



**UNIVERSIDADE DO ESTADO DE MATO GROSSO  
CAMPUS DE SINOP  
FACULDADE DE CIÊNCIAS EXATAS E TECNOLÓGICAS  
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE  
NACIONAL PROFMAT**



**ALESSANDRO DOS SANTOS GOES**

**COMPARAÇÃO ENTRE MÉTODOS DE  
APROXIMAÇÃO/INTERPOLAÇÃO POLINOMIAL: UMA APLICAÇÃO  
AOS HISTOGRAMAS DE IMAGENS DIGITAIS.**

Sinop

2017



**ALESSANDRO DOS SANTOS GOES**

**COMPARAÇÃO ENTRE MÉTODOS DE  
APROXIMAÇÃO/INTERPOLAÇÃO POLINOMIAL: UMA APLICAÇÃO  
AOS HISTOGRAMAS DE IMAGENS DIGITAIS.**

Dissertação apresentada ao Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, da Universidade do Estado de Mato Grosso - UNEMAT como requisito parcial para obtenção do grau de Mestre em Matemática.

Prof. Dr. Giovane Maia do Vale  
Orientador

Prof. Me. João Gabriel Ribeiro  
Co-Orientador

Sinop

2017

## CIP – CATALOGAÇÃO NA PUBLICAÇÃO

G598c Goes, Alessandro dos Santos.

Comparação entre métodos de aproximação/interpolação polinomial:  
uma aplicação aos histogramas de imagens digitais / Alessandro dos Santos  
Goes. – Sinop, 2017.

106 p.

Orientador: Dr. Giovane Maia do Vale.

Co-orientador: Me. João Gabriel Ribeiro.

Dissertação (Mestrado) – Universidade do Estado de Mato Grosso,  
*Campus* Universitário de Sinop, Faculdade de Ciências Exatas e  
Tecnológicas, Programa de Pós-graduação Profissional em Matemática.

1. Histogramas. 2. Funções Polinomiais. 3. Métodos de Aproximação.  
4. Mestrado Profissional em Matemática. I. Vale, G. M. do, Dr. II. Ribeiro,  
J. G., Me. III. Título. IV. Título: uma aplicação aos histogramas de imagens  
digitais.

CDU 311.218:512.62



ESTADO DE MATO GROSSO  
SECRETARIA DE ESTADO DE CIÊNCIA E TECNOLOGIA  
UNIVERSIDADE DO ESTADO DE MATO GROSSO  
CAMPUS UNIVERSITÁRIO DE SINOP  
FACET – FACULDADE DE CIÊNCIAS EXATAS E TECNOLÓGICAS.  
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL  
PROFMAT UNEMAT - SINOP



**ALESSANDRO DOS SANTOS GOES**

COMPARAÇÃO ENTRE MÉTODOS DE APROXIMAÇÃO/INTERPOLAÇÃO  
POLINOMIAL: UMA APLICAÇÃO AOS HISTOGRAMAS DE IMAGENS  
DIGITAIS

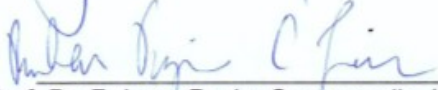
Dissertação apresentada ao Programa de Mestrado Profissional em Matemática em Rede Nacional da Universidade do Estado de Mato Grosso - UNEMAT no *Campus* Universitário de Sinop, para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Giovane Maia do Vale

Aprovado em: 08/06/2017

BANCA EXAMINADORA:

  
\_\_\_\_\_  
Prof. Dr. Giovane Maia do Vale - UNEMAT

  
\_\_\_\_\_  
Prof. Dr. Rubens Pazim Carnevarollo Júnior- UFMT

  
\_\_\_\_\_  
Prof. Dr. Oscar Antonio González Chong - UNEMAT

SINOP – JUNHO - 2017



Programa de Mestrado Profissionalizante em Matemática em Rede Nacional  
UNEMAT- Sinop Avenida dos Ingás, nº 3001 - Centro – CEP: 78.555-000 – Sinop-  
MT.Tel./Fax: (66)9601-8925 – Cx. Postal: 680 – profmat-unemat@unemat-net.br



Ativar o  
Acesse Co

*Dedico este trabalho à minha esposa Ludmilla e ao meu filho Davi. Ambos foram minha fonte de inspiração e força para que ele se concretizasse.*

## **AGRADECIMENTOS**

À Deus, que nos dá a vida e a ânsia por sempre aprender mais.

À minha esposa Ludmilla, pela paciência nos momentos que estive ausente e ao meu filho Davi, que cresceu juntamente com essa conquista.

Aos meus familiares, pelo apoio incondicional.

Ao professor Ricardo Andrade, parceiro e amigo de quase 30 mil Km de estrada, que foram rodados para que chegássemos até aqui.

À UNEMAT, por oportunizar esse Mestrado nesta região, de modo que eu pude ser beneficiado.

À SBM, CAPES e IMPA pela coordenação do PROFMAT.

À todos os professores envolvidos com o PROFMAT no Campus da UNEMAT em Sinop, pela dedicação e ensino.

Aos professores Giovane Maia do Vale e João Gabriel Ribeiro, pelo apoio e orientação.

Aos amigos e companheiros de curso. Juntos, lutamos e vencemos.





*“Não há ramo da Matemática, por mais abstrato que seja,  
que não possa um dia vir a ser aplicado  
aos fenômenos do mundo real.”*

*Nikolai Lobachevsky*



## RESUMO

Uma das principais e mais difíceis tarefas do professor é tornar palpável para o aluno o conhecimento ensinado. Essa tarefa pode ser ainda mais complexa quando os conteúdos são matemáticos. Assim, geralmente o professor apresenta, por exemplo, certas formulações funcionais sem mostrar sua dedução ou explicar os métodos que geraram tal conteúdo. Portanto, no sentido de, mesmo que pontualmente, se tentar sanar tais vacuidades, tem-se a pesquisa aqui descrita. Nestes termos, este trabalho teve como objetivo o estudo e a aplicação de métodos de aproximação/interpolação aos histogramas, obtidos a partir de imagens digitais, a fim de obter, mesmo que aproximadamente, uma forma funcional polinomial que os descrevessem. Ou seja, os histogramas foram utilizados como funções de domínio discreto a partir das quais se gerou formulações funcionais polinomiais que, considerando um domínio contínuo, descrevessem as oscilações de frequência, mediante as variações das intensidades de brilho. Para tanto, foram utilizados como dados de entrada os histogramas das componentes vermelha, verde e azul de uma imagem digital colorida de entrada, bem como, o histograma relativo à imagem de luminância inerente à imagem colorida. Os métodos de aproximação/interpolação utilizados foram o Método dos Mínimos Quadrados Ordinários (MMQ), o método de aproximação *B-Splines* acoplado ao MMQ, o método da Máxima Verossimilhança (MV), e, por fim, o método de aproximação *B-Splines* acoplado ao método MV. Os resultados gerados foram avaliados visualmente, com o auxílio de gráficos, e numericamente, por meio de parâmetros estatísticos estabelecidos com base nas discrepâncias (erros) entre os valores originais dos histogramas e os valores calculados por meio das formas polinomiais geradas. De modo geral, os polinômios gerados foram eficazes em modelar os dados de entrada. Porém, os resultados advindos do método MV foram os que apresentaram menor erro. O método de aproximação *B-Splines* gerou um acréscimo de qualidade aos resultados somente quando acoplado ao MMQ.

**Palavras-chave:** Histogramas; Funções Polinomiais; Métodos de Aproximação e Interpolação.



## ABSTRACT

One of the main and most difficult teacher tasks is to make taught knowledge palpable to student. This task can be even more complex when contents are mathematical. Thus, teacher usually presents, for example, some functional formulations without showing their deduction or explaining methods that generated such content. Therefore, even if punctually, in sense of attempting to heal such vacuities, it has research described here. So, this work aimed to study and apply approximation/interpolation methods to histograms, obtained from digital images, in order to, obtain even if approximately a polynomial functional form that would describe them. That is, histograms were used as discrete domain functions from which polynomial functional formulations were generated. Those formulations, considering a continuous domain, would describe frequency oscillations, in reason of brightness intensity variations. To do so, histograms from red, green, and blue components of an input digital color image were used as input data, as well as, histogram related to luminance image inherent in color image. Approximation/interpolation methods used were Ordinary Least Squares Method (OLS), B-Splines approximation method coupled to OLS, Maximum Likelihood (ML) method, and, finally, B-Splines approximation method Coupled to ML method. Generated results were evaluated visually, with help of graphs, and numerically, by means of statistical parameters established on basis of discrepancies (errors) between original values of histograms and values calculated through generated polynomial forms. In general, generated polynomials were effective in modeling input data. However, results from MV method were those that presented smallest error. B-Splines approximation method generated a quality increase to results only when coupled to OLS.

**Key-words:** Histograms; Polinomial Functions; Approximation/Interpolation Metolds.



## LISTA DE FIGURAS

Figura 1 – Modelo esquemático de Imagem Digital .....	29
Figura 2 – Codificação da imagem em 8 <i>bits</i> ilustrada na Figura 1. ....	30
Figura 3 – Imagem colorida.....	31
Figura 4 – Imagem da Luminância gerada a partir das bandas R, G e B da Imagem colorida	32
Figura 5 – Imagem aérea .....	33
Figura 6 – Histogramas das bandas espectrais da imagem aérea (Figura 5).....	33
Figura 7 – Gráfico ilustrativo do tipo <i>Boxplot</i> da variável SQE. ....	49
Figura 8 – Imagem em tons de cinza: componente R. ....	53
Figura 9 – Histograma da imagem em tons de cinza (Figura 8).....	54
Figura 10 – Dados de Entrada: <b>a)</b> Imagem colorida; <b>b)</b> Componente R; <b>c)</b> Componente G; <b>d)</b> Componente B; <b>e)</b> Imagem de Luminância. ....	59
Figura 11 – Comportamento da componente R: <b>a)</b> Gráfico de Dispersão (PxR); <b>b)</b> Histograma de R (RxP). ....	60
Figura 12 – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio. ....	61
Figura 13 – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.	63
Figura 14 – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.....	64
Figura 15 – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente R. ....	64
Figura 16 – Gráfico <i>Boxplot</i> dos resíduos quadráticos dos ajustes para a componente R. ....	67
Figura 17 – Comportamento da componente G: <b>a)</b> Gráfico de Dispersão (PxG); <b>b)</b> Histograma de G (GxP). ....	69
Figura 18 – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio. ....	70
Figura 19 – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.	71
Figura 20 – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.....	72
Figura 21 – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente G. ....	73
Figura 22 – Gráfico <i>Boxplot</i> dos resíduos quadráticos dos ajustes para a componente G. ....	76

Figura 23 – Comportamento da componente B: <b>a)</b> Gráfico de Dispersão (PxB); <b>b)</b> Histograma de B (BxP). .....	78
Figura 24 – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio. ....	78
Figura 25 – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.	80
Figura 26 – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.....	81
Figura 27 – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente B. ....	81
Figura 28 – Gráfico <i>Boxplot</i> dos resíduos quadráticos dos ajustes para a componente B. ....	84
Figura 29 – Comportamento da imagem de luminância L: <b>a)</b> Gráfico de Dispersão (PxL); <b>b)</b> Histograma de L (LxP). .....	85
Figura 30 – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio. ....	86
Figura 31 – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.	87
Figura 32 – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.....	88
Figura 33 – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da imagem de luminância L. ....	89
Figura 34 – Gráfico <i>Boxplot</i> dos resíduos quadráticos dos ajustes para a componente L.....	91



## LISTA DE TABELAS

Tabela 1 – Modelos analisados para cada componente da imagem. ....	48
Tabela 2 – Comparação de Qualidade do ajuste de cada modelo.....	51
Tabela 3 – Rol de comparações entre modelos funcionais.....	56
Tabela 4 – Comparação das Estimativas de cada ajuste.....	57
Tabela 5 – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Componente R. ....	65
Tabela 6 – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente R.....	68
Tabela 7 – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Componente G. ....	73
Tabela 8 – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente G.....	77
Tabela 9 – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Componente B. ....	82
Tabela 10 – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente B. ....	84
Tabela 11 – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Imagem de luminância L.....	89
Tabela 12 – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a imagem de luminância L.....	92



## SUMÁRIO

LISTA DE FIGURAS.....	15
LISTA DE TABELAS .....	17
<b>1 INTRODUÇÃO.....</b>	<b>21</b>
1.1 CONSIDERAÇÕES INICIAIS.....	21
1.2 TRABALHOS RELACIONADOS.....	23
1.3 OBJETIVO.....	24
<b>1.3.1 Objetivo Geral .....</b>	<b>24</b>
<b>1.3.2 Objetivos Específicos.....</b>	<b>24</b>
1.4 ESTRUTURA DO TRABALHO .....	25
<b>2 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>27</b>
2.1 IMAGEM DIGITAL .....	27
<b>2.1.1 Luminância.....</b>	<b>31</b>
<b>2.1.2 Histograma de uma Imagem Digital .....</b>	<b>32</b>
2.2 MÉTODO DOS MÍNIMOS QUADRADOS ORDINÁRIOS (MMQ OU MQO) .....	34
2.3 MÉTODO DA MÁXIMA VEROSSIMILHANÇA (MV) .....	37
2.4 MÉTODO <i>SPLINES</i> .....	41
<b>2.4.1 Spline Cúbica.....</b>	<b>42</b>
<b>2.4.2 Basis Splines(B-Spline) .....</b>	<b>45</b>
2.5 MÉTODOS DE INFERÊNCIA ESTATÍSTICA .....	48
<b>3 METODOLOGIA .....</b>	<b>52</b>
3.1 GERAÇÃO DOS DADOS DE ENTRADA: HISTOGRAMAS .....	52
3.2 DETERMINAÇÃO DE FÓRMULAS PARA AS FUNÇÕES POLINOMIAIS.....	54
3.3 AVALIAÇÃO DOS RESULTADOS.....	56
<b>4 RESULTADOS E ANÁLISES.....</b>	<b>58</b>
4.1 ASPECTOS COMPUTACIONAIS: .....	58
<b>4.1.1 Softwares.....</b>	<b>58</b>
<b>4.1.2 Dados .....</b>	<b>59</b>
4.2 RESULTADOS EXPERIMENTAIS E ANÁLISE.....	60
<b>4.2.1 Experimento com a componente R .....</b>	<b>60</b>
<b>4.2.2 Experimento com a componente G.....</b>	<b>69</b>
<b>4.2.3 Experimento com a componente B .....</b>	<b>77</b>
<b>4.2.4 Experimento com a imagem de Luminância (L) .....</b>	<b>85</b>
<b>5 CONCLUSÕES E RECOMENDAÇÕES.....</b>	<b>93</b>

5.1 CONCLUSÕES.....	93
5.2 RECOMENDAÇÕES .....	96
<b>REFERÊNCIAS</b> .....	98
<b>APÊNDICE I</b> .....	101

## 1 INTRODUÇÃO

### 1.1 CONSIDERAÇÕES INICIAIS

Um dos grandes desafios que o educador encontra atualmente, não somente na disciplina de Matemática, mas em todas as áreas do saber, é a árdua tarefa de mostrar a seus alunos aplicações cotidianas palpáveis de assuntos abordados em sala de aula.

A tarefa de materializar conceitos e exemplificá-los coerentemente fica mais complexa quando o conteúdo explanado não é tão usual, como é o caso das funções de domínio discreto. Isso acontece porque ao se deparar com certas formulações matemáticas distribuídas ao longo do Ensino Fundamental e Médio, os alunos, e infelizmente muitos professores, não possuem a noção de como tais formulações foram geradas. O que se supõem, a partir do discurso de muitos docentes, é que os aspectos discretos da Matemática são pouco privilegiados em suas formações e, conseqüentemente, na formação de seus alunos. Assim, o elo entre certas formulações de domínio contínuo, que permeiam o ensino formal, e a sua forma de obtenção, muitas vezes provenientes de dados discretos, é obscuro.

Um dos conteúdos no qual a desconexão citada acima ocorre freqüentemente é aquele voltado às “Funções”, as quais consistem de um dos mais importantes conteúdos no âmbito do Ensino Médio. Legalmente as Orientações Curriculares para o Ensino Médio citam que:

O estudo das funções permite ao aluno adquirir a linguagem algébrica como a linguagem das ciências, necessária para expressar a relação entre grandezas e modelar situações-problema, construindo modelos descritivos de fenômenos e permitindo várias conexões dentro e fora da própria Matemática. (BRASIL,2006,p.121)

Assim, é intuito do próprio sistema educacional que o aluno venha a se tornar hábil em modelar situações-problema e construir modelos descritivos funcionais de tais fenômenos.

Considerando particularmente o ensino de Matemática e diante deste problema, é coerente nortear a atividade docente pelos objetivos expressos nos Parâmetros Curriculares Nacionais (PCN), que rezam o seguinte:

Ao final do ensino médio, espera-se que os alunos saibam usar a Matemática para resolver problemas práticos do cotidiano; para modelar fenômenos em outras áreas do conhecimento; compreendam que a Matemática é uma ciência com características próprias, que se organiza via teoremas e demonstrações; percebam a Matemática como um conhecimento social e historicamente construído; saibam apreciar a importância da Matemática no desenvolvimento científico e tecnológico. A forma de trabalhar os conteúdos deve sempre agregar um valor formativo no que

diz respeito ao desenvolvimento do pensamento matemático. Isso significa colocar os alunos em um processo de aprendizagem que valorize o raciocínio matemático – nos aspectos de formular questões, perguntar-se sobre a existência de solução, estabelecer hipóteses e tirar conclusões, apresentar exemplos e contra-exemplos, generalizar situações, abstrair regularidades, criar modelos, argumentar com fundamentação lógico-dedutiva. Também significa um processo de ensino que valorize, tanto a apresentação de propriedades matemáticas acompanhadas de explicação, quanto a de fórmulas acompanhadas de dedução, e que valorize o uso da Matemática para a resolução de problemas interessantes, quer sejam de aplicação ou de natureza simplesmente teórica. (BRASIL, 1999, p.69-70)

Sendo assim, o trabalho aqui apresentado consiste do estudo e aplicação de métodos de aproximação/interpolação polinomial às funções de domínio discreto, para a geração de fórmulas matemáticas que representem, mesmo que de modo aproximado, tais funções. Além disso, apresenta-se a avaliação comparativa dos métodos utilizados no cumprimento da tarefa executada. Logo, foi estudada e mostrada uma aplicação que materializa um conhecimento muito útil em diversas áreas do conhecimento como, por exemplo, nas Engenharias. Como funções de domínio discreto, propôs-se o uso de histogramas de imagens digitais. No entanto, os métodos aqui apresentados podem ser utilizados em quaisquer tipos de dados discretos, desde que as observações se apresentem em quantidade suficiente.

