

## **Análise estatística preliminar de dados pluviométricos mensais, sazonais e anuais para o Estado do Paraná**

Yara Rúbia de Mello <sup>(a)</sup>, Tony Vinicius Moreira Sampaio <sup>(b)</sup>

<sup>(a)</sup> Programa de Pós-Graduação em Geografia, Universidade Federal do Paraná, yarademello@gmail.com

<sup>(b)</sup> Programa de Pós-Graduação em Geografia, Universidade Federal do Paraná, tony2sampaio@gmail.com

### **Eixo: 3. Climatologia em diferentes níveis escalares: mudanças e variabilidades**

#### **Resumo**

Existem milhares de pluviômetros instalados no Brasil, os quais possibilitam o desenvolvimento de uma gama de estudos. Porém, ao se utilizar os dados são necessários alguns cuidados para averiguar a confiabilidade dos mesmos. Neste sentido, o objetivo deste trabalho foi analisar estatisticamente uma série histórica de 30 anos de dados pluviométricos mensais, sazonais e anuais para o Estado do Paraná. Para isso, foram utilizados dados disponíveis no site da ANA, selecionado o período de 30 anos com o maior número de estações sem falhas; e analisados os dados quanto a sua consistência, presença de *outliers*, homogeneidade e tendência. O período selecionado foi de 1977 a 2006, a quantidade de estações adequadas por categoria de análise variou de 404 a 467. A média de precipitação anual foi de 1.684,5 mm, a estação mais chuvosa foi o verão com 533,6 mm, e o mês mais chuvoso foi janeiro com 192,4 mm.

**Palavras chave:** Precipitação; pluviômetros; homogeneidade.

#### **1. Introdução**

Existem dezenas de instituições que monitoram o tempo e/ou o clima no Brasil. Muitas delas disponibilizam os dados gratuitamente aos usuários, como é o caso da Agência Nacional de Águas – ANA. Esses dados são frequentemente utilizados em estudos técnicos e científicos para os mais diversos fins. Porém, muitos desses estudos não descrevem, ou mesmo não analisam os dados estatisticamente, o que é fundamental, uma vez que os mesmos estão sujeitos a diversos fatores que causam inconsistência.

Este trabalho analisa uma série temporal de dados sobre a precipitação no estado do Paraná, aplicando testes específicos para detecção de *outliers* (Grubbs), homogeneidade (Pettitt e Standard Normal Homogeneity Teste - SNHT) e tendência dos dados (Mann-Kendall), testes estes amplamente utilizados na literatura (GONZÁLES-ROUCO *et al.*, 2001; LONGOBARDI *et al.*, 2009; KANG *et al.*, 2012; SENA *et al.*, 2013; MIRZAEI *et al.*, 2014; VIVEKANANDAN, 2015).

*Outliers* são observações de uma amostra que se desviam acentuadamente dos outros elementos, os quais podem causar inconsistências. Estes dados poderão afetar a homogeneidade da série. Os dados serão homogêneos se as variações forem causadas pelo tempo e clima, uma vez que a homogeneidade é



representada por uma série de dados climáticos que tenham variações causadas por características físicas. Atividades como a mudança de local da estação, desenvolvimento urbano no entorno e erros de medição, podem interferir na homogeneidade dos dados, e como consequência na qualidade e validade dos mesmos. A série de dados pode ser considerada tendenciosa quando, ao longo do tempo, apresentar um comportamento de crescimento ou decrescimento (TUCCI, 2001; CARAM, 2007; SILVA, 2012).

Este trabalho é a fase inicial de um estudo no qual se pretende realizar um mapeamento geostatístico da precipitação para o estado do Paraná. Tendo em vista que uma análise estatística preliminar de dados climatológicos é fundamental, seja qual for sua utilidade, o presente estudo objetivou analisar uma série pluviométrica de dados mensais, sazonais e anuais para o Estado, compreendendo um período de 30 anos.

## 2. Materiais e Métodos

### 2.1 Área de estudo

A área de estudo compreende o estado Paraná - PR, o qual está localizado na Região Sul do Brasil, entre as latitudes 22°30'S e 26°43'S e longitudes 48°05'W e 54°37'W (Figura 1).

