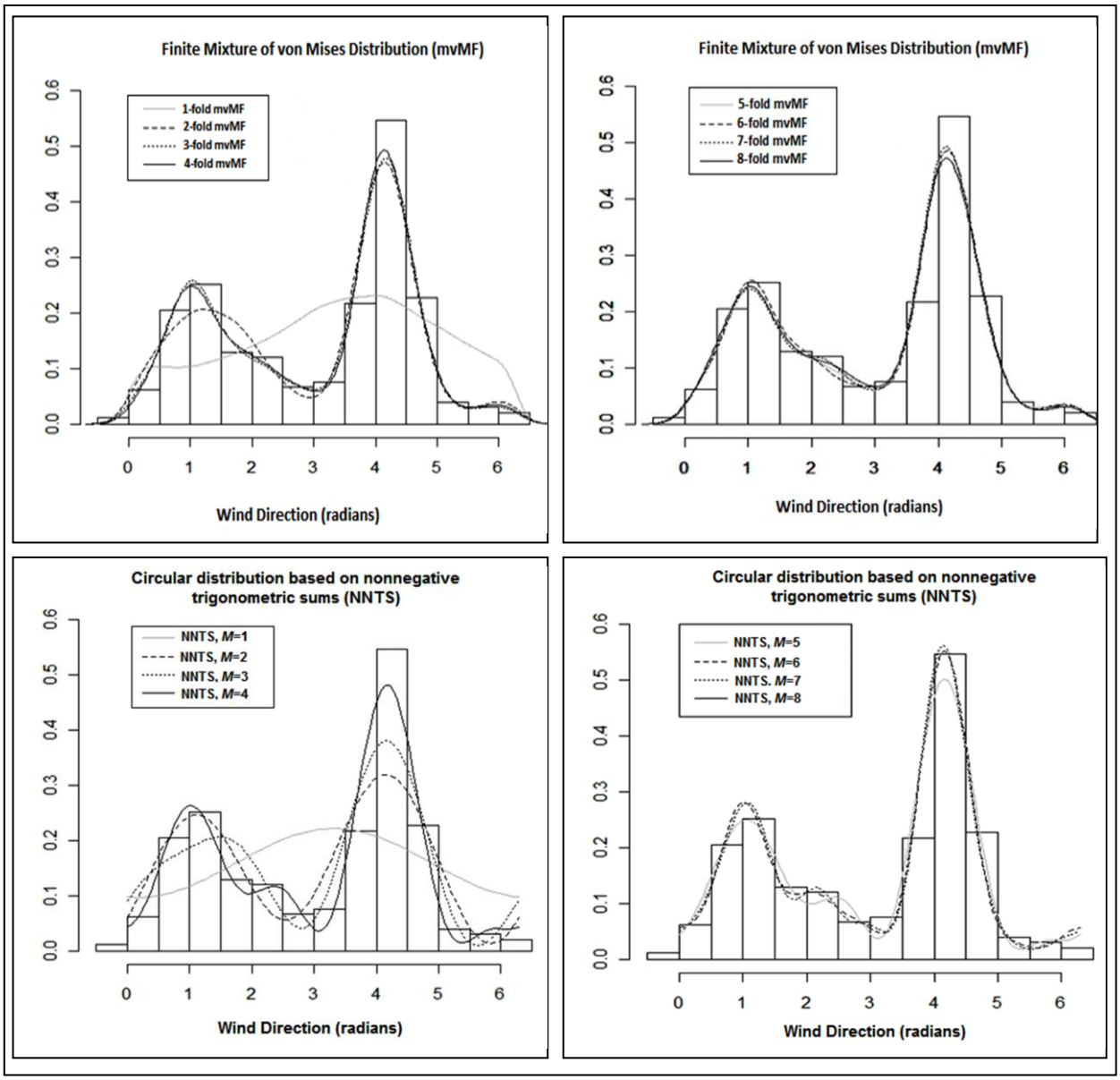

Figure 3: The fitted mvMF and NNTS model respectively for wind direction in Kudat.

