

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO



MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL



**PROFMAT**

**ISRAEL CARDOSO**

**ANÁLISE DA PRECIPITAÇÃO MENSAL PROVÁVEL PARA A FAZENDA  
ESMERIL - PATROCÍNIO-MG**

UBERABA-MG

2016

**ISRAEL CARDOSO**

**ANÁLISE DA PRECIPITAÇÃO MENSAL PROVÁVEL PARA A FAZENDA  
ESMERIL - PATROCÍNIO-MG**

Dissertação apresentada ao curso de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, da Universidade Federal do Triângulo Mineiro, como requisito para obtenção do Título de Mestre em Matemática.

Orientador: Prof. Dr. Flávio Molina Da Silva

UBERABA-MG

2016

A minha companheira, amada e linda esposa Léa,  
meus filhos Isabele, Davi e Benjamim, por quem  
busco sempre o meu melhor em tudo.

## AGRADECIMENTOS

A Deus primeiramente por ser o meu Senhor e salvador e possibilitar com que esse sonho se tornasse realidade.

A minha família, minha querida esposa Léa, pelo incentivo, compreensão, auxílio, amor e companheirismo, meus filhos Isabele, Davi e Benjamim pelo amor, carinho e compreensão pelos momentos de ausência durante o curso, vocês são minhas fontes de estímulo e dedicação.

A Universidade Federal do Triângulo Mineiro – UFTM pela oportunidade de realizar esse curso.

A Capes pelo financiamento deste projeto.

Ao prof. Dr. Flávio pela orientação deste trabalho.

Aos professores da banca, Dr. Flávio Molina da Silva, Dr. Rafael Henrique Palma Lima e Ms. Josimeire Maximiano dos Santos, pelas sugestões e contribuições.

Ao Arinaldo Oliveira por fornecer os dados pluviométricos coletados na fazenda Esmeril sem os quais não seria possível a confecção deste trabalho.

Ao colega e amigo José Henrique Bizinotto pela hospitalidade na cidade de Uberaba. Sem sua cooperação teria sido mais difícil à conclusão deste curso.

Ao colega Geraldo Henrique Alves Pereira pela colaboração nos estudos, companheirismo e amizade.

Ao colega Edmilson Nahass Franco pelo apoio nas viagens, parceria nos estudos, pelas conversas e grande amigo que se tornou.

A Andréia Alves de Souza e Raíssa Goulart pela ajuda no abstract.

## RESUMO

O ajuste de modelos probabilístico aos dados diários de chuva além de fornecer um resumo sucinto destes dados, representa uma técnica eficiente para a análise dessas informações. Este trabalho tem como objetivo verificar se a média diária de cada mês da região da fazenda Esmeril situada no município de Patrocínio-MG, do período de 2000 a 2012, se ajusta a uma função de distribuição Normal, prever a quantidade pluviométrica de determinados meses utilizando a distribuição Normal e verificar os dados da série temporal estudando seu comportamento pluviométrico neste intervalo de treze anos usando os conceitos de Séries Temporais. Primeiramente foi verificado que os dados se ajustaram a função de distribuição pelo teste de Kolmogorov-Smirnov com nível de significância de 5%. Após verificar a aderência dos valores à distribuição normal, foi feita uma previsão usando o Teorema Central do Limite para os dez meses iniciais de cada intervalo, porém este método não se mostrou muito eficiente. O estudo da série temporal mostrou que os dados pluviométricos do período do intervalo considerado para as variáveis total mensal, média diária de cada mês, total semanal e média diária de cada semana não possuem tendência, tem sazonalidade e apresentam variações irregulares.

**Palavras-chave:**Séries Temporais.Teorema Central do Limite.Probabilidade.Precipitação.

## ABSTRACT

The adjustment of probabilistic models to daily rainfall data and provide a brief summary of these data is an efficient technique for the analysis of this information. This study aims to determine if the daily average for each month of Emery Farm area in the municipality of Patrocínio-MG, from 2000 to 2012, adjusts a Normal distribution function, predict the rainfall amount of certain months using Normal distribution and verify the data of the time series studying its rainfall behavior in this interval of thirteen years using the concepts of time series. First it was found that the data set the distribution function using the Kolmogorov-Smirnov test with 5% significance level. After checking the adherence of values to the normal distribution, a prediction using the Central Limit Theorem for the first ten months of each interval was made, but this method was not very efficient. The study of the time series showed that the rainfall data of the range of the period considered for the total monthly variable daily average of each month, weekly total and daily average of each week have no tendency has seasonality and have irregular variations.

**Keywords:** Time Series. Central Limit Theorem. Probability. Precipitation.

## LISTA DE FIGURAS

1. Aspecto do gráfico da curva normal	14
2. Gráfico da variável média diária de cada mês	23
3. Gráfico da variável total mensal	27
4. Gráfico da variável total mensal e sua tendência	28
5. Gráfico da variável total mensal e gráfico sem a sua tendência	29
6. Gráfico da variável média diária de cada mês	30
7. Gráfico da variável média diária de cada mês e sua tendência	30
8. Gráfico da variável média diária de cada mês e gráfico sem a sua tendência	31
9. Gráfico da variável total semanal	32
10. Gráfico da variável total semanal e sua tendência	32
11. Gráfico da variável total semanal e gráfico sem a sua tendência	33
12. Gráfico da variável média diária de cada semana	34
13. Gráfico da variável média diária de cada semana e sua tendência	34
14. Gráfico da variável média diária de cada semana e gráfico sem a sua tendência	35
15. Gráfico da variável total mensal sem a sua sazonalidade	36
16. Gráfico da variável média diária de cada mês sem a sua sazonalidade	37
17. Gráfico da variável total semanal sem a sua sazonalidade	37
18. Gráfico da variável média diária de cada semana sem a sua sazonalidade	38
19. Gráfico das variações cíclicas e irregulares da variável total mensal	38
20. Gráfico das variações cíclicas e irregulares da variável média diária de cada mês	39
21. Gráfico das variações cíclicas e irregulares da variável total semanal	40
22. Gráfico das variações cíclicas e irregulares da variável média diária de cada semana	40

## **LISTA DE TABELAS**

1. Aderência da variável média diária de cada mês à distribuição Normal 24
2. Previsão pluviométrica do valor médio diário acumulado para os dez primeiros meses 26



## SUMÁRIO

1. Introdução	11
2. Conceitos básicos de probabilidade e séries temporais	13
3. Apresentação do problema	18
4. Resultados	22
4.1. Previsão pluviométrica da média diária de cada mês usando a função de distribuição Normal	22
4.2. Análise da série temporal	26
5. Conclusão e propostas futuras	41
6. Referências	42

## 1. INTRODUÇÃO

A água tem fundamental importância para a manutenção da vida e falar da relevância dos conhecimentos sobre a água, em suas diversas dimensões, é falar da sobrevivência da espécie humana, da conservação e do equilíbrio da biodiversidade e das relações de dependência entre seres vivos e ambientes naturais (BACCI& PATACA, 2008).

Bacci e Pataca (2008) afirmam ainda que na sociedade em que vivemos, a água passou a ser vista como recurso hídrico e não mais como um bem natural. A exploração dos recursos naturais, dentre eles a água, de forma bastante agressiva e descontrolada, levou a uma crise socioambiental bastante profunda. Que hoje nos deparamos com uma situação na qual estamos enfrentando um dos mais graves problemas neste século.

Segundo Diniz (2013), a precipitação pluviométrica é uma das variáveis meteorológicas de maior importância para a sociedade em consequência de sua influência direta sobre vários aspectos e atividades relacionadas ao homem como agricultura, pecuária, etc. O excesso ou escassez de chuvas acarretam danos socioeconômicos e ambientais de grande magnitude. E uma importante ferramenta que auxilia na previsão desses eventos extremos é o monitoramento diário do regime pluviométrico sobre uma dada localidade.

Ribeiro e Lunardi (1997) também salientam a importância da caracterização da precipitação em um local para o planejamento de atividades agrícolas, sendo imprescindível no dimensionamento de reservatórios de água, na elaboração de projetos de proteção e conservação de solos e em atividades de lazer e esportivas.

Deste modo, o ajuste de modelos probabilísticos aos dados diários de chuva, além de fornecerem um resumo sucinto destes dados, representa uma técnica eficiente para a análise dessas informações (ALMEIDA, 1995).

Sendo que a distribuição Normal é uma das mais importantes distribuições de probabilidade para variáveis aleatórias contínuas, cuja distribuição é chamada distribuição Normal ou Gaussiana, a qual serve como modelo de distribuição para diversas aplicações (MORETTIN e BUSSAB, 2003).

Diante disso, percebe-se a importância dos trabalhos relacionados à precipitação pluviométrica. Pois, através de análises estatísticas, podem-se tomar decisões mais acertadas com relação ao planejamento e desenvolvimento das atividades agrícolas, bem como seu melhor aproveitamento. E ainda fazer previsões sobre o seu comportamento futuro.

Oliveira (2013) analisou o ajuste das distribuições de probabilidade Exponencial, Gama, Weibull, Log-normal e Gumbel, e análise descritiva da série temporal de precipitações

pluviométricas diárias da fazenda Esmeril, município de Patrocínio-MG, para as variáveis média mensal, total mensal e máximo total. Resultando que os dados da série de precipitações, não apresentam tendência, apenas sazonalidade com periodicidade anual e com alta variabilidade. O teste Kolmogorov-Smirnov mostrou que a média mensal é ajustada pelas distribuições Exponencial, Weibull e Log-normal, no caso do máximo mensal, o ajuste foi verificado para as distribuições: Gama e Weibull. E que as distribuições: Exponencial, Gama e Weibull ajustaram-se aos dados da série de precipitações pluviométricas em estudo, para o total mensal.

Este trabalho tem como objetivo verificar os dados pluviométricos coletados diariamente no período de janeiro de 2000 a dezembro de 2012 da fazenda Esmeril, situada no Município de patrocínio – MG, citada em Oliveira (2013). Primeiramente, analisar se a variável média diária de cada mês para cada ano se ajusta a uma função de distribuição normal, prever a quantidade pluviométrica de determinados meses utilizando a distribuição normal acumulada e justificado pelo Teorema do Limite Central e verificar os dados da série temporal estudando o seu comportamento pluviométrico através dos conceitos de Séries Temporais.

O trabalho foi estruturado em cinco seções, em que, além desta primeira que traz a introdução, tem-se a Seção 2, que trata de conceitos básicos de função de distribuição de probabilidade, a função de distribuição Normal, o Teorema Central do Limite, o teste de aderência Kolmogorov-Smirnov e conceitos de séries temporais referentes a esse estudo. A apresentação do problema está na Seção 3 onde estão contidos os objetivos, que são verificar se a média pluviométrica diária de cada mês se ajusta a uma função de distribuição Normal, verificar a aderência usando o teste de Kolmogorov-Smirnov, fazer uma previsão justificada pelo Teorema Central do Limite, fazer um estudo dos dados da série temporal no intervalo de janeiro de 2000 a dezembro de 2012 da fazenda em estudo. Ainda nesta seção, vários referenciais teóricos de diversas regiões do Brasil são apresentados com exemplos de aplicações. A Seção 4 apresenta os resultados obtidos neste trabalho, em que a variável média diária de cada mês se ajustou à distribuição Normal, a previsão da média pluviométrica diária de cada mês usando a distribuição Normal mostrou-se pouco eficiente e ainda que a série de dados não tenha tendência, é sazonal e possui irregularidades. Por fim, a Seção 5 traz as considerações finais sobre os resultados obtidos neste trabalho, e propostas futuras para previsão de dados pluviométricos em uma dada localidade.

## 2. CONCEITOS BÁSICOS

Nesta seção, será apresentado um referencial teórico para melhor entendimento das aplicações neste trabalho. Para Meyer (2012) todas as vezes que empregarmos Matemática a fim de estudar alguns fenômenos de observação, devemos essencialmente começar pela construção de um modelo matemático determinístico ou probabilístico para esses fenômenos. Neste trabalho, são estudados modelos probabilísticos a partir de estatísticas, uma vez que as variáveis de estudo são as variáveis aleatórias: total mensal, média diária de cada mês, total semanal e média diária de cada semana da série de precipitações pluviométricas da fazenda Esmeril no município de Patrocínio-MG.

Para Meyer (2012) se  $E$  é um experimento aleatório e  $S$  um espaço amostral associado e se  $X$  é uma variável aleatória definida em  $S$  tal que  $X(S)$  seja infinito não-enumerável, isto é,  $X(S)$  seja um intervalo de números reais, então  $X$  é dita uma variável aleatória contínua. Dessa forma, considerando que nossa variável de estudo é a quantidade pluviométrica de um determinado local, temos um experimento aleatório cuja variável aleatória está definida em um espaço amostral contínuo, com valores reais não negativos. Diz-se que  $X$  é uma variável aleatória contínua, se existir uma função  $f$ , denominada função densidade de probabilidade f.d.p. de  $X$ , para  $-\infty < x < +\infty$  que satisfaça às seguintes condições:

(a)  $f(x) \geq 0$ , para todo  $x$ ;

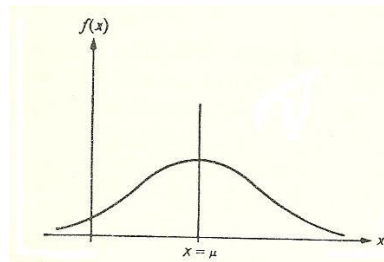
(b)  $\int_{-\infty}^{+\infty} f(x) dx = 1$ ;

(c) para quaisquer  $a, b$ , com  $-\infty < a < b < +\infty$ , teremos  $P(a \leq X \leq b) = \int_a^b f(x) dx$ .

