

ANÁLISE DESCRITIVA DE DADOS DE VENTOS VISANDO APLICAÇÕES EM PLATAFORMAS

G. M. G. Leal¹, F. A. M. Souza²

¹ UFCG/CCT/DME/ Av. Aprígio Veloso, 882 Bodocongó, 58109-970
Campina Grande PB, Tel: 310 1030 e-mail: graycimary@terra.com.br

² UFCG/CCT/DME– e-mail: fsouza@dme.ufcg.edu.br

Resumo – Na indústria de exploração de Petróleo e Gás Natural, alguns fenômenos geram informações que podem ser caracterizadas como dados direcionais, entre eles podemos citar as direções de ventos. Mardia e Jupp (2000) revisam que estas direções são responsáveis por fornecer uma fonte natural de dados direcionais. Para análise de dados dessa natureza, necessita-se do emprego de uma metodologia estatística específica e diferente daquela utilizada para variáveis lineares. O objetivo deste trabalho é estudar as técnicas de análises de dados de direcionais com vista à aplicação na indústria de Petróleo e Gás Natural.

Palavras-Chave: Dados direcionais; Direção de ventos; Distribuição von Mises.

Abstract – In the industry of exploration of Oil and Natural Gas, some phenomena generate information that can be characterized as directional data, among them we have directions of winds. Mardia and Jupp (2000) revise that these directions are responsible for supplying a natural source of directional data. For analysis this type of data, specific and different statistics of that one used for linear variables is needed, that is, we need another methodology. The objective of this work is to study the techniques of analysis of directional data of with sight to the application in the industry of Oil and Natural Gas.

Keywords: Directional data; Direction of winds; von Mises distribution

1. Introdução

Na indústria de exploração de Petróleo e Gás Natural, ocorrem fenômenos que são abordados como dados direcionais, tais como: direcionamentos de sedimentos, direções de ventos, direções de ondas e correntes marítimas, mudanças na direção de uma broca durante a perfuração de um poço, direções nas fraturas de uma rocha, entre outros. A instalação de uma plataforma de petróleo não é uma atividade simples, uma vez que elas estão submetidas a constantes ações de ventos. Por isso, é de fundamental importância o conhecimento de informações sobre direções de ventos, na construção e implantação de plataformas tanto terrestres quanto marítimas, uma vez que os ventos provocam impactos de turbulências nas estruturas.

Para análise de dados direcionais, necessita-se de uma metodologia estatística diferente daquela utilizada para variáveis lineares, uma vez que medidas angulares são obtidas a partir de uma origem arbitrária, enquanto observações lineares têm uma referência bem definida.

O objetivo deste trabalho é estudar algumas técnicas para análises de dados direcionais. Especificamente, fazer uma análise descritiva de direções de ventos, além de estudar algumas distribuições de probabilidades e fazer inferências relativas a médias direcionais.

2. Metodologia

Numa primeira etapa, foram obtidos dados sobre ventos, relativos a um período de 10 anos, de diversos postos de estações meteorológicas dos estados do Rio Grande do Norte, Ceará, Paraíba, Bahia, Sergipe e Alagoas, contendo informações de direções e velocidades dos ventos. As medições foram diárias e em três horários distintos (12:00h, 18:00h e 24:00h). Em seguida, fizemos a análise descritiva desses dados relativos a algumas localidades que fossem próximas de unidades de produção e exploração de petróleo da Petrobrás bem como localizadas na faixa litorânea. Em seguida aplicamos um teste estatístico para testar igualdade entre médias direcionais. Para este trabalho escolhemos as seguintes cidades: Macau-RN, Mossoró-RN, Aracati-CE, Fortaleza-CE, Aracajú-SE, Maceió-AL e Coruripe-AL.

3. Fundamentação Teórica

3.1. Introdução a Dados Direcionais

Existe interesse por medidas na forma de ângulos em diversas áreas do conhecimento. De modo particular, encontramos exemplos de dados direcionais em Geologia, Biologia, Meteorologia, Oceanografia, Medicina e em Engenharia.

Em Geologia, uma atenção predominante é dedicada às propriedades escalares dos objetos, que evidentemente fornecem importantes informações sobre as condições de formações de rochas. Existe, porém, um outro conjunto de propriedades, denominadas vetoriais, cujo estudo em diversas circunstâncias é de fundamental importância para o entendimento de fenômenos geológicos. Um grande número de atributos geológicos podem ser representados por linhas ou planos, e a sua medida resulta em dados angulares.