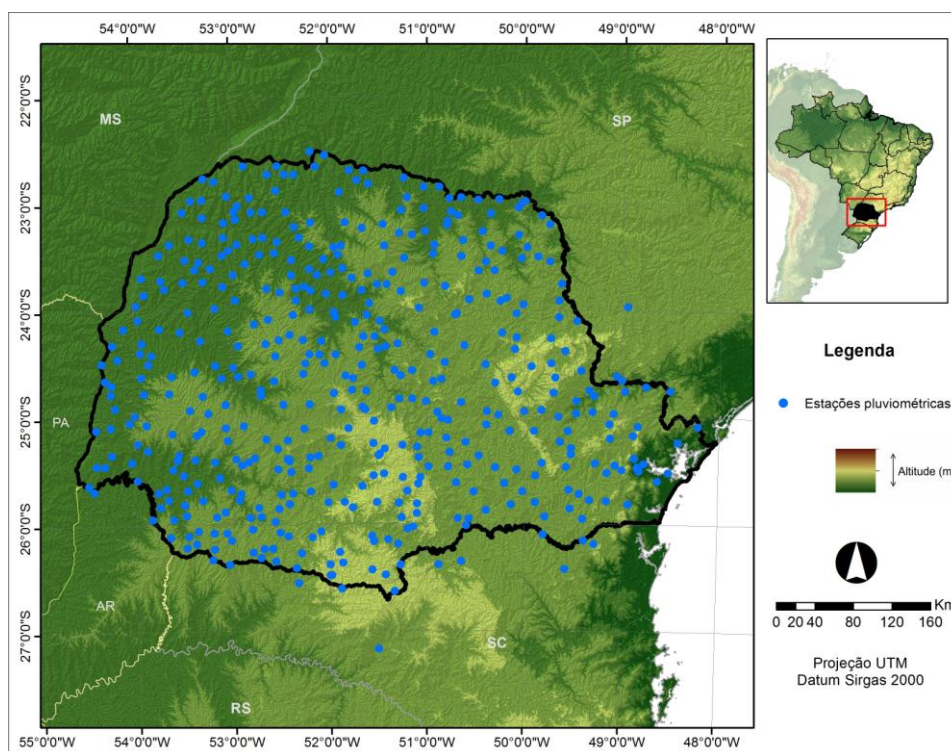


Figura 1 – Localização do Estado do Paraná e da totalidade das estações pluviométricas utilizadas na pesquisa, referente ao período mensal, sazonal e anual. Fonte: ANA; SRTM90.

O Clima do Paraná, segundo a classificação de Köppen, pode ser dividido em Cfa (Clima subtropical com tendência de concentração das chuvas nos meses de verão, sem estação seca definida) na porção norte, oeste, grande parte do sudoeste e parte da região sul do Estado; e Cfb (Clima temperado sem estação seca definida) na região sul do PR (FRITZSONS, 2011).

## 2.2 Dados pluviométricos

Foram utilizados dados pluviométricos disponíveis no site da Agência Nacional de Águas, através do link: [hidroweb.ana.gov.br](http://hidroweb.ana.gov.br). Os intervalos de tempo avaliados foram: mensal, sazonal e anual.

Os dados diários foram organizados em totais mensais, referente aos anos analisados, através do *software* Hidro 1.2. Foram utilizados também dados de estações dos estados vizinhos (Santa Catarina e São Paulo), para reduzir o efeito de borda na geração de modelos de precipitação para o Paraná.

## 2.3 Análises estatísticas

As estações com mais de 10% de falhas foram eliminadas. Por exemplo, se uma estação (na análise anual) possuísse mais de três anos com falhas, seria eliminada; se uma estação (na análise mensal e sazonal) com mais de três meses janeiro, ou três meses fevereiro, etc., possuísse falhas, seria eliminada. Esta análise com o limiar de 10% foi baseada no estudo de Teixeira (2010).

Foi realizada análise da consistência dos dados pluviométricos, com a validação dos valores aceitáveis. Para isso, os valores de cada período (mês ou ano) foram avaliados em relação à média, mínimo e máximo da estação, e do período (mês ou ano) analisado, em relação a todas as estações. Os valores mínimos e máximos aceitáveis para os dados mensais variaram de acordo com o período do ano, sendo divididos de janeiro a maio, e setembro a dezembro como uma categoria; e junho a agosto como outra. Isso baseado nas estatísticas descritivas destes dados (média, mínimo, máximo, frequência).

Avaliaram-se os *outliers* (se necessário excluídos) pelo teste de Grubbs, a homogeneidade pelo teste de Pettitt e/ou SNHT, e a tendência dos dados pelo teste de Mann-Kendall. Adotou-se o nível de significância de 5%.

A sazonalidade foi mensurada através dos meses de janeiro, fevereiro e dezembro (verão); março, abril e maio (outono); junho, julho e agosto (inverno); e setembro, outubro e novembro (primavera). Essa divisão foi baseada na similaridade da distribuição das chuvas ao longo do ano. A mesma categorização foi utilizada por Fritzsons *et al.* (2011), em estudo sobre a definição de zonas homogêneas de precipitação no estado do Paraná; e por Jorge (2009), em estudo climático da fachada Atlântica Sul do Brasil.

Nesta primeira etapa foram utilizados os *softwares ArcGIS, Microsoft Excel, ProGrid, Statistica e XLStat*.