Cabe esclarecer que, para cada imagem colorida utilizada, foram gerados quatro histogramas. Ou seja, foram gerados três histogramas relativos aos canais vermelho(*Red*), verde(*Green*) e azul(*Blue*), que compõem a imagem, e um quarto histograma relativo à imagem de luminância, aqui denominada *L*, gerada a partir da combinação linear dos canais (RGB) da imagem original.

Para a realização da pesquisa utilizou-se o *software* RStudio, no qual foram implementados os métodos de aproximação/interpolação Polinomial *B-Splines*, Método dos Mínimos Quadrados (MMQ) e Máxima Verossimilhança (MV), que foram estudados a fim de se compreender conceitos, definições e algoritmos.

Com a realização da pesquisa, buscou-se preencher a lacuna conceitual ora existente entre as funções de domínio contínuo e as funções de domínio discreto, apresentando ferramentas matemáticas densas e eficazes na transição entre os conteúdos.

Enfatiza-se que tal metodologia possui potencial para fundamentar análises de tendências e previsão de resultados dos fenômenos modelados.

Cabe esclarecer que, mesmo que os conteúdos estudados sejam considerados avançados se comparados aos conteúdos de Ensino Médio, tal fato não impede que qualquer professor venha a apreendê-los e que, adequadamente, tais conteúdos sejam utilizados em suas aulas. Entende-se que tal exercício científico possa ser encarado como uma oportunidade

para a aquisição de experiência no uso de *softwares* e para o aprofundamento nos conhecimentos matemáticos.

## 1.2 TRABALHOS RELACIONADOS

Quando se efetua uma busca por relatos científicos (artigos, dissertações, teses etc..) a respeito do tema abordado neste trabalho, pode-se encontrar inúmeros documentos que trazem conteúdos relacionados. Tem-se que a maioria deles trata especificamente sobre apenas um dos métodos de interpolação/aproximação que aqui foram abordados. Isto é, alguns documentos tratam do Método dos Mínimos Quadrados, outros abordam o Método da Máxima Verossimilhança e, frequentemente, encontra-se trabalhos relacionados ao Método *Splines*, nas suas mais diversas variações. No entanto, artigos que tenham os três métodos em foco são menos frequentes. A fim de se fornecer uma contextualização ao leitor sobre o tema e propiciar a comparação entre o trabalho aqui apresentado e o que vem sendo cientificamente produzido, são apresentados alguns trabalhos relacionados.

Nunhez (1990) propõem o Método *Spline* Modificado. Tal método é baseado no acoplamento do Método *Spline* ao Método da Máxima Verossimilhança e foi desenvolvido para a suavização de dados experimentais com incertezas em todas as variáveis medidas. Segundo o autor, este método é uma extensão do Método *Spline* Estendido no qual o Método *Spline* Cúbico é acoplado ao Método dos Mínimos Quadrados.

Menezes *et al.*(2010) apresentam a teoria da interpolação *Spline* aplicada à análise do crescimento da cana-de-açúcar, com o objetivo de verificar a precisão desta técnica de interpolação. Para tanto, são inseridos pontos em intervalos com extremos conhecidos ou calculando valores que já foram obtidos em medidas reais.

Oliveira (2011), utiliza funções *Splines* para o estudo de curvas de crescimento em ovinos cruzados. A autora também aborda os outros dois métodos utilizados neste trabalho, fazendo ao final uma comparação entre os métodos utilizados.

De Souza(2014) utilizou o MMQ para trabalhar de forma diferenciada com alunos de ensino Médio o tratamento de produto matricial. O autor faz uma explanação inicial dos conteúdos relacionados às matrizes e, posteriormente, baseando-se em Anton (2001), trata do assunto “Método dos Mínimos Quadrados”. O autor consegue traçar de forma didática um paralelo entre um assunto rotineiramente abordado em toda sala de Ensino Médio (Matrizes) e um assunto não tão usual, como é o Método dos Mínimos Quadrados.

Ribeiro(2014) realiza uma análise temporal de diversos dados coletados sobre a população da cidade de Iporá - Goiás, tais como, óbitos provocados por câncer, quantidade de alunos matriculados no ensino Básico e parâmetros da dinâmica populacional, entre outros. Para realizar essa análise, o autor apresenta inicialmente o conceito de interpolação polinomial e, posteriormente, utiliza o Método dos Mínimos Quadrados (MMQ) e outros métodos muito utilizados, como o Método de Lagrange e de Newton.

Pereira e Carvalho (2014) apresentam um método baseado em *B-Splines* fechada com pesos. O método consiste da reconstrução da seção de um objeto a partir de um conjunto de pontos conhecidos, obtidos por um digitalizador tridimensional. No método, cada ponto tem um peso representando sua importância na definição da curva que descreve o objeto. Pereira e Carvalho (2014) especificam que a forma de definição do peso de cada ponto do objeto e os ajustes feitos na curva foram minuciosamente analisados a fim de obter a melhor reconstrução da seção do objeto.

### 1.3 OBJETIVO

#### 1.3.1 Objetivo Geral

Para fins de esclarecimento, o objetivo geral deste trabalho foi instituído como segue. A partir de funções de domínio discreto, estabelecidas com base em histogramas dos canais vermelho, verde e azul de imagens digitais e em imagens de luminância, objetivou-se obter fórmulas matemáticas que representassem, mesmo que de modo aproximado, tais funções (histogramas). As referidas formulações foram obtidas por meio de métodos de aproximação/interpolação e foram estatisticamente comparadas.

#### 1.3.2 Objetivos Específicos

A fim de se atingir o objetivo geral expresso, os seguintes objetivos específicos foram estabelecidos e alcançados:

- Efetuar estudos e obter a compreensão dos conceitos e definições relativas ao método de Aproximação Polinomial *B-Splines*, e aos métodos de interpolação por Mínimos Quadrados (MMQ) e por Máxima Verossimilhança(MV);



- Implementar um programa de computador que efetue a decomposição de imagens digitais coloridas, obtendo, separadamente, os vetores vermelho, verde e azul (componentes das imagens). Tal programa deve ainda gerar os histogramas das componentes das imagens e da imagem de luminância calculada a partir das componentes da imagem de entrada;
- Familiarizar-se e aprender comandos do *software* RStudio, utilizado no processamento dos dados dos histogramas;
- Por meio do *software* RStudio e considerando os Métodos de aproximação/interpolação citados, gerar fórmulas polinomiais que representem, mesmo que aproximadamente, as funções de domínio discreto (histogramas).
- Efetuar a comparação estatística dos métodos estudados por meio da análise inferencial dos respectivos resultados gerados.

#### 1.4 ESTRUTURA DO TRABALHO

Este trabalho foi dividido em 5 capítulos. Como pôde ser visto, o capítulo 1 contém a introdução do trabalho, destinada a gerar uma breve contextualização ao leitor e fornecer uma ideia básica acerca do trabalho, bem como, expressar os principais objetivos da pesquisa.

O capítulo 2 contém os fundamentos teóricos mínimos necessários à compreensão da metodologia elaborada e executada. Neste capítulo são dadas as definições relacionadas às imagens digitais, a partir das quais foram gerados os histogramas relacionados aos canais R, G e B e à imagem de luminância (L). Além disso, constam, de forma sucinta, os métodos utilizados na pesquisa, a saber, MMQ, MV e *B-Splines*. Em tempo, ainda foram explanados os métodos estatísticos utilizados na avaliação e comparação dos resultados advindos dos métodos de aproximação/interpolação acima citados.

O capítulo 3 destinou-se à apresentação da metodologia utilizada para alcançar os objetivos propostos para este trabalho. Assim, discorreu-se sobre os passos para a geração dos histogramas dos canais vermelho (R), verde(G) e azul (B), além da luminância (L) da imagem digital utilizada. Também foram abordadas as etapas para a formulação de funções de domínio contínuo provenientes de funções de domínio discreto (histogramas).

O capítulo 4 contém a análise dos resultados obtidos a partir do processamento dos dados de 4 histogramas. As comparações estatísticas dos resultados são apresentadas e pormenorizadas.

O capítulo 5, por fim, contém as considerações finais depreendidas da realização do trabalho aqui exposto, bem como as principais recomendações.

## 2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo destina-se à apresentação dos conteúdos básicos necessários à execução do trabalho aqui descrito. Na seção 2.1 são apresentadas as definições relativas às imagens digitais, que serão tomadas como dados de entrada e a partir das quais os histogramas serão computados. Assim, na referida subseção estão presentes os conceitos de luminância e histograma, no âmbito do Processamento Digital de Imagens. A partir da seção 2.2, estão descritos os métodos de aproximação/interpolação que foram utilizados na pesquisa. Para a construção destas subseções foram pesquisados preliminarmente os textos clássicos de, Hartley e Rao (1967), Searle (1971), Graybil (1976), De Boor (1977), Hastie e Tibshirani (1990), Draper e Smith (1998), Nunhez (1990) e De Boor (2001), que fornecem informações pertinentes às metodologias do MMQ, MV, *Basis Splines (B-Splines)*, *B-Splines* acoplados ao MMQ, e *B-Splines* acoplados à MV, que foram utilizados na obtenção das funções polinomiais representantes dos histogramas das componentes da imagem, R, G e B, e da imagem de luminância L. Para fins de materialização das definições e algoritmos, a explanação da teoria foi feita com vistas à aplicação trabalhada. Assim, os histogramas e as informações a eles pertinentes foram utilizados na explanação que segue. Por fim, na seção 2.5 estão presentes os métodos estatísticos que foram utilizados na comparação dos resultados obtidos mediante a execução da pesquisa.

### 2.1 IMAGEM DIGITAL

Segundo Geladi e Grahn (1996), uma imagem é uma cópia/reprodução de uma figura ou objeto em seu ambiente natural. Quando, no entanto, o assunto se volta às imagens digitais em tons de cinza, pode-se defini-las como uma estrutura bidimensional “quadriculada”, na qual cada uma de suas parcelas (quadrículas) armazena um valor de intensidade de brilho, comumente denominado *digital number (DN)*. A cada quadrícula de uma imagem digital dá-se o nome de *pixel*, que é uma palavra derivada da expressão *picture element*. No caso de uma imagem digital colorida, cada quadrícula ou *pixel* armazenará três valores de intensidade de brilho, relativas às componentes vermelho, verde e azul, necessárias à geração de uma cor. Crósta (1992) afirma que, no âmbito do Processamento Digital de Imagens, as componentes

da imagem coloridas são chamadas de bandas espectrais da imagem (ou, coloquialmente, bandas) e são referidas por meio dos respectivos nomes em inglês (*Red*, *Green* e *Blue*).

Matematicamente, considerando um domínio contínuo, pode-se definir uma imagem como sendo uma função na qual cada ponto do plano da imagem corresponde a uma única intensidade de brilho ou cor (MENEGUETTE *et. al.*, 2003), ou seja:

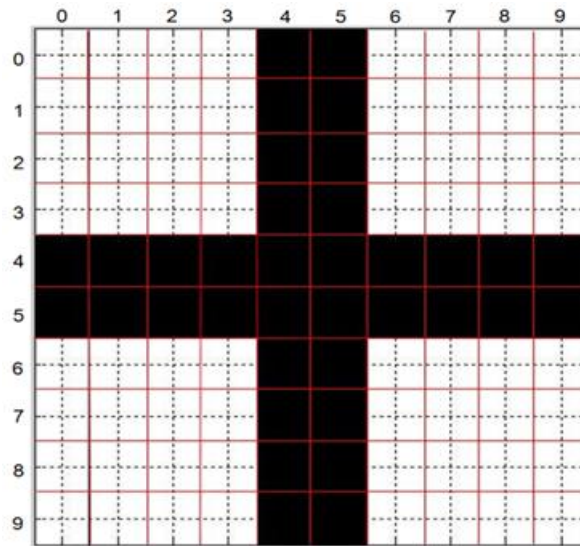
$$\begin{aligned} f: U \subset \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow c \end{aligned} \quad (1)$$

onde:

- $f$  denota a imagem;
- $U$  é uma sub-região da cena;
- $\mathbb{R}$  é o conjunto dos números Reais ao qual pertence a intensidade de brilho  $c$  (ou cor).

Apesar da precisão e do rigor da definição matemática de imagem, usualmente elas são tratadas ou trabalhadas do ponto de vista discreto como sendo uma matriz de valores representantes das intensidades de brilho ou cores. E mais, os níveis de brilho, seja para a codificação dos tons de cinza, seja para a definição das cores, são dados em escalas de valores naturais.

Segundo Gonzalez e Woods (2000), existem dois tipos básicos de codificação computacional para as imagens: o tipo matricial ou *raster* e o tipo vetorial. A Figura 1, que representa a imagem de uma cruz, ilustra o primeiro tipo de codificação. Nesta figura, adaptada de Silva (2010), o tipo matricial é apresentado de modo esquemático. É possível verificar que as linhas e colunas de *pixels*, delimitadas por linhas vermelhas, são numeradas, respectivamente, a partir do 0 (zero), em ordem crescente, de cima para baixo e da esquerda para a direita. Além disso, as coordenadas dos *pixels*, dadas no sistema (linha, coluna), correspondem ao seu centróide (marcados pelos cruzamentos das linhas pretas tracejadas).

**Figura 1** – Modelo esquemático de Imagem Digital.

Fonte: Adaptado de Silva (2010).

Em imagens em tons de cinza, frequentemente as intensidades de brilho podem assumir valores que vão de 0 a 255 (256 valores), sendo que a intensidade 0 (zero) corresponde ao preto absoluto e o 255 ao branco absoluto. Valores intermediários correspondem às tonalidades de cinza (CRÓSTA, 1992). Este intervalo de representação de tonalidades é dito representação em 8 *bits*. Ou seja, estão disponíveis  $2^8 = 256$  intensidades de tons de cinza. Segundo Silva (2010), existem imagens que possuem uma quantidade diferente de níveis de brilho disponíveis. Exemplo destas imagens, são as imagens geradas pelos sensores dos satélites RapidEye, que possuem 12 *bits* em cada banda espectral (4096 intensidades de brilho disponíveis). A esta quantidade (ou escala) de níveis de brilho, os quais os *pixels* de uma imagem podem assumir, dá-se o nome de resolução radiométrica (CRÓSTA, 1992). Assim, infere-se que, uma imagem em 12 *bits* possui melhor resolução radiométrica que uma imagem em 8 *bits*.

Nestes termos, o quadro apresentado na Figura 2, considerando uma resolução radiométrica de 8 *bits*, representa as intensidades de brilho presentes no esquema apresentado na Figura 1.

**Figura 2** – Codificação da imagem em 8 *bits* ilustrada na Figura 1.

	0	1	2	3	4	5	6	7	8	9
0	255	255	255	255	0	0	255	255	255	255
1	255	255	255	255	0	0	255	255	255	255
2	255	255	255	255	0	0	255	255	255	255
3	255	255	255	255	0	0	255	255	255	255
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	255	255	255	255	0	0	255	255	255	255
7	255	255	255	255	0	0	255	255	255	255
8	255	255	255	255	0	0	255	255	255	255
9	255	255	255	255	0	0	255	255	255	255

Fonte: Adaptado de Silva (2010).

Meneguette *et. al.* (2003) especificam que, se a imagem for colorida, cada quadrícula da codificação possuirá 3 (três) valores numéricos, relativos aos canais R (*Red*), G (*Green*) e B (*Blue*), que determinarão, por composição, a cor do *pixel*.

Quando a imagem for do tipo vetorial as intensidades de cada *pixel* tenderão a não coincidir com a verdadeira intensidade de brilho/cor, amostrada da cena imageada. Isto ocorre, pois, para fins de compactação, a representação de intensidade de brilho/cor sofre uma distorção (GONZALES e WOODS, 2000). Ou seja, nesta forma de codificação as intensidades de brilho/cor serão tomadas, muitas vezes como aproximações baseadas nas informações originais dos *pixels* e de seus entornos.

Decorre do exposto acima que, as imagens do tipo vetorial não são frequentemente utilizadas no âmbito do Processamento Digital de Imagens, isto porque, é necessário que cada *pixel* possua a informação original obtida quando de seu imageamento (VALE, 2007). Decorre da explanação que, o formato mais freqüentemente utilizado é o matricial ou *raster*, que foi adotado na execução deste trabalho.

Associado às imagens digitais encontra-se o conceito de luminância. Segundo Gomes e Velho (1994), a luminância é definida como sendo a intensidade física da luz na região do espectro visível ponderado por uma curva de sensibilidade espectral que corresponde à sensação de brilho do sistema de visão humano. Nesse caso, a luminância provém dos bastonetes, ao passo que a crominância (informação de cor) advém dos cones (CRÓSTA, 1992).

### 2.1.1 Luminância

No âmbito do Processamento Digital de Imagens, a luminância de uma imagem pode ser computada a partir das componentes primárias RGB como (GOMES e VELHO, 1994):

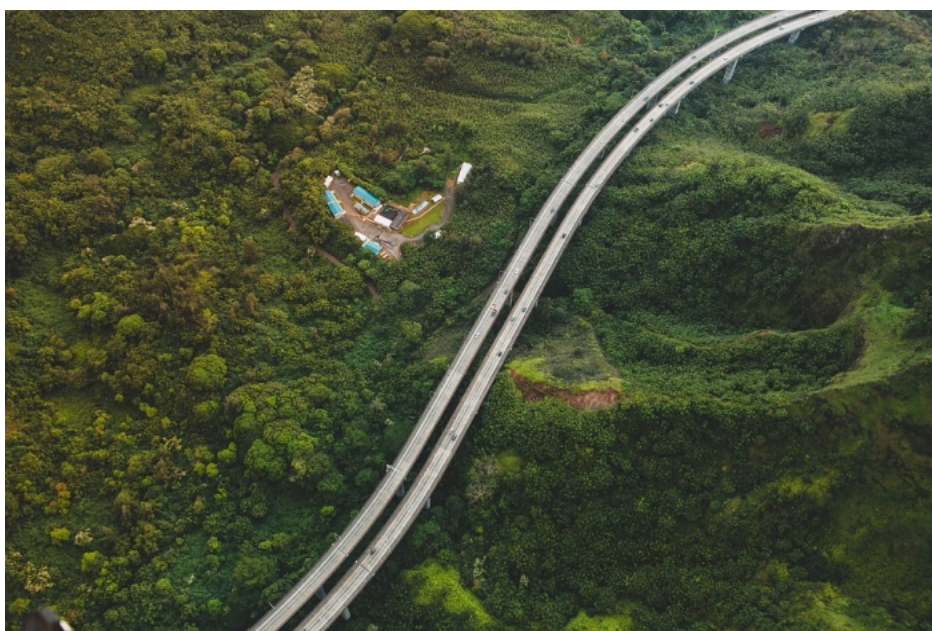
$$Y = 0,125.R + 0,7154.G + 0,0721.B \quad (2)$$

onde:

- $Y$  denota a luminância;
- $R, G$  e  $B$  são as componentes (bandas espectrais) da imagem colorida de entrada

Para fins de materialização do conceito de luminância tem-se a Figura 3, na qual aparece uma imagem colorida e a Figura 4, que é a sua respectiva imagem de luminância.