Uma das mais importantes distribuições de probabilidade para variáveis aleatórias contínuas é a distribuição Normal (ou Gaussiana) que é definida por uma f.d.p. da forma:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right], \text{ para } -\infty < x < +\infty \quad (1)$$

A média  $\mu$  e o desvio padrão  $\sigma$  devem satisfazer as seguintes condições  $-\infty < \mu < +\infty$  e  $\sigma > 0$ . Geralmente é usada a seguinte notação:  $N(\mu, \sigma^2)$  para a variável  $X$  que tem distribuição Normal com média  $\mu$  e desvio padrão  $\sigma$ . A Figura 1 mostra um exemplo de um gráfico dessa distribuição. Ele apresenta a forma de sino e é simétrico em relação à média  $\mu$  (MEYER, 2012).

**Figura 1-Distribuição Normal**

Aspecto do gráfico da curva normal da variável aleatória  $X$  com Média  $\mu$  e desvio padrão  $\sigma$ . Fonte: Meyer [(2001)]

Neste trabalho, para o estudo da quantidade de precipitação pluviométrica será considerado a média diária de cada mês dos valores pluviométricos. A distribuição Normal será usada, não apenas por representar diversos fenômenos naturais, mas principalmente ao fato da propriedade dada pelo Teorema Central do Limite. Assim, pode-se relacionar a média mensal de cada ano à função de distribuição Normal e para isso será utilizado o Teorema Central do Limite, de acordo com seguinte conceito.

Segundo Meyer (2012), o Teorema Central do Limite diz que se uma variável aleatória  $X$  puder ser representada pela soma de quaisquer  $n$  variáveis aleatórias independentes, então esta soma, para  $n$  suficientemente grande, terá distribuição aproximadamente normal. Um dos enunciados desse teorema pode ser apresentado da seguinte maneira:

Teorema Central do Limite: Seja  $X_1, X_2, \dots, X_n$ , uma sequência de variáveis aleatórias independentes, com  $E X_i = \mu_i$  e  $V X_i = \sigma_i^2$ ,  $i = 1, 2, \dots, n$  Façamos  $X = X_1 + X_2 + \dots + X_n$ . Então,

$$Z_n = \frac{X - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \rightarrow N(0,1) . \quad (2)$$

Para validar a relação entre a função de distribuição Normal com a média diária de cada mês, será utilizado o teste de Kolmogorov-Smirnov para verificar a aderência dos dados a distribuição. Para Costa Neto (2002), uma importante classe de testes não-paramétricos é constituída pelos chamados testes de aderência, em que a hipótese testada refere-se à forma da distribuição da população. Nesses testes, admitimos, por hipótese, que a distribuição da variável de interesse na população seja descrita por determinado modelo de distribuição de probabilidade e testamos esse modelo, ou seja, verificamos a boa ou má aderência dos dados da amostra ao modelo. Se obtivermos uma boa aderência e a amostra for razoavelmente grande, poderemos, em princípio, admitir que o modelo a um dado nível de significância forneça uma boa idealização da distribuição populacional. Kolmogorov e Smirnov

desenvolveram um teste, em geral mais poderoso que o do Qui-quadrado ( $\chi^2$ ) para testar a aderência, em que a variável de teste é a maior diferença observada entre a função de distribuição acumulada do modelo e a da amostra. A função de distribuição acumulada do modelo testado, ou função de repartição, dá as probabilidades acumuladas em cada ponto, ou seja,  $F(x) = P(X \leq x)$ . A função de distribuição acumulada da amostra corresponderá ao gráfico das frequências relativas acumuladas. Essa segunda função será denotada por  $G(x)$ . O teste consta simplesmente da verificação do valor  $d = \max F(x) - G(x)$  e da comparação com um valor crítico tabelado em função de  $\alpha$  e  $n$ , em que  $\alpha$  é o nível de significância e  $n$  é o tamanho da amostra. Se  $d$  for maior que o valor crítico, definido segundo o nível de significância e o tamanho da amostra, rejeita-se a hipótese de que os dados seguem a distribuição que está sendo feita o teste de aderência. Valores críticos para os níveis de significância usuais são dados em uma tabela (COSTA NETO, 2002).

Neste trabalho também é feito um estudo de toda a série para observar o comportamento da variável de índice pluviométrico da fazenda Esmeril no município de Patrocínio-MG ao longo do tempo, utilizando os conceitos de séries temporais. Segundo Reis (2015) séries temporais são um conjunto de observações sobre uma variável, ordenados no tempo, e registrados em períodos regulares, como a precipitação pluvial diária de uma localidade ou região. A suposição básica que norteia a análise de séries temporais é que há um sistema previsível, relacionado com o tempo, que exerceu influência sobre os dados no passado e pode continuar a fazê-lo no futuro. Este sistema causal costuma atuar criando padrões não aleatórios que podem ser detectados em um gráfico da série temporal, ou mediante algum outro processo estatístico permitindo fazer previsões sobre o futuro, orientando a tomada de decisões. Assim, o problema fundamental é utilizar um modelo que permita incluir os vários tipos de padrões, possibilitando realizar previsões. O ponto de partida é realizar a decomposição da série em padrões.

Reis (2015) afirma que o modelo clássico de todas as séries temporais é composto de quatro padrões como a tendência ( $T$ ), que é o comportamento de longo prazo da série, que pode ser causada pelo crescimento demográfico, mudança gradual de hábitos de consumo, ou qualquer outro aspecto que afete a variável de interesse no longo prazo; as variações cíclicas ou ciclos ( $C$ ), que são flutuações nos valores da variável com duração superior a um ano, e que se repetem com certa periodicidade, que podem ser resultado de fenômenos climáticos como o *El Niño* (que se repete com periodicidade superior a um ano); variações sazonais ou sazonalidade ( $S$ ), flutuações nos valores da variável com duração inferior a um ano, e que se

repetem todos os anos, geralmente em função das estações do ano; e variações irregulares (*I*), que são as flutuações inexplicáveis, resultado de fatos fortuitos e inesperados como catástrofes naturais, tempestades, tornados, furacões, etc.

A decomposição da série permitirá identificar quais componentes estão atuando naquele conjunto em particular, além de possibilitar obter índices ou equações para realizar previsões para períodos futuros da série. A questão crucial do modelo clássico é decidir entre o modelo aditivo e multiplicativo, que é a equação que relaciona as componentes com a variável. No modelo aditivo valor da série (*Y*) será o resultado da soma dos valores das componentes (que apresentam a mesma unidade da variável):  $Y = T + C + S + I$  ou  $Y = T + C + I$  (se os dados forem registrados anualmente). Todas as componentes têm a mesma unidade da série: se esta for dada em milímetros todas também terão tal unidade. Já no modelo multiplicativo, o produto das componentes resultará na variável da série:  $Y = T.C.S.I$  ou  $Y = T.C.I$  (se os dados forem registrados anualmente). Há, porém, uma diferença crucial: apenas a tendência tem a mesma unidade da variável. As demais componentes têm valores que modificam a tendência: assumem valores em torno de 1 (se maiores do que 1 aumenta a tendência, se menores diminuem a tendência, se exatamente iguais a 1 não causam efeito) (REIS,2015). Neste trabalho, foi utilizado o modelo aditivo em virtude da quantidade de zeros dos índices pluviométricos diários, que gerou indeterminações ao calcular os índices sazonais no modelo multiplicativo, inviabilizando o seu uso.

A tendência descreve o comportamento da variável retratada na série temporal no longo prazo, sendo que há três objetivos básicos na sua identificação: avaliar o seu comportamento para utilizá-lo em previsões, removê-la da série para facilitar a visualização das outras componentes, ou ainda identificar o nível da série (o valor ou faixa típica de valores que a variável pode assumir se não for observado comportamento crescente ou decrescente no longo prazo). A obtenção da tendência pode ser feita através de um modelo de regressão (como o modelo linear), onde a reta de tendência será  $T = a + b.t$  com  $T$  igual ao valor da tendência,  $t$  o valor do tempo,  $b$  o coeficiente angular da reta (se positivo indica tendência crescente, se negativo a tendência é decrescente) e  $a$  o coeficiente linear da reta. As equações dos coeficientes são:

$$b = \frac{n \cdot \sum_{i=1}^n t_i \cdot y_i - \sum_{i=1}^n t_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i\right)^2}, \quad e \quad a = \frac{\sum_{i=1}^n y_i - b \cdot \sum_{i=1}^n t_i}{n} \quad (3)$$

em que  $y_i$  é um valor qualquer da variável registrada na série temporal,  $t_i$  é o período associado a  $y_i$  e  $n$  é o número de períodos da série (REIS, 2015).

A obtenção da tendência também pode ser feita através de médias móveis que são uma forma alternativa de obtenção da mesma. Para tal calcula-se inicialmente a média dos primeiros  $n$  períodos da série, colocando o resultado no período exatamente no centro deles. Progressivamente, acrescenta-se um período seguinte e se despreza o primeiro da média imediatamente anterior, calculando novas médias, que vão se movendo até o fim da série. O número de períodos  $n$  é chamado de ordem da série. Quanto maior o número de períodos da série agrupados pela média móvel mais suavizada fica a linha de tendência, representando melhor o comportamento de longo prazo. E quanto menor, o número de períodos, mais a tendência acompanham os dados originais. Por este motivo, quando uma série apresenta muitas irregularidades é comum suavizá-la através de médias móveis (REIS, 2015).

Uma vez identificada a tendência, seja por equações ou por médias móveis, ela pode ser removida da série para facilitar a visualização das outras componentes, que em um modelo aditivo é feito pela equação da forma  $I + S + C = Y - T$ .

As variações sazonais são oscilações de curto prazo, que ocorrem sempre dentro do ano, e que se repetem sistematicamente ano após ano. Os índices sazonais modificam a tendência ao serem somados (modelo aditivo). Se todos os índices forem próximos ou exatamente iguais à zero, então as componentes sazonais parecem não exercer grande efeito sobre a série (REIS, 2015).

Há vários métodos para a obtenção dos índices sazonais, entre eles o método da razão para a média móvel que consiste inicialmente em obter médias móveis de ordem igual ao número de períodos sazonais, que neste trabalho para os casos mensais foram de 12 meses e para os casos semanais foram de 52 semanas. Como as ordens são pares, é necessário obter médias móveis centradas de dois períodos. Para isto, primeiramente calcula-se as médias móveis de 12 períodos (mensal) ou 52 períodos (semanal) e, em seguida, obter as médias móveis de dois períodos centradas. Após obter as médias móveis, para encontrar os índices sazonais para cada período no modelo aditivo, subtraem-se dos valores originais da série as médias móveis centradas calculadas, para em seguida obter medidas de síntese dos índices, que representarão cada período sazonal, calculando a média aritmética simples dos valores correspondentes ao período sazonal (média dos índices obtidos em todos os janeiros, fevereiro e assim sucessivamente até dezembro). Para fazer as correções necessárias para que a soma dos índices seja coerente (igual à zero para o modelo aditivo), soma-se todos os índices calculados na etapa anterior e divide-se esta soma pela ordem da sazonalidade (12 se forem mensal e 52 se for semanal). O resultado é subtraído de cada um dos índices,



garantindo que a soma deles seja igual à zero. Dessa forma encontram-se os índices sazonais de janeiro a dezembro que será aplicado a toda série de dados (REIS, 2015).

A sazonalidade da série pode ser removida, subtraindo os valores originais de cada período por seu respectivo índice sazonal, pelo modelo aditivo  $Y - S = T + C + I$ , fazendo assim a análise do resultado.

Para Reis (2015) as variações cíclicas e irregulares, geralmente são avaliadas em conjunto, pois as variações irregulares são aleatórias. Conforme visto anteriormente, as variações cíclicas são padrões de longo prazo (superiores há um ano), como por exemplo, fenômenos climáticos, já as variações irregulares são resultado de fatos fortuitos, inesperados. Elas são obtidas através da remoção das componentes tendência e sazonalidade (esta última apenas se os dados não forem anuais), o que no modelo aditivo é feito pela equação na forma  $Y - S - T = C + I$  em que  $Y$  é o valor original da série,  $T$  é a tendência, e  $S$  é a componente sazonal (representada através dos índices sazonais). Para identificar se há ciclos na série, há padrões que devem ser observados no gráfico das variações cíclicas e irregulares. Com isso, é possível fazer o estudo da série temporal da quantidade pluviométrica da fazenda.

### 3. APRESENTAÇÃO DO PROBLEMA

Os dados utilizados foram fornecidos por Oliveira (2013), obtidos nas observações referentes às precipitações pluviométricas diárias na fazenda Esmeril, município de Patrocínio-MG. Estas foram expressas em altura de lâmina d'água (mm), abrangendo o período de janeiro de 2000 a dezembro de 2012, num total de treze anos. O município de Patrocínio está localizado na latitude  $18^{\circ} 56' S$  e longitude  $46^{\circ} 59' W$  e altitude média de 972 metros acima do mar. Segundo Mendes (2001) o clima da região segundo a classificação Koppen é tropical com verão quente e úmido e inverno mais frio e seco.

O objetivo inicial deste trabalho é verificar se a média pluviométrica diária de cada mês de cada ano se ajusta a uma função de distribuição normal. E diferente da maioria dos demais trabalhos sobre este tema, que em sua maioria, só verificam a aderência dos índices a uma determinada distribuição, neste trabalho foi feita uma previsão da quantidade pluviométrica acumulada em determinados meses usando a distribuição Normal justificado pelo Teorema Central do Limite. Para isso, foram consideradas para fins de análise, as precipitações médias diárias de cada mês em planilhas de treze em treze meses (junho a junho), tendo um total de doze anos, pois dessa forma o gráfico se parece com a forma de um

sino. Sendo que a aderência as probabilidade foi verificada via teste de Kolmogorov-Smirnov ao nível de 5 % de significância.