### 3. Resultados e Discussão

Foram localizadas no site da ANA um total de 1.444 estações com dados para o estado do Paraná. O período de 30 anos consecutivos com o maior número de estações com dados foi de 1977 a 2006, o que reduziu o número de estações para um total de 476 estações. Foram averiguadas as estações nos municípios próximos dos estados vizinhos, com a série de dados relativa ao mesmo período, totalizando 492 estações.

Foi realizada a eliminação das estações com mais de 10% de falhas, restando 419 estações para a análise anual; e 471 estações para a análise mensal e sazonal. Do total analisado, 169 estações não apresentaram nenhuma falha para o período no Paraná. Esse número é considerável, porém, quando plotadas as coordenadas das estações em ambiente SIG, verificou-se que algumas áreas apresentaram baixa densidade de estações. Sendo assim, optou-se por utilizar algumas das estações que continham menos de 10% de falhas.

Em seguida foram analisados os dados anuais, quanto à presença de *outliers*, homogeneidade e tendência. Das 419 estações, 16,7% (70 estações) apresentaram valores extremos, destas optou-se por eliminar apenas 7 estações, pois o restante coincidiu com anos chuvosos em outras estações. Quanto a homogeneidade dos dados, 2,1% (9 estações) mostraram haver alteração na série histórica, destas, optou-se por eliminar 8. Com a eliminação destas estações, restou um total de 404 para a análise anual (Figura 1).

A análise de tendência foi realizada com essas 404 estações, sendo que 2,7% (11 estações) apresentaram tendência, sendo 3 negativas e 8 positivas. Estas estações estão destacadas na Figura 2, de forma a verificar se existe um padrão espacial dessas tendências. As estações em vermelho apresentam tendência negativa, as estações em azul tendência positiva, nesta figura também estão destacadas a tendência dos dados sazonais, a qual será discutida posteriormente no texto.

Para os dados mensais as análises exigiram muito mais tempo, por conta da quantidade de dados. Foram analisadas 471 estações, e para cada uma delas os doze meses do ano, gerando um total de 5.652 análises para cada categoria de teste (*outlier* e homogeneidade). Primeiramente foi aplicado o teste para detecção de *outliers*, o qual detectou em 26,6% da amostra a presença de dados extremos. O número foi bastante alto, desta forma foram realizadas conferências em 10% da amostra (dados com *outliers*), sendo um total



de 150 análises. Destas, apenas 2,7% (4 eventos) não coincidiram com períodos chuvosos, e ambos os 4 eventos da amostra se mostraram homogêneos, segundo o teste de Pettitt.

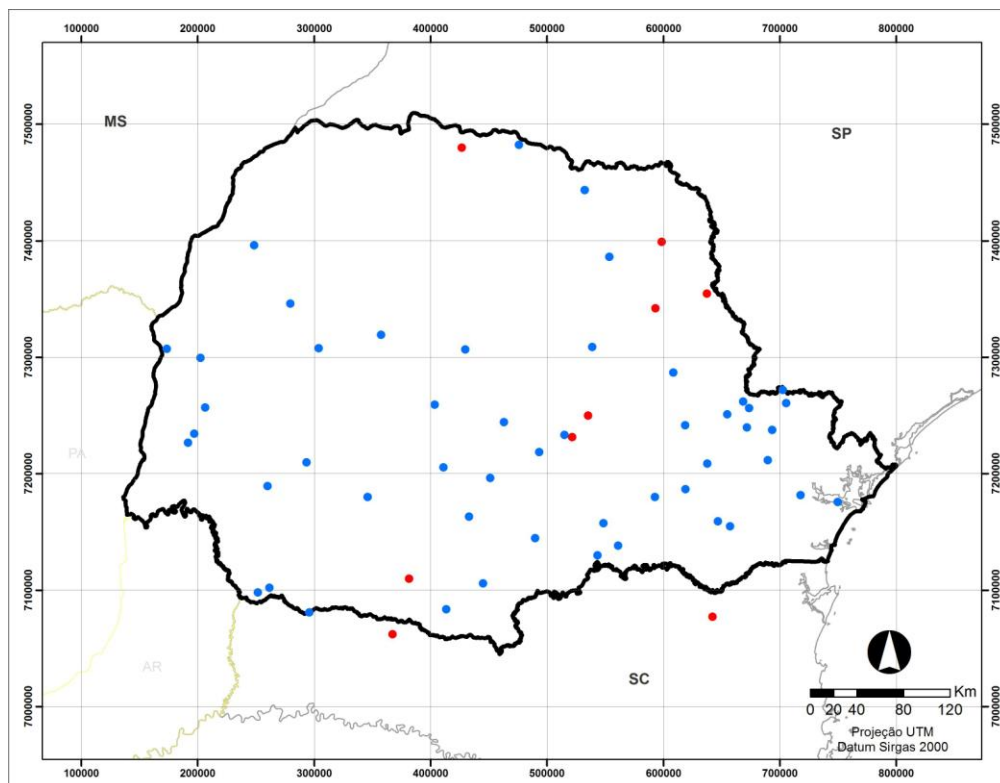


Figura 2 – Tendência estatística das estações referente aos períodos de análise anual e sazonal (pontos vermelhos = estações com tendência negativa; pontos azuis = estações com tendência positiva).