**Figura 3** – Imagem colorida.



Fonte: Resultados da pesquisa.

**Figura 4** – Imagem da Luminância gerada a partir das bandas R, G e B da Imagem colorida.



Fonte: Resultados da pesquisa.

### 2.1.2 Histograma de uma Imagem Digital

Conforme Silva (2010), em processamento de imagens, trabalha-se sempre com os tons de cinza (*digital numbers* ou DN's) atribuídos aos *pixels* de uma imagem. De forma idêntica, pode-se trabalhar com as intensidades de brilho de uma dada banda espectral, uma vez que a imagem relativa a esta conterá apenas as intensidades de brilho relativas àquela dada matiz.

Crósta (1992) afirma que, o histograma é uma das formas mais comuns de se representar a distribuição dos DN's de uma imagem e, possivelmente, a mais útil em Processamento Digital de Imagens. O histograma fornece informação sobre a frequência de cada DN no plano da imagem. No caso das imagens de 8 *bits*, os DN's variam de 0 a 255.

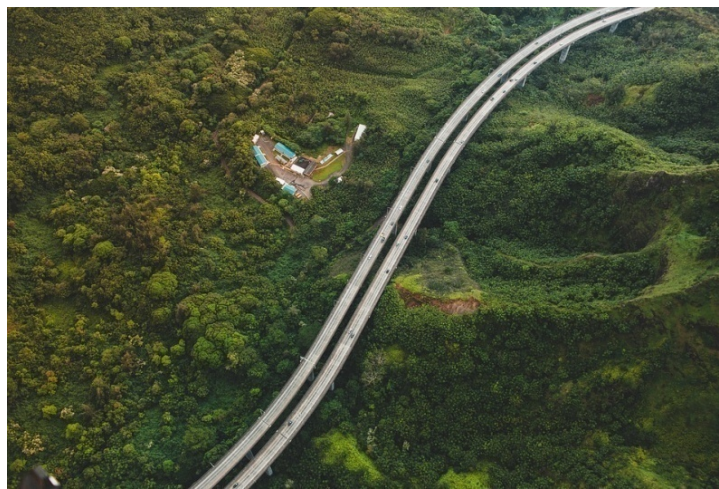
Os histogramas são também conhecidos como distribuição de intensidades e/ou Função de Densidade de Probabilidade (PDF). Esse último termo advém do fato de que, estatisticamente, o histograma representa, neste caso, a probabilidade de se achar um DN de um dado valor dentro de uma imagem.

Outro ponto importante com relação aos histogramas é que eles representam dados discretos. Logo, a distribuição de intensidades é representada graficamente por colunas discretas, contáveis e enumeráveis. Os histogramas normalmente tem a sua importância associada ao realce de contraste em imagens.



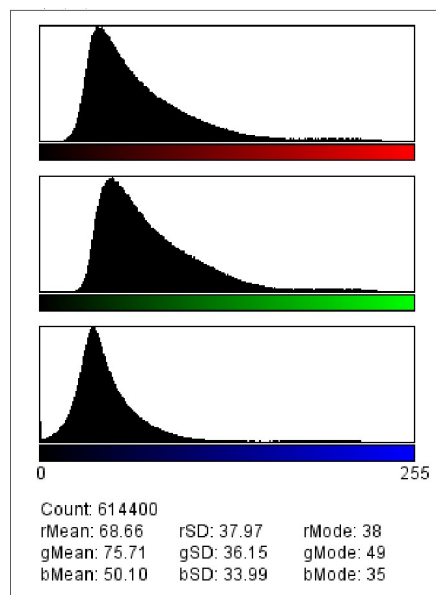
Ao se observar o histograma de uma imagem, tem-se uma noção sobre algumas das características desta imagem. Ou seja, a forma do histograma e a distribuição de suas colunas fornecem informações importantes sobre a imagem, tais como, intensidade média e espalhamento dos valores de DN, considerando sua frequência. Quanto ao espalhamento, muitas vezes quantificado pelo desvio-padrão, cabe dizer que, este dá a medida do contraste de uma imagem, de tal forma que, quanto maior o espalhamento ao longo do eixo dos DN's, (eixo x) maior será o contraste da imagem.

**Figura 5** – Imagem aérea.



Fonte: Resultados da pesquisa.

**Figura 6** – Histogramas das bandas espectrais da imagem aérea (Figura 5).



Fonte: Resultados da pesquisa.

Para fins de materialização dos conceitos, na Figura 5 é apresentada uma imagem colorida. Trata-se de uma imagem aérea na qual predomina vegetação em tom escuro. Assim, se pode verificar na Figura 6, os histogramas das bandas espectrais  $R, G$  e  $B$  desta imagem, que as intensidades de brilho que possuem maior frequência são aquelas que se encontram próximas da origem. Em contrapartida, pode-se verificar que a frequência para as maiores intensidades de brilho, nas três bandas, não é nula. Estes valores dizem respeito às feições claras presentes no plano da imagem, tais como, rodovias e edificações.

Da simples análise visual do histograma, pode-se depreender que a média e a moda dos DN's seja baixa, visto que, a maior frequência de intensidades encontra-se próxima da origem. Tal conclusão é confirmada pelos valores de média ( $rMean, gMean$  e  $bMean$ ) e moda ( $rMode, gMode$  e  $bMode$ ), presentes na porção inferior da Figura 6. Ainda na porção inferior da Figura 6, pode-se verificar também os valores de desvio-padrão (*Standard – Deviation*) para as três bandas espectrais ( $rSD, gSD$  e  $bSD$ ).

## 2.2 MÉTODO DOS MÍNIMOS QUADRADOS ORDINÁRIOS (MMQ OU MQO)

Considerando um conjunto de 256 intensidades de brilho de  $pixels(P_i)$ , com ( $i = 1, 2, \dots, 256$ ) para cada histograma gerado a partir de uma imagem analisada, tem-se os pontos experimentais para cada componente  $(P_i, R_i), (P_i, G_i), (P_i, B_i)$  e  $(P_i, L_i)$ , com ( $i = 1, 2, \dots, 256$ ). Dessa forma, a partir destes dados, buscou-se descrever de forma aproximada os histogramas das componentes por meio de funções polinomiais estimadas do tipo  $\hat{R}_i = f_R(P_i, \beta_{R0}, \dots, \beta_{Rn}), \hat{G}_i = f_G(P_i, \beta_{G0}, \dots, \beta_{Gn}), \hat{B}_i = f_B(P_i, \beta_{B0}, \dots, \beta_{Bn})$  e  $\hat{L}_i = f_L(P_i, \beta_{L0}, \dots, \beta_{Ln})$ , nas quais os coeficientes  $\beta_{Rj}, \beta_{Gj}, \beta_{Bj}$  e  $\beta_{Lj}$ , com  $j = (0, \dots, n)$  de cada modelo foram determinados nos termos da metodologia proposta, sendo  $n < 256$ .

Assim, supõe-se que a variável independente  $P_i$  de cada histograma (domínio) seja isenta de erros, e todo erro esteja associado às variáveis dependentes ( $R_i, G_i, B_i$  e  $L_i$ ). Considerando que, cada função polinomial gerada estabeleça uma aproximação a um histograma, conclui-se que exista um erro (desvio ou resíduo ( $e_i$ )) entre a função e o histograma. Assim, o modelo deve ser construído de forma que este erro seja mínimo, considerando que, sendo randômico, este erro possua distribuição normal, média nula ( $\mu = 0$ ) e variância ( $\sigma^2$ ) constante.

Logo, pode-se estabelecer o método para um histograma genérico qualquer como segue. Estabelecendo  $(Y_i)$  como sendo um valor de frequência associado a uma intensidade  $(P_i)$ , ambos tomados em histograma, e  $(\hat{Y}_i)$  um valor de imagem estimado por meio da função polinomial  $f$  (função polinomial para um histograma genérico) calculada em  $(P_i)$ , então o erro pode ser definido como  $e_i = (Y_i - \hat{Y}_i) \sim N(0, \sigma^2)$ . Assim o objetivo principal deste método é a estimação dos coeficientes  $\beta_0, \beta_1, \dots, \beta_n$  do modelo polinomial, minimizando o somatório dos quadrados de  $e_i$ , como mostra a equação abaixo.

$$Y = \sum_{i=1}^{256} e_i^2 = \sum_{i=1}^{256} (Y_i - \hat{Y}_i)^2 = \text{mínimo} \cong 0. \quad (3)$$

onde:

- $\hat{Y}_i = f(P_i, \beta_0, \dots, \beta_n)$ , são valores estimados de  $\hat{R}_i, \hat{G}_i, \hat{B}_i$  e  $\hat{L}_i$ , para  $i = 1, 2, \dots, 256$ .

O ajuste é feito minimizando a equação (3) em relação aos coeficientes  $(\beta_0, \dots, \beta_n)$  da função  $f$  e é dado pela fórmula:

$$\frac{\partial e_i}{\partial \beta_n} = -2 \sum_{i=1}^{256} (Y_i - \hat{Y}_i) \cdot \frac{\partial \hat{Y}_i}{\partial \beta_n} = 0, n = 1, 2, \dots, A. \quad (4)$$

onde:

- O valor 256 é o número de graus de liberdade;
- A constante “A” é o número de coeficientes estimados no MMQ, com  $A < 256$ . Assim será possível obter  $(\beta_0, \dots, \beta_n)$ . Cabe esclarecer que, com o aumento de “A” espera-se um aumento decrescente assintótico dos erros.

Considerando que o modelo (função polinomial) para o ajustamento pelo MMQ é  $\hat{Y}_i = \beta_0 + \beta_1 P_i + \beta_2 P_i^2 + \dots + \beta_n P_i^n$ , com  $i = 1, 2, \dots, 256$  e os histogramas fornecem os valores  $(P_i, Y_i)$ , então o problema recai sobre a determinação dos coeficientes  $\beta_k$ , para  $k = 0, 1, 2, \dots, n$ . Portanto, o modelo é linear e o método a ser utilizado para a determinação dos coeficientes é o ajustamento com modelo paramétrico (sem iterações).

Assim, para cada um dos histogramas das componentes da imagem, estima-se um modelo que o represente, baseado na seguinte forma estrutural geral:

$$Y_i = f(P_i, \beta_0, \dots, \beta_n) = \beta_0 + \beta_1.P_i + \dots + \beta_n.P_i^n + e_i. \quad (5)$$

Nestes termos, partindo-se deste modelo, a forma funcional para cada histograma poderá ser estimada, mediante a minimização da soma dos quadrados dos erros  $(\sum_{i=1}^{256} e_i^2)$ , chegando-se à:

$$\hat{Y}_i = f(X_{ij}, \beta_0, \dots, \beta_n) = \beta_0 + \beta_1.X_{i1} + \dots + \beta_n.X_{in} + e_i \quad (6)$$

onde:

- $i = 1, 2, \dots, 256$ .
- $j = 1, 2, \dots, n$  e
- $(X_{i1}, X_{i2}, \dots, X_{in}) = (P_i, P_i^2, \dots, P_i^n)$ , são as variáveis regressoras, assim consideradas para fins de representação.

Portanto, este modelo pode ser expresso como um sistema de 256 equações, que pode ser representado matricialmente como:

$$Y = X.B + E \quad (7)$$

Sendo as respectivas matrizes da equação acima dadas por:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{256} \end{bmatrix}; \quad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1n} \\ 1 & X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2561} & X_{2562} & \cdots & X_{256n} \end{bmatrix}; \quad B = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \quad \text{e} \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{256} \end{bmatrix}$$

Com o intuito de minimizar os erros da equação 7, tem-se que  $E = Y - X.B$ , e a soma dos quadrados dos resíduos, matricialmente, pode ser efetuada da seguinte maneira:

$$Z = E^t.E = (Y - X.B)^t.(Y - X.B) \quad (8)$$

Desenvolvendo-se a equação 8 tem-se que:

$$Z = Y^t \cdot Y - Y^t \cdot X \cdot B - B^t \cdot X^t \cdot Y + B^t \cdot X^t \cdot B \cdot X \quad (9)$$

Observa-se que  $Y^t \cdot X \cdot B$  e  $B^t \cdot X^t \cdot Y$  são iguais, pois uma é a transposta da outra e o resultado de ambas é o mesmo escalar, na sentença acima. Rearranjando a equação 9, tem-se:

$$Z = Y^t \cdot Y - 2 \cdot B^t \cdot X^t \cdot Y + B^t \cdot X^t \cdot B \cdot X \quad (10)$$

A fim de minimizar a função  $Z$ , em busca dos valores de  $B$  que tornem sua diferencial identicamente nula, deve-se derivar  $Z$  em relação  $B$  e igualar a expressão resultante a zero. Procedendo desta forma obtêm-se:

$$\frac{\partial Z}{\partial B} = -2 \cdot Y^t \cdot X + 2 \cdot B^t \cdot X^t \cdot X = 0 \quad (11)$$

resultando em

$$\hat{B} = (X^t \cdot X)^{-1} \cdot X^t \cdot Y = (X' \cdot X)^{-1} \cdot X' \cdot Y \quad (12)$$

onde  $\hat{B}$  é o estimador que torna mínima a soma dos quadrados dos resíduos  $Z = E^t \cdot E = (Y - X \cdot B)^t \cdot (Y - X \cdot B)$ , obtendo-se dessa forma o vetor  $B$  dos parâmetros  $(\beta_0, \beta_1, \dots, \beta_n)$  da equação 7.

### 2.3 MÉTODO DA MÁXIMA VEROSSIMILHANÇA (MV)

O Método da Máxima Verossimilhança parte da definição de erro ( $e_i$ ) e dos mesmos pressupostos estabelecidos para o MMQ ( $e_i \sim N(0, \sigma^2)$ ) e ele, igualmente, é utilizado para estimar os parâmetros  $\beta_0, \beta_1, \dots, \beta_n$  de cada função polinomial, a partir da maximização da função de verossimilhança  $L$  (equação 13) para os erros, da seguinte maneira:

$$L = L(e_1, \dots, e_{256}; \beta_0, \dots, \beta_n; \sigma^2) = L(e_1) \cdot L(e_2) \dots L(e_{256}) \quad (13)$$

Desta forma tem-se que:

$$\begin{aligned} L &= \prod_{i=1}^{256} \left\{ \frac{1}{(2\pi\sigma^2)} \cdot \exp\left[-\frac{1}{(2\sigma^2)} \cdot (e_i)^2\right] \right\} = \\ &= \prod_{i=1}^{256} \left\{ \frac{1}{(2\pi\sigma^2)} \cdot \exp\left[-\frac{1}{(2\sigma^2)} \cdot (Y_i - \hat{Y}_i)^2\right] \right\} \end{aligned} \quad (14)$$

onde:

- $\Pi$  é o produto de 256 fatores (produtório);
- $\pi \cong 3,14159 \dots$
- $\exp$ : base do logaritmo neperiano (número “ $e \cong 2,71828 \dots$ ” de Euler)
- $\sigma$ : desvio padrão associado ao erro.

Esta expressão é conhecida na literatura como função de Verossimilhança e, neste caso, está estabelecida para os desvios quadráticos. Assim, aplicando o logaritmo neperiano em ambos os membros da equação, tem-se:

$$\begin{aligned} \ln(L) &= \ln \left( \prod_{i=1}^{256} \left\{ \frac{1}{(2\pi\sigma^2)} \cdot \exp\left[-\frac{1}{(2\sigma^2)} \cdot (e_i)^2\right] \right\} \right) \\ &= \ln \left( \prod_{i=1}^{256} \left\{ \frac{1}{(2\pi\sigma^2)} \cdot \exp\left[-\frac{1}{(2\sigma^2)} \cdot (Y_i - \hat{Y}_i)^2\right] \right\} \right) \\ &= -\frac{n}{2} \ln(2\pi) - (n) \ln(\sigma) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{256} (Y_i - \hat{Y}_i)^2 \end{aligned} \quad (15)$$

Assim, maximiza-se a equação 15, por meio de derivação em relação aos parâmetros e igualando tal derivada a zero:

$$\frac{\partial \ln(L)}{\partial \beta_i} = 0, \quad i = 1, 2, \dots, A \quad (16)$$

Cabe esclarecer que, 256 é o número de graus de liberdade e “A” é número de coeficientes que neste trabalho será estimado pelo MMQ e também pelo MV. Assim será possível obter  $\beta_0, \beta_1, \dots, \beta_n$  por ambos os métodos.

Analogamente ao MMQ, chega-se a um sistema de equações dado por  $Y = X.B + E$ , no qual a soma dos quadrados dos resíduos  $Z = E^t.E = (Y - X.B)^t.(Y - X.B)$ , será maximizada matricialmente, assumindo que os erros são normalmente distribuídos com média zero e matriz de variância-covariância  $V$  constante, isto é,  $E \sim N(0, V)$ . Assim tem-se que:

$$\begin{aligned} \hat{B} &= (X'.X)^{-1}.X'.Y = (X'.X)^{-1}.X'(X.B + E) = \\ &= (X'.X)^{-1}.X'.X.B + (X'.X)^{-1}.X'.E = B + (X'.X)^{-1}.X'.E \end{aligned} \quad (17)$$

Resultando em:

$$\hat{B} - B = (X'.X)^{-1}.X'.E \cong 0 \quad (18)$$

Caso a diferença acima seja nula, tem-se que:

$$\hat{B} = B$$

O que revela que  $\hat{B} = (X'.X)^{-1}.X'.Y$  apresenta-se como um bom estimador não tendencioso de  $B$ .