Além disso, após o estudo de previsão em cada ano, usando a distribuição Normal, tem-se também como objetivo estudar os dados da série temporal, ao longo deste intervalo de treze anos, verificando o comportamento pluviométrico dos dados considerados, para as variáveis: total mensal, média diária de cada mês, total semanal e média diária de cada semana. Utilizando o modelo clássico das Séries Temporais com a verificação da tendência linear, médias móveis, sazonalidade e ciclos para ajudar nos planejamentos dos trabalhos da fazenda Esmeril.

Vários trabalhos semelhantes a este, de autores diferentes, em diversas regiões do Brasil, serão apresentados nesta seção como referenciais teóricos e exemplos de aplicações, como o estudo feito por Neto et al. (2005), que analisou as distribuições Normal, Log-normal, Gama, Exponencial e Weibull para verificar qual delas se ajustaria às séries históricas de precipitação pluvial diária em Piracicaba-SP, sendo consideradas, para fins de análise, apenas cada série de precipitação pluvial diária de dias semelhantes. Utilizou-se o teste de Qui-quadrado para verificar a aderência das probabilidades estimadas às frequências observadas com 5 % de significância, sendo possível observar a superioridade do ajustamento da distribuição Gama e Weibull, com predominância da Gama em relação à Weibull. Verificou-se para algumas datas que a aderência das distribuições Log-normal e exponencial, as quais predominam nos meses de fevereiro, março, outubro e novembro para o modelo Exponencial, e março, novembro e dezembro para o modelo Log-normal. Em algumas séries, nenhum dos modelos estudados se ajustou aos dados. Concluindo, conforme resultados obtidos, para Piracicaba-SP recomenda-se a distribuição Gama para as estimativas diárias de probabilidade de precipitação pluvial para os meses de janeiro, fevereiro, março, outubro, novembro e dezembro.

Em Minas Gerais, Catalunha *et al.* (2002) analisou as distribuições de probabilidade Exponencial, Gama, Log-normal (a dois e três parâmetros), Normal e Weibull para as séries de precipitações pluviais no estado. Os testes, não-paramétricos, de Qui-quadrado e de Kolmogorov-Smirnov foram utilizados para verificar a aderência das probabilidades estimadas aos dados observados. Foi considerada para fins de análise, a precipitação diária e total para períodos decendiais e mensais de janeiro a dezembro. Como período chuvoso foi considerado os meses de janeiro, fevereiro, março, outubro, novembro e dezembro, e como período seco, os meses de abril, maio, junho, julho, agosto e setembro. Para os valores diários de precipitação, observou-se a superioridade do ajustamento da distribuição Weibull. No caso

dos valores totais de precipitação para o período seco, houve predominância da distribuição exponencial, no período chuvoso prevaleceram às distribuições Weibull, Exponencial, Gama e Normal, respectivamente sendo que esta última apareceu somente em dois meses. Portanto, o autor do artigo, não recomenda para o estado de Minas Gerais, a distribuição Gama, amplamente utilizada por pesquisadores.

Moreira *et al.* (2010), utilizou a distribuição Gama Incompleta para definir diferentes níveis de probabilidade de precipitação no município de Nova Maringá-MT. Foram utilizados dados pluviométricos disponibilizados pela Agência Nacional de Águas (ANA), do período de 1983 a 2007, totalizando um histórico ininterrupto de 25 anos. Neste caso, a função Gama Incompleta expressou adequadamente à probabilidade de precipitação na região em estudo e o teste de Kolmogorov-Smirnov foi adequado para representar as precipitações mensais em função de distribuição de probabilidade. A região apresenta duas estações definidas, uma estação seca de março a julho e outra chuvosa de outubro a janeiro, sendo caracterizados os outros meses com períodos de transição entre as estações.

A análise de frequência hidrológica dos dados de precipitação pluvial de algumas estações agroclimatológicas da região sul do Rio Grande do Sul, feita por Damé *et al.* (1996), teve como objetivo ajustar os dados observados de precipitação pluvial anual e máximo e mínimo mensal, para diversos níveis de probabilidade das estações agroclimatológicas de Pelotas, Morro Redondo, Domingos Petrolina e Rio Grande. A série de valores de precipitação pluvial anual das estações de Pelotas, Morro Redondo, Domingos Petrolina e Rio Grande ajustaram-se adequadamente as distribuições de probabilidade Normal, Log-normal, Log-normal e Gumbel, respectivamente. Já as séries dos máximos anuais adequaram-se, para as estações descritas anteriormente, às distribuições de probabilidades Gumbel, Gumbel e Log-normal, respectivamente. A série de mínimos anuais de Rio Grande não foi considerada homogênea pelo “Run-Test”.

Testar o ajuste de diferentes funções de distribuição de frequência à precipitação mensal e anual e estimar a precipitação pluviométrica esperada em diferentes níveis de probabilidade para a cidade de Iguatu-CE, foram objetivos de Silva, I. *et al.* (2009). Para isso foram testadas as distribuições de probabilidade Exponencial, Gumbel I, Gama, Log-normal, Normal, Weibull e Beta, sendo que a série de dados utilizados abrange o período de 1913 e 1989 obtidos da base de dados da SUDENE. A aderência das distribuições foi testada pelos testes de qui-quadrado ao nível de 5 % de significância e Kolmogorov-Smirnov ao nível de 20 % sendo que a distribuição que obteve melhor desempenho no ajuste à precipitação deste local para a série mensal foi a Gumbel, sendo que as funções Gama, Log-normal, Normal e

Weibull foram as que obtiveram o menor número de aderência nos testes. Para a série anual recomenda-se o uso da distribuição Normal sendo que todas as distribuições proporcionaram uma boa adequação aos dados de precipitação com exceção da distribuição Exponencial que não demonstrou aderência segundo os testes.

Um estudo da precipitação efetiva para o município de Lavras - MG foi feito por Sampaio *et al.* (2000), utilizando a distribuição Log-normal com dados da estação climatológica da Universidade Federal de Lavras, no período de 1914 a 1991, com os métodos para se estimar a precipitação efetiva: percentagem fixa (PF), fórmula empírica (AGLW/FAO) e o do *Soil Conservation Service Method (USDA)*. Considerando-se o nível de probabilidade de 75 %, foram encontrados valores de precipitação efetiva anual de 125,4; 313,6; 501,8; 330,6 e 485,0 mm para métodos PF 20, 50 e 80%, FAO e USDA, respectivamente. Para o nível de probabilidade de noventa por cento, encontraram-se os menores totais anuais de precipitação efetiva, sendo: 79,5; 198,8; 318,1; 188,8 e 327,8 mm para os métodos PF 20, 50 e 80%, FAO e USDA, respectivamente. Concluindo, portanto que os valores mensais obtidos pelos métodos FAO e USDA se aproximam dos estimados pelo método PF 50 e 80%, respectivamente, para os períodos intermediários entre as estações de seca e de chuvas principalmente.

Visando determinar a precipitação pluviométrica mensal provável para Lages-SC, Soccol *et al.* (2010) analisou o ajuste da distribuição Gama aos dados de série de 80 anos, sendo que em todos os meses a distribuição Gama ajustou-se aos dados observados ( $p < 0,05$ ). Para a estimação dos parâmetros  $\alpha$  e  $\beta$  da referida distribuição utilizaram o método da máxima verossimilhança, e o teste de Kolmogorov-Smirnov para verificar a aderência do modelo aos dados, obtendo a partir da distribuição ajustada a estimativa da precipitação mensal para os níveis de probabilidade de 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90 e 95%, correspondentes à ocorrência da precipitação mínima provável  $P(X > x)$ . Os resultados deste trabalho mostraram que a precipitação pluviométrica média mensal, na média dos 12 meses do ano, ocorreu em nível de probabilidade de 31,93% com coeficiente de variação de 3,63%, o que comprova que seu valor não deve ser adotado como referência em projetos agrícolas.

Objetivando estimar a precipitação mensal provável para o município de Boa Vista-RR, Araújo *et al.* (2001), usando os dados de precipitação pluviométrica mensal de quarenta e oito anos da série histórica compreendida entre os anos de 1923 e 1997, obteve as estimativas dessa precipitação em níveis de 10, 20, 25, 30, 50, 60, 70, 75, 80 e 90% de probabilidade utilizando as funções de distribuição Normal e Gama Mista. Verificou-se um

bom ajuste dos valores mensais principalmente à distribuição Gama Mista, exceto para os meses secos de janeiro e fevereiro. E, concluiu que o período de maior precipitação pluviométrica em Boa Vista-RR, tem início em abril e se prolonga até setembro, indicando um período viável de cultivo por seis meses com uso eventual de irrigação complementar.

Após a apresentação do problema e alguns referenciais, nesta próxima seção serão apresentados todos os resultados encontrados neste estudo.

## 4. RESULTADOS

Nesta seção serão apresentados todos os resultados encontrados neste trabalho. Inicialmente será exposto o ajuste da variável média diária de cada mês à distribuição Normal, uma previsão da média pluviométrica diária de cada mês usando a distribuição Normal, tendo como base o histórico da série e justificado pelo Teorema Central do Limite. E em seguida apresentado o estudo de Séries Temporais da série de dados pluviométricos para os treze anos do período de 2000 a 2012, verificando tendência, sazonalidade, ciclos e irregularidade da série utilizando médias móveis e o software Excel para elaboração das tabelas e gráficos.

### 4.1. Previsão da Média Pluviométrica Diária de cada mês usando a Distribuição Normal

Primeiramente, para fazer a estimativa da quantidade média pluviométrica diária acumulada até um determinado mês, há a necessidade de organizar os dados e garantir que os mesmos possuam uma distribuição Normal para uma determinada média e desvio-padrão. Os dados pluviométricos considerados são diários, do período de 2000 a 2012 da fazenda Esmeril em Patrocínio-MG, apresentados em Oliveira (2013). Os dados foram organizados em planilhas contendo treze meses, por exemplo, de junho de 2000 a junho de 2001, junho de 2001 a junho de 2002, assim sucessivamente até junho de 2011 a junho de 2012, pois dessa forma o gráfico da variável média diária de cada mês tem a forma idêntica a um sino como no gráfico da distribuição Normal, totalizando doze planilhas. Para cada planilha foi estimado o valor esperado e o desvio padrão das médias diárias de cada mês usando do método de estimação pontual do momento cuja fórmula geral é:

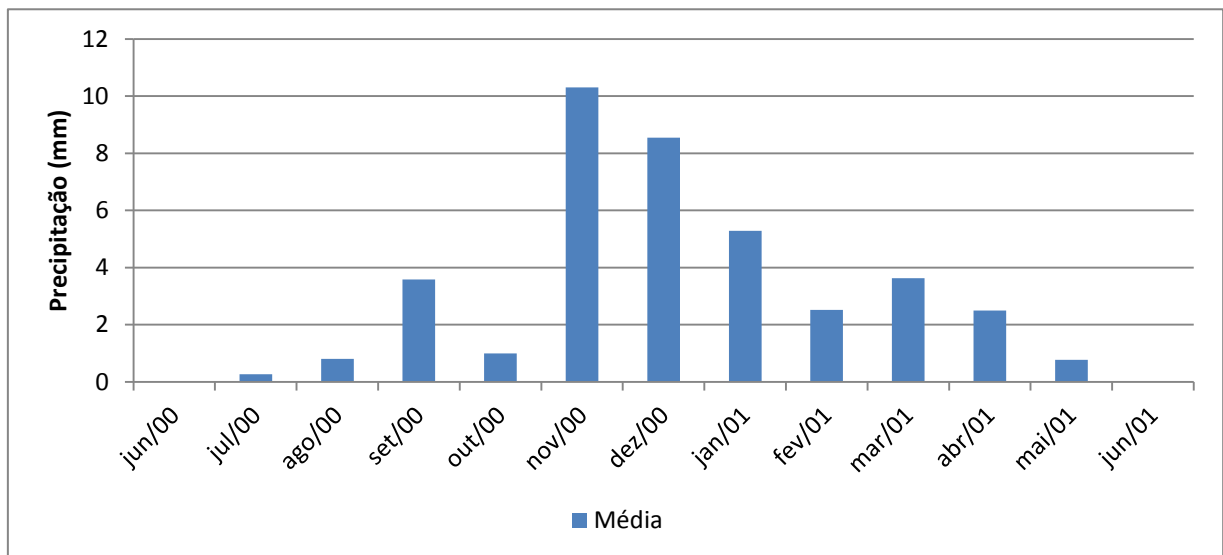
$$m_t = \frac{\sum_{i=1}^n x_i^t}{n} \quad (4)$$

em que  $x_i$  são os dados pluviométricos,  $i = 1, \dots, n$ ,  $t$  é a ordem do momento e  $n$  é a frequência dos dados. Depois, encontrou-se o valor da distribuição Normalizada  $z$  em relação

a cada média diária de cada mês e sua respectiva probabilidade utilizando a tabela Normal Reduzida. Usando o teste de Kolmogorov-Smirnov com um nível de 5% de probabilidade para um período de  $n$  igual há treze meses cujo valor crítico seria 0,361, verificou-se que a média pluviométrica diária de cada mês se adere a uma distribuição Normal. Pois, os valores dos desvios máximos  $d$ , em cada período de treze meses foram todos inferiores ao valor crítico de 0,361 como mostra a Tabela 1. Cabe mencionar que para cada intervalo considerado as médias se aderiram à distribuição Normal com médias e desvios padrão diferentes.

No gráfico apresentado na Figura 2 estão representadas as médias diárias de cada mês do período de junho de 2000 a junho de 2001.