Com isso, percebeu-se que o teste de Grubbs aponta como *outliers* valores que são aceitáveis, em sua maioria, por se tratar de períodos chuvosos. Desta forma, determinou-se que para a análise dos dados mensais, seria utilizado o teste de homogeneidade para conferência da qualidade dos dados.

O teste de homogeneidade de Pettitt, por sua vez, apontou apenas 4,8% da amostra como sendo não homogênea. Optou-se nesta segunda etapa da análise dos dados mensais por não aplicar outro teste de homogeneidade (SNHT) para verificar se os valores-r eram significativos, e sim fazer uma comparação gráfica entre os dados da estação analisada referente ao período “x” analisado, e a média de todas as estações. Com isso, decidiu-se por eliminar apenas 20% da amostra dos dados não homogêneos.

O que se observou, de modo geral, é que o teste de homogeneidade acusou principalmente os meses onde a precipitação é mais irregular (verão e primavera) como não homogêneos, 46% dos dados referiram-se ao verão, 42,9% a primavera, 9,5% ao outono e apenas 1,5% ao inverno. Por outro lado, o teste para detecção



de *outlier* encontrou valores-*r* não significativos mais similares em frequência entre as estações do ano, com destaque para o outono (29,4%) e inverno (33,5%).

Após essa conferência, foram eliminados os meses das estações não homogêneas, e o restante dos meses para a mesma estação, referente à estação do ano em que o mês não homogêneo estava inserido. Por exemplo, se a estação “x” em janeiro não apresentasse homogeneidade, os dados dos meses de fevereiro e dezembro (verão) também seriam eliminados. Sendo assim, o total final de estações para os meses de verão foi de 454, para o outono de 457, para o inverno de 467, e para a primavera de 455 (Figura 1).

Quanto à tendência dos dados, o verão apresentou 4,2% das estações com tendências (19 estações/positivas), o outono 0,9% (4 estações/negativas), o inverno 0,2% (uma estação/negativa) e a primavera 6,2% (27 estações/positivas; uma estação/negativa) (Figura 2). A primavera e o verão foram respectivamente os períodos com o maior número de estações com tendência positiva, apesar de o número ser baixo. Justamente nesse período que as precipitações intensas são mais frequentes. Esta análise corrobora a afirmativa de Cavalcanti *et al.* (2009), a qual expõe que estudos indicam tendências de aumento na intensidade e na frequência de eventos de chuvas intensas em partes da região Sul e Sudeste do Brasil.

De modo geral, as tendências de aumento de precipitação se espacializaram por todo o território do Paraná, exceto em parte da porção centro-norte e noroeste (Figura 2). E as tendências de decréscimo de precipitação espacializaram-se principalmente na porção oriental do território. Para uma análise mais detalhada faz-se necessário a elaboração de mapas mensais e sazonais separadamente e, o relacionamento dos dados com os sistemas atmosféricos atuantes na região.

Para obter uma melhor observação espacial da dinâmica da distribuição de precipitação entre as estações pluviométricas, foram desenvolvidos cartogramas contendo os totais de precipitação por categoria (anual e sazonal) dividida em quantis (Figuras 3 e 4); conforme a seguinte classificação: super-seco (<5%); seco (5% - 25%); normal (25%-75%); úmido (75%-95%); e super-úmido (>95%).