A partir da equação 18, pode-se construir a matriz de variância-covariância dos estimadores dos parâmetros  $V$ . Tem-se que:

$$\begin{aligned} \hat{B} - B &= (X'.X)^{-1}.X'.E \\ V = Cov(\hat{B}) &= (\hat{B} - B).(\hat{B} - B)' = (X'.X)^{-1}.X'.E.E'.X'(X'.X)^{-1} \\ &= (X'.X)^{-1}.X'.\sigma^2 I.X'(X'.X)^{-1} = (X'.X)^{-1}.\sigma^2 \end{aligned} \quad (19)$$

Em que  $(X'.X)^{-1}.\sigma^2$  é a matriz de variância-covariância dos estimadores. Na diagonal principal desta matriz encontram-se as variâncias dos estimadores dos parâmetros e, fora da diagonal, as covariâncias dos estimadores dos parâmetros. Ou seja:

$$\begin{aligned}
V = Cov(\hat{B}) &= \begin{bmatrix} (\hat{\beta}_0 - \beta_0)^2 & (\hat{\beta}_0 - \beta_0) \cdot (\hat{\beta}_1 - \beta_1) \cdots & (\hat{\beta}_0 - \beta_0) \cdot (\hat{\beta}_n - \beta_n) \\ (\hat{\beta}_0 - \beta_0) \cdot (\hat{\beta}_1 - \beta_1) & (\hat{\beta}_1 - \beta_1)^2 & \cdots & (\hat{\beta}_1 - \beta_1) \cdot (\hat{\beta}_n - \beta_n) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ (\hat{\beta}_0 - \beta_0) \cdot (\hat{\beta}_n - \beta_n) & (\hat{\beta}_1 - \beta_1) \cdot (\hat{\beta}_n - \beta_n) \cdots & & (\hat{\beta}_n - \beta_n)^2 \end{bmatrix} \\
&= \begin{bmatrix} \hat{V}(\hat{\beta}_0) & \widehat{Cov}(\hat{\beta}_0, \hat{\beta}_1) \cdots \widehat{Cov}(\hat{\beta}_0, \hat{\beta}_n) \\ \widehat{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \hat{V}(\hat{\beta}_1) & \vdots & \widehat{Cov}(\hat{\beta}_1, \hat{\beta}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{Cov}(\hat{\beta}_0, \hat{\beta}_n) & \widehat{Cov}(\hat{\beta}_1, \hat{\beta}_n) & & \hat{V}(\hat{\beta}_n) \end{bmatrix} = \begin{bmatrix} \sigma_{00} & \sigma_{01} \cdots \sigma_{0n} \\ \sigma_{10} & \sigma_{11} & \vdots & \sigma_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n0} & \sigma_{n1} \cdots \sigma_{nn} \end{bmatrix} \quad (20)
\end{aligned}$$

Desta maneira, construindo a função de verossimilhança novamente, tem-se que:

$$\begin{aligned}
L &= \prod_{i=1}^{256} \left\{ \frac{1}{(2\pi\sigma^2)} \cdot \exp\left[-\frac{1}{(2\sigma^2)}(e_i)^2\right] \right\} = \\
&= 2\pi^{-\frac{1}{2}(256)} \cdot |V|^{-\frac{1}{2}(256)} \cdot \exp\left\{-\frac{1}{2}(Y - X \cdot B)' \cdot V^{-1} \cdot (Y - X \cdot B)\right\} \quad (21)
\end{aligned}$$

A fim de maximizar o logaritmo neperiano da função  $L$ , em busca dos valores de  $B$  que tornem sua diferencial identicamente nula, deve-se derivar  $\ln(L)$  em relação  $B$  e igualar a expressão resultante a zero. Procedendo desta forma obtém-se:

$$\frac{\partial \ln(L)}{\partial B} = 0$$

ou

$$\hat{B} = (X' \cdot V^{-1} \cdot X)^{-1} \cdot X' \cdot V^{-1} \cdot Y \quad (22)$$

Sendo  $\hat{B}$  um bom estimador que torna máximo os parâmetros  $(\beta_0, \dots, \beta_n)$  da equação 6 (em MMQ). Sob as mesmas condições de estimação dos parâmetros e com  $V = I\sigma^2$ , este estimador nada mais é do que o estimador dos mínimos quadrados ordinários visto anteriormente.



## 2.4 MÉTODO *SPLINES*

Por vezes, para algumas aplicações é necessária uma representação de curvas complexas mais consisas e eficientes que as formas discretas. Assim, por exemplo, representar uma curva por uma sucessão de segmentos de reta poderia ser uma saída. Porém, tal representação pode não ser suficientemente boa para certas aplicações. Surge então a necessidade de se lançar mão de ferramentas matemáticas mais rebuscadas.

Neste contexto estão as *Splines* e, por meio delas, pode-se abordar o problema considerando duas vertentes. Pode-se definir uma curva que passe por um conjunto determinado de pontos nodais (problema de interpolação) ou buscar por formulações de curvas que passem próximas ao conjunto determinado de pontos (problema de aproximação).

Nesta seção serão apresentados métodos baseados em *Splines* que seguem ambas as vertentes. *Grosso modo*, pode-se dizer que um problema de interpolação via *Splines* se resume em conectar pontos nodais espaçados por curvas, exibindo uma representação matemática condizente que seja capaz de gerar novos pontos em uma densidade adequada. As curvas, definidas entre pontos nodais consecutivos, são dependentes entre si. Essa dependência das funções *Splines* é expressa pela igualdade das derivadas em ambos os lados dos pontos nodais, exceto no primeiro e último ponto nodal. Assim as funções descrevem trajetórias que passam pelos pontos nodais que as definem, formando uma curva lisa ou um novo ajuste para os tais pontos.

Por outro lado, se não houver a injunção de que a curva passe obrigatoriamente por um conjunto pré-definido de pontos, então tem-se os métodos *Splines* de aproximação. De qualquer forma, a escolha para as funções de ajuste de curva recai sobre funções polinomiais e, assim como nos métodos anteriores, estes métodos destinam-se ao ajuste de curvas com resíduo mínimo (caso de suavização).

Pode-se citar como exemplo de *Splines* de interpolação a *Spline* linear, a *Spline* quadrática e a *Splines* cúbica que, aliadas aos polinômios ortogonais de Legendre, são as mais utilizadas. Como exemplo de *Splines* de aproximação tem-se as *Basis Splines (B-Splines)* e *Penalized B-Splines (P-Splines)*, sendo que esta última utiliza-se de técnicas bayesianas para seus ajustes. (De BOOR, 1977; HASTIE, 1990; De BOOR, 2001).

Neste trabalho optou-se pelo uso das *Basis Splines*, justamente por estas apresentarem vantagens em relação aos polinômios ortogonais de Legendre e à *Spline* cúbica. Ou seja, as *B-Splines* apresentam propriedades numéricas superiores por promover maior esparsialidade e

quebra de multicolinearidade na matriz de termos covariáveis, priorizam o ajuste local em detrimento do global, fato que as tornam mais robustas justamente pelo reduzido número de dados para certas faixas, considerando algumas análises, e proporcionam melhor controle da influência de observações individuais (MEYER, 2005; WELHAM et al., 2005; WELHAM et al., 2007; SÁNCHEZ et al., 2008; MENEZES et al., 2012; RESENDE et al., 2014).

### 2.4.1 *Spline* Cúbica

Quando se tem uma função  $f(x)$  estabelecida em  $(n + 1)$  pontos de um domínio discreto e se tenta aproximar tal função por um polinômio de grau  $n$ , capaz de gerar uma interpolação entre os  $(n + 1)$  pontos tabelados, pode-se chegar a um resultado inconsistente (popularmente conhecido como fenômeno de Runge). Assim, alternativamente opta-se por interpolar  $f(x)$  considerando subconjuntos de seu domínio que contenham poucos pontos, obtendo-se dessa forma um polinômio de menor grau, impondo ainda que a função de interpolação seja contínua e possua derivadas contínuas até uma certa ordem.

Seguindo esta ideia, dada a função  $f(x)$  citada acima, pode-se tomar subconjuntos com dois pontos e se chegar a um polinômio de grau 1. Porém, se for gerada uma aproximação de  $f(x)$  por meio de uma função linear por partes ( $S_1(x)$ ), verifica-se que  $S_1(x)$  é contínua, mas não possui derivada em todos os pontos de seu domínio. Ou seja, a função  $S_1(x)$  apresentará angulosidades nos pontos do conjunto discreto que definem o domínio de  $f(x)$ . Logo,  $S_1(x)$  não terá derivadas nestes pontos, pois os limites laterais da razão incremental que definem a integral de  $S_1(x)$  nos ditos pontos serão distintos, logo não existirá o limite (ou derivada) nestes pontos.

Optando-se por utilizar subconjuntos com 3 pontos e se chegar a polinômios de grau 2, mais uma vez se tem um problema semelhante ao caso anterior. Isto é, tem-se mais uma vez a garantia de continuidade da função que aproxima  $f(x)$  de ( $S_2(x)$ ) mas existirão pontos para os quais  $S'_2(x)$  não existirá.

Ruggiero e Lopes (1996) especificam que, no caso das funções *Splines*, a opção é aproximar a função discreta (tabelada), em cada subintervalo  $[x_i, x_{i+1}]$ , por um polinômio de grau  $p$ , impondo algumas condições sobre a função. A definição a seguir estabelece os pormenores desta ideia.

Considere a função  $f(x)$  tabelada nos pontos  $a = x_0 < x_1 < \dots < x_n = b$ .

Segundo Ruggiero e Lopes (1996), uma função  $S_p(x)$  é denominada *Spline* de grau  $p$  com nós nos pontos  $x_i, i = 0, 1, \dots, n$ , se satisfazer as seguintes condições:

- a) Em cada subintervalo  $[x_i, x_{i+1}], i = 0, 1, \dots, n$ ,  $S_p(x)$  é um polinômio de grau  $p$ :  $S_p(x)$ .
- b)  $S_p(x)$  é contínua e tem derivada contínua até ordem  $(p - 1)$  em  $[a, b]$ .
- c) Se, além disto,  $S_p(x)$  também satisfaz a condição:  $S_p(x_i) = f(x_i)$  com  $i = 0, 1, \dots, n$  então  $S_p(x)$  será denominada *Spline* interpolante.

Em conformidade com o exposto acima, quando  $p = 3$ , então a *Spline* é chamada de *Spline* cúbica. Ruggiero e Lopes (1996) especificam que as *Splines* cúbicas são as mais utilizadas e possuem vantagens quando comparadas com as lineares e com as quadráticas, pois:

- As *Splines* lineares não possuem derivadas contínuas nos nós;
- As *Splines* quadráticas possuem apenas a derivada primeira contínua. Ou seja, como as derivadas segundas são descontínuas nos nós, a curvatura das *Splines* quadráticas podem mudar abruptamente nos referidos nós.

Assim, pode-se expressar a definição de *Spline* cúbica como segue:

Supondo que  $f(x)$  seja função de domínio discreto e esteja tabelada nos pontos  $x_i, i = 0, 1, \dots, n$ . A função  $S_3(x)$  é chamada *Spline* cúbica interpolante de  $f(x)$  nos nós  $x_i, i = 0, 1, \dots, n$  se existem  $n$  polinômios de grau 3,  $s_k(x), k = 0, 1, \dots, n$ , tais que:

- i)  $S_3(x) = s_k(x)$  para  $x \in [x_i, x_{i+1}], k = 0, 1, \dots, n$ ;
- ii)  $S_3(x) = s_k(x), i = 0, 1, \dots, n$ ;
- iii)  $s_k(x_k) = s_{k+1}(x_k), i = 0, 1, \dots, (n - 1)$ ;
- iv)  $s'_k(x_k) = s'_{k+1}(x_k), i = 0, 1, \dots, (n - 1)$ ;
- v)  $s''_k(x_k) = s''_{k+1}(x_k), i = 0, 1, \dots, (n - 1)$ ;

Considerando as condições acima, o objetivo recai sobre a determinação de um polinômio de terceiro grau para cada intervalo entre os pontos amostrados. Isto é, a interpolação entre pontos consecutivos se dá por uma função como a mostrada abaixo:

$$s_k(x) = a_k(x - x_k)^3 + b_k(x - x_k)^2 + c_k(x - x_k) + d_k \quad (23)$$

Onde,  $k = 0, 1, \dots, n$ , considerando que estão disponíveis  $n + 1$  nós. Assim, tem-se  $n$  intervalos, e deverão ser estabelecidas  $4n$  constantes ( $a_k, b_k, c_k, e d_k$ ), ou seja, quatro constantes para cada uma das  $s_k(x)$ .

Impõem-se as seguintes condições para que  $S_3(x)$  seja *Spline* interpolante de  $f(x)$  em  $x_0, x_1, \dots, x_n$  (RUGGIERO e LOPES, 1996):

- $(n + 1)$  condições para que  $S_3(x)$  interpole  $f(x)$  nos nós;
- $(n - 1)$  condições para que  $S_3(x)$  esteja bem definida nos nós (continuidade de  $S_3(x)$  em  $[x_0, x_n]$ );
- $(n - 1)$  condições para que  $S'_3(x)$  seja contínua em  $[x_0, x_n]$ ; e
- $(n - 1)$  condições para que  $S''_3(x)$  seja contínua em  $[x_0, x_n]$ .

Calculando-se a quantidade total de condições expressas acima, chega-se à  $4n - 2$  condições. Porém, como  $4n$  coeficientes devem ser calculados, se faz necessário estabelecer mais duas condições para se chegar a  $4n$  condições (quatro condições por subintervalo).

Se as condições complementares forem tomadas como  $S''_3(x_0) = g_0 = 0$ ;  $S''_3(x_n) = g_n = 0$ , onde  $g_0$  e  $g_n$  são imagem da derivada segunda de  $S_3(x)$  nos pontos inicial e final do domínio de  $f(x)$ , então a *Spline* interpolante é chamada de *Spline natural*. Adotar tais condições implica em supor que os polinômios cúbicos tenderão à funções lineares nos intervalos extremos.

Para satisfazer as condições (i) à (v) de *Spline* interpolante e usando as notações  $s''_k(x_k) = g_k, f(x_k) = y_k$  e  $h(x_k) = x_k - x_{k-1}$  (para  $k = 0, 1, \dots, (n - 1)$ ), chega-se à:

$$a_k = \frac{g_k - g_{k-1}}{6h_k} \quad (24)$$

$$b_k = \frac{g_k}{2} h_k \quad (25)$$

$$c_k = \frac{y_k - y_{k-1}}{h_k} + \frac{2h_k g_k + g_{k-1} h_k}{6} \quad (26)$$

$$d_k = y_k \quad (27)$$

Pode-se verificar nas equações acima que os coeficientes ( $a_k, b_k, c_k, e d_k$ ) são dependentes de  $g_k$ . Assim, para se obter  $g_k$ , via condição (iv), tem-se:

$$h_k g_{k-1} + 2(h_k + h_{k+1})g_k + h_{k+1}g_{k+1} = 6 \left( \frac{y_{k+1} - y_k}{h_{k+1}} - \frac{y_k - y_{k-1}}{h_k} \right) \quad (28)$$

Considerando a variação de  $k$  ( $k = 1, \dots, (n - 1)$ ), tem-se que a equação 28 estabelece um sistema de equações lineares nas incógnitas  $g_0$  a  $g_n$ . Porém, cabe lembrar que a igualdade  $g_0 = g_n = 0$  foi estabelecida pelas condições complementares expressas anteriormente (*Spline* cúbica natural). Logo, matricialmente, pode-se estabelecer um sistema indeterminado  $AX = B$ , com:

$$X = [g_0 \quad g_1 \quad \dots \quad g_n]^T \quad (29)$$

$$A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ h_1 & 2(h_1 + h_2) & h_2 & & \\ & h_2 & 2(h_2 + h_3) & h_3 & \\ & & \ddots & \ddots & \ddots \\ 0 & 0 & & h_{n-1} & 2(h_{n-1} + h_n) & h_n \\ & & & & \dots & 1 \end{bmatrix}_{(n+1) \times (n+1)} \quad (30)$$

$$B = 6 \cdot \begin{bmatrix} 0 \\ \frac{y_2 - y_1}{h_2} - \frac{y_1 - y_0}{h_1} \\ \frac{y_3 - y_2}{h_3} - \frac{y_2 - y_1}{h_2} \\ \vdots \\ \frac{y_n - y_{n-1}}{h_n} - \frac{y_{n-1} - y_{n-2}}{h_{n-1}} \\ 0 \end{bmatrix}_{(n+1) \times 1} \quad (31)$$

### 2.4.2 *Basis Splines(B-Spline)*

O termo *B-Splines*, que decorre da contração da expressão *Basis Splines*, foi criado por Isaac Jacob Schoenberg. Segundo De Boor (2001), a *B-Splines* é um tipo de *Spline* na qual se tem controle local da curva, de forma que a alteração de um ponto de controle modifica a curva apenas em uma pequena vizinhança do referido ponto, em função da ordem de continuidade. As curvas *B-Splines* são curvas compostas, visto que são oriundas da junção de outras curvas que, neste caso, são de caráter polinomial e mantêm sua continuidade e regularidade.

As funções de base da curva *B-Splines* possuem duas propriedades dignas de nota: 1) Seu domínio é subdividido em nós; e 2) As funções de base não são diferentes de zero em todo o intervalo;

Para a obtenção de uma curva *B-Spline* são necessários três itens: um vetor de nós, as funções de base e um conjunto de pontos de controle.

Assim, genericamente, uma curva *B-Spline* de grau  $p$  qualquer (isto é, ordem  $p + 1$ ) é definida por:

$$C(u) = \sum_{i=0}^n N_{i,p}(u)P_i \quad (32)$$

Onde:

- $i$  é contador para os pontos de controle  $P_i (i = 1, \dots, n)$ ;
- $N_{i,p}$  denota as funções de base polinomiais de grau  $p$  (ou ordem  $p + 1$ ), definidas em um vetor com  $m + 1$  nós (tal que  $m \geq p$ ). O vetor de nós estabelece uma partição do domínio de tal modo que  $u_i \leq u_{i+1}, \forall_i$ . Cada  $N_{i,p}$  depende somente de  $p + 1$  valores do vetor de nós e independe da quantidade de pontos de controle.
- $P_i$  são pontos de controle que definem o polígono de controle e integram o vetor de pontos de controle.

Em relação ao vetor de nós (lista com valores  $u_j, j = 1, \dots, m + 1$ ), cabe dizer que, se  $u_i - u_{i+1} = a, \forall_i$ , tal que “ $a$ ” é constante, então o vetor de nós é dito uniforme. E mais, uma vez que a curvatura depende das primeiras e segundas derivadas, e a função *Spline* possui caráter polinomial, parece ser imediato que o grau a ser adotado pelo polinômio deva ser 3 ( $p = 3$ ).

Assim, as funções de base são computadas de forma recursiva, seguindo o algoritmo de De Boor (2001), da seguinte forma:

$$N_{i,p}(u) = \begin{cases} 1 & \text{para } u_i \leq u \leq u_{i+1} \\ 0 & \text{caso contrário} \end{cases} \quad (33)$$

$$N_{i,p} = \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u) \quad (34)$$

Pereira e Carvalho (2014) especificam que:

- $N_{i,0}(u)$  é uma função pulso, que é sempre igual a zero exceto para valores de  $u$  (nó) no intervalo  $[u_i, u_{i+1})$ ;
- Para  $p > 0$ ,  $N_{i,p}(u)$  uma combinação linear de duas funções de base de grau  $p - 1$ ;
- O cálculo de um conjunto de funções de base requer a especificação de um vetor de nós  $U$  e o grau  $p$ ;
- A equação que estabelece  $N_{i,p}(u)$  pode levar a quocientes do tipo 0/0 (indeterminação). Esses quocientes são definidos como zero;
- As funções de base  $N_{i,p}(u)$  são polinomiais por partes, definidas em todo domínio real. Geralmente somente o intervalo  $[u_0, u_m]$  é de interesse;
- O intervalo  $[u_i, u_{i+1})$  é denominado  $i$ -ésimo intervalo do nó. Este intervalo pode ter comprimento nulo, desde que os nós não necessitem ser distintos.