Medidas angulares são obtidas a partir de uma origem arbitrária, e assim os métodos estatísticos baseados em escalas numéricas lineares nem sempre podem ser aplicados às medidas angulares. Nesse caso, os dados devem ser tratados com técnicas diferentes. As medidas resumo que são usadas para variáveis lineares, não fazem sentido quando se trata de uma variável direcional, conforme mostraremos mais adiante.

De uma forma geral, quando se deseja estudar dados que representam *direções*, eles podem ser tratados em um espaço de qualquer dimensão, entretanto nas situações práticas geralmente o interesse está nos espaços de dimensão 2 ou 3 (bi- ou tri- dimensional), onde as direções podem ser representadas geometricamente por pontos sobre o círculo unitário, ou sobre a superfície de uma esfera, respectivamente.

3.2. Dados Direcionais

No caso bi-dimensional, uma observação de uma variável direcional pode ser vista como um ponto P sobre o círculo unitário, centrado na origem O do sistema de coordenadas cartesianas e, portanto, tem-se que \overrightarrow{OP} é um vetor unitário em \mathcal{R}^2 . Essa observação direcional pode ainda ser vista como um ângulo y , formado pelo vetor \overrightarrow{OP} e o semi-eixo O_x , medido no sentido anti-horário. Assim, uma observação de uma variável circular pode ser representada pelo par ordenado $(\cos(y), \sin(y))$, em coordenadas cartesianas, ou pelo par $(1, y)$, em coordenadas polares.

Embora os dados circulares sejam coletados em diferentes tipos de unidades, a conversão destes dados para graus é feita por uma simples transformação linear, obtendo-se valores no intervalo $[0^\circ, 360^\circ)$ ou $[-180^\circ, 180^\circ)$ dependendo do interesse na análise. Em princípio, podemos trabalhar com dados circulares expressos em qualquer medida de ângulo, tais como: graus, grado, radiano, etc. Mas, por razões teóricas, adota-se trabalhar em radianos, com intervalo de variação $[0, 2\pi)$ ou $[-\pi, \pi)$, dependendo do interesse da análise.

Em alguns casos, como por exemplo, no estudo de fraturas em rochas, não há como distinguir qual é o início ou o fim dessa fratura. Um dado desse tipo é denominado de *axial* diferente dos dados de direções de ventos que denominamos de *vetorial* (Fisher, 1993).

Para trabalharmos com distribuições de probabilidades referentes a variáveis aleatórias circulares precisaremos estabelecer algumas medidas descritivas relacionadas com os conceitos de posição, dispersão, etc.

3.3. Medidas Descritivas no Círculo

Consideremos uma amostra aleatória de n observações circulares y_1, \dots, y_n , às quais podemos associar n vetores unitários $\overrightarrow{OP_1}, \dots, \overrightarrow{OP_n}$, conforme foi mencionado anteriormente. Uma maneira natural e intuitiva de resumir esses vetores unitários é através da sua soma. Do ponto de vista algébrico, calculemos as seguintes quantidades:

$$S = \sum_{i=1}^n \text{sen}(y_i), \quad C = \sum_{i=1}^n \text{cos}(y_i), \quad R^2 = C^2 + S^2 \quad (1)$$

A média circular, ou média direcional, de y_1, \dots, y_n é definida como sendo o ângulo $\bar{\mu}$ correspondente ao vetor resultante da soma $\overrightarrow{OP_1} + \dots + \overrightarrow{OP_n}$. Esse ângulo $\bar{\mu}$ deve satisfazer as condições:

$$\begin{cases} \text{cos}(\bar{\mu}) = C/R \\ \text{sen}(\bar{\mu}) = S/R \\ \bar{\mu} = \text{arctg}(S/R) \end{cases}$$

A quantidade $R = \sqrt{C^2 + S^2}$ representa o comprimento do vetor resultante e fornece informação sobre o grau de concentração dos ângulos observados. Isto porque, se todas as n observações circulares forem iguais a um mesmo ângulo, então $R=n$. Esse caso representa uma concentração máxima possível, e corresponde a uma variabilidade mínima. Outra situação extrema é quando os dados encontram-se uniformemente distribuídos sobre o círculo. Nesse caso tem-se uma concentração mínima e uma variabilidade máxima. No caso de concentração mínima, que significa $R=0$, a média circular não está definida. É importante salientar também que $R=0$ não significa necessariamente uma distribuição uniforme em torno do círculo, pois pode acontecer algum tipo de estrutura de agrupamentos, fornecendo $R=0$.