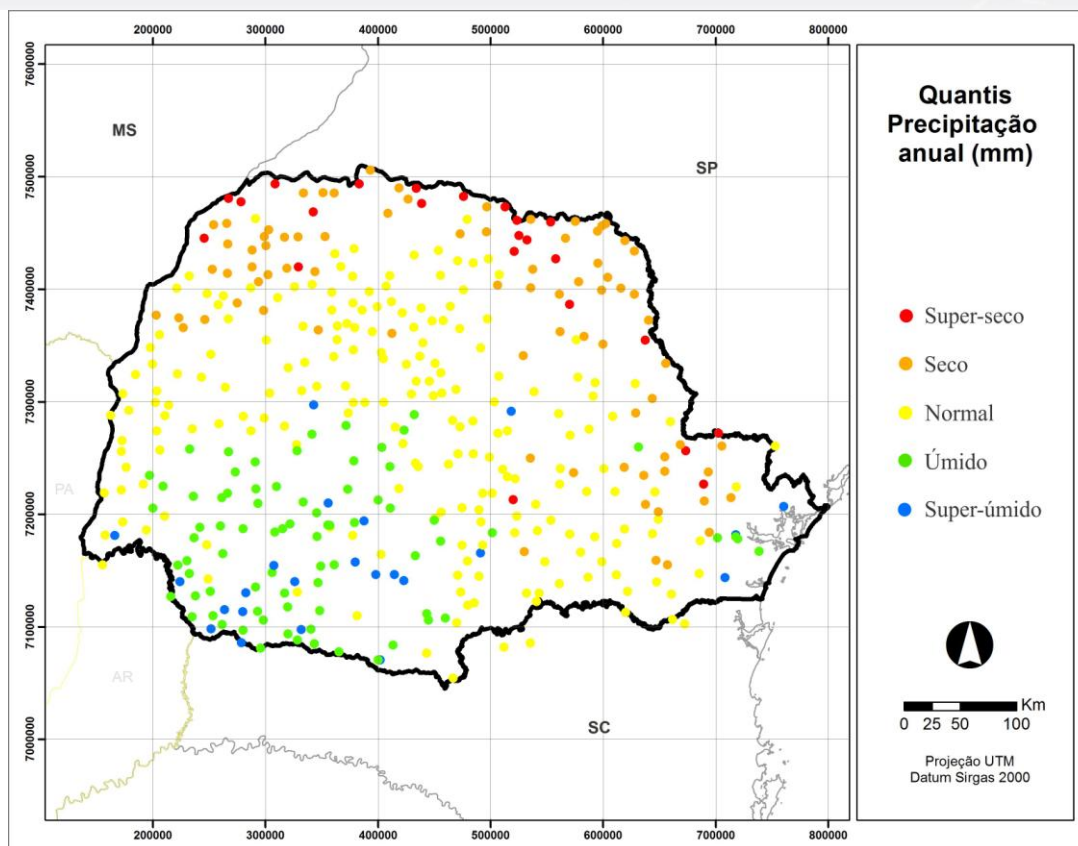


Figura 3 – Classificação da precipitação anual das 404 estações a partir da técnica dos quantis.

Conforme Figura 3, é possível observar a existência de um gradiente de precipitação no sentido sul-norte, com os maiores índices ao sul do Estado e os menores ao norte, com uma transição bem definida. A região leste, nas proximidades da Serra do Mar também apresenta valores altos de precipitação.

De modo geral, a espacialização dos quantis de precipitação em relação as estações do ano seguiu o padrão anual, o verão configurou-se mais distinto (Figura 4). A primavera e o outono apresentaram um padrão mais similar, com uma aparente região de sombra de chuva configurada na porção leste. Já no inverno ficou nítida a diminuição nos índices pluviométricos para as estações próximas a Serra do Mar, em relação as outras estações pluviométricas.

A média anual de precipitação para o conjunto das estações analisadas foi de 1.684,5 mm, a estação pluviométrica com o maior índice é a Bananal, localizada em Guaraqueçaba com 2.576,3 mm. Em relação ao período de 1977 a 2006, os anos mais chuvosos em média foram 1983 (2.026,2 mm), 1997 (1.811,3 mm) e 1998 (1.755,3 mm), já os anos mais secos foram respectivamente 1985 (1.074,8 mm), 1978 (1.120,4 mm) e 1988 (1.172,5 mm) (Figura 5).