Para o cálculo de uma curva *B-Spline*, Piegl e Tiller (1997) especificam que são necessárias três etapas:

1. Encontrar o vão no qual o nó  $u_i$  se encontra, dentro do vetor de nós  $U$ ;
2. Calcular as funções de base  $N_{i,p}(u)$  não nulas;
3. Multiplicar as funções de base não nulas pelos respectivos pontos de controle.

De Boor (2001) especifica ainda que uma curva *B-Spline* uniforme apresenta alguns elementos que podem ser mudados, conforme a necessidade:

- Os pontos de controle podem ser alterados a fim de gerar alterações de forma necessárias ou convenientes;
- Colocar pontos de controle repetidos implica em fazer com que a curva se aproxime destes pontos;
- Colocar pontos de controle repetidos em quantidade suficiente faz com que a curva passe por estes pontos;
- Aumentar o grau do polinômio, eleva a continuidade da curva nos nós, e acrescenta a suavidade da curva e tende a mover a curva para além do polígono definido pelos pontos de controle.
- Pode-se fechar o polígono definido pelos pontos de controle repetindo ordenadamente, ao final da lista de pontos do vetor de controle, os “ $p$ ” pontos do início deste vetor.

## 2.5 MÉTODOS DE INFERÊNCIA ESTATÍSTICA

Neste trabalho, para cada componente de imagem R, G e B e para a imagem de luminância (L) são estimados 4 modelos polinomiais, um para cada método utilizado. Cada modelo possui um conjunto com  $n+1$  coeficientes  $\beta_j$  ( $j = 0, \dots, n$ ) particulares. Especifica-se que,  $n < 256$  e que 256 é o número de observações amostradas, ou seja, os graus de liberdade do problema abordado. Para fins de avaliação de desempenho, cada conjunto de coeficientes é submetido à validação via “teste t”. Neste caso, o teste possui as seguintes hipóteses a serem consideradas:

$$H_0: B_j = 0, \text{ com } j = 0, \dots, n$$

$$H_1: B_j \neq 0, \text{ com } j = 0, \dots, n$$

O seguinte teste geral é aplicado a cada coeficiente:

$$t(\hat{\beta}_j) = \frac{(\hat{\beta}_j - \beta_j)}{dp(\beta_j)} \quad (35)$$

Deve-se considerar que a sentença acima possui distribuição t-Student, com  $n - j$  graus de liberdade associados à estimativa da variância residual de cada modelo.

O resumo dos modelos distintos produzidos e analisados é apresentado pela tabela 1.

**Tabela 1** – Modelos analisados para cada componente da imagem.

VARIÁVEL INDEPENDENTE	VARIÁVEL RESPOSTA	AJUSTES POLINOMIAIS
P	COMPONENTES (R, G, B e L)	MMQ
P	COMPONENTES (R, G, B e L)	<i>B-SPLINES</i> + MMQ
P	COMPONENTES (R, G, B e L)	MV
P	COMPONENTES (R, G, B e L)	<i>B-SPLINE</i> + MV

Fonte: O autor.

São capturados de cada modelo os resíduos quadráticos  $e_i^2$ , por meio da diferença quadrática de valores amostrados e estimados dos ajustes de cada componente para 256 intensidades de brilho, conforme a equação:



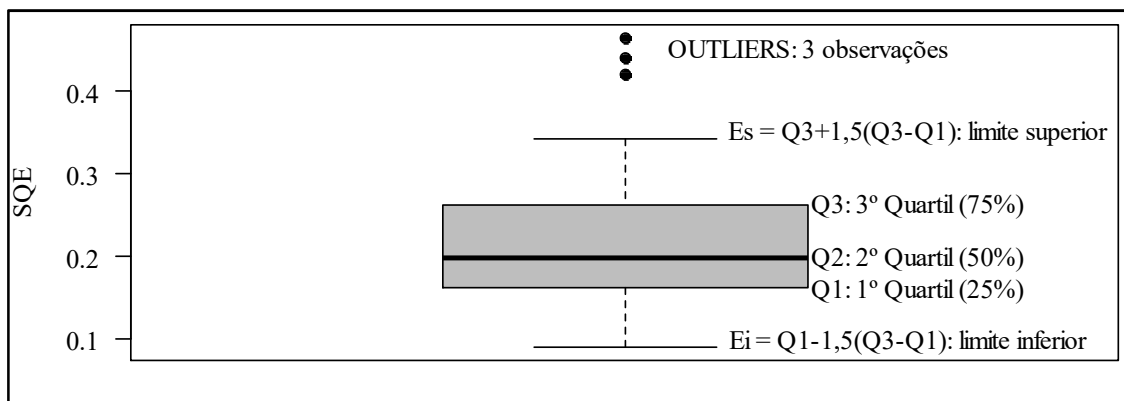
$$e_i^2 = (Y_i - \hat{Y}_i)^2 \quad (36)$$

Onde:

- $Y_i$  são os valores de frequência dos níveis de brilho observados nos histogramas;
- $\hat{Y}_i$  são valores de frequência dos níveis de brilho estimados a partir das fórmulas polinomiais geradas.

Para fins de análise visual, são construídos gráficos do tipo *Boxplot* para as variáveis “Soma Quadrática dos Erros” (*SQE*) de cada um dos modelos, afim de se ter uma ideia do comportamento e distribuição dos resíduos quadráticos dos ajustes para cada componente. A Figura 7 ilustra o gráfico *Boxplot* e resume as informações nele contidas:

**Figura 7** – Gráfico ilustrativo do tipo *Boxplot* da variável *SQE*.



Fonte: O autor.

Para apoiar a análise numérica, a partir da variável ( $e_i^2$ ), são calculados a soma dos quadrados dos erros (*SQE*), a média do quadrado dos erros (*MÉDIA*), o desvio-padrão do quadrado dos erros (*DP*), o coeficiente de variação (*CV*) e  $R^2$ , conforme as respectivas equações:

$$SQE = \sum_{i=1}^{256} e_i^2 = \sum_{i=1}^{256} (Y_i - \hat{Y}_i)^2 \quad (37)$$

$$MÉDIA = \frac{SQE}{256} \quad (38)$$

$$DP = \sqrt{\frac{\sum_{i=1}^{256} (e_i^2 - MÉDIA)^2}{256 - 1}} \quad (39)$$

$$CV = \left( \frac{DP}{MÉDIA} \right) \cdot 100 \quad (40)$$

$$R^2 = \frac{SQMODELO}{SQTOTAL} \quad (41)$$

Cabe lembrar que, na equação 41, o numerador (*SQMODELO*) e o denominador (*SQTOTAL*), são derivados de expressões advindas do ajuste clássico do MMQ. Assim, tem-se que:

$$\sum_{i=1}^{256} (Y_i - \bar{Y})^2 = \sum_{i=1}^{256} (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{256} (Y_i - \hat{Y}_i)^2 \quad (42)$$

Desta forma, na expressão 41 tem-se:

$$SQTOTAL = SQMODELO + SQE \quad (43)$$

Com:

$$SQTOTAL = \sum_{i=1}^{256} (Y_i - \bar{Y})^2 \quad (44)$$

$$SQMODELO = \sum_{i=1}^{256} (\hat{Y}_i - \bar{Y})^2 \quad (45)$$

$$SQE = \sum_{i=1}^{256} (Y_i - \hat{Y}_i)^2 \quad (46)$$

Logo, a partir das estimativas relativas aos erros quadráticos de cada modelo e considerando as equações (37), (38), (39), (40) e (41), é possível se ter noção do comportamento do resíduo de cada ajuste, quanto a sua variabilidade e sua influência/perturbação em cada modelo, como também, o grau da representatividade de cada ajuste. Ou seja, pode-se ter noção de qual modelo possuirá o melhor ajuste. Para tanto, as comparações listadas na tabela 2 foram consideradas.

**Tabela 2** – Comparação de Qualidade do ajuste de cada modelo.

<b>ESTATÍSTICAS BASEADAS EM <math>e_i^2</math></b>	<b>MMQ</b>	<b><i>B-SPLINES</i> +MMQ</b>	<b>MV</b>	<b><i>B-SPLINES</i> +MV</b>
<b>SQE</b>	P x Frequência	P x Frequência	P x Frequência	P x Frequência
<b>MÉDIA</b>	P x Frequência	P x Frequência	P x Frequência	P x Frequência
<b>DP</b>	P x Frequência	P x Frequência	P x Frequência	P x Frequência
<b>CV</b>	P x Frequência	P x Frequência	P x Frequência	P x Frequência
<b>R<sup>2</sup></b>	P x Frequência	P x Frequência	P x Frequência	P x Frequência

Fonte: O autor.

### 3 METODOLOGIA

Este capítulo destina-se à apresentação dos procedimentos metodológicos realizados a fim de que os objetivos expressos fossem alcançados. Para tanto, na seção 3.1 estão os procedimentos para a geração dos histogramas dos canais vermelho, verde e azul de uma imagem colorida de entrada e de sua respectiva imagem de luminância, calculada com base nos canais particionados. Na seção 3.2 estão descritos os pormenores da geração de formulação de funções de domínio contínuo a partir de funções de domínio discreto (histogramas). As principais formas de avaliação dos resultados alcançados estão na seção 3.3.

#### 3.1 GERAÇÃO DOS DADOS DE ENTRADA: HISTOGRAMAS

Cabe inicialmente explicar que, o objetivo de um método de ajuste, quando aplicado a um modelo matemático, é encontrar estimativas dos parâmetros de uma variável resposta qualquer, a partir de valores médios de uma variável independente, minimizando os seus desvios, popularmente conhecidos como erros ou resíduos das estimativas de ajustes.

Assim convém recordar que, o presente trabalho teve como escopo principal a geração de formulação matemática de funções de domínio contínuo a partir de funções de domínio discreto por meio de alguns métodos de aproximação/interpolação a saber: *B-Splines*, Método dos Mínimos Quadrados (MMQ) e Método da Máxima Verossimilhança (MV). Para tanto, é necessário que se estabeleça quais foram os dados de entrada para os referidos métodos.

Para a pesquisa, foram utilizadas como funções de domínio discreto (dados de entrada) os histogramas dos canais vermelho, verde e azul (ou em inglês: *Red*, *Green* e *Blue*) de imagens coloridas de entrada, e também o histograma da imagem de luminância, gerado a partir dos canais da respectiva imagem de entrada. Para tanto, foi realizada inicialmente a implementação de um *software* em linguagem C++ que fosse capaz de:

- 1) Ler uma imagem colorida de tipo RAW a partir do sistema de arquivos. Tais imagens tem resolução radiométrica de 8 *bits* por canal. Assim, foram gerados histogramas com 256 intensidades de brilho (cardinalidade do domínio da função discreta);
- 2) Particionar a imagem lida em canais vermelho, verde e azul, que a compõem;

- 3) Gerar um histograma para cada canal por meio da contagem das quantidades de *pixels* relativas a cada uma das 256 intensidades de brilho (Imagem da função de domínio discreto);
- 4) Gerar a imagem de luminância relativa à imagem de entrada;
- 5) Gerar o histograma da imagem de luminância (como o proposto no passo 3).
- 6) Armazenar os dados relativos aos histogramas em arquivos do tipo ASCII.

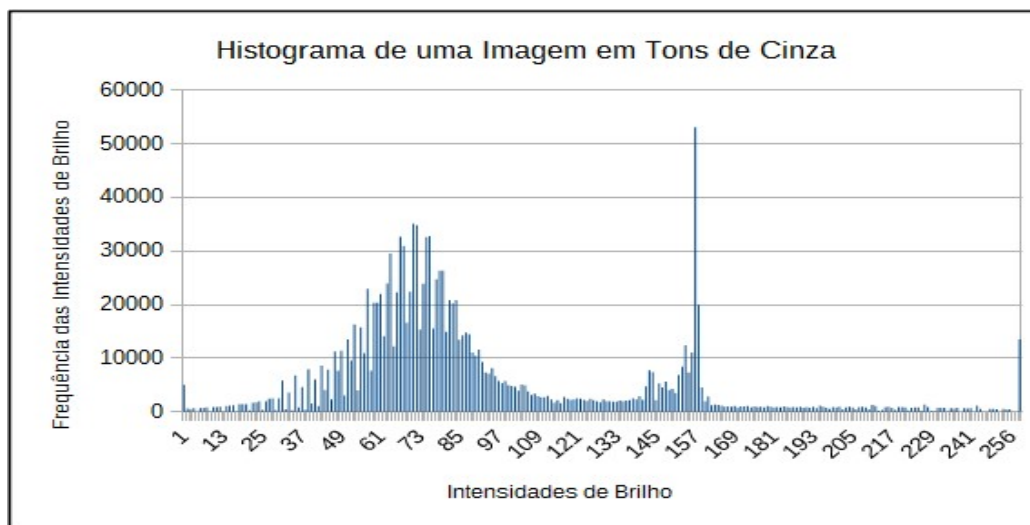
Os histogramas gerados constituiram os dados de entrada para os métodos de aproximação/intrepolação utilizados. Para fins de explanação, as componentes vermelho, verde e azul da imagem de entradas serão referidas, respectivamente, por R, G e B. Da mesma forma, far-se-á referência à imagem de Luminância por meio da letra L. Para fins de ilustração dos dados de entrada tem-se as Figuras 8 e 9, que apresentam, respectivamente, uma imagem em tons de cinza (componente R) e seu histograma (representação gráfica).

**Figura 8** – Imagem em tons de cinza: componente R.



Fonte: Resultados da pesquisa.

Por fim, os dados gerados foram organizados em planilhas eletrônicas e preparados para posterior processamento no *software* RStudio. Para tanto, apesar das componentes R, G, B e L possuírem 256 níveis intensidades de brilho, graduadas por valores inteiros que variam de 0 (zero) a 255, para fins de processamento utilizou-se uma graduação iniciada por 1 (um) e que variou em valores inteiros até 256, como mostra a Figura 9.

**Figura 9** – Histograma da imagem em tons de cinza (Figura 8)

Fonte: O autor.

Expressa-se que, apesar de se ter utilizado histogramas de imagens digitais como dados de entrada, a metodologia aqui descrita pode ser aplicada a quaisquer funções de domínio discreto provenientes das mais variadas origens e com cardinalidade de domínio apropriada.

Para fins de explanação, os pontos originais provenientes dos histogramas serão ditos pontos de controle.

### 3.2 DETERMINAÇÃO DE FÓRMULAS PARA AS FUNÇÕES POLINOMIAIS

Cabe esclarecer inicialmente que, para os ajustes polinomiais foram utilizados isoladamente os Métodos dos Mínimos Quadrados (MMQ) e da Máxima Verossimilhança (MV). Já o método *B-Splines* foi utilizado em conjunto com os dois métodos anteriores. Ou seja, foram geradas formulações polinomiais a partir da combinação *B-Splines* e MMQ, e da combinação *B-Splines* e MV. Assim, nos testes que envolveram *B-Splines*, houve o intuito de causar inicialmente um adensamento no conjunto de pontos a ser processado pelas metodologias que a sucederam. Tal decisão objetivou a verificação de uma possível melhora das estimativas geradas pelo MMQ ou pelo método MV.

A decisão por esta estratégia foi motivada pela constatação do caráter estocástico que o MMQ e o método MV possuem. Ou seja, segundo Pereira e Carvalho (2014), nestes métodos é possível ponderar os dados de entrada de modo a “forçar” que a curva obtida a partir do polinômio gerado passe ou se aproxime de um determinado ponto. Tal fato não ocorre tão

facilmente por meio do método *B-Splines*, que gera curvas que não passam necessariamente pelos pontos de controle.

Assim, inicialmente, por meio do método clássico do MMQ, na sua forma paramétrica, investigou-se qual o grau do polinômio que melhor se ajustaria à série de 256 intensidades de brilho ( $P_i$ ) de *pixels*, considerando os histogramas/variação para cada uma das 3 componentes da imagem de entrada (R, G e B) e para a imagem de luminância (L). E mais, os experimentos consideraram e testaram a super especificação do modelo polinomial utilizado. As seguintes equações mostram as formas polinômiais consideradas:

$$\hat{R}_i = f_R(P_i, \beta_{R0}, \dots, \beta_{Rn}) = \beta_{R0} + \beta_{R1} \cdot P_i + \dots + \beta_{Rn} \cdot P_i^n + e_{Ri} \quad (47)$$

$$\hat{G}_i = f_G(P_i, \beta_{G0}, \dots, \beta_{Gn}) = \beta_{G0} + \beta_{G1} \cdot P_i + \dots + \beta_{Gn} \cdot P_i^n + e_{Gi} \quad (48)$$

$$\hat{B}_i = f_B(P_i, \beta_{B0}, \dots, \beta_{Bn}) = \beta_{B0} + \beta_{B1} \cdot P_i + \dots + \beta_{Bn} \cdot P_i^n + e_{Bi} \quad (49)$$

$$\hat{L}_i = f_L(P_i, \beta_{L0}, \dots, \beta_{Ln}) = \beta_{L0} + \beta_{L1} \cdot P_i + \dots + \beta_{Ln} \cdot P_i^n + e_{Li} \quad (50)$$

Onde:

- $\hat{R}_i, \hat{G}_i, \hat{B}_i$  e  $\hat{L}_i$  são os valores estimados das componentes *R, G, B* e *L* (frequências reais das intensidades verificadas nos histogramas) a partir da formas polinômiais das funções  $f_R, f_G, f_B$  e  $f_L$ , respectivamente, considerando a intensidade de brilho  $P_i$ ;
- $P_i$  são as intensidades dos níveis de brilho para cada componente de cor e da luminância;
- $\beta_{Rj}, \beta_{Gj}, \beta_{Bj}$  e  $\beta_{Lj}$  são os coeficientes dos polinômios que constituem as formulações das funções que aproximam *R, G, B* e *L*. Tem-se que  $j = 0, \dots, n$ , sendo  $n < 256$  o grau do polinômio;
- $e_{Ri}, e_{Gi}, e_{Bi}$  e  $e_{Li}$ , são os erros ou discrepâncias entre os valores de cada uma das funções discretas e sua respectiva aproximação dada pela função polinomial encontrada. Tem-se que  $i = 1, \dots, 256$ ;

Uma vez que, o melhor ajuste de cada polinômio (determinação do grau  $n$ ) para cada histograma em estudo tenha sido encontrado, realizou-se novas estimativas dos coeficientes dos polinômios que definem cada função por meio dos demais métodos citados, considerando o mesmo grau  $n$ , determinado inicialmente pelo MMQ.