**Figura 2 – Gráfico da variável média diária de cada mês de junho de 2000 a junho de 2001**



Fonte: **Elaborado pelo próprio autor**

Para exemplificar, ao verificar a aderência das precipitações pluviométricas em estudo à função de distribuição de probabilidade Normal do período de junho de 2000 a junho de 2001 pelo teste de Kolmogorov-Smirnov, com significância de  $\alpha$  igual a cinco por cento, encontrou-se o desvio máximo  $d = 0,2004$ , que é menor do que o valor crítico tabelado 0,361 para  $n$  igual há treze meses. Daí conclui-se que os dados nesse período se ajustam à distribuição Normal. No gráfico da Figura 2 é possível perceber que em treze meses, no início tem-se uma baixa quantidade de recursos naturais, no meio do período há uma alta quantidade e no fim do intervalo dispõe-se novamente de uma baixa quantidade de recursos para empregar nos planos de atividades agrícolas na fazenda.

**Tabela 1 - Aderência da variável média diária de cada mês à distribuição de probabilidade Normal no período de junho de 2000 a junho de 2012.**

Resultados dos testes de aderência				
Períodos	Média do período	Desvio padrão	Valor de d	p-valor
Junho 2000 a junho 2001	3,02	3,16	0,20	0,46
Junho 2001 a junho 2002	4,44	4,71	0,24	0,36
Junho 2002 a junho 2003	3,60	4,65	0,31	0,26
Junho 2003 a junho 2004	3,97	4,45	0,22	0,40
Junho 2004 a junho 2005	3,40	4,04	0,22	0,42
Junho 2005 a junho 2006	5,41	5,66	0,23	0,40
Junho 2006 a junho 2007	5,34	6,07	0,24	0,38
Junho 2007 a junho 2008	3,58	3,82	0,20	0,46
Junho 2008 a junho 2009	3,96	4,39	0,26	0,36
Junho 2009 a junho 2010	5,00	4,80	0,18	0,5
Junho 2010 a junho 2011	5,02	5,28	0,26	0,36
Junho 2011 a junho 2012	3,29	3,97	0,33	0,24

Fonte: Elaborado pelo próprio autor

Na Tabela 1 os dados estão organizados do seguinte modo: na primeira coluna são apresentados os intervalos de meses que foram calculados a média diária de cada mês e verificada a aderência à distribuição normal. A segunda coluna apresenta a média estimada pelo método do momento para a distribuição normal. A terceira coluna contém o desvio padrão para a distribuição Normal, que foi obtido usando o primeiro e segundo momento, na quarta coluna têm-se os valores de  $d$  da diferença máxima do teste de Kolmogorov-Smirnov e na quinta, e última coluna está o seu p-valor. Como o valor de teste é o mesmo para todos os intervalos, sendo igual a 0,361, percebe-se que para cada intervalo de treze meses as médias diárias de cada período têm aderência a uma distribuição normal com médias e desvios padrão apresentados. Note que apenas o terceiro e o último intervalo têm valor de teste próximo ao máximo para não rejeitar a aderência da distribuição Normal, ou seja, estão próximos do limite para rejeição da hipótese.

Após verificar a aderência dos valores à distribuição Normal, é possível verificar a possibilidade de prever a média pluviométrica diária de cada mês acumulada, considerando apenas a média pluviométrica diária de cada mês acumulada de alguns meses iniciais. Para tal tentativa de previsão, foi usado o Teorema Central do Limite com a equação:

$$Z = \frac{X - \mu}{\sigma}. \quad (5)$$

em que,  $Z$  é o valor da distribuição normalizada,  $X$  é a soma dos  $m$  primeiros meses,  $\mu$  é a média da distribuição Normal e  $\sigma$  é o desvio padrão. Na Tabela 2, observam-se todos os valores da média da distribuição Normal, o seu desvio padrão e a previsão de quantidade de chuva para dez meses iniciais, foi escolhido dez meses, mas poderia ter sido escolhido para qualquer quantidade de meses. Para calcular a média da distribuição Normal  $\mu$ , o desvio padrão  $\sigma$  e a previsão para um período de dez meses foi utilizada a fórmula:

$$x = z \cdot \sigma + \mu. \quad (6)$$

com a qual foi montado um sistema de duas equações em que  $x_1$  é igual à soma das médias diárias dos cinco primeiros meses (de junho a outubro), e  $z_1$  o valor da distribuição normalizada da razão de  $x_1$  pela soma das médias diárias dos treze meses de cada planilha. Sendo  $x_2$  igual à soma das médias diárias dos seis primeiros meses (de junho a novembro), e  $z_2$  o valor da distribuição normalizada da razão de  $x_2$  pela soma das médias diárias dos treze meses de cada planilha, conforme mostra a equação 7:

$$\begin{aligned} z_1\mu + \sigma &= x_1 \\ z_2\mu + \sigma &= x_2 \end{aligned} \quad (7)$$

Para exemplificar, no período de junho de 2000 a junho de 2001 temos que  $x_1 = 5,66$ ,  $x_2 = 15,96$ ,  $z_1 = -1,06$  e  $z_2 = -0,24$ . Substituindo no sistema encontrou-se  $\mu = 18,98$  e  $\sigma = 12,56$ . Em seguida, determinou-se o valor da distribuição normalizada para dez meses que é dado pelo valor da distribuição normalizada da razão de 10 meses por 13 meses, sendo igual a 0,74. Depois usando a equação (6) multiplicou-se 0,74 pelo desvio padrão e somou-se a média a esse resultado. O valor obtido é a previsão da soma dos dez primeiros meses que nesse intervalo é de 28,27 mm. E como a soma dos 10 primeiros meses é de 35,9 mm, encontrou-se uma aproximação neste intervalo de 0,79 que está exposto na primeira linha da tabela 2.

A seguir, a Tabela 2, apresenta todos os resultados obtidos das aproximações das doze planilhas de junho de 2000 a junho de 2001, junho de 2001 a junho de 2002 até junho de 2011 a junho de 2012 .



**Tabela 2 – Previsão pluviométrica do valor médio diário acumulado para os dez primeiros meses de cada intervalo para a variável média mensal**

Período	Média da Distribuição Normal	Desvio Padrão	Aproximação
Junho 2000 a junho 2001	18,98	12,56	0,79
Junho 2001 a junho 2002	25,14	15,71	0,65
Junho 2002 a junho 2003	16,77	9,37	0,55
Junho 2003 a junho 2004	20,03	11,40	0,61
Junho 2004 a junho 2005	15,88	8,79	0,54
Junho 2005 a junho 2006	30,30	16,88	0,66
Junho 2006 a junho 2007	30,86	19,91	0,71
Junho 2007 a junho 2008	18,24	10,70	0,60
Junho 2008 a junho 2009	13,14	6,24	0,39
Junho 2009 a junho 2010	28,80	18,60	0,74
Junho 2010 a junho 2011	30,30	19,95	0,71
Junho 2011 a junho 2012	16,55	9,63	0,60

Fonte: Elaborado pelo próprio autor

Na Tabela 2 os dados estão organizados do seguinte modo: na primeira coluna são apresentados os intervalos de meses que será calculada a média da distribuição Normal, o desvio padrão e calculado uma previsão para os dez primeiros meses. A segunda coluna apresenta a média da distribuição Normal. A terceira coluna contém o desvio padrão, sendo que este, assim como a média, foi calculado usando as equações (6) e (7), por fim, na última coluna têm-se os valores da previsão em porcentagem para os primeiros dez meses, calculados utilizando a equação (6) e o Teorema Central do Limite.

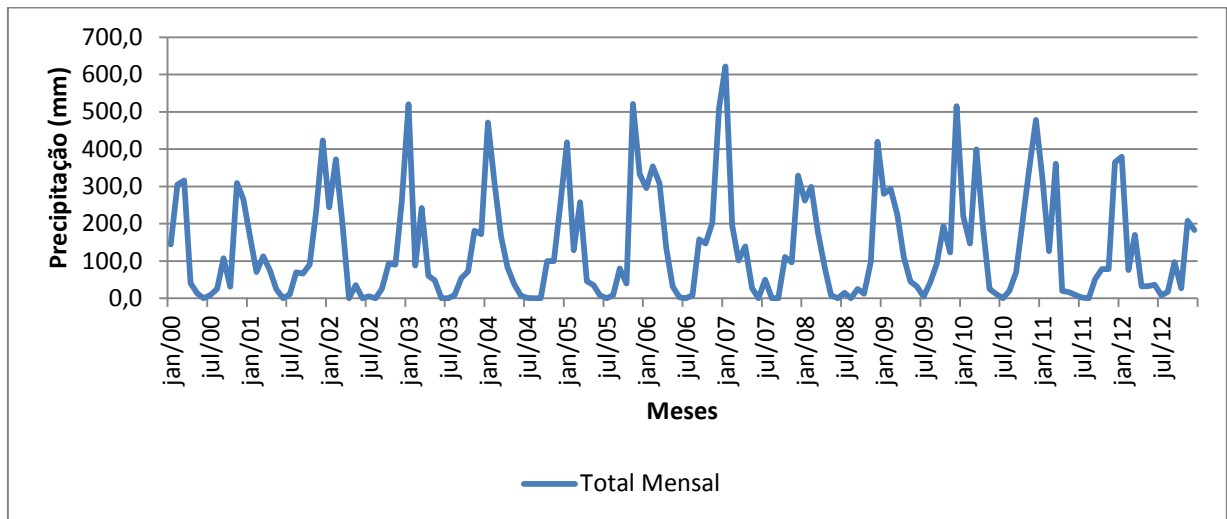
Entretanto, segundo Bernardo (1995) o nível de 75 % a 80 % de probabilidade de ocorrência de chuva é o mais confiável para dimensionamento de projetos agrícolas ou de irrigação, o que mostra que essa previsão não se mostrou muito eficaz como revela a Tabela 2. Concluindo que para o intervalo de treze meses os dados pluviométricos se ajustam a função de distribuição Normal e que ao utilizar o Teorema Central do Limite encontrou-se uma aproximação acumulada para dez meses iniciais. Isso pode ser feito para qualquer quantidade de meses, mas o método se mostra menos eficiente quando é usado um valor muito pequeno de meses para estimar a média e o desvio padrão da distribuição.

#### 4.2. Análise da Série Temporal

Os dados pluviométricos diários apresentados em Oliveira (2013) organizados em quatro planilhas para as variáveis: total mensal, média diária de cada mês, total semanal e média diária de cada semana, foram verificados utilizando o modelo clássico das Séries

Temporais com investigação das Médias Móveis, Tendência Linear, Sazonalidade, Ciclos e Irregularidades. No gráfico apresentado na Figura 3 está representada a variável total mensal do período de janeiro de 2000 a dezembro de 2012 dos dados pluviométricos diários da fazenda Esmeril em Patrocínio-MG.

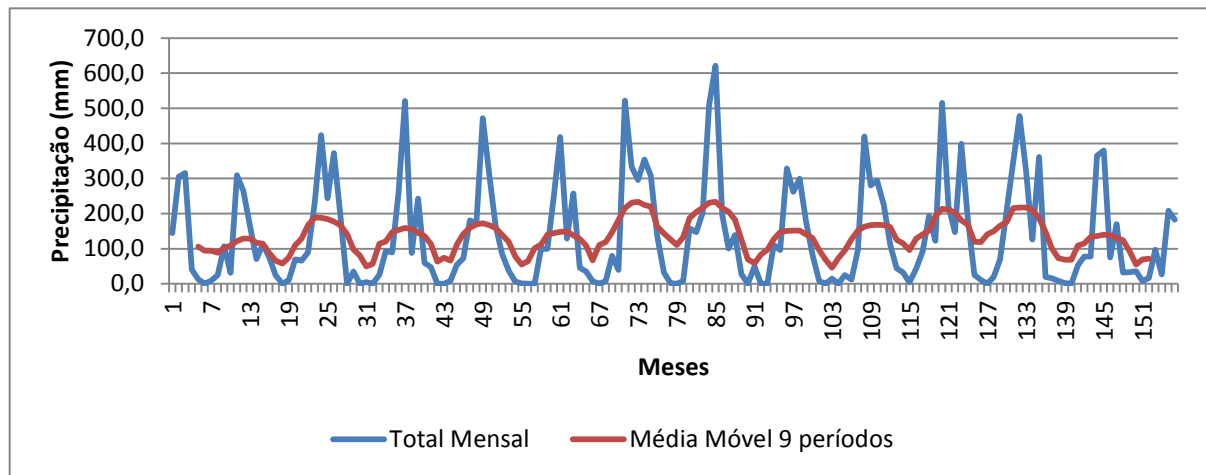
**Figura 3 – Gráfico da variável total mensal no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Sugere-se que esta série de dados pluviométricos neste intervalo não tem tendência, possui sazonalidade, além disso, foi verificado se a série dispõe de ciclos e irregularidades. Para estas averiguações, inicialmente foi feito o cálculo das médias móveis para verificara existência ou não de uma tendência. Para tal se calculou a média móvel de nove períodos, que consistiu em determinar a média dos primeiros nove períodos da série, colocando o resultado exatamente no centro deles. E progressivamente acrescentou-se um período seguinte e se desprezou o primeiro da média imediatamente anterior. Em seguida, calcularam-se as novas médias que foram se movendo até o fim da série, resultando na Figura 4, que contém o gráfico da variável total mensal e a sua tendência usando média móvel. Cabe mencionar que foram feitas suavizações com 5, 7, 9, 11e 13 meses, sendo escolhida a média móvel de 9 períodos para ser apresentada neste trabalho, pois a mesma nos permite fazer uma melhor análise da série de dados.