Em vez de R , é mais comum usar o comprimento médio do vetor resultante, definido por $\bar{R} = R/n$, o qual tem a vantagem de variar no intervalo $[0,1]$. A variância circular amostral é definida por $V=1 - \bar{R}^2$, enquanto que o desvio padrão circular é definido por $\sigma = \{-2 \log(1-V)\}^{1/2}$, diferentemente da definição usual para dados lineares, onde o desvio padrão corresponde à raiz quadrada da variância, Fisher (1993).

3.4. Modelos Probabilísticos para Dados Direcionais

Uma distribuição de probabilidade de direções de vento pode surgir como uma distribuição marginal obtida a partir da distribuição conjunta da velocidade e direção do vento ou como uma distribuição condicional a uma determinada velocidade do vento, Mardia e Jupp (2000, p.8).

Entretanto, muitos modelos probabilísticos para dados circulares têm sido derivados a partir de transformações de modelos probabilísticos lineares usuais, ou mesmo de vetores aleatórios bi-dimensionais dando origem às distribuições arqueadas.

De uma maneira geral, dada uma variável aleatória linear Z , podemos construir uma variável circular Y , assumindo valores no intervalo $[0, 2\pi)$ através da transformação $Y=Z \text{mod}(2\pi)$. São as chamadas distribuições arqueadas.

Se Z for absolutamente contínua, com função densidade de probabilidade $f_Z(z)$, então a correspondente função densidade de probabilidade de Y será dada pela Equação 2:

$$f_Y(Y) = \sum_{k=-\infty}^{\infty} f_Z(y + 2k\pi) \quad \forall y \in [0, 2\pi) \quad (2)$$

3.4.1. A Distribuição Uniforme Circular u_c

Dizemos que uma variável aleatória Y tem distribuição uniforme sobre um círculo unitário se sua função densidade de probabilidade for dada pela Equação 3:

$$f_y(Y) = \frac{1}{2\pi} \quad 0 \leq y < 2\pi \quad (3)$$

A função característica da distribuição uniforme é dada pela Equação 4:

$$\phi_p = \frac{e^{2ip\pi} - 1}{2ip\pi} \quad p = \pm 1, \pm 2, \dots \quad (4)$$

Portanto,

$$\phi_p = \begin{cases} 0, & \text{se } p \neq 0 \\ 1, & \text{se } p = 0 \end{cases}$$

A média angular, não está definida para este modelo e $V=1$. Isto significa que a concentração é nula em torno de qualquer direção.

3.4.3. A Distribuição von Mises – $v\mathcal{M}(\mu, \lambda)$

A distribuição von Mises é tão importante para dados circulares, quanto o modelo Gaussiano é para dados lineares em geral. Esse modelo foi proposto pelo físico alemão R. von Mises, em 1918, ao investigar desvio de pesos atômicos de valores inteiros, Mardia e Jupp (2000, p.36). Seu propósito era mostrar que os pesos atômicos são números inteiros sujeitos a erros. Transformando a parte fracionária dos pesos atômicos em desvios angulares, então, ele passou do problema físico original para um problema estatístico, a partir do qual originou essa distribuição.

Dizemos que uma variável aleatória circular Y tem distribuição von Mises com parâmetros μ e λ se sua função densidade de probabilidade for dada pela Equação 6 :

$$f_y(y; \mu, \lambda) = \frac{1}{2\pi I_0(\lambda)} \exp[\lambda \cos(y - \mu)] \quad 0 \leq y < 2\pi, \quad 0 \leq \mu < 2\pi \text{ e } \lambda > 0$$

(6)

onde $I_0(\lambda)$ é a função de Bessel modificada do 1º tipo e ordem zero, que é dada pela série

$$I_0(\lambda) = \sum_{r=0}^{\infty} \frac{1}{r!^2} \left(\frac{1}{2} \lambda \right)^{2r}$$

Algumas características da distribuição von Mises envolvem a função $A_p(\lambda) = \frac{I_p(\lambda)}{I_0(\lambda)}$. O parâmetro μ

representa a média circular de Y , enquanto λ é denominado parâmetro de concentração, e é tal que o comprimento resultante médio $\rho = A_1(\lambda)$. A dispersão circular de Y é dada por $\delta = \frac{1}{\lambda A_1(\lambda)}$

3.5. Teste Estatístico

Para testar a igualdade entre médias direcionais foram utilizados dois tipos de teste, conhecidos com o método P e método M.