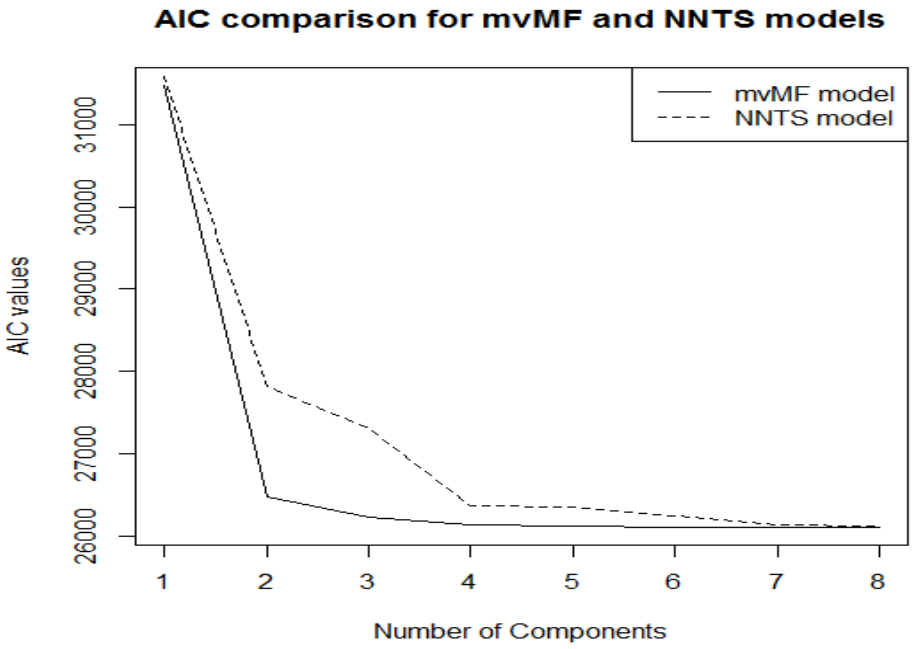

Figure 4: AIC comparison for both of the models mvMF and NNTS.

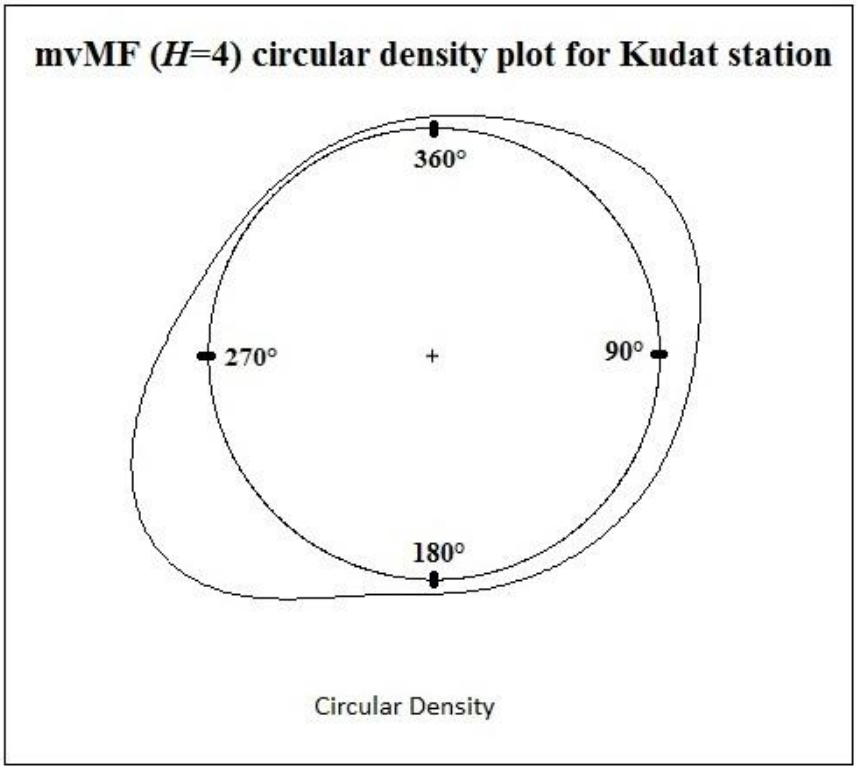

Figure 5: Circular density plot for the mvMF (H=4) model at the Kudat station.

Table 1: Descriptive statistics for Kudat wind direction

Kudat wind station	
Mean direction	218.21°
Circular variance	0.978
25^{th} percentile	80°
50^{th} percentile	220°
75^{th} percentile	240°

Table 2: The parameter estimates for the mvMF model up to six numbers of components