A escolha do “melhor ajuste” deu-se por meio da análise dos erros ( $e_{Ri}$ ,  $e_{Gi}$ ,  $e_{Bi}$  e  $e_{Li}$ ). Assim, foi considerado o grau do polinômio que apresentou menor erro, quando da comparação de suas imagens com os dados originais (histogramas).

Desta forma, houve a possibilidade de comparação de cada modelo e suas nuances explicativas, conforme o resumido na tabela 3.

**Tabela 3** – Rol de comparações entre modelos funcionais.

<b>VARIÁVEL INDEPENDENTE (Intensidades de brilho)</b>	<b>VARIÁVEL REPOSTA (Frequência das intensidades)</b>	<b>AJUSTES POLINOMIAIS (Métodos de ajuste)</b>
P	R	MMQ
P	R	<i>B-SPLINES</i> +MMQ
P	R	MV
P	R	<i>B-SPLINES</i> +MV
P	G	MMQ
P	G	<i>B-SPLINES</i> +MMQ
P	G	MV
P	G	<i>B-SPLINES</i> +MV
P	B	MMQ
P	B	<i>B-SPLINES</i> +MMQ
P	B	MV
P	B	<i>B-SPLINES</i> +MV
P	L	MMQ
P	L	<i>B-SPLINES</i> +MMQ
P	L	MV
P	L	<i>B-SPLINES</i> +MV

Fonte: O autor.

As funções polinomiais geradas (tabela 3) foram avaliadas quanto a sua qualidade de ajuste por meio dos coeficientes polinomiais estimados e sua validade. As avaliações também se deram por meio de comparações de parâmetros de qualidade estimados com base em cada função polinomial obtida por cada um dos métodos citados.

### 3.3 AVALIAÇÃO DOS RESULTADOS

A análise dos resultados se deu de forma visual e numérica. A análise visual foi apoiada por gráficos relativos aos ajustes e aos erros. Os parâmetros numéricos utilizados para a comparação de desempenho das funções polinomiais estabelecidas foram calculados com



relação aos erros (discrepâncias entre os valores estimados e os valores advindos dos histogramas), para cada modelo gerado, e foram os seguintes: a soma dos quadrados dos erros (*SQE*), a média aritmética do quadrado dos erros (*MÉDIA*), o desvio-padrão do quadrado dos erros (*DP*), o coeficiente de variação (*CV*) e o coeficiente  $R^2$ . Cabe especificar que, o  $R^2$  quantifica/qualifica, percentualmente, a qualidade do ajuste para cada curva. Cada componente da imagem e a imagem de luminância tiveram sua respectiva forma funcional analisada objetivando a inferência da qualidade do ajuste por meio dos parâmetros acima citados, conforme ilustra a tabela 4.

**Tabela 4** – Comparação das Estimativas de cada ajuste.

ÍNDICES DE VARIAÇÃO	MMQ	<i>B-SPLINES</i> + MMQ	MV	<i>B-SPLINE</i> +MV
<b>SQE</b>	P contra R	P contra R	P contra R	P contra R
<b>MÉDIA</b>	P contra R	P contra R	P contra R	P contra R
<b>DP</b>	P contra R	P contra R	P contra R	P contra R
<b>CV</b>	P contra R	P contra R	P contra R	P contra R
<b>R<sup>2</sup></b>	P contra R	P contra R	P contra R	P contra R
<b>SQE</b>	P contra G	P contra G	P contra G	P contra G
<b>MÉDIA</b>	P contra G	P contra G	P contra G	P contra G
<b>DP</b>	P contra G	P contra G	P contra G	P contra G
<b>CV</b>	P contra G	P contra G	P contra G	P contra G
<b>R<sup>2</sup></b>	P contra G	P contra G	P contra G	P contra G
<b>SQE</b>	P contra B	P contra B	P contra B	P contra B
<b>MÉDIA</b>	P contra B	P contra B	P contra B	P contra B
<b>DP</b>	P contra B	P contra B	P contra B	P contra B
<b>CV</b>	P contra B	P contra B	P contra B	P contra B
<b>R<sup>2</sup></b>	P contra B	P contra B	P contra B	P contra B
<b>SQE</b>	P contra L	P contra L	P contra L	P contra L
<b>MÉDIA</b>	P contra L	P contra L	P contra L	P contra L
<b>DP</b>	P contra L	P contra L	P contra L	P contra L
<b>CV</b>	P contra L	P contra L	P contra L	P contra L
<b>R<sup>2</sup></b>	P contra L	P contra L	P contra L	P contra L

Fonte: O autor

Por fim, da análise de todas as informações geradas sobre os ajustes, foi possível depreender qual o melhor modelo/método para cada componente em foco.

## 4 RESULTADOS E ANÁLISES

Neste capítulo encontram-se os principais resultados e análises advindos da execução da pesquisa. Inicialmente, na seção 4.1 são apresentados os *softwares* utilizados para a geração dos resultados apresentados e são descritos os dados utilizados para a computação da metodologia apresentada. Na seção 4.2 os principais resultados são apresentados e discutidos.

### 4.1 ASPECTOS COMPUTACIONAIS

#### 4.1.1 *Softwares*

Para o processamento da imagem de entrada e a geração dos dados (histogramas) foi implementado um programa em linguagem C++. O *software* implementado materializa os passos expressos na seção 3.1 e trata-se de uma aplicação console que, após a leitura da imagem RAW de entrada, constrói os histogramas citados anteriormente e os armazena em arquivos do tipo ASCII. A implementação do *software* de processamento das imagens se deu no Compilador Borland C++ Builder - versão 6.0 – *Enterprise Suite* (domínio público).

Uma vez gerados os histogramas, estes foram organizados para fins de processamento no *software* RStudio. O *software* livre RStudio é um ambiente de desenvolvimento integrado (ou, em inglês, IDE - *Integrated Development Environment*) que trabalha com a linguagem R. O R é uma linguagem de programação destinada inicialmente à geração de gráficos e cálculos estatísticos. No entanto, as potencialidades do R não se restringem a cálculos estatísticos, uma vez que, se pode fazer o *download* gratuito de “pacotes” que contém ferramentas computacionais voltadas às mais diferentes áreas do conhecimento.

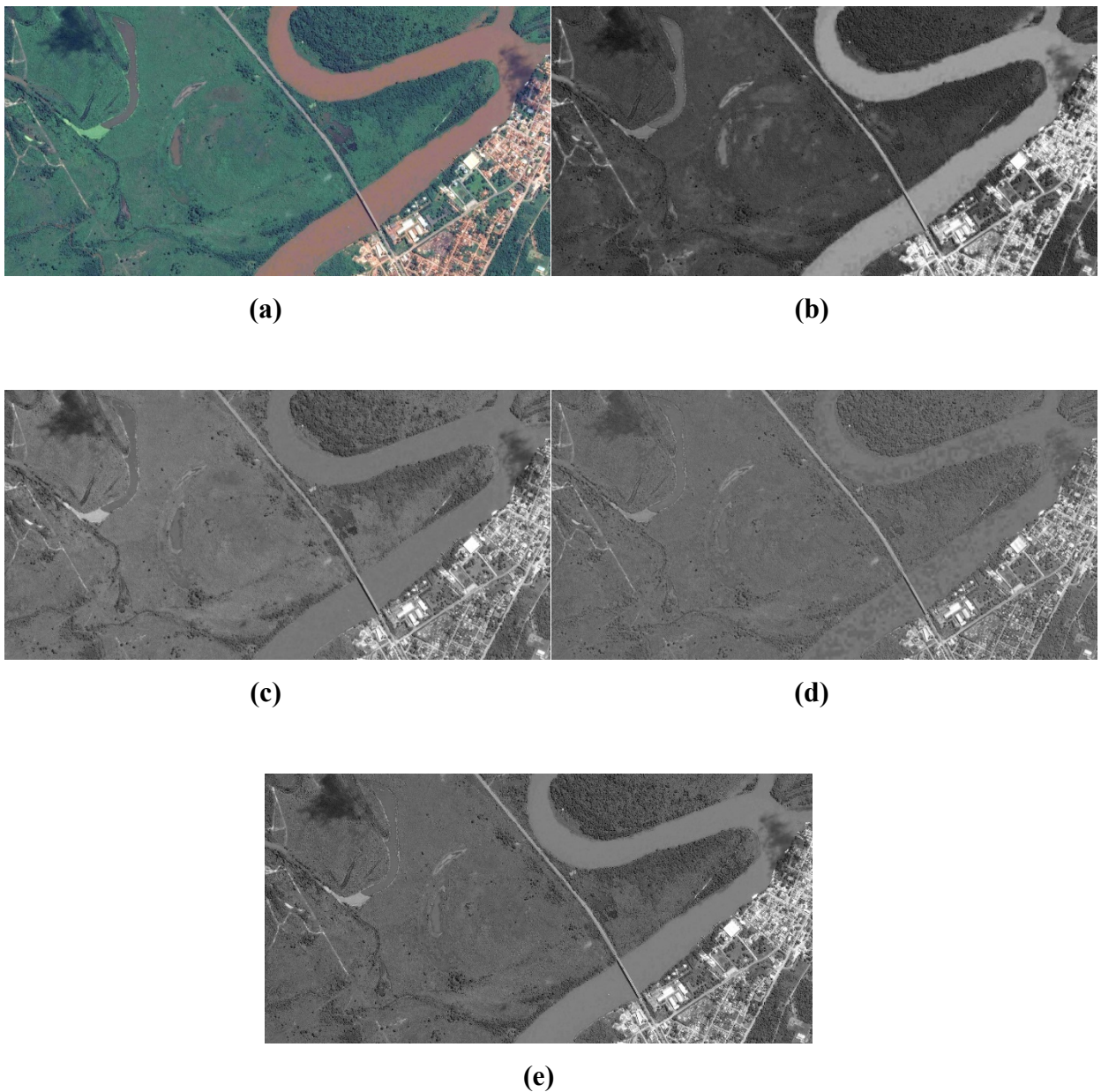
Para que o RStudio funcione adequadamente, deve-se instalar primeiramente o *software* R, que materializa a linguagem R, e posteriormente o programa RStudio. O programa R, que é escrito em linguagem de programação C++, é disponibilizado no site do CRAN (*The Comprehensive R Archive Network*). Já o RStudio tem uma versão *desktop* gratuita disponível em seu *site* (<https://www.RStudio.com/products/RStudio/#Desk>).

### 4.1.2 Dados

Considerando que, para cada imagem de entrada são gerados quatro histogramas, optou-se por se tomar apenas uma imagem para fins de experimentação. Tal imagem foi obtida do *Google Earth*.

A imagem escolhida, suas componentes e a respectiva imagem de luminância são apresentadas na Figura 10.

**Figura 10** – Dados de Entrada: **a)** Imagem colorida; **b)** Componente R; **c)** Componente G; **d)** Componente B; **e)** Imagem de Luminância.



Fonte: Resultados da pesquisa.

Especifica-se que a imagem utilizada possui dimensão de 1667x820 *pixels* e mostra a tomada aérea de uma rodovia do município de Cáceres – MT. A dita rodovia transpõe o rio Paraguai e é ladeada por edificações (cidade de Cáceres), vegetação e o próprio rio.

Os valores relativos aos histogramas das imagens estão no Apêndice I.

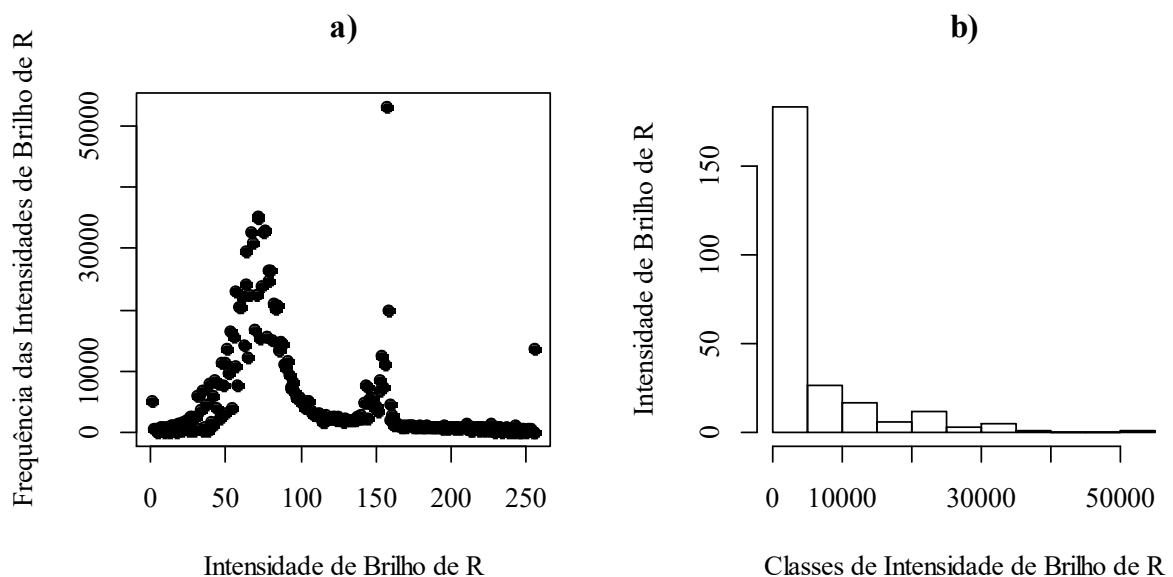
## 4.2 RESULTADOS EXPERIMENTAIS E ANÁLISE

Nas subseções que seguem são apresentados e discutidos os principais resultados advindos da pesquisa. Tais resultados encontram-se distribuídos conforme a ordem predita na metodologia. Assim, inicialmente são apresentados e discutidos os resultados obtidos a partir do histograma relativo à componente R e, na sequência, o mesmo é feito para as componentes G e B, e para a imagem de luminância (L).

### 4.2.1 Experimento com a componente R

Primeiramente, foi analisado o comportamento da variável independente “intensidade de brilho” (P) com relação à variação das frequências destas intensidades de brilho relativas à componente R. Na Figura 11 são apresentados o gráfico de dispersão dos valores de P contra R (Figura 11 a)) e também o histograma de R contra P (Figura 11 b)).

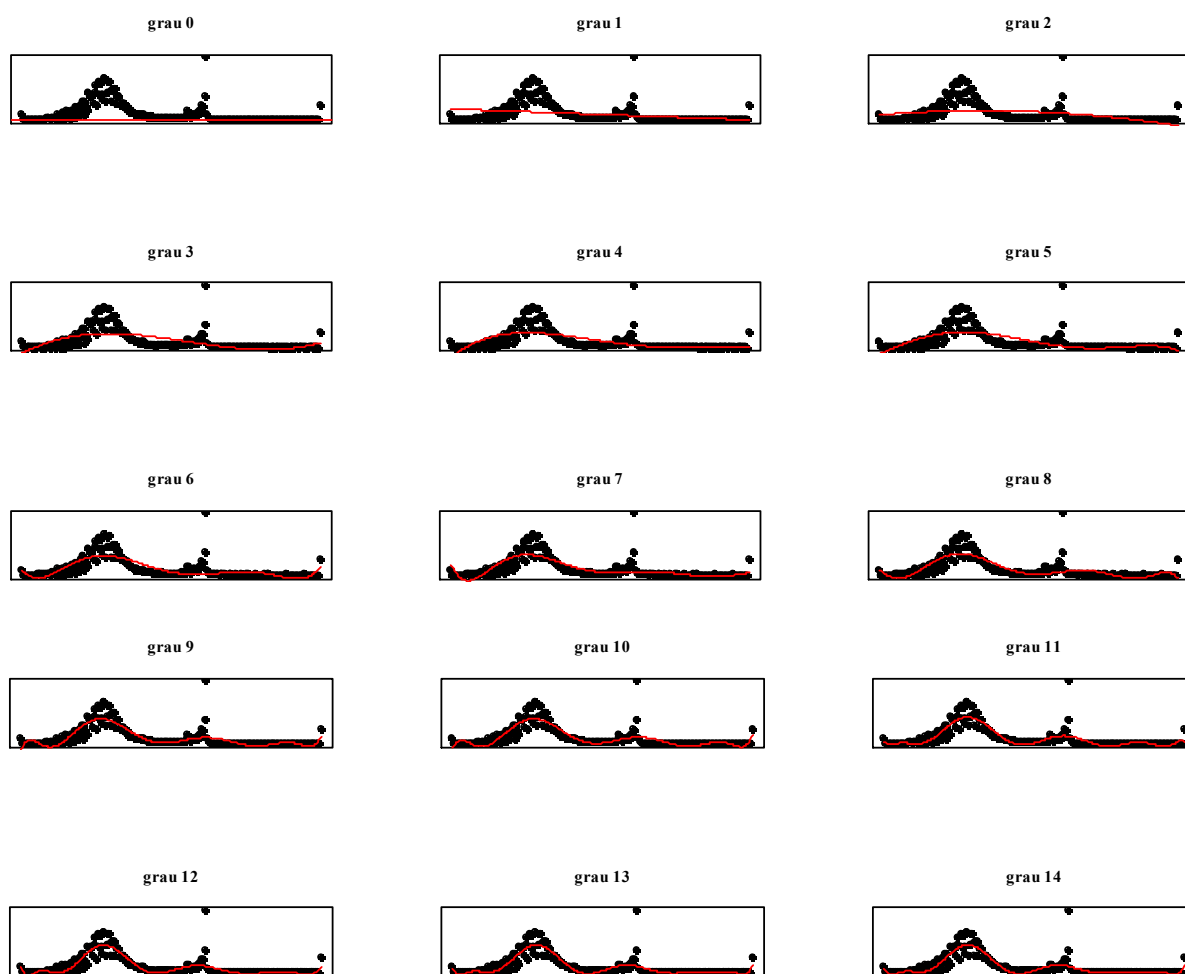
**Figura 11** – Comportamento da componente R: **a)** Gráfico de Dispersão (PxR); **b)** Histograma de R (RxP).