3.5.1. Teste de Igualdade entre Médias Direcionais

É um teste de comparação entre médias direcionais que tem como critério inicial o tamanho da amostra. Em amostras grandes ($n > 25$), a análise é feita admitindo-se que as populações possuem dispersões comparáveis, conduzindo à adoção de um dos procedimentos P e M descritos a seguir. Se a maior das dispersões não for mais que quatro vezes a menor das dispersões, deve-se adotar o método P. Caso contrário deve-se adotar o método M.

3.5.1.1. Método P

Sendo $\hat{\delta}_1, \dots, \hat{\delta}_r$ as estimativas das dispersões das r populações, se $\hat{\delta}_{\max} / \hat{\delta}_{\min} \leq 4$ calcule,

$$\hat{C}_p = \sum_{i=1}^r n_i \cos \hat{\mu}_i \quad \hat{S}_p = \sum_{i=1}^r n_i \sin \hat{\mu}_i \quad \text{e} \quad \hat{R}_p = \sqrt{\hat{C}_p^2 + \hat{S}_p^2}$$

E uma estimativa média $\hat{\delta}_0 = \sum_{i=1}^n \frac{n_i \hat{\delta}_i}{N}$. Assim, a estatística de teste é:

$$Y_r = \frac{2(N - R_p)}{\hat{\delta}_I} \quad (7)$$

A hipótese de igualdade entre as médias direcionais é rejeitada se Y_r for demasiadamente grande.

3.5.1.2. Método M

Se $\hat{\delta}_{\max} / \hat{\delta}_{\min} > 4$ calcule:

$$\hat{C}_M = \sum_{i=1}^r (\cos \hat{\mu}_i) / \hat{\sigma}_i^2 \quad \hat{S}_M = \sum_{i=1}^r (\sin \hat{\mu}_i) / \hat{\sigma}_i^2 \quad \text{logo,} \quad R_M = \sqrt{\hat{C}_M + \hat{S}_M}$$

A estatística de teste é:

$$Y_r = 2 \left(\sum_{i=1}^r 1 / \hat{\sigma}_i^2 - R_M \right) \quad (8)$$

Novamente, A hipótese de igualdade entre médias direcionais é rejeitada se Y_r for demasiadamente grande.

Em ambos os métodos o valor calculado da estatística é comparado com o percentil da distribuição qui-quadrado com $(r-1)$ graus de liberdade.

4. Resultados e Conclusões

Através do Departamento de Ciências Atmosféricas da Universidade Federal de Campina Grande obtivemos dados, contendo informações de direções e velocidades dos ventos como resultado de medições diárias em três horários distintos (12h, 18h e 24h) num período de dez anos. Esses dados foram coletados em postos de estações meteorológicas dos estados do Rio Grande do Norte, Ceará, Bahia, Paraíba, Alagoas e Sergipe.

Após a obtenção dos dados prosseguimos nosso estudo aplicando as técnicas para análise descritiva dos mesmos. Para este trabalho, utilizamos os dados relativos às cidades de Aracati-CE, Fortaleza-CE, Macau-RN e Mossoró-RN em paralelo com Aracajú-SE, Coruripe-AL e Maceió-AL.

Como resultado da análise descritiva dos dados de direção de ventos nas cidades acima mencionadas, obtivemos as medidas apresentadas nas Tabelas 1 e 2.

Tabela 1. Medidas Descritivas da Direção dos Ventos

VARIÁVEL	MEDIDAS	ARACATI	FORTALEZA	MACAU	MOSSORÓ
DIREÇÃO	Nº de Observações	9025	10955	9944	284
	Média Direcional	120.18°	63.07°	107.39°	75.97°
	Concentração	0.66	0.58	0.74	0.63
	I.C. (95%)	(119°; 121°)	(62°; 64°)	(107°; 108°)	(69°; 82°)

Tabela 2. Medidas Descritivas Direção dos Ventos

VARIÁVEL	MEDIDAS	ARACAJÚ	MACEIÓ	CORURIFE
DIREÇÃO	Nº de Observações	10544	10213	10491
	Média Direcional	104.84°	103.46°	56.61°
	Concentração	0.72	0.66	0.35
	I.C. (95%)	(103°-105.71°)	(102°-104.48°)	(54.50°-58.73°)

Na análise descritiva da velocidade dos ventos, que é uma variável linear, obtivemos as medidas representadas nas tabelas 3 e 4.