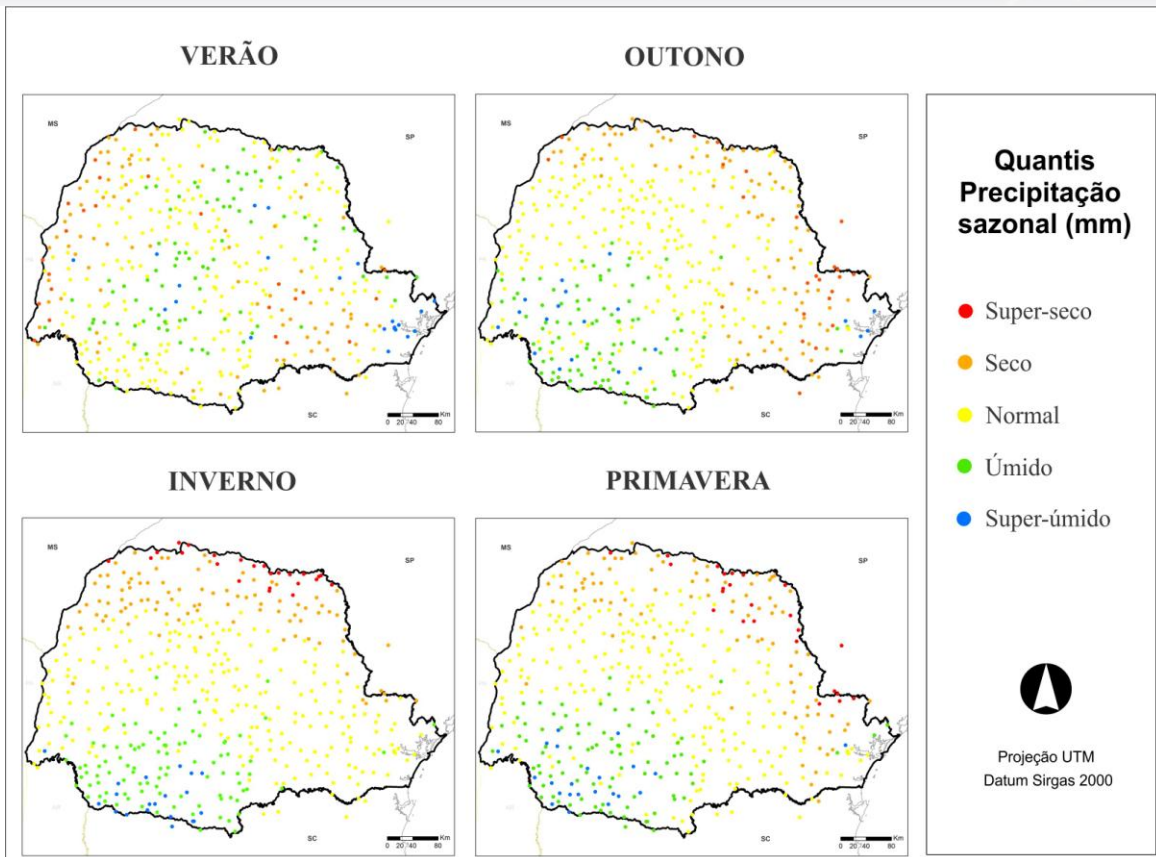


Figura 4 – Classificação da precipitação sazonal das estações correspondentes a cada período a partir da técnica dos quantis.

Em relação à sazonalidade, a estação do ano com o maior acumulado médio de precipitação foi o verão (533,6 mm), seguido pela primavera (472,5 mm), outono (399,8) e inverno (287,1 mm). A Figura 6 mostra os acumulados médios por estação do ano para o período analisado (1977-2006).



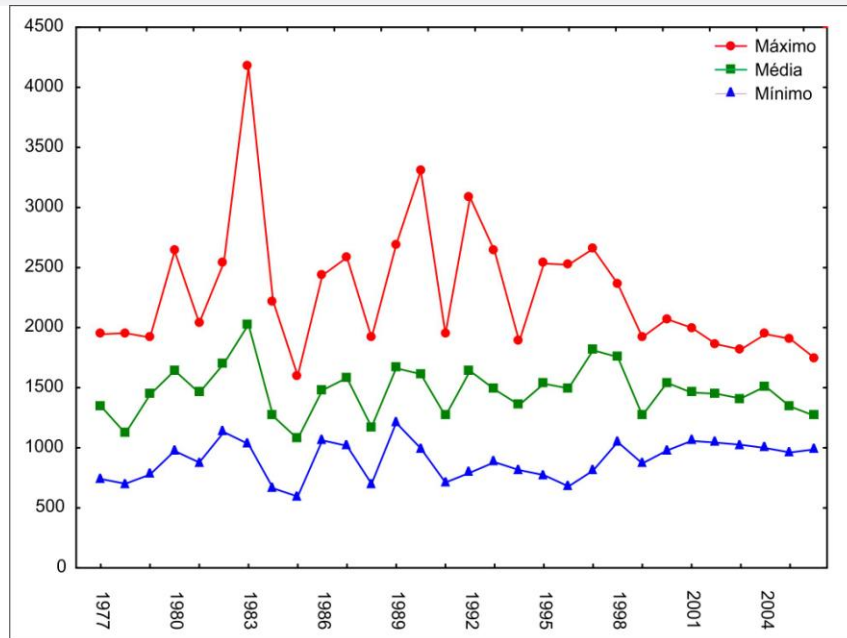


Figura 5 – Gráfico com a precipitação anual máxima, média e mínima das estações pluviométricas analisadas para o período de 1977 a 2006.