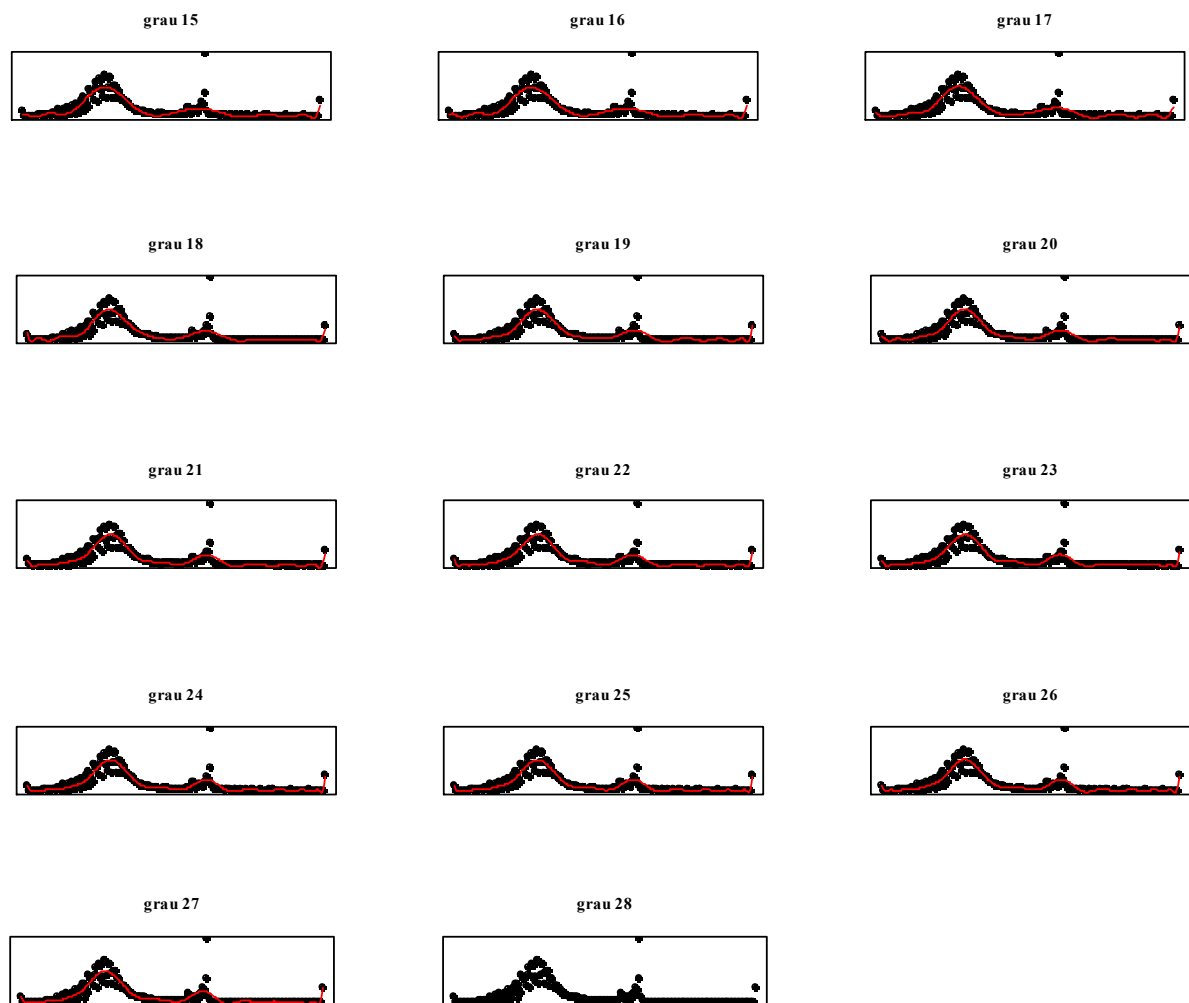


Fonte: Resultados da pesquisa.

A partir da análise visual dos gráficos presentes na Figura 11 pôde-se ter uma ideia do comportamento variacional das frequências das 256 intensidades de brilho (P) que integram a componente R da imagem de entrada. Assim, por meio de exaustivos ajustes pelo método dos mínimos quadrados ordinários (MMQ), verificou-se que um polinômio de grau 27 poderia vir a fornecer uma curva que representasse adequadamente os dados de entrada. A Figura 12 mostra os esboços dos gráficos relativos aos ajustes realizados via MMQ para a determinação do grau do polinômio a ser utilizado nos experimentos posteriores. Nestes esboços, as curvas em vermelho representam os ajustes polinomiais e os pontos pretos simbolizam os pontos de controle iniciais. Cabe esclarecer que, a realização de sucessivos ajustes objetivaram a super especificação de um grau de polinômio, considerando a minimização das discrepâncias (erros) entre os valores obtidos por meio do dito polinômio e os dados originais.

**Figura 12** – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio.





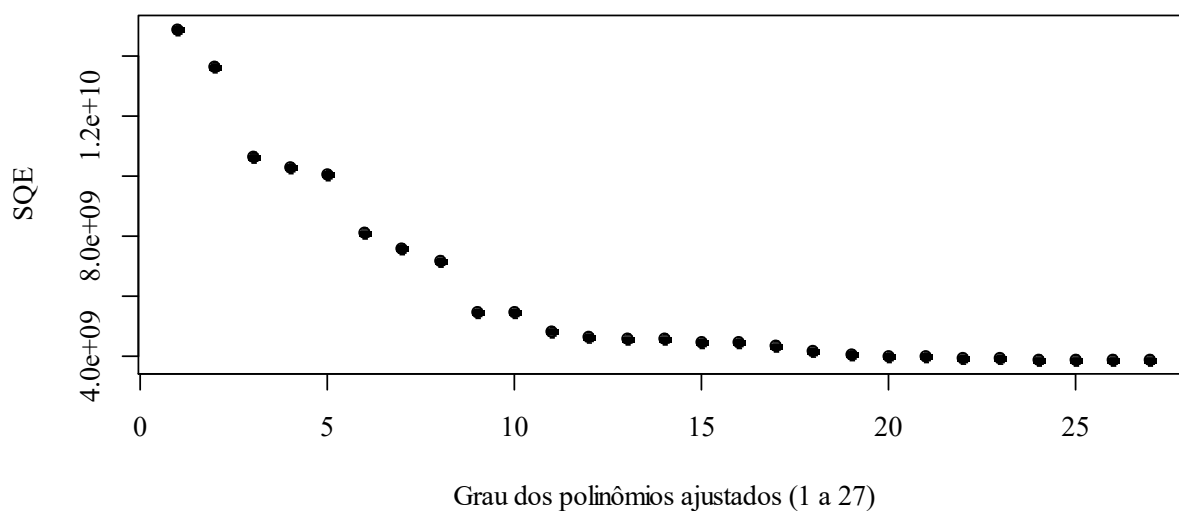
Fonte: Resultados da pesquisa.

Observa-se que, aparentemente, quanto maior a especificação do modelo, ou seja, quanto maior for o número das componentes dos polinômio que explicam a variação de  $R$ , melhor será o ajuste da curva. Cabe expressar que o *software* RStudio efetuou ajustamentos somente até o polinômio de grau 27 (ver dois últimos esboços da Figura 12). A partir deste limite não há ajuste polinomial via MMQ. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada pelo RStudio. Acredita-se, especulativamente, que tal fato tenha ocorrido devido à extrapolação dos recursos de *hardware* ou da existência de alguma limitação de *software*.

Dizer que um melhor ajuste ocorreu, equivale a expressar que o erro ou resíduo, associado a cada modelo subsequentemente calculado, diminuiu. Este comportamento de decréscimo do erro pode ser avaliado pela soma dos erros quadráticos de cada modelo e não apenas pela simples soma dos resíduos, pois, em ajustes de curvas, a soma total do resíduo de

cada modelo será sempre nula. O gráfico da Figura 13 mostra o comportamento assintótico de redução da soma do erro quadrático capturado de cada modelo, mediante o aumento do grau do polinômio ajustado.

**Figura 13** – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.

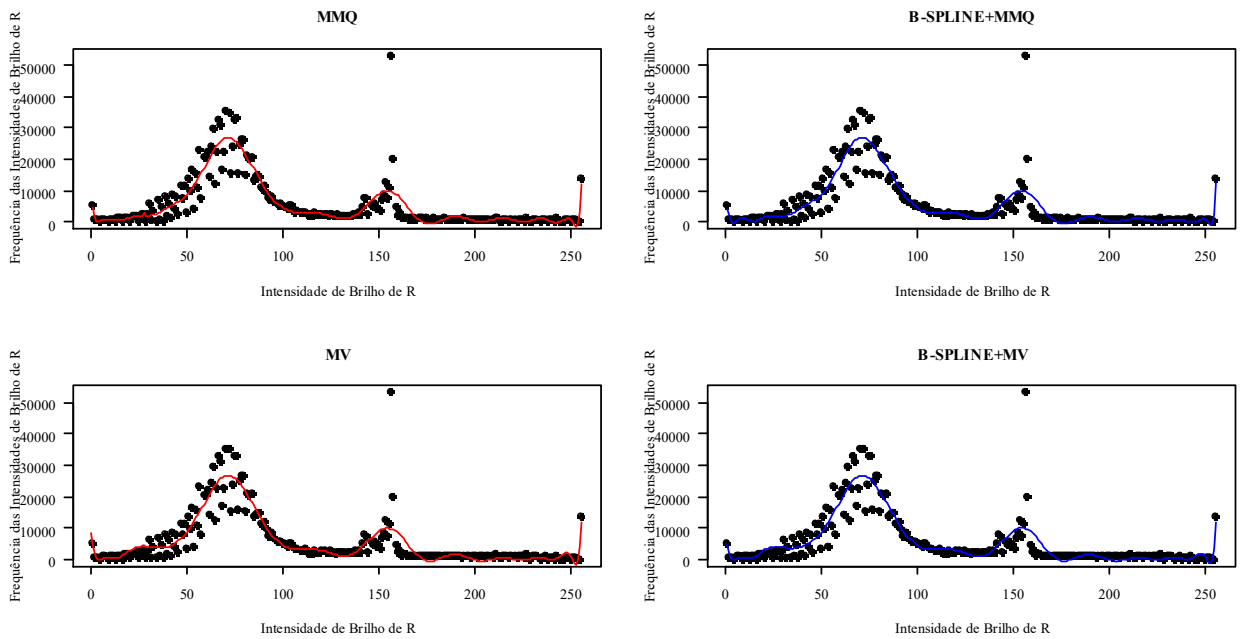


Fonte: Resultados da pesquisa.

De certa forma, com aumento do grau do polinômio observa-se uma melhora do ajuste. Tal acréscimo de qualidade é indicado pela diminuição do valor dos resíduos gerados pelo ajuste via MMQ. Assim verifica-se que, tomar o polinômio com grau 27 implica em trabalhar com um modelo que possui uma componente erro bastante reduzida, em comparação com as demais componentes representadas graficamente por seus erros na Figura 13.

Desta forma, mediante a adoção de um polinômio de grau 27 como modelo, foram analisados os outros tipos de ajustes previstos na metodologia, com intuito de averiguar suas respectivas precisões, vantagens e desvantagens. Cabe recordar que, os métodos de ajustes efetuados e analisados foram: o Método dos Mínimos Quadrados Ordinários (MMQ), *B-Splines* acoplado ao Método dos Mínimos Quadrados Ordinários (*B-SPLINES*+MMQ), o método da Máxima Verossimilhança (MV) e *B-Splines* acoplado ao método de Máxima Verossimilhança (*B-SPLINES*+MV). Os gráficos da Figura 14 apresentam as curvas obtidas destes ajustes (curvas em vermelho) sobrepostas aos gráficos de dispersão para fins de análise visual comparativa.

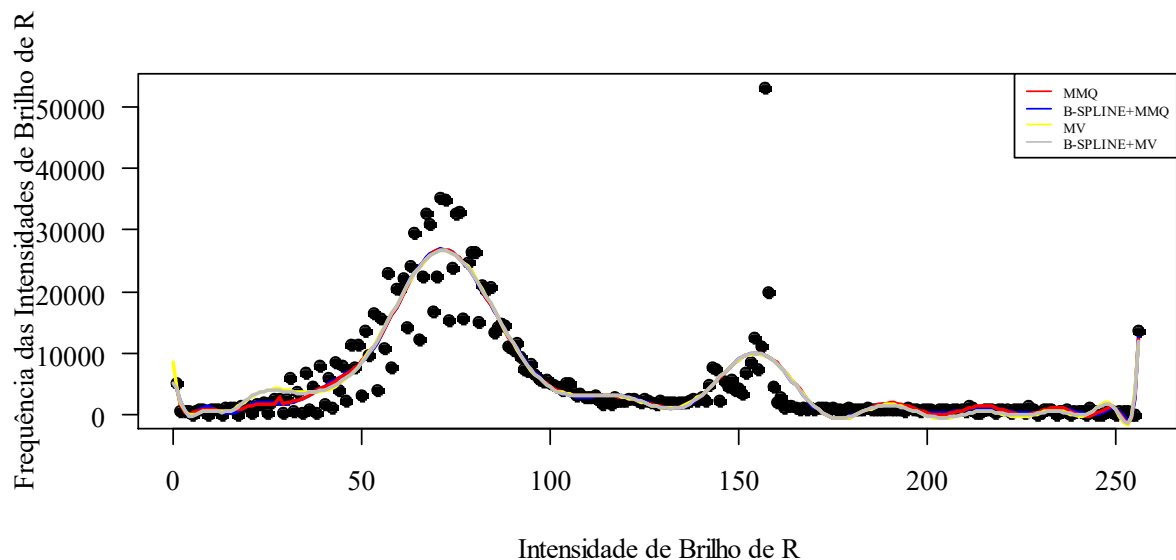
**Figura 14** – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.



Fonte: Resultados da pesquisa.

Da análise visual dos gráficos presentes na Figura 14, infere-se que haja alguma semelhança entre as curvas obtidas dos diferentes ajustes. A fim de propiciar uma melhor análise visual e favorecer a identificação das discrepâncias entre as curvas, tem-se o gráfico da Figura 15, no qual as curvas estão plotadas conjuntamente e em diferentes cores.

**Figura 15** – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente R.



Fonte: Resultados da pesquisa.



Observando o gráfico da Figura 15, pode-se exarar que realmente as curvas geradas a partir das formas polinomiais obtidas são bastante semelhantes, apesar da distinção dos métodos utilizados. Porém, a escolha do melhor modelo gerado deve ainda considerar a quantificação da significância estatísticas dos parâmetros dos ajustes (p-valor/probabilidade de se errar nas estimativas geradas) e a análise dos erros. Assim, pode-se chegar a um ranqueamento dos ajustes.

As diferenças entre os modelos obtidos podem ser melhor avaliadas por meio da análise/comparação numérica dos coeficientes gerados para cada polinômio representativo da componente R, juntamente com as respectivas significâncias estatísticas (p-valor), atreladas a cada coeficiente. Para tanto, tem-se a tabela 5, na qual tais valores são apresentados:

**Tabela 5** – Valores dos coeficientes polinomiais e suas significâncias estatísticas:  
Componente R.

GRAU	Coefficientes -MMQ-	P- VALOR (%)	Coefficientes -B- SPLINES +MMQ-	P- VALOR (%)	Coefficientes -MV-	P- VALOR (%)	Coefficientes -B-SPLINES +MV-	P- VALOR (%)
0	5.339,60	<2,00E-16	4,30E+03	0,29	5.337,57	<2,2E-16	4.253	0,30
1	-42.358,50	<2,00E-16	-1,24E+04	0,88	-45.862,00	<2,2E-16	-17.571	0,83
2	-35.125,40	2,35E-15	-2,62E+05	0,78	-33.888,74	3,49E-14	-122.604	0,89
3	55.029,20	2,00E-16	3,58E+06	0,61	55.904,30	<2,2E-16	1.955.169	0,76
4	-18.611,80	1,04E-05	-2,62E+07	0,51	-17.755,90	1,76E-05	-14.579.432	0,67
5	-15.213,20	2,85E-04	1,32E+08	0,45	-11.871,91	2,53E-03	73.296.438	0,61
6	44.212,90	<2,00E-16	-5,09E+08	0,41	41.803,92	<2,2E-16	-275.788.507	0,58
7	-22.672,80	1,05E-07	1,55E+09	0,38	-21.975,32	2,38E-07	817.998.901	0,57
8	-20.607,90	1,18E-06	-3,90E+09	0,37	-22.274,48	1,20E-08	-1.968.633.051	0,58
9	40.461,70	<2,00E-16	7,94E+09	0,37	38.765,95	<2,2E-16	3.906.259.181	0,59
10	5.105,10	2,20E-01	-1,37E+10	0,39	7.324,83	5,00E-02	-6.443.720.468	0,62
11	-24.991,00	5,75E-09	2,00E+10	0,41	-27.505,84	8,66E-12	8.859.243.572	0,66
12	13.807,20	9,62E-04	-2,47E+10	0,44	15.724,80	4,42E-05	-10.096.467.269	0,71
13	7.787,10	6,00E-02	2,60E+10	0,48	7.318,83	5,12E-02	9.366.131.434	0,76
14	-1.300,80	7,50E-01	-2,32E+10	0,53	-2.546,35	4,82E-01	-6.773.106.567	0,83
15	10.531,10	1,10E-02	-1,75E+10	0,58	1.2931,06	5,63E-04	3387.854.684	0,90
16	-916,70	8,20E-01	-1,12E+10	0,64	-2.402,52	5,30E-01	-571.909.365	0,98
17	-10.525,30	1,10E-02	6,00E+09	0,70	-8.548,81	1,71E-02	-899.871.944	0,95
18	13.716,40	1,00E-03	-2,77E+09	0,76	14.636,61	5,73E-05	1.150.802.630	0,88
19	11.274,70	7,00E-03	1,12E+09	0,79	10.584,16	3,83E-03	-799.940.931	0,83
20	-5.626,90	1,70E-01	-4,21E+08	0,81	-4.183,14	2,47E-01	388.459.642	0,80
21	-4.834,00	2,40E-01	1,53E+08	0,80	-5.927,95	8,77E-02	-136.185.644	0,80
22	6.156,80	1,40E-01	-5,33E+07	0,76	5.019,32	1,70E-01	32.407.355	0,83

continua

continuação

23	5.623,90	1,70E-01	1,67E+07	0,68	5.706,37	1,02E-01	-3.680.698	0,92
24	4.745,90	2,50E-01	-4,37E+06	0,53	2.213,30	5,44E-01	-685.420	0,91
25	-1.494,00	7,20E-01	8,79E+05	0,34	-2.213,47	5,26E-01	423.576	0,61
26	-4.115,30	3,20E-01	-1,32E+05	0,09	-4.062,11	2,52E-01	-100.928	0,17
27	1.145,50	7,80E-01	8,18E+03	0,16	-832,66	8,40E-01	7.765	0,18
<b>Quantidade total de coeficientes significantes de cada estimativa</b>								
<b>coef.sig</b>	<b>16</b>		<b>1</b>		<b>18</b>		<b>0</b>	

**Obs.:** Todas as células em cinza remetem à presença de significância estatística do seu respectivo coeficiente. Os “coeficientes significantes” (**coef.sig**) são aqueles cujos valores de significância estatística são iguais ou inferiores a 10%.

