

2 REVISÃO BIBLIOGRÁFICA

O estudo climatológico de uma determinada região sem levar em consideração um único ponto de observação (estação meteorológica), representa conhecer características de determinada área envolvendo dados de outros pontos. Esses pontos, que representam regiões, podem ser obtidos através de técnicas estatísticas apropriadas. A mais utilizada, em diversas áreas do conhecimento científico, é a análise multivariada, através dela podemos obter grupos homogêneos das variáveis em questão e investigar o seu comportamento espacial e temporal.

Algumas técnicas são utilizadas dentro da análise multivariada como, por exemplo, a (ACP) Análise de Componentes Principais, a (AA) Análise de Agrupamento e a (AF) Análise Fatorial.

2.1 Análise de Agrupamento

A análise de agrupamento, a qual é utilizada em processos de classificação, consiste em determinar o nível de similaridade ou dissimilaridade entre indivíduos aplicando uma função de agrupamento a uma determinada variável. Agrupamentos também podem ser obtidos através da análise de componentes

principais, mais especificamente na retenção dos escores dos componentes que explicam um percentual aceitável da variância dos dados.

De acordo com Everitt (1974) existem várias técnicas de Análise de Agrupamento como as de partição-otimização, de densidade, técnicas de agrupamentos com interseção e outras.

A análise de agrupamento inicia-se pela determinação de uma função de agrupamento, onde esta é usada como critério para medir a distância entre dois elementos ou para estabelecer o quanto eles são parecidos. Alguns autores a chama de “coeficiente de parecença”, que se divide em duas categorias: medida de similaridade e dissimilaridade.

Sabe-se, ainda, que existe na literatura vários métodos de agrupamento associados às várias funções de agrupamento, gerando diferentes formas de grupos. Os relacionados por Wilks (1995) são: Método da Ligação Simples, Método da Ligação Completa, Método da Centróide e Método de Ward.

2.2 Determinação de regiões homogêneas de variáveis meteorológicas usando técnicas de agrupamento

Antes de tratar dos agrupamentos de estações meteorológicas usando variáveis meteorológicas, deve-se salientar a utilização dessa prática em outras áreas de conhecimento, na biologia, por exemplo, onde são efetuadas diversas medidas em um conjunto de unidades, não previamente particionadas, a análise de agrupamento é usada para classificar animais e plantas, chamada de taxonomia numérica. O uso da técnica em biologia é visto em Curi (1985), em um exemplo onde são estudados os aminoácidos livres do tecido cerebral de ratos usando a análise de agrupamento, complementada com a ordenação das unidades por componentes principais e com análise de variância multivariada, que serviu para indicar o nível de similaridade mais adequado para a definição dos agrupamentos. Em medicina é usada para identificar as doenças e seus estágios assim como, também, para avaliar as condições de saúde pública (Possoli, 1984).

Em estudos de poluição atmosférica pesquisadores como, por exemplo, Dorling et al. (1992) combinaram a análise de agrupamento de dados isobáricos (1000mb) das linhas de trajetória do ar para estabelecer modelos sinóticos e relacioná-los com a concentração de poluentes (aerossóis e concentrações iônicas) na precipitação pluvial em Eskadalemuir, sul da Escócia. Yu & Chang (2001) analisaram séries temporais de dados de poluentes do ar, O₃ e PM₁₀, para determinar a divisão das bacias hidrográficas (em termos de qualidade do ar) em Taiwan; empregaram métodos estatísticos multivariados, especificamente o método rotacional varimax e análise de agrupamento. O período estudado foi de julho de 1993 a junho de 1998 e as técnicas de agrupamento usadas geraram cinco grupos homogêneos.

Em meteorologia sinótica, Fernau & Samson (1990 a, b) aplicaram a análise de agrupamento aos padrões de transportes atmosféricos a fim de definir uma climatologia sinótica representativa do movimento das massas de ar. Cheng & Wallace (1993) usaram o método hierárquico de agrupamento de Ward para gerar grupos homogêneos da altura geopotencial de 500 hPa. Os dados usados no estudo foram do inverno começando em janeiro de 1945 até dezembro de 1984 usando pontos de grade desde 20° N até o Polo.

O conhecimento dos padrões climáticos de determinadas regiões, tem despertado o interesse de pesquisadores (Richman & Lamb, 1985) e levou Wolter (1987) ao uso de agrupamento de variáveis meteorológicas, como pressão ao nível médio do mar, vento superficial, cobertura de nuvens e temperatura da superfície do mar, sobre os oceanos Atlântico tropical, Pacífico leste e Índico, visando o entendimento da circulação geral da atmosfera e da climatologia.

Com o objetivo de determinar regiões homogêneas através de técnicas estatísticas como: ACP e AA, utilizando como variável de agrupamento dados de precipitação e como método de agrupamento o de Ward, muitos trabalhos já foram realizados. Gadgil & Iyengar (1980) utilizaram a ACP para obter modelos de variação temporal da chuva sobre 53 estações na península Indiana.

Baptista da Silva et al., (2001), através da AA, estudaram o comportamento dos totais trimestrais das chuvas no estado do Ceará, a partir de 20 estações meteorológicas distribuídas por toda extensão deste Estado, tendo como medida de similaridade a distância euclidiana ao quadrado e a inércia intra-classe como critério de agregação. Braga et al. (1998) determinaram regiões homogêneas desta variável no Estado da Bahia para estudar o seu ciclo médio anual. Os dados utilizados foram séries diárias de 140 estações meteorológicas cada uma com, no mínimo, 30 anos de dados. A mesma autora em 1992, usando as duas técnicas e o mesmo método estabeleceu 5 regiões homogêneas de precipitação pluvial. Usou dados decendiais de 50 localidades no Estado do Rio Grande do Norte para um período de 30 anos.