mvMF	Parameter Estimates			
	$\boldsymbol{\mu}^T$		κ	ω
N=1	-0.7854526	-0.6189218	0.42187	1
N=2	-0.5153732	-0.8569659	7.242308	0.4787608
	0.2909280	0.9567449	1.628116	0.5212392
N=3	0.02426039	0.9997057	8.096920	0.4945873
	0.56316151	0.8263469	1.002905	0.3759604
	-0.50926732	-0.8606084	12.27827	0.1294524
N=4	0.56772170	0.8232205	11.658204	0.3474064
	-0.31381192	-0.9494852	9.2272398	0.1439757
	-0.63765746	-0.7703200	10.752359	0.1412261
	0.02245815	0.9997478	0.970675	0.3673919
N=5	0.5274866	0.8495633	10.086927	0.1753316
	0.9273593	0.3741721	2.207358	0.1015669
	-0.3347130	-0.9423201	7.297946	0.2486590
	-0.6281035	-0.7781298	12.444567	0.2712363
	-0.5343394	0.8452700	2.124835	0.2032062
N=6	-0.7545886	0.6561982	0.9392541	0.1013886
	0.5624254	0.8268481	15.2835080	0.1124414
	-0.2662294	-0.9639097	12.0485284	0.1329380
	-0.4584617	-0.8887142	7.8910636	0.1663857
	0.2389779	0.9710250	1.2310384	0.2985423
	-0.6692827	-0.7430078	20.1925421	0.1883040
N=7	0.9742367	0.2255278	3.069639	0.08087669
	-0.5527640	0.8333378	2.834249	0.17782239
	-0.3763575	-0.9264745	8.225171	0.15866292
	-0.3802596	-0.9248798	5.966826	0.08184980
	-0.6323128	-0.7747132	15.641954	0.19080226
	0.5208237	0.8536643	15.641954	0.20578241
	-0.5310920	-0.847314	15.6419	0.10420354
N=8	-0.6038800	-0.79707526	3.419402	0.12369715
	0.9959767	0.08961304	1.205793	0.03612474
	-0.6425706	-0.76622647	18.202829	0.23905701
	-0.2556487	-0.96676975	12.510171	0.17256913
	0.5312111	0.84723950	12.639344	0.13472241
	0.8595642	0.51102773	2.555491	0.06450745
	0.3521644	0.93593815	2.438121	0.09955324
	-0.6015076	0.79886709	3.233072	0.12976885

Table 3: The parameter estimates for the NNTS model up to six numbers of components

NNTS	Parameter Estimates				
	$\hat{\theta}_k = \hat{\theta}_{rk} + i\hat{\theta}_{ck}$			$\hat{\theta}_k = \hat{\theta}_{rk} + i\hat{\theta}_{ck}$	
M=1	k=0;	0.39088770 + i(0.000)	***M=6***	k=0;	0.33051696 + i(0.000)
	k=1;	-0.07851115 + i(0.01406219)		k=1;	-0.06448836 – i(0.0147715)
				k=2;	-0.08979539 – i(0.14949561)
M=2	k=0;	0.35145494 + i(0.000)		k=3;	0.06867154 + i(0.00497386)
	k=1;	-0.07058781 + i(0.00018935)		k=4;	-0.05872581 + i(0.07351352)
	k=2;	-0.08954087- i(0.15044643)		k=5;	-0.00175609 – i(0.01752988)
				k=6;	0.03374423 + i(0.00904722)
M=3	k=0;	0.34473199 + i(0.000)			
	k=1;	-0.06519571 - i(0.01050682)	***M=7***	k=0;	0.32966310 + i(0.000)
	k=2;	-0.08975569 – i(0.15251008)		k=1;	-0.06524948 - i(0.01568993)
	k=3;	0.06761674 + i(0.00815430)		k=2;	-0.08992444 – i(0.15000096)
				k=3;	0.06880256 + i (0.00497416)
M=4	k=0;	0.33206489 + i(0.000)		k=4;	-0.05858657 + i(0.07387359)
	k=1;	-0.06721701 - i(0.01422902)		k=5;	-0.00161553 – i (0.01720893)
	k=2;	-0.08899561 - i(0.15149014)		k=6;	0.0339152 0+ i(0.00906594)
	k=3;	0.06842034 + i(0.00487464)		k=7;	-0.01385732 - i (0.00392186)
	k=4;	-0.05831522 + i(0.07205559)			
			M=8	k=0;	0.32927979 + i(0.000)
M=5	k=0;	0.33191670 + i(0.000)		k=1;	-0.06543936 – i(0.01564708)
	k=1;	-0.06720818- i(0.01503347)		k=2;	-0.09033585 – i(0.14995585)
	k=2;	-0.08942703 – i (0.15065709)		k=3;	0.06905995 + i(0.00473962)
	k=3;	0.06766157 + i(0.00499270)		k=4;	-0.05825125 + i(0.07392021)
	k=4;	-0.05874579 + i(0.07219662)		k=5;	-0.00165490 – i(0.01733769)
	k=5;	-0.00101357 – i(0.01676556)		k=6;	0.03366950 + i(0.00892418)
				k=7;	-0.01402637 – i(0.00393321)
				k=8;	0.00447990 – i(0.01254067)

Table 4: AIC values for comparison between mvMF model and NNTS model

Model	AIC	Model	AIC
mvMF (*H*=1)	31472.46	**NNTS (*M*=1)**	31576.27
mvMF (*H*=2)	26468.94	**NNTS (*M*=2)**	27811.47
mvMF (*H*=3)	26231.88	**NNTS (*M*=3)**	27311.61
mvMF (*H*=4)	26141.58	**NNTS (*M*=4)**	26368.29
mvMF (*H*=5)	26123.57	**NNTS (*M*=5)**	26356.16
mvMF (*H*=6)	26113.17	**NNTS (*M*=6)**	26239.38
mvMF (*H*=7)	26111.30	**NNTS (*M*=7)**	26135.64
mvMF (*H*=8)	26110.60	**NNTS (*M*=8)**	26120.32