**Figura 4 – Gráfico da variável total mensal e sua tendência utilizando médias móveis de nove períodos, no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

A Figura 4 mostra que a tendência parece manter-se constante em intervalos de doze meses, o que justifica a escolha do período de treze meses na previsão no Tópico 4.1. Nota-se que a cada período de doze meses o período com maior quantidade de chuva vai de setembro a fevereiro, sendo que o maior volume acontece nos meses de novembro, dezembro e janeiro. Nos meses de março, abril, maio, junho, julho e agosto chove muito pouco, e junho, julho e agosto praticamente não chove nesta localidade. Com essas informações é possível que se faça um planejamento melhor na fazenda com relação à plantação e irrigação. A partir do gráfico da figura 4, as médias móveis também sugerem que não há tendência na série, e para verificar isso foi calculado a tendência linear usando a equação  $T = a + b.t$ , em que  $T$  é o valor da tendência,  $t$  o valor do tempo,  $b$  o coeficiente angular da reta e  $a$  é o coeficiente linear da reta e as equações dos coeficientes denotadas por (3).

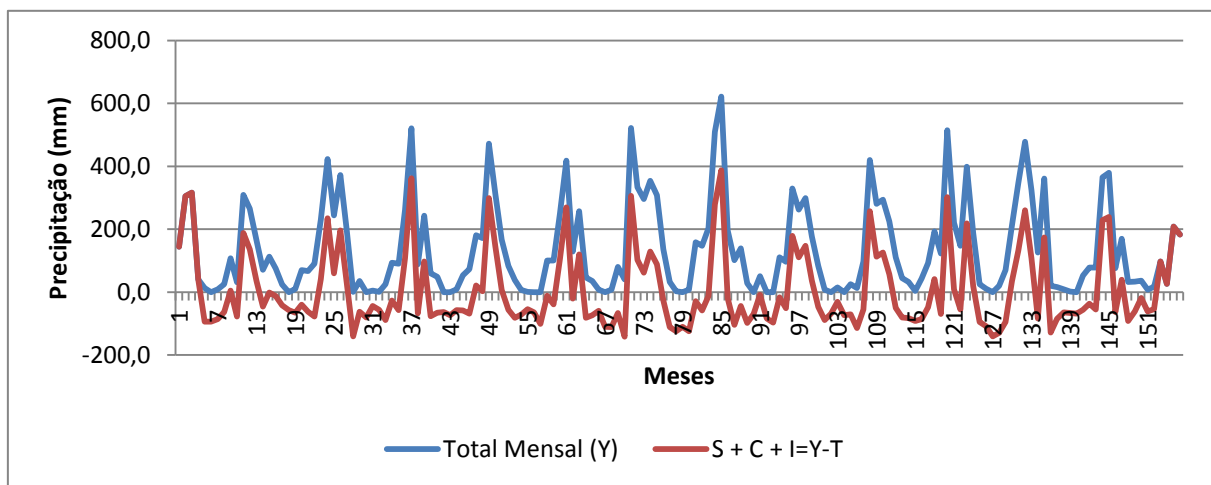
Para a variável total mensal, sendo  $n$  igual a 156 meses equivalente há 13 anos, usando as fórmulas acima, encontrou-se a equação da tendência linear  $T = 138,2 - 0,049.t$ . Esta última equação tem o valor do coeficiente angular igual à  $- 0,049$ , que denota um comportamento ligeiramente decrescente (quase imperceptível) no longo prazo. E como o valor de  $b$  é muito próximo de zero, realça a sugestão inicial de que não haja tendência e que está oscilando em torno de 138,2 mm. Para verificar que não há tendência na série é necessário realizar o teste de hipótese para a inclinação da reta. Para isso, basta obter o valor de:

$$t_{n-2} = \frac{(b-\beta_0)}{S_R S_{yy}} \quad (8)$$

Em que  $b = -0,049$ ,  $\beta_0 = 0$ ,  $S_R = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{n-2}$ ,  $y_i$  é o valor da amostra, e  $\hat{y}_i$  é o valor estimado pela reta obtida. Considerando uma significância de 5%, e o valor calculado de  $t_{n-2} = -0,275$ , cujo p-valor é 0,784, conclui-se a não rejeição da hipótese nula de que a série de dados não tem tendência.

Após a obtenção da tendência é possível analisar o comportamento da série sem a mesma. Como mencionado, este estudo foi feito pelo método aditivo, pois em vários momentos a série possui valor nulo. Para tal foi usado a equação  $Y - T = S + C + I$  em que a tendência foi calculada utilizando médias móveis de nove períodos. Na Figura 5 é exibido o gráfico da variável total mensal juntamente com a total mensal sem a sua tendência.

**Figura 5 – Gráfico da variável total mensal e gráfico da variável total mensal sem a tendência, no período de janeiro de 2000 a dezembro de 2012**

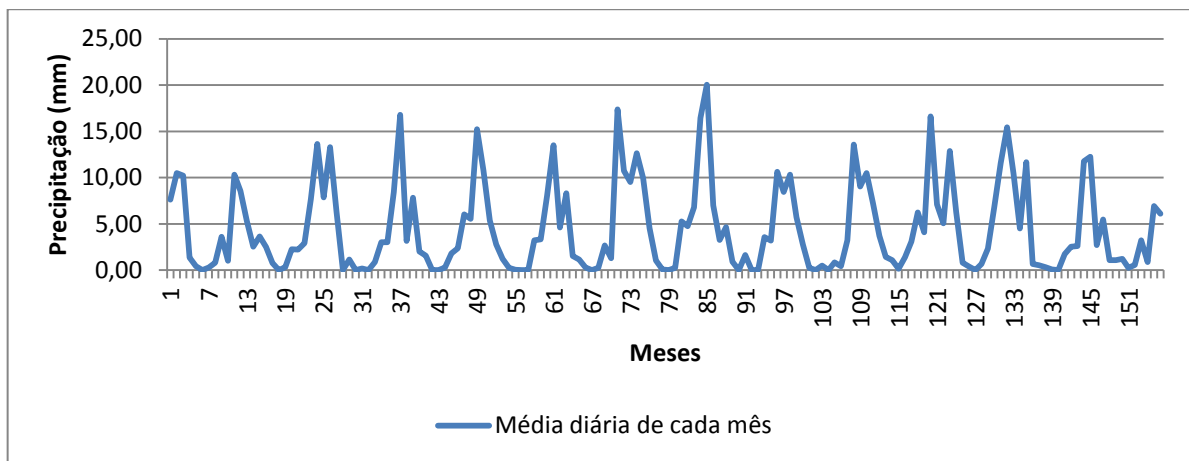


Fonte: Elaborado pelo próprio autor

Ao observar a Figura 5 percebe-se que após a retirada da tendência, o gráfico resultante quando comparado ao da série original foi deslocado para baixo, aparentemente retirando de cada valor da série a constante 138,2 mm, deixando agora a constante da tendência sobre o eixo dos  $x$ , sugerindo novamente que não haja tendência. Mais uma vez percebe-se que para a variável total mensal a cada doze meses têm-se picos e vales periodicamente que representam respectivamente uma quantidade maior de chuva no início e final do ano e falta dela no meio do ano periodicamente. Indicando assim a presença da sazonalidade e ajudando nos planejamentos do trabalho na fazenda e no seu aprimoramento.

Assim como foram feitos os cálculos da tendência para a variável total mensal, estes também foram feitos para as variáveis média diária de cada mês, total semanal e média diária de cada semana. A seguir estão os resultados obtidos para essas variáveis. Sendo que inicialmente, as Figuras 6, 7 e 8 representam os gráficos dos resultados obtidos da variável média diária de cada mês.

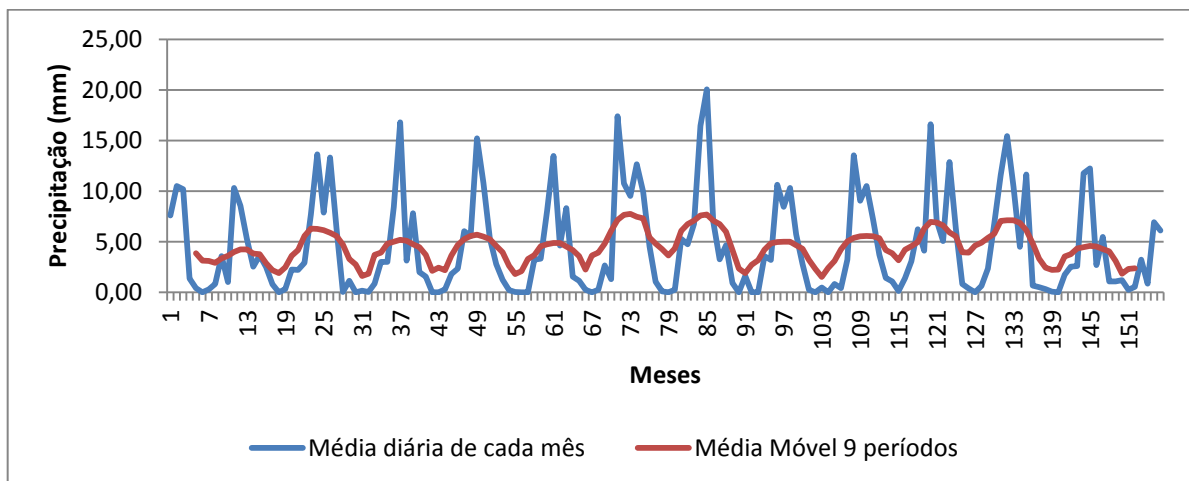
**Figura 6 – Gráfico da variável média diária de cada mês no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Neste gráfico da Figura 6, fica mais claro que não há tendência por se tratar de maneira mais específica os dados pluviométricos com a variável média diária de cada mês e não mais com a variável total mensal.

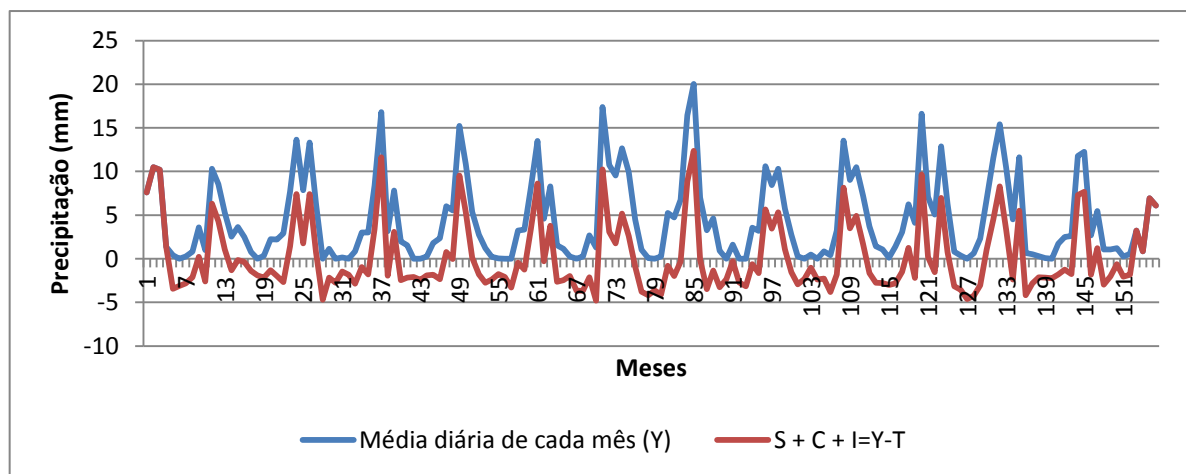
**Figura 7 – Gráfico da variável média diária de cada mês e sua tendência utilizando médias móveis de nove períodos no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Na Figura 7 as médias móveis também sugerem que não há tendência na série, e para averiguar isso foi calculado a tendência linear. Utilizando as expressões denotadas por (3) para os cálculos da tendência linear, obteve-se a seguinte equação da variável média diária de cada mês:  $T = 4,6 - 0,002.t$ . Assim, denota-se novamente que não há tendência e está em torno de 4,6 mm para a variável média diária de cada mês. Esta última equação tem o valor do coeficiente angular igual à  $-0,002$ , que denota, um comportamento ligeiramente decrescente. Fazendo o teste hipótese, de modo análogo ao anterior, temos o valor  $t_{n-2} = 0,235$  com p-valor 0,815 e assim, pelo teste de hipótese a série de dados não possui tendência.

**Figura 8 – Gráfico da variável média diária de cada mês e da variável média mensal sem a sua tendência no período de janeiro de 2000 a dezembro de 2012**

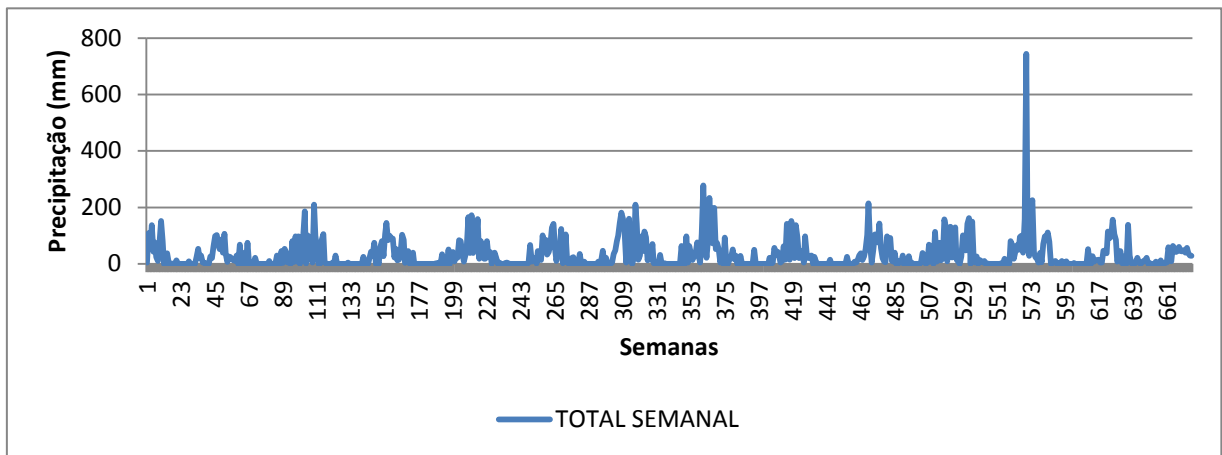


Fonte: Elaborado pelo próprio autor

Ao observar as Figuras 6, 7 e 8 percebe-se que a variável média diária de cada mês não tem tendência, com valor aparentemente constante de 4,6 mm aproximadamente. Percebeu-se também que após a retirada da tendência, o gráfico resultante quando comparado ao da série original foi deslocado para baixo, ao se retirar de cada valor da série a constante 4,6 mm. Deixando agora a constante da tendência sobre o eixo  $x$ , sugerindo mais uma vez que não há tendência. E que, em intervalos de doze meses, tem-se novamente períodos chuvosos e de escassez de chuva, mostrando a presença da sazonalidade.