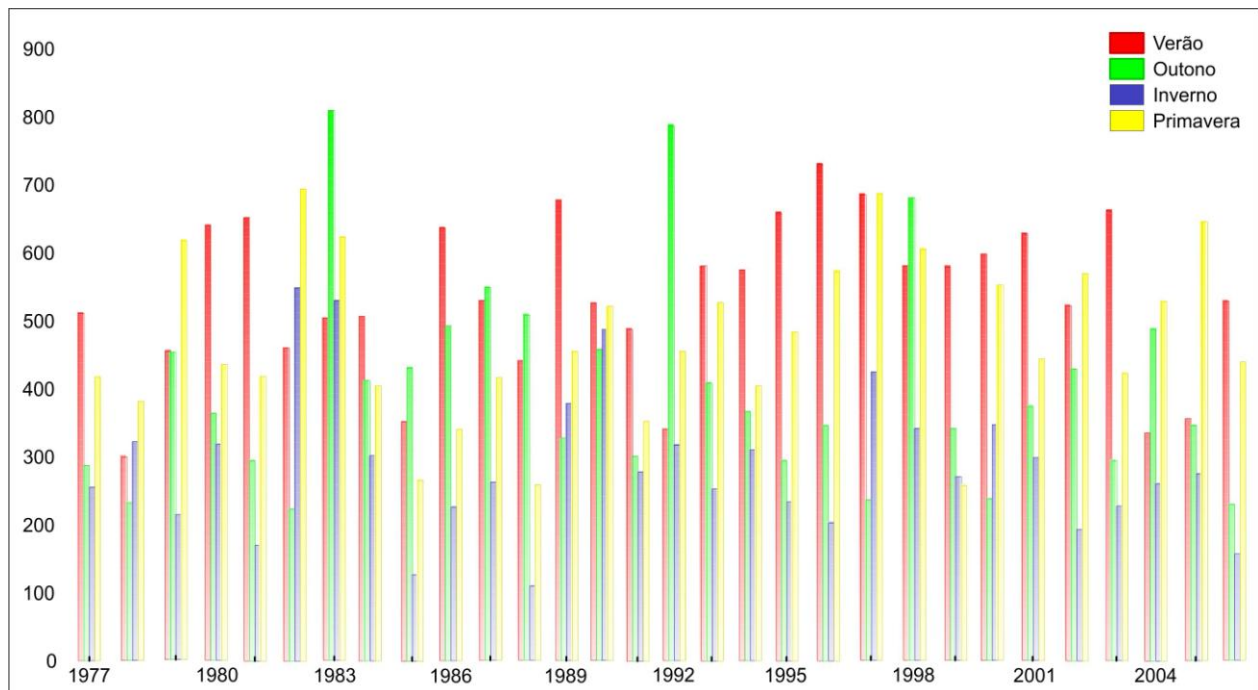


Figura 6 – Gráfico com os acumulados médios sazonais das estações pluviométricas analisadas para o período de 1977 a 2006.



Quanto a distribuição mensal da precipitação, observou-se que o mês mais chuvoso em média foi o mês de janeiro (192,4 mm), e o mês menos chuvoso foi o mês de agosto (76,8 mm), conforme Figura 7. Na mesma figura não foram destacados os valores mínimos absolutos, pois para todos os meses foi igual a zero. Por máximo absoluto entende-se o valor máximo encontrado entre todas as estações e todos os anos para cada mês separadamente.

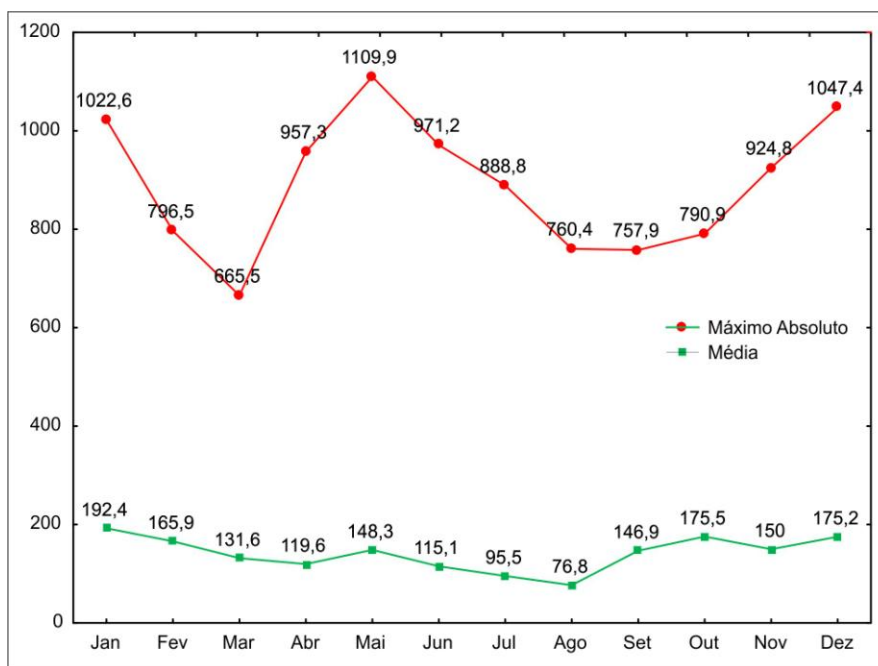


Figura 7 – Gráfico com a precipitação mensal máxima absoluta e média das estações pluviométricas analisadas para o período de 1977 a 2006.