Tabela 3. Medidas Descritivas da Velocidade dos Ventos

VARIÁVEL	MEDIDAS	ARACATI	FORTALEZA	MACAU	MOSSORÓ
VELOCIDADE	Média	4.4	3.0	5.5	1.7
	Desvio Padrão	2.3	3.0	2.1	2.8
	I.C. (95%)	(1.4; 4.5)	(2.9; 3.0)	(5.5; 5.6)	(1.4; 2.0)

Tabela 4. Medidas Descritivas da Velocidade dos Ventos

VARIÁVEL	MEDIDAS	ARACAJÚ	MACEIÓ	CORURIFE
VELOCIDADE	Média	2.6	4.1	1.0
	Desvio Padrão	1.8	2.2	1.0
	I.C. (95%)	(2.6; 2.7)	(4.1; 4.2)	(0.9; 1.0)

Nas Figuras 1 e 2 temos representações gráficas dos resultados referentes às velocidades. Por sua vez, na Figura 3 temos histogramas representando as direções dos ventos.

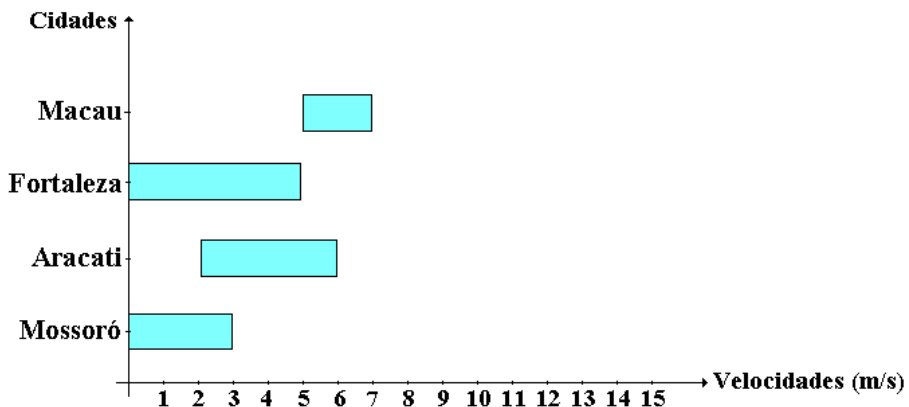


Figura 1 Variação entre os quartis das velocidades dos ventos.

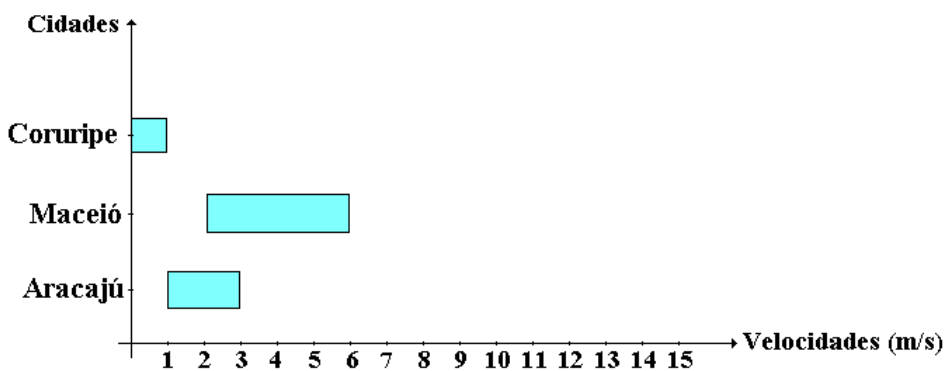


Figura 2 Variação entre os quartis das velocidades dos ventos.