Nas Figuras 9, 10 e 11 estão expostos os resultados encontrados para a variável total semanal em relação à tendência. A figura nove apresenta a variável total semanal.

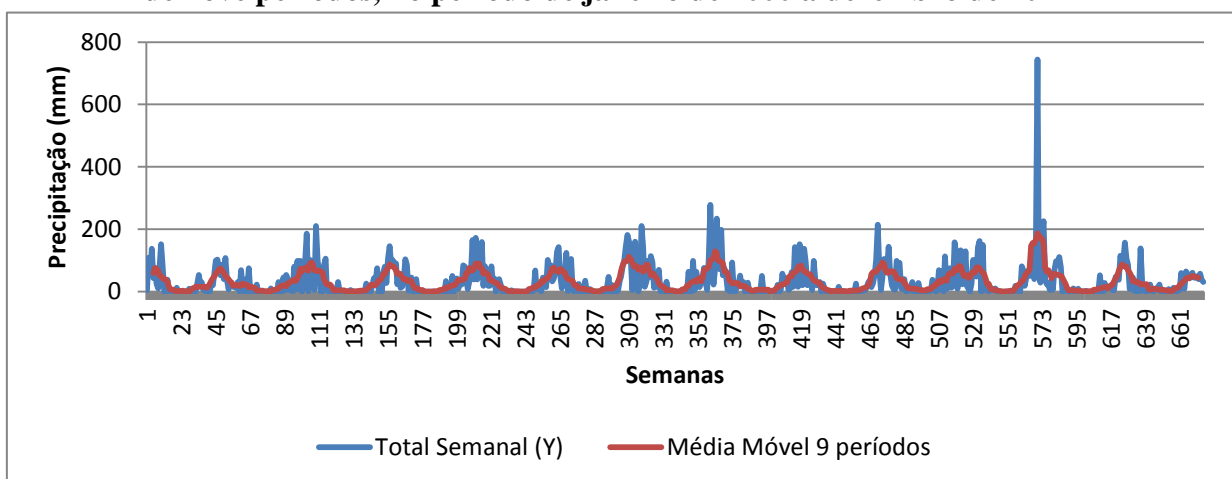
**Figura 9 – Gráfico da variável total semanal no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Observe que na semana 571 há um pico no volume pluviométrico. Este valor pode influenciar toda a análise por estar muito fora de todo o padrão da série. Porém este não será desconsiderado da análise, pois não há como precisar se foi um fator isolado.

**Figura 10 – Gráfico da variável total semanal e sua tendência utilizando médias móveis de nove períodos, no período de janeiro de 2000 a dezembro de 2012**



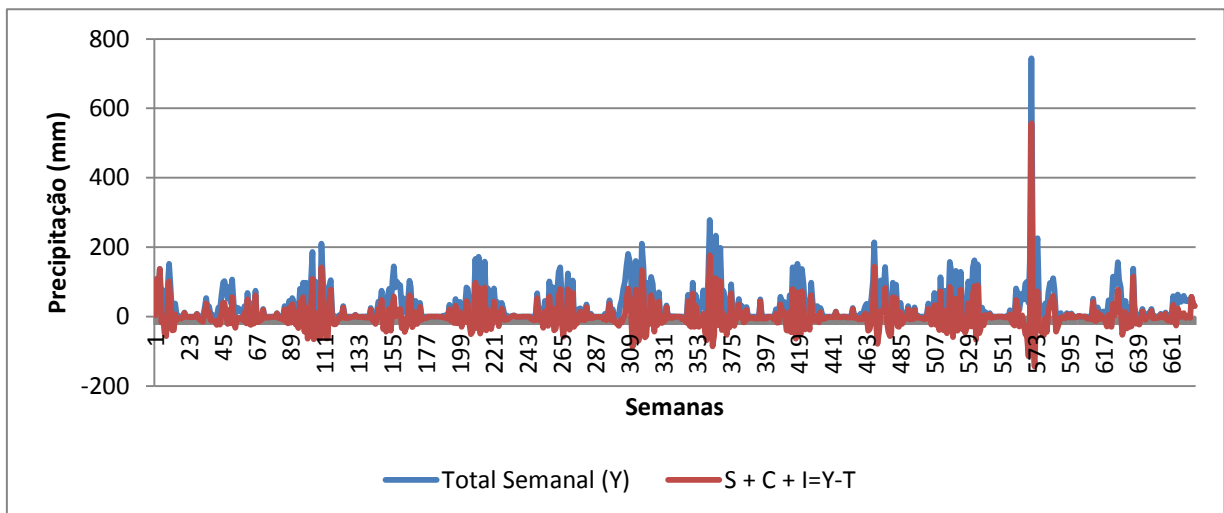
Fonte: Elaborado pelo próprio autor

Novamente é possível perceber que não existe tendência na série, que fica bem mais evidente com a média móvel da variável total semanal. Para verificar isso usando as expressões denotadas por (3), foi obtida a equação da tendência linear para a variável total semanal  $T = 34,5 - 0,006 \cdot t$ . Assim, percebe-se que não há tendência e está em torno de 34,5 mm. Esta última equação tem o valor do coeficiente angular igual à  $-0,006$ , que denota um

comportamento ligeiramente decrescente. E como o valor de  $b$  é muito próximo de zero, realça a sugestão inicial de que a tendência seja constante e está oscilando em torno de 34,5 mm. Fazendo o teste hipótese, de modo análogo ao anterior, temos o valor  $t_{n-2} = 1,177$  com p-valor 0,239 e assim, pelo teste de hipótese não é possível excluir a existência da tendência da série.

Do mesmo modo que foi analisada a série para as variáveis total mensal e média diária de cada mês sem a tendência, pode-se também retirá-la da série para verificar o comportamento da variável total semanal sem a influência de sua tendência, como mostra a Figura 11.

**Figura 11 – Gráfico da variável total semanal e da variável total semanal sem a sua tendência, no período de janeiro de 2000 a dezembro de 2012**



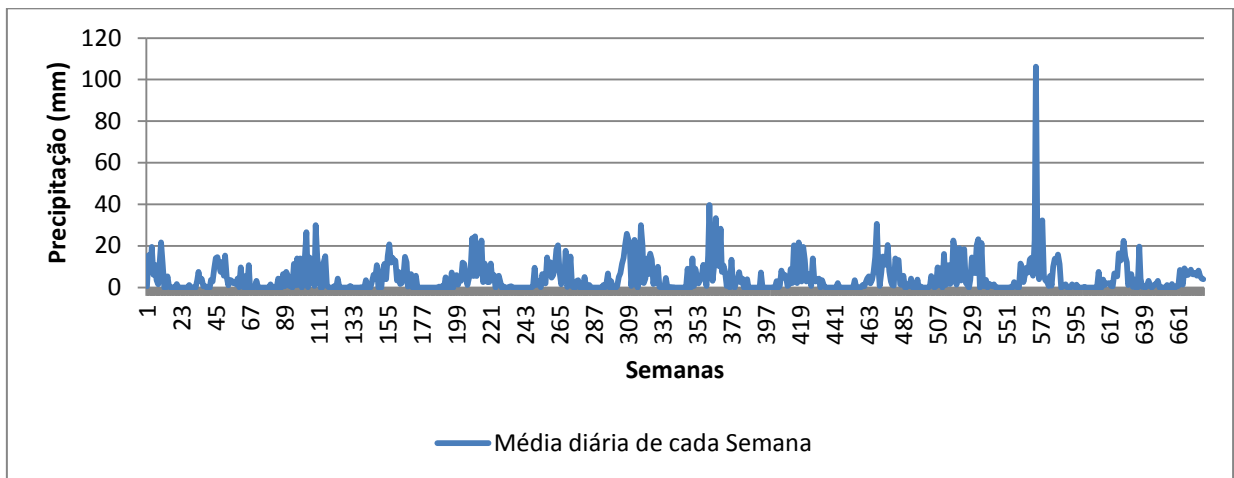
Fonte: Elaborado pelo próprio autor

Percebe-se que após a retirada da tendência, o gráfico resultante quando comparado ao da série original foi deslocado para baixo, aparentemente retirando de cada valor da série a constante 34,5 mm. Deixando agora a constante da tendência sobre o eixo dos x, sugerindo novamente que a tendência é constante. Outra vez nota-se a presença da sazonalidade em virtude dos períodos regulares de chuva e escassez dela dentro do período de cinquenta e duas semanas, equivalente há um ano.

Agora, nas Figuras 12, 13 e 14 estão representados os resultados alcançados da variável média diária de cada semana. A Figura 12 representa os dados da variável média diária de cada semana no intervalo considerado.



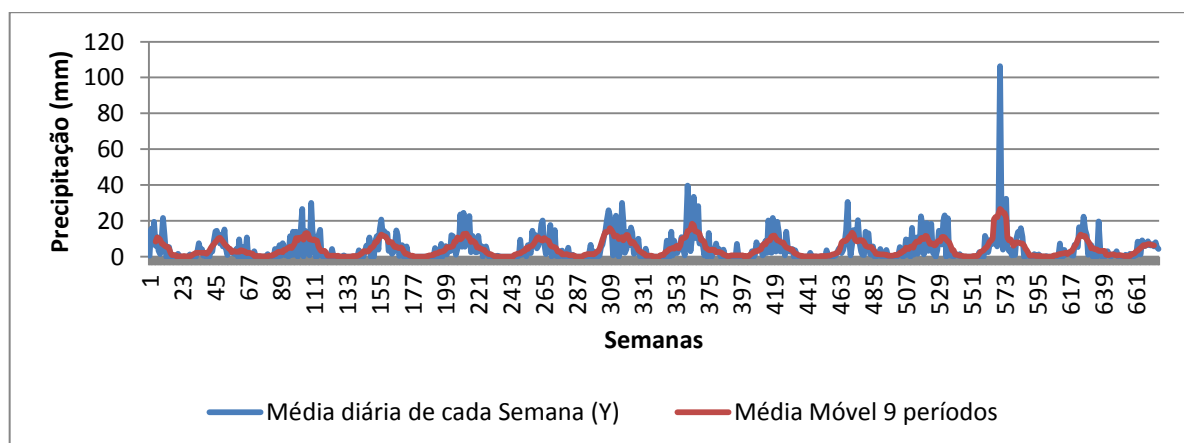
**Figura 12 – Gráfico da variável média diária de cada semana no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Na Figura 12 está representada a variável média diária de cada semana e pode-se fazer uma análise mais detalhada dos dados pluviométricos considerados. Novamente, é possível observar que da mesma maneira que na variável total semanal, na semana 571 há um pico no volume pluviométrico. E que este valor pode influenciar toda a análise por estar muito fora de todo o padrão da série, mas também este não será desconsiderado da análise.

**Figura 13 – Gráfico da variável média diária de cada semana e sua tendência utilizando médias móveis de nove períodos, no período de janeiro de 2000 a dezembro de 2012**



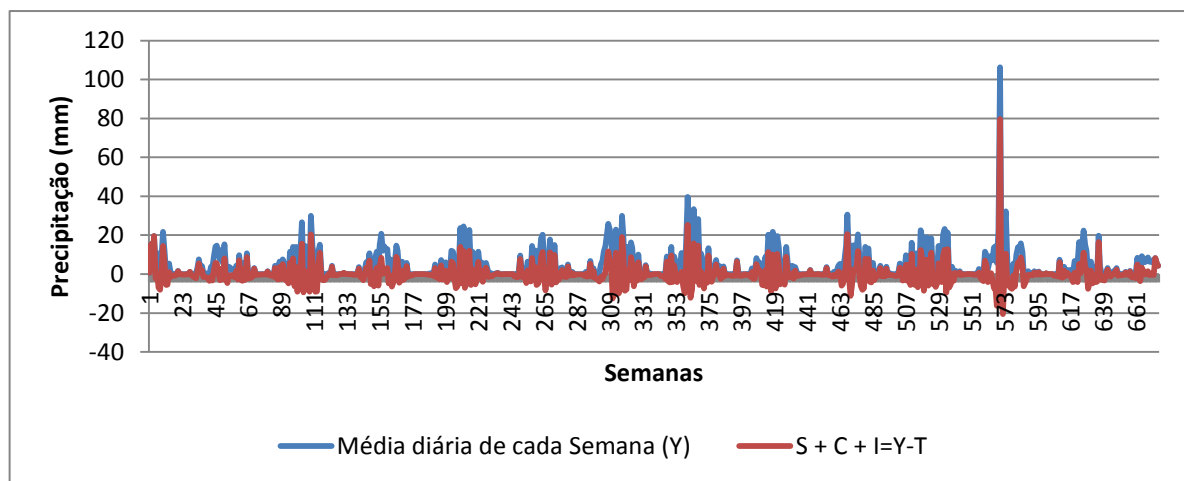
Fonte: Elaborado pelo próprio autor

É possível verificar que não existe uma tendência na série, isso fica bastante evidenciado com o gráfico da variável média diária de cada semana. Empregando as expressões denotadas por (3) para obter a tendência linear, foi obtida a equação da tendência

para a variável média diária de cada semana  $T = 4,9 - 0,001.t$ . Esta última equação tem o valor do coeficiente angular igual à  $-0,001$ , que indica um comportamento ligeiramente decrescente. E como o valor de  $b$  é muito próximo de zero, torna claro a sugestão inicial de que não há tendência e que esta está oscilando em torno de  $4,9$  mm. Fazendo o teste hipótese, de modo análogo aos anteriores, temos o valor  $t_{n-2} = 1,177$  com p-valor  $0,239$  e assim, pelo teste de hipótese a série de dados possui tendência, mesmo que mínima.