Os gráficos e mapas gerados são úteis para estudos mais detalhados, por exemplo, sabe-se a partir dos resultados encontrados que o ano de 1983 foi um ano chuvoso, com um valor máximo registrado acima dos 4.000 mm, e que a chuva se concentrou principalmente no outono. Como o objetivo deste trabalho foi realizar análise estatística dos dados, estas considerações foram feitas de forma resumida.

#### 4. Conclusão

De modo geral, o modelo de análise dos dados adotado demandou muito tempo para sistematização, processamento e análise dos resultados, o que resultou em muitas horas de trabalho. Porém, os resultados



encontrados se mostraram adequados, e evidenciou-se a necessidade de realizá-los para identificação dos valores válidos para a região e para evitar análises errôneas acerca dos padrões observados para o Paraná.

A fim de otimizar o tempo de processamento e a sequência dos estudos, sugere-se uma mudança na sequência dos trabalhos desenvolvidos. Por exemplo, pretende-se ao invés de aplicar o teste de homogeneidade para a sequência de cada mês separadamente para cada estação, exemplo, janeiro de 1977 a janeiro de 2006 para as 471 estações; dividir a área de estudo em regiões com pluviosidade similar e, na sequência, selecionar uma estação de referência e fazer a média de todas as estações para a mesma região com a aplicação do teste de homogeneidade de dupla massa, ou o teste de Pettitt e SNHT para a média por região de pluviosidade similar. Isso porque ao longo das análises notou-se que muitos meses de outubro de estações próximas apresentaram valores-r não significativos, sendo que a distribuição em ambas às estações seguia um padrão, mostrando que os dados eram sim homogêneos.

A primeira etapa de análise de consistência dos dados foi muito importante, pois diversos valores de precipitação foram eliminados; esta etapa incluiu uma análise de valores discrepantes, que se mostrou mais eficiente do que a aplicação do teste de Grubbs, pois este indicou períodos de extremos chuvosos como *outliers*.

## 5. Bibliografia

- CARAM, R. de O. **Reconstrução de séries e análise geostatística da precipitação no Estado de Minas Gerais**. Dissertação (*Magister Scientiae*) – Universidade Federal de Viçosa, 2007.
- CAVALCANTI, I. F. A.; FERREIRA, N. J.; DIAS, M. A. F. da S.; SILVA, M. G. A. J. da (Org.:). **Tempo e clima no Brasil**. São Paulo: Oficina de Textos, 2009.
- FRITZSONS, E.; MANTOVANI, L. E.; WREGGE, M. S.; NETO, A. C. Análise da pluviometria para definição de zonas homogêneas no Estado do Paraná. **RA'E GA**, v. 23, p. 555-572, 2011.
- GONZÁLES-ROUCO, J. F.; JIMÉNEZ, J. L.; QUESADA, V.; VALERO, F.. Quality control and homogeneity of precipitation data in the southwest of Europe. **Journal of Climate**, v.14, p.964-978, 2001.
- JORGE, F. V. **Fachada Atlântica Sul do Brasil: dinâmica e tendências climáticas regionais no contexto das mudanças globais**. Dissertação (Mestrado em Geografia). Universidade Federal do Paraná, 2009.
- KANG, H. M.; YUSOF, F.. Homogeneity tests on daily rainfall series in Peninsular Malaysia. **Int. J. Contem. Math. Sciences**, v.7, n.1, p. 9-22, 2012.
- LONGOBARDI, A.; VILLANI, P. Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. **International Journal of Climatology**, 2009.
- MIRZAEI, S. M.; RAOOF, M.; GHASEMI, A.; ESMALI, A.; ETAATI, H.; MORADNEZHADI, M.; MIRZAEI, Y. Determination of a some simple methods for outlier detection in maximum daily rainfall (case study: Baliglichay Watershed Basin – Ardebil Province – Iran). **Bulletin of Environment, Pharmacology and Life Sciences**, v.3, p.110-117, 2014.
- SENA, J. P. de O.; LUCENA, D. B. Identificação de tendência da precipitação na microrregião do Cariri Paraibano. **Revista Brasileira de Geografia Física**, v.6, n.5, p.1400-1416, 2013.



SILVA, A. N. **Detecção de outliers em séries espaço-temporais: análise de precipitação em Minas Gerais.** Dissertação (Mestrado em Estatística Aplicada e Biometria). Universidade Federal de Viçosa, 2012.

TEIXEIRA, M. da S. **Caracterização física e dinâmica de episódios de chuvas intensas nas regiões Sul e Sudeste do Brasil.** Tese (Doutorado em Meteorologia). Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2010.

TUCCI, C. E. M. **Hidrologia: ciência e aplicação.** 2. ed.; 2. reimpr. – Porto Alegre: Ed. Universidade/ UFRGS: ABRH, 2001.

VIVEKANANDAN, N. Estimation of probable maximum precipitation using statistical methods. **World Journal of Research and Review**, v.1, I.2, p.13-16, 2015.