Fazendo uso também da AA Freitas & Grimm (1998) aplicaram o método em questão aos dados de precipitação pluvial diárias para o período de 1967 a 1992, para várias estações pluviométricas localizadas no Paraná além de algumas estações meteorológicas dos Estados de São Paulo e Santa Catarina, de forma a determinar as datas de início e final da estação chuvosa no Paraná, em cada região homogênea. Souza et al. (1992) fizeram a homogeneização dessa variável para o Estado de Alagoas, através da AA, cujo objetivo foi avaliar o comportamento das precipitações pluviais médias e máximas utilizando dados de 60 estações com um período mínimo de 14 e um máximo de 72 anos.

Silva et al. (1996) determinaram grupos homogêneos de precipitação no setor leste da Região Nordeste. Para este trabalho utilizaram dados totais mensais de 127 postos meteorológicos localizados nos tabuleiros costeiros e baixadas litorâneas de municípios dos Estados do Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia que apresentaram registros contínuos para um período superior a 20 anos. Os grupos foram obtidos através do mesmo método, citado anteriormente, sendo selecionados quatro com características pluviometricamente homogêneas.

Visando estudar a variabilidade da precipitação pluvial no Estado de Santa Catarina e correlacioná-la com as anomalias da temperatura da superfície do mar,

Baldo et al. (2000) recorreram à esta técnica e método para a determinação das regiões homogêneas dessa variável.

Utilizando a mesma técnica, Lopes et al. (1996) através do método da ligação completa e do método da ligação simples obtiveram cinco grupos pluviometricamente homogêneos no Estado do Rio Grande do Norte.

Rodrigues et al. (1996) usaram o método da classificação hierárquica na identificação de regiões pluviometricamente homogêneas no Estado do Ceará. Foram usados totais diários de precipitação pluvial coletados em 84 postos pluviométricos pertencentes à rede de estações meteorológicas da SUDENE e do Estado do Ceará, referente ao período de 1913 a 1983. Os métodos utilizados foram: ligação simples, ligação completa e centróide; 6 grupos foram obtidos e o método que gerou os grupos mais representativos foi o da centróide.

As regiões homogêneas, obtidas da análise de agrupamento, são usadas para vários fins como, por exemplo, classificação climatológica, conforme feita por Becker (1992). A autora usou regiões homogêneas de precipitação pluvial e de temperatura decendiais e mensais para estabelecer padrões climatológicos do Rio Grande do Sul. No mesmo trabalho, esses grupos foram também usados para simular séries temporais e preencher valores faltantes no conjunto de dados.

Objetivando comparar o método de classificação climática de Köppen (1931) e um método multivariado, Herrera et al. (1996) utilizaram a análise de agrupamento usando médias, de dez variáveis meteorológicas que caracterizam o clima, de 16 estações do Estado de São Paulo para um período de 13 anos de observações. Foram obtidos grupos tanto pelo método de agrupamento quanto pela técnica de componentes principais. Segundo os autores, os grupos obtidos tanto por um método quanto por outro, foram concordantes com a classificação climática para o Estado de São Paulo.

Vários pesquisadores têm feito uso de uma variedade de dados meteorológicos para definir tipos climáticos e delinear zonas de similaridade climática. Fovell & Fovell (1993) fizeram uma regionalização de zonas climáticas nos Estados Unidos aplicando técnicas hierárquicas de agrupamento, mais especificamente o método da ligação média, a um conjunto de dados de

temperatura média e total acumulado de precipitação pluvial para um período de 50 anos (1931 a 1980) em 34 regiões climáticas. Foram obtidos 14 grupos não só pelo método de agrupamento, mas também através da análise de componentes principais.

Considerando as técnicas aglomerativa e divisiva da análise de agrupamento, Ramos (2001) utilizou as duas formas de classificação para estudar as variações das precipitações pluviais no Mediterrâneo, as quais são responsáveis pelo clima da região. Da técnica divisiva, foi usado o método "K-média" e da hierárquica aglomerativa, o método de Ward. Os métodos foram aplicados às precipitações pluviais diárias do período de 1889 a 1999 do Alt Penedès, região Nordeste da Espanha. Foram analisados, também, valores de precipitações pluviais anuais das estações da primavera e outono. Segundo conclusão do autor, a utilização de mais de um método de agrupamento, permite extrair um número maior de informações. O uso da técnica hierárquica permite visualizar a combinação das observações para formar os grupos, o que não é possível na técnica divisiva. Dentro da técnica hierárquica, o método de Ward permite interpretar melhor os resultados.

Kim et al. (2000), através da mesma técnica, determinaram nove regiões homogêneas no Estado do Paraná usando dados de temperaturas máximas e mínimas diárias de vinte e nove estações meteorológicas.

Com relação a dados de temperatura, Diniz (2002) utilizou a técnica de AA e como função de agrupamento a distância euclidiana para analisar as temperaturas máximas e mínimas no Estado do Rio Grande do Sul, com dados obtidos em 40 estações meteorológicas no período de 1913 a 1998. Para este trabalho foram testados 4 métodos hierárquicos de agrupamento, tendo sido escolhido o da ligação completa por apresentar uma estrutura climatologicamente coerente do Estado. Foram obtidas 4 regiões homogêneas, todas com um número satisfatório de estações meteorológicas.