Da mesma maneira, na Figura 14 está representado a retirada da tendência da série para verificar o comportamento desta sem a influência da tendência.

**Figura 14 – Gráfico da variável média diária de cada semana e gráfico da variável média semanal sem a sua Tendência, no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

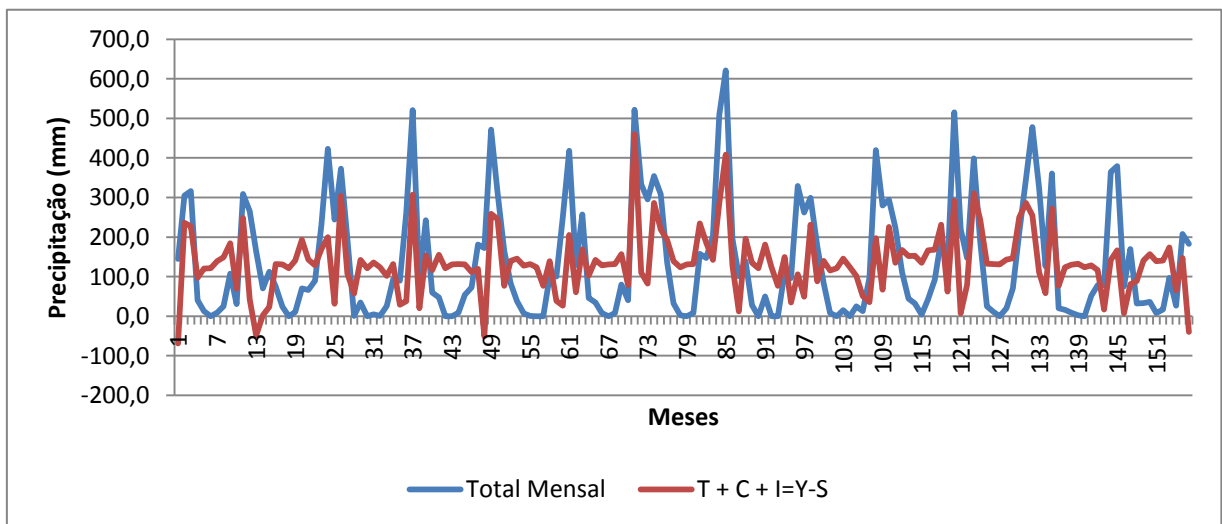
É possível observar novamente que depois da remoção da tendência, o gráfico obtido quando comparado ao da série original foi deslocado para baixo, retirando de cada valor da série a constante  $4,9$  mm. Deixando agora a constante da tendência sobre o eixo  $x$ , indicando que não tem tendência. Assim, tem-se que na série não possui tendência e que em intervalos de cinquenta e duas semanas há picos que representam os períodos chuvosos e vales que mostram escassez de recursos pluviométricos. Sendo que nos meses de novembro, dezembro e janeiro chovem mais, e junho, julho e agosto quase não chovem. Com os gráficos de médias de cada semana é possível fazer um bom planejamento agrícola e de irrigação nesta fazenda.

Indica-se então que há sazonalidade na série, pois em todos os gráficos (total mensal e semanal, média diária de cada mês e semana) é possível perceber que há um comportamento que se repete periodicamente a cada intervalo de um ano (o que justifica também a separação dos treze meses do método de previsão do tópico 4.1). Para verificar a influência desta

sazonalidade será utilizado o cálculo de médias móveis de doze períodos (para variável total mensal e média diária de cada mês) e de cinquenta e dois períodos, equivalente há um ano (para variável total semanal e média diária de cada semana).

Foram calculados os índices sazonais utilizando o cálculo das médias móveis de doze e cinquenta e dois períodos respectivamente no modelo aditivo, e fez-se a retirada da sazonalidade usando a equação  $Y - S = T + C + I$  e médias móveis, obtendo os gráficos 15, 16, 17 e 18.

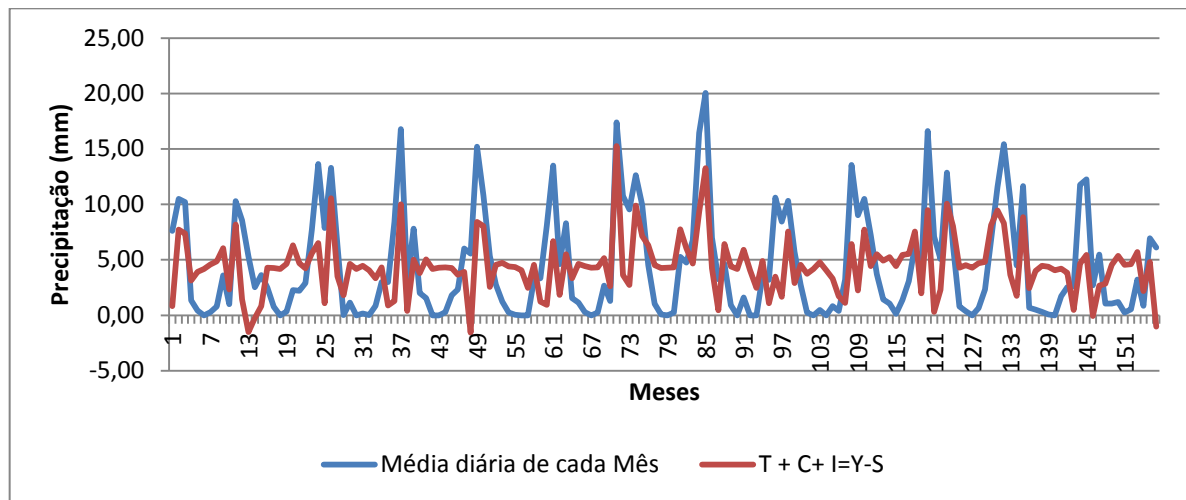
**Gráfico 15 – Gráfico da variável total mensal sem a sua sazonalidade no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Após a retirada da sazonalidade, nota-se que os valores pluviométricos se desenvolvem em torno de um valor que é igual a 138,2, conforme encontrado na equação de tendência. Os pontos em vermelho abaixo de 138,2 mm são períodos em que o volume pluviométrico fica abaixo da média do total mensal, e os pontos muito acima são momentos em que a chuva fica acima da média. Além disso, há intervalos de tempo não fixos em que existem pequenas oscilações próximas da média.

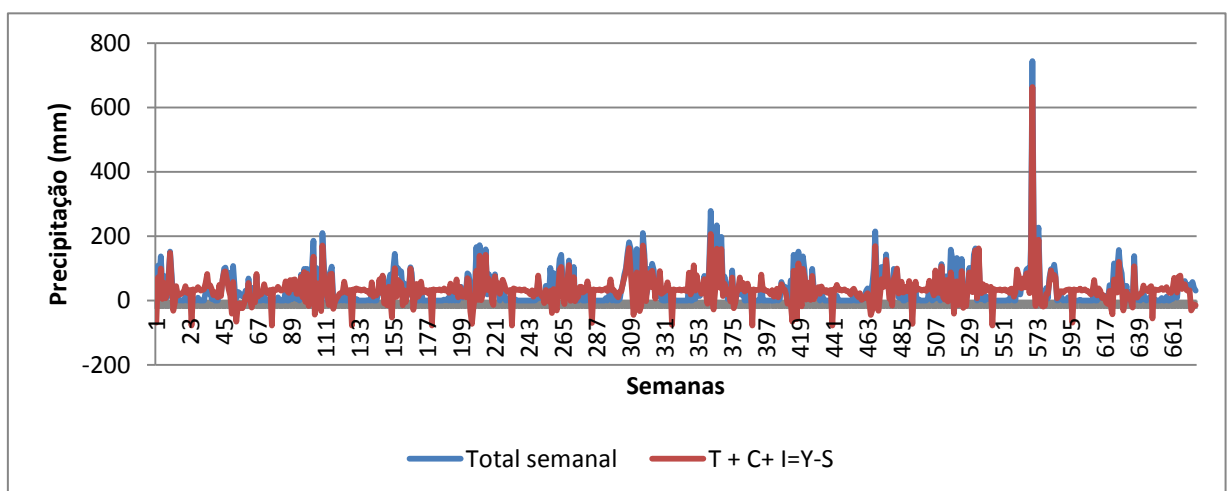
**Figura 16 – Gráfico da variável média diária de cada mês sem a sua sazonalidade no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Removida a sazonalidade, observa-se com mais evidência que os valores pluviométricos se desenvolvem em torno de um valor que é igual a 4,6, conforme encontrado na equação de tendência. Os pontos em vermelho abaixo de 4,6 mm são períodos em que o volume pluviométrico fica abaixo da média, os pontos muito acima são momentos em que a chuva fica acima da média. Como na variável total mensal, há intervalos de tempo não fixos que possuem pequenas oscilações próximas da média.

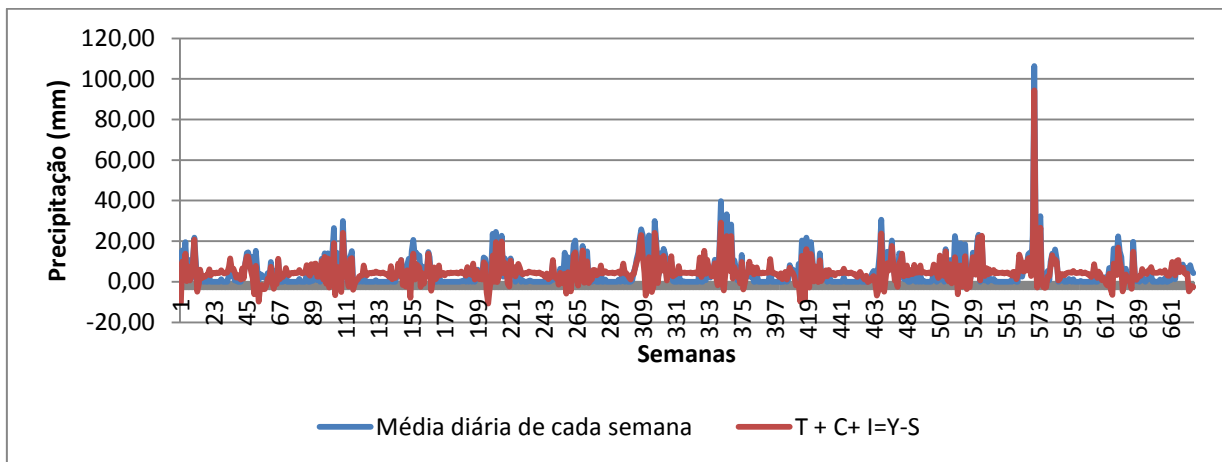
**Figura 17 – Gráfico da variável total semanal sem a sua sazonalidade no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Na Figura 17, é mais perceptível que após a retirada da sazonalidade os valores pluviométricos se desenvolvem em torno de um valor que é igual a 35,4, conforme encontrado na equação de tendência.

**Figura 18 – Gráfico da variável média diária de cada semana sem a sua sazonalidade no período de janeiro de 2000 a dezembro de 2012**

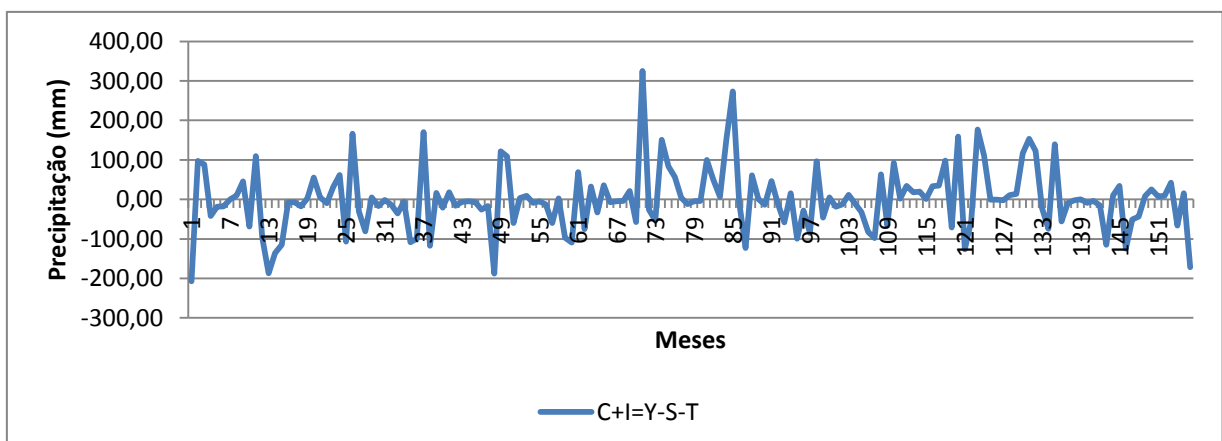


Fonte: Elaborado pelo próprio autor

Após a retirada da sazonalidade, ficou claro com a variável média diária de cada semana, que os valores pluviométricos se desenvolvem em torno de um valor constante e igual a 4,9, conforme encontrado na equação de tendência. Há intervalos de tempo não fixos em que há pequenas oscilações próximas da média.

Para a obtenção das variações cíclicas e irregulares foi feita a retirada da tendência e da sazonalidade. Usando a equação  $C + I = Y - T - S$  e médias móveis obteve-se os gráficos da figura 19, 20, 21 e 22.