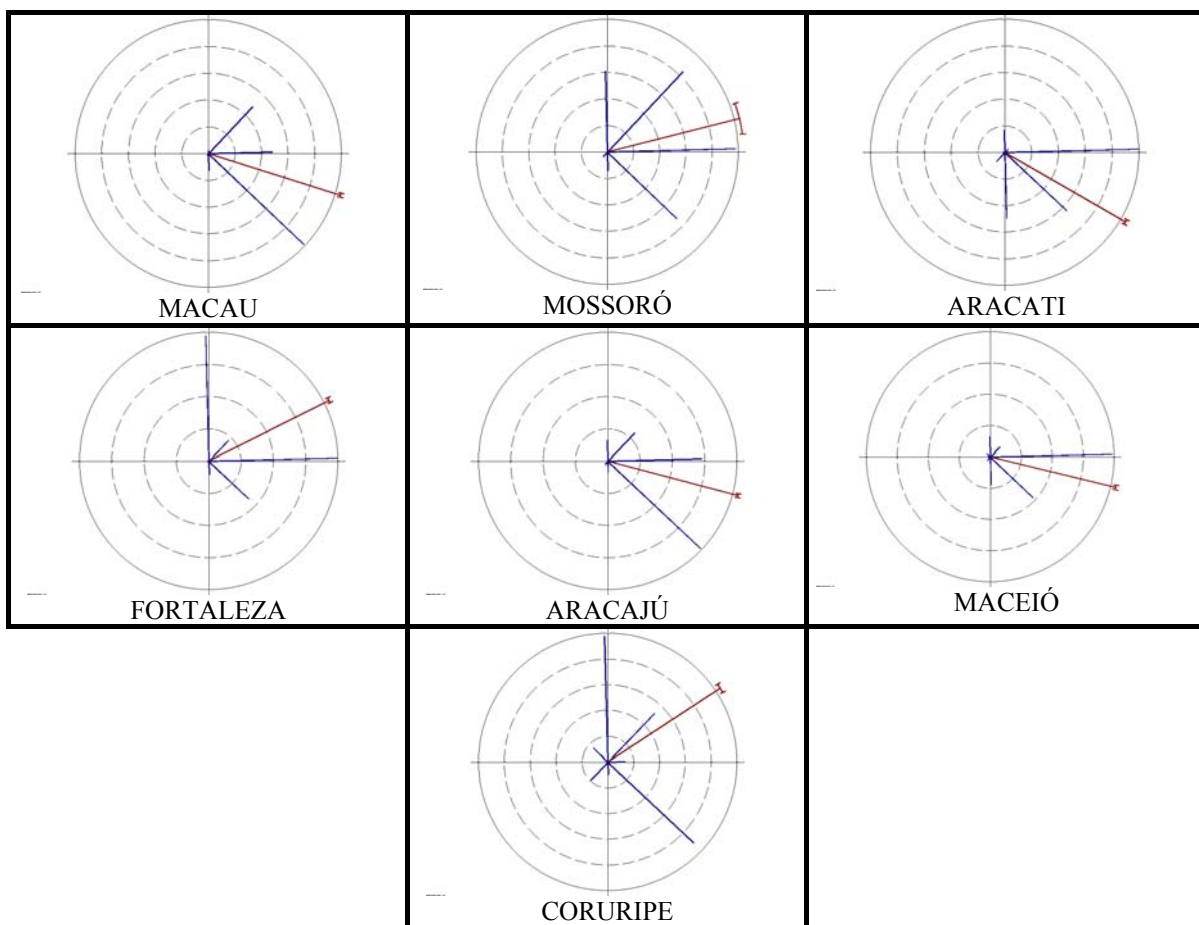


Figura 3 Histogramas das direções dos ventos por localidade.

A partir das Figuras 1 e 2 verificamos que os dados de Coruripe são os mais concentrados e de menor velocidade média, enquanto Fortaleza apresentou maior variabilidade nos dados. .

Nos histogramas da Figura 3, as linhas em azul representam ocorrências de direções enquanto que a linha em vermelho representa a direção média dos ventos. Analisando os histogramas observa-se que as médias direcionais das cidades Aracati, Fortaleza, Macau e Mossoró, mesmo sendo próximas umas das outras, as médias direcionais possuem direções diferentes. Esta diferença mostrou-se ser significativa ao nível de 5% de significância, através do método P. Com relação às cidades Maceió, Coruripe e Aracajú o mesmo método P não detectou diferença significativa entre as respectivas médias direcionais.

Segundo a classificação de Beaufort, em todas as cidades, os ventos podem ser classificados entre fracos e moderados. Desta forma, podemos concluir que os ventos nas cidades que foram consideradas no presente trabalho, não são danosos às estruturas de plataformas instaladas nestas localidades.

5. Agradecimentos

Agradecemos à Agência Nacional do Petróleo – ANP através do Programa de Recursos Humanos – PRH (25) pelo apoio financeiro e incentivo a esta pesquisa.

6. Referências

FISHER, N. I. *Statistical Analysis of Circular Data*. Cambridge.: University Press.1993

MARDIA, K. V. *Statistics of Directional Data*. London: Academic Press..1972.

MARDIA K.V. e JUPP, P.E. *Directional Statistics*. London: John Wiley & Sons LTD. 2000