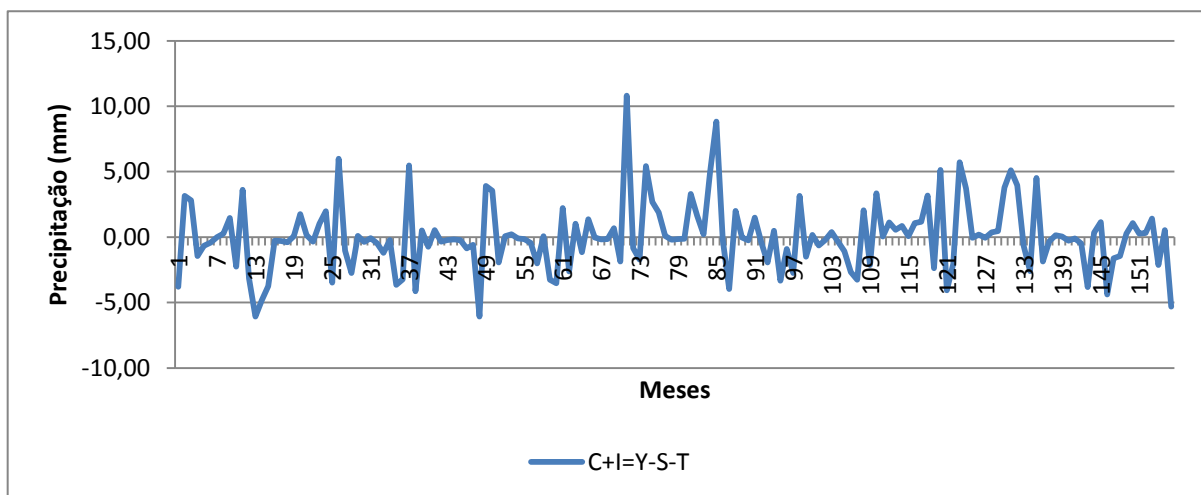
**Figura 19 - Gráfico das variações cíclicas e irregulares da variável total mensal no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Observando o gráfico 19 não foi possível perceber visualmente a parte cíclica, pois não se verificou comportamentos que se repetem periodicamente com prazos superiores há um ano. Essas variações cíclicas, quando existem são perceptíveis apenas em séries mais longas. Como a média está centrada em zero percebe-se que há meses em que as amplitudes estão bem acima e meses que estão bem abaixo da média, denotando a presença de variações irregulares. Mas não se conseguiu explicar a parte irregular na série, visto que esta parte é aleatória. Assim, podemos obter o erro de aproximação do valor obtido pela regressão da série. O Erro Quadrático Médio do estimador é dado pela variância da diferença entre o valor real e o valor obtido pela tendência, uma vez que o estimador é não tendencioso. Neste caso o  $EQM = 6393,8$ .

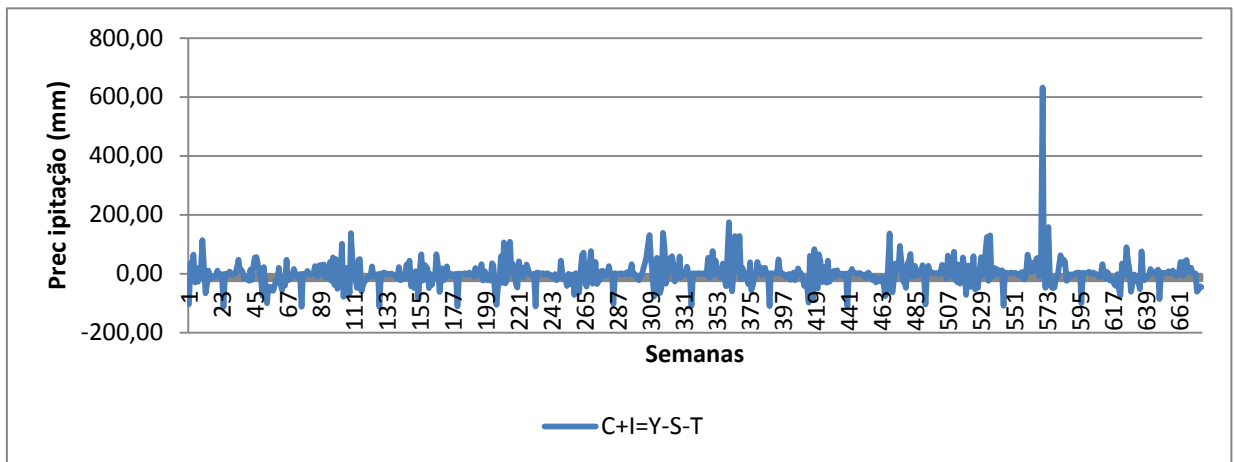
**Figura 20 - Gráfico das Variações cíclicas e irregulares da variável média diária de cada mês no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Na figura 20 percebe-se que a série de dados pluviométricos possui variações irregulares e não tem variações cíclicas. Fazendo o cálculo do Erro Médio Quadrático, de modo análogo ao anterior, temos que o  $EQM = 6,77$ .

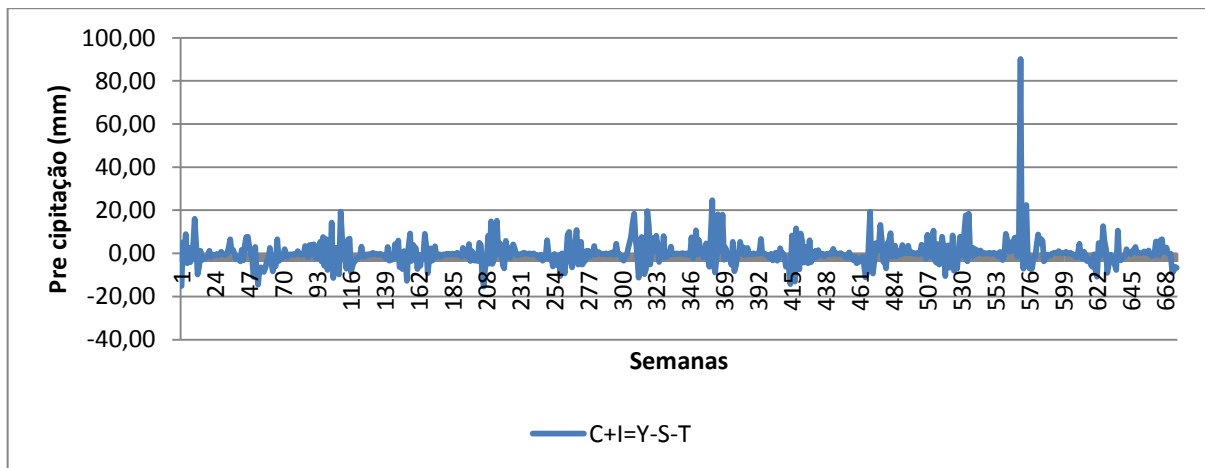
**Figura 21 - Gráfico das variações cíclicas e irregulares da variável total semanal no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Não se verificou comportamentos que se repetem periodicamente com prazos superiores há um ano na Figura 21. Novamente não foi possível perceber a parte cíclica. Fazendo o cálculo do Erro Médio Quadrático, de modo análogo ao feito anteriormente, temos que o  $EQM = 2066,5$ .

**Figura 22 - Gráfico das variações cíclicas e irregulares da variável média diária de cada semana no período de janeiro de 2000 a dezembro de 2012**



Fonte: Elaborado pelo próprio autor

Na figura 22 da variável média semanal fica ainda mais evidente que a série de dados pluviométricos analisada não é cíclica e que possui variações irregulares. Fazendo o cálculo do Erro Médio Quadrático, de modo análogo, temos que o  $EQM = 37,6$ .

Assim nos gráficos das Figuras de 3 a 22 é percebe-se que a série de dados em estudo não possui tendência, há intervalos que se repetem periodicamente em períodos de doze meses, ou seja, a série tem sazonalidade e tem variações irregulares.

## **5. CONCLUSÃO E PROPOSTAS FUTURAS**

Este trabalho teve como objetivo mostrar se a variável média diária de cada mês dos índices pluviométricos diários da fazenda Esmeril em Patrocínio-MG, no período de 2000 a 2012, se ajustaria a função de distribuição Normal, prever a quantidade de chuva para determinados meses utilizando a distribuição Normal, e fazer um estudo do comportamento desta série temporal em análise. A série de dados de precipitações que foi organizada em planilhas de junho de 2000 a junho de 2001, junho de 2001 a junho de 2002, assim sucessivamente até junho de 2011 a junho de 2012, utilizando o teste de aderência de Kolmogorov-Smirnov com nível de significância de 5%, se ajustaram a função de distribuição Normal.

Após verificar a aderência dos valores a uma distribuição Normal, foi feita uma previsão da quantidade de chuva usando o Teorema Central do Limite para dez meses iniciais de cada intervalo, podendo ser feito para qualquer quantidade de meses. Entretanto, o método se mostrou menos eficiente quando é usado um valor muito pequeno de meses para estimar a média, o desvio padrão da distribuição e conseqüentemente sua previsão como mostrou os resultados obtidos.

No estudo da série temporal em análise foi observado que os dados pluviométricos do período de 2000 a 2012 para as variáveis total mensal, média diária de cada mês, total semanal e média diária de cada semana não têm tendência, possui sazonalidade e apresentam variações irregulares.

Propõe-se para trabalhos futuros fazer, fazer um estudo da série comparando apenas os meses de cada ano para verificar o comportamento dos dados ano a ano para cada mês, explicar no que isso pode influenciar na previsão da quantidade pluviométrica, usar redes neurais para descrever a série e verificar o uso das mesmas para previsão dos dados.



## 6. REFERÊNCIAS

- ALMEIDA, R. M. B. **Características climatológicas do regime de chuva em Minas Gerais**. 1995, 64p. Dissertação (Mestrado em Engenharia Agrícola)- Curso de Pós-graduação em Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa. 1995.
- ARAÚJO, W. F., et al. Precipitação pluviométrica mensal provável em Boa Vista-RR. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.5, n.3, Set/Dez 2001. Disponível em: <<http://www.scielo.br>>. Acesso em: 16 Nov. 2015
- BACCI, D. DE LA C.; PATACA, E. M. Educação para a água. **Estudos Avançados**, São Paulo, v. 22, n. 63, p. 211-226, jan. 2008. Disponível em: <<http://www.revistas.usp.br>>. Acesso em: 07 Out. 2015.
- BERNARDO, S. **Manual de irrigação**. 6ª ed. Viçosa: Imprensa Universitária, 1995. 657p.
- CATALUNHA, M. J., et al. Avaliação de cinco funções densidade de probabilidade a séries de precipitação pluvial no estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v.10, n.1, p.153-162,2002.
- COSTA NETO, P. L. de O. **Estatística**. 3ª edição. São Paulo: Blucher, 2002.
- DAMÉ, R. DE C., et al. Análise de frequência hidrológica dos dados de precipitação pluvial de algumas estações agroclimatológicas da região sul do Rio Grande do Sul. **Ciência Rural**, Santa Maria, v.26, n.3, p.351-355, 1996.
- DINIZ, J. M. T. Variabilidade da precipitação e do número de dias com chuvas de duas cidades distintas da Paraíba. **HOLOS**, v. 3, p. 171-180, ago. 2013. Disponível em: <<http://www2.ifrn.edu.br>>. Acesso em: 07 out. 2015.
- MENDES, P. C. **A gênese espacial das chuvas na cidade de Uberlândia-MG**. 2001, 237p., Dissertação (Mestrado em Geografia) - Universidade Federal de Uberlândia, Uberlândia. 2001.
- MEYER, P. L. **Probabilidade: aplicações à estatística**. 2ª edição. Rio de Janeiro: LTC, 2012.
- MOREIRA, P. S. P., et al. Distribuição e probabilidade de ocorrência de chuvas no município de Nova Maringá-MT. **Revista de Ciências Agroambientais**, Alta Floresta, v.8, n.1, p.9-20, 2010.
- MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. 5ª edição, São Paulo: Saraiva, 2003.
- NETO, D. D., et al. Ajuste de modelos de distribuição de probabilidade a séries históricas de precipitação pluvial diária em Piracicaba-SP. **Revista Brasileira de Agrometeorologia**, v.13, n.2, p. 273-283, 2005.
- OLIVEIRA, A. de. **Ajuste de distribuição de probabilidade a dados de precipitações pluviométricas mensais na Fazenda Esmeril – Patrocínio - MG**. 2013, 65 p. Dissertação

(Mestrado profissional em matemática em rede nacional - PROFMAT)-Universidade Federal do Triângulo Mineiro, Uberaba. 2013.

REIS, M. M. **Análise de séries temporais**. Disponível em: <<http://www.inf.ufsc.br/~marcelo/cap.4.pdf>>. Acesso em: 04 jan. 2016.

RIBEIRO, A. M.; LUNARDI, D. M. C. A precipitação mensal provável para Londrina-PR, através da função Gama. **Energia na Agricultura**, Botucatu, v.12, n.4, p.37-44, 1997.

SAMPAIO, S. C., et al. Estudo da precipitação efetiva para o município de Lavras, MG. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.4, n.2, p.210-213, 2000.

SILVA, Í. N., et al. Análise de distribuição de frequência da precipitação mensal e anual para a cidade de Iguatu-CE. In: CONGRESSO BRASILEIRO de AGROMETEOROLOGIA, XVI, 2009, Belo Horizonte. **Anais...** Belo Horizonte, 2009.

SOCCOL, O. J.; CARDOSO, C. O.; MIQUELUTTI, D. J. Análise da precipitação mensal provável para o município de Lages-SC. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.14, n.6, p.569-574, 2010.