Fonte: Resultados da pesquisa.

Analisando a tabela 5, observa-se que as formas polinomiais funcionais geradas via MMQ e MV possuem coeficientes com valores numéricos semelhantes. Tal semelhança se estende aos sinais de tais valores e se reflete no valor das respectivas significâncias estatísticas calculadas. Fato análogo ocorre entre os modelos gerados via *B-SPLINES*+MMQ e *B-SPLINES*+MV.

A fim de facilitar a abstração das formas funcionais geradas, as respectivas formulações são apresentadas abaixo, considerando os coeficientes presentes na tabela 5:

1) **MMQ:**

$$\hat{R}_i = 5.339,60 - 42.358,50P_i + \dots + 1.145,50P_i^{27} + e_{Ri} \quad (51)$$

2) ***B-SPLINES*+MMQ:**

$$\hat{R}_i = (4,30E + 03) - (1,24E + 04)P_i + \dots + (8,18E + 08)P_i^{27} + e_{Ri} \quad (52)$$

3) **MV:**

$$\hat{R}_i = 5.337,57 - 45.862P_i + \dots - 832P_i^{27} + e_{Ri} \quad (53)$$

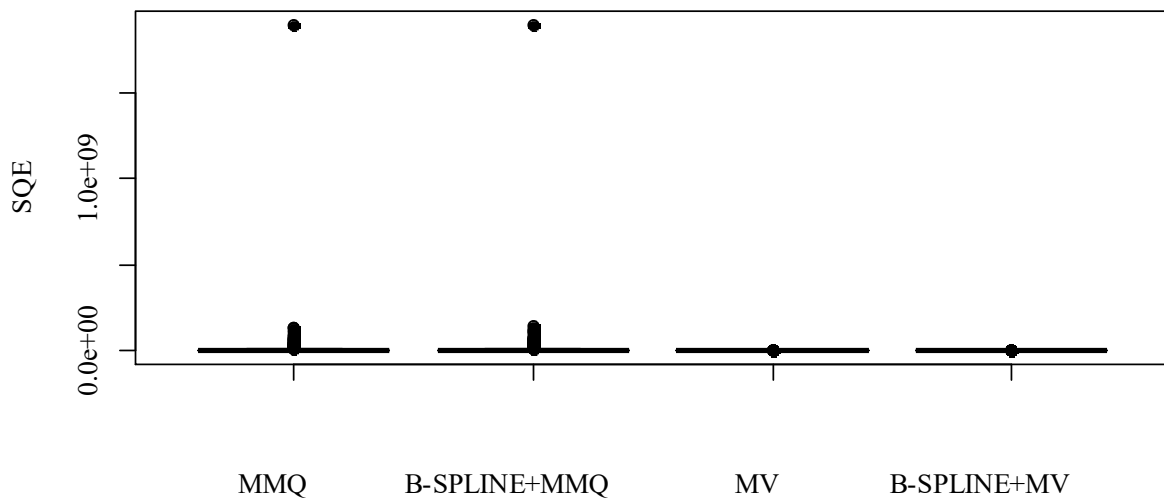
4) ***B-SPLINES*+MV:**

$$\hat{R}_i = 4.253 - 17.571P_i + \dots - 7.765P_i^{27} + e_{Ri} \quad (54)$$

Quando se considera a quantidade de coeficientes significantes relativos aos ajustes efetuados por cada método, verifica-se que aqueles gerados apenas via MMQ e MV são em maior quantidade. Este fato poderia induzir à conclusão precipitada sobre uma possível superioridade dos métodos e sobre a baixa influência do uso de *B-Splines* no acréscimo de qualidade dos ajustes. Porém, a conclusão prematura não é aconselhável, cabendo ainda a análise dos erros relativos a cada ajuste.

Da análise dos resíduos ( $e_{Ri}$ ) de cada um dos modelos ajustados, pôde-se chegar à resultados conclusivos quanto ao nível de eficácia das metodologias utilizadas, com relação à componente R. Assim, analisando o gráfico da Figura 16, verifica-se a semelhança no comportamento dos erros, de modo que, as maiores semelhanças são estabelecidas entre o par de métodos MMQ e *B-SPLINES*+MMQ e entre MV e *B-SPLINES*+MV. De qualquer forma, os métodos baseados em MV tiveram desempenho superior.

**Figura 16** – Gráfico *Boxplot* dos resíduos quadráticos dos ajustes para a componente R.



Fonte: Resultados da pesquisa.

Pela análise numérica, efetuada com base nos valores apresentados na tabela 6, percebe-se que os ajustes baseados em MMQ são ligeiramente inferiores aos ajustes obtidos via MV. Ou seja, os ajustes obtidos com base em MV apresentam menores valores de SQE, MÉDIA, DP e CV(%), e maiores valores de  $R^2$ (%). Além disso, o ajuste gerado somente por MV apresenta 19 coeficientes significantes validados por testes t, que geraram os p-valores apresentados na tabela 5. E mais, neste caso particular, não se pôde perceber numericamente a influência relevante positiva ou negativa do adensamento de pontos gerados via *B-Splines*, quando de seu acoplamento às metodologias baseadas em MMQ e MV.

**Tabela 6** – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente R.

PARÂMETROS ESTATÍSTICOS	MMQ	B-SPLINES+ MMQ	MV	B-SPLINES+ MV
SQE	3,88E+09	3,88E+09	3,49E+06	3,49E+09
MÉDIA	1,52E+07	1,52E+07	1,36E+04	1,36E+04
DP	1,19E+08	1,19E+08	7,23E+04	7,23E+04
CV(%)	7,87E+02	7,87E+02	5,31E+02	5,31E+02
R <sup>2</sup> (%)	7,67E+01	7,68E+01	9,99E+01	9,99E+01

Fonte: Resultados da pesquisa.

Cabe esclarecer que, dos parâmetros apresentados na tabela acima, o  $R^2$  (Coeficiente de Determinação) é o responsável por qualificar percentualmente o ajuste efetuado. Porém, Rawlings *et al.* (1998) especificam que é necessário tomar cuidado com as conclusões tomadas a partir dos valores de  $R^2$ , pois, apesar do coeficiente de determinação ser uma medida descritiva da qualidade do ajuste obtido, ele não deve ser considerado isoladamente em uma análise, mas sempre aliado a outros parâmetros de diagnóstico. Tal conselho se justifica pelo fato do valor de  $R^2$  sofrer a influência de alguns elementos como, por exemplo: a quantidade de observações consideradas no ajuste e o número de termos presentes no modelo de ajustamento.

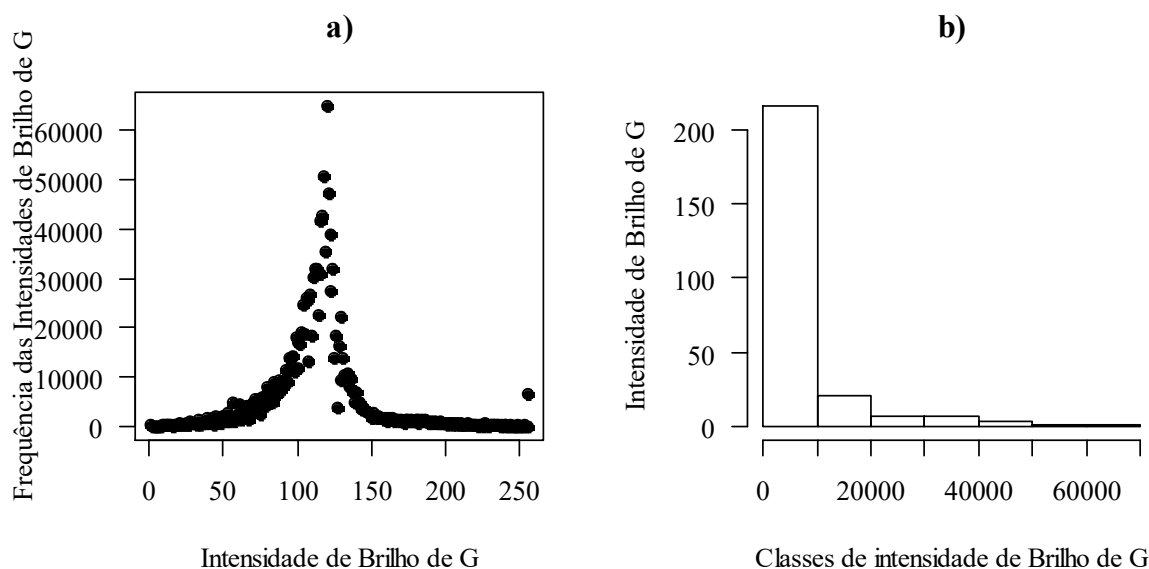
Assim, Rawlings *et al.* (1998) explicam que quando o número de observações diminui, o valor do coeficiente de determinação tende a crescer. E mais, expressam que o valor de  $R^2$  pode se tornar maior ao se adicionar um número suficiente de termos ao modelo. Assim, por exemplo, se não houver dados repetidos em um conjunto com “n” dados e o modelo for um polinômio de grau (n-1), então  $R^2$  indicará um ajuste perfeito, mesmo que este ajustamento não explique perfeitamente os dados. Da afirmação anterior, conclui-se que, embora o coeficiente tenha seu valor aumentado mediante a adição de termos ao modelo, isto não significa necessariamente que o novo modelo é superior a um modelo anterior com menos termos (RAWLINGS *et al.* 1998).

Considerando o recorrido até o momento, verifica-se que a atitude de se estimar a qualidade dos ajustamentos gerados pelos métodos propostos por meio dos parâmetros presentes na tabela 6 foi uma decisão coerente, baseada na literatura pertinente.

#### 4.2.2 Experimento com a componente G

A fim de se ter uma ideia do comportamento da variação das frequências das intensidades de brilho relativas à componente G, considerando a variação da variável independente (intensidade de brilho - P), tem-se o gráfico de dispersão dos valores de P contra G (Figura 17 a)) e também o histograma de G contra P (Figura 17 b)).

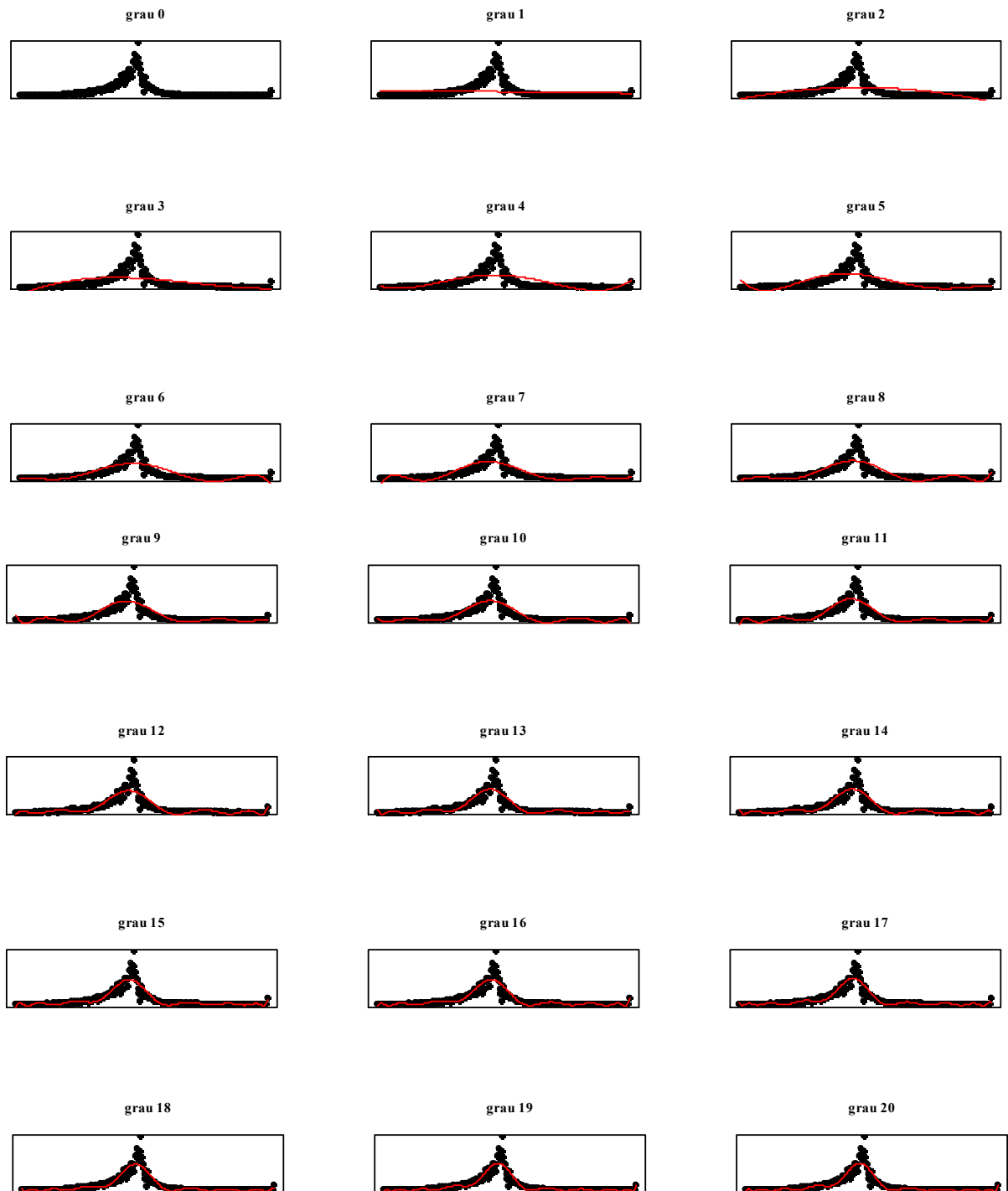
**Figura 17** – Comportamento da componente G: **a)** Gráfico de Dispersão (PxG); **b)** Histograma de G (GxP).

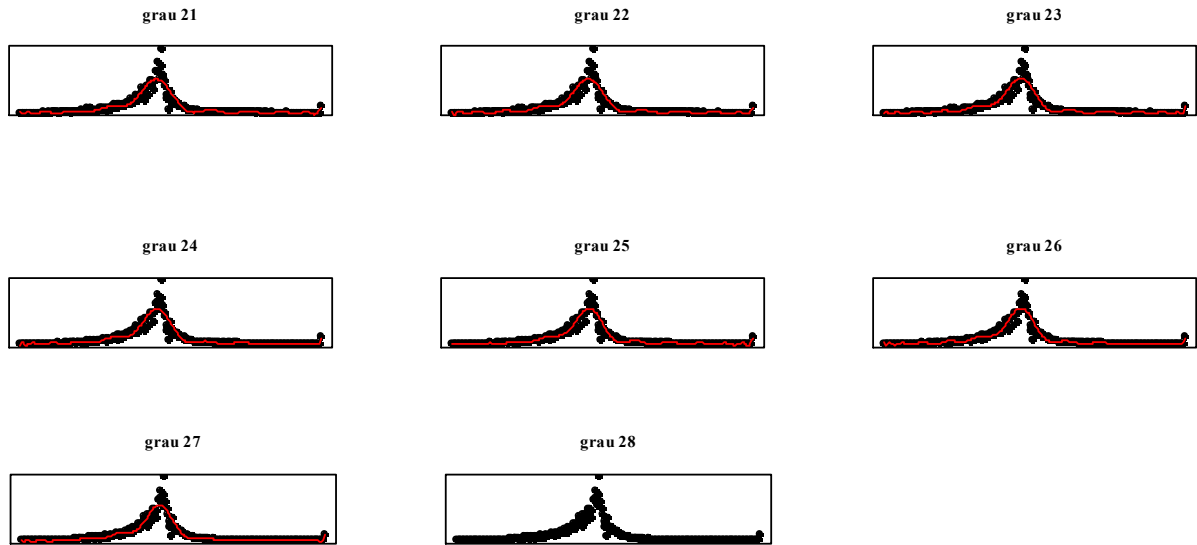


Fonte: Resultados da pesquisa.

Analogamente ao que foi feito para a componente R, foram realizados sucessivos ajustes com os dados relativos à componente G via MMQ. Tais ajustes destinaram-se à determinação do grau do polinômio a ser utilizado como modelo pelas metodologias de ajuste propostas. O esboço dos gráficos relativos a estes ajustes são apresentados na Figura 18. Em cada esboço da Figura 18, os pontos pretos representam os pontos de controle iniciais e as curvas em vermelho representam os ajustes polinomiais. Por meio destes esboços pode-se ter ideia da qualidade de cada ajustamento em representar o conjunto de dados original

**Figura 18** – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio.



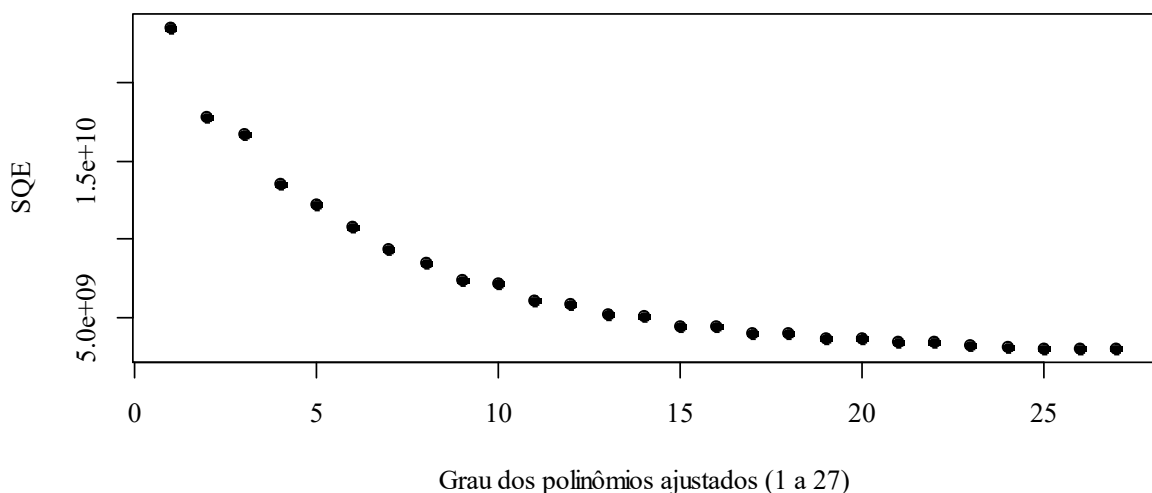


Fonte: Resultados da pesquisa.

De modo semelhante ao que ocorreu para os testes feitos para a componente R, o *software* R-Studio efetuou ajustes polinomiais apenas até o grau 27 (ver dois últimos esboços da Figura 18). Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada pelo RStudio. Mais uma vez não se pôde determinar as causas pelas quais o *software* não gerou os coeficientes do polinômio de grau 28, restando apenas citar as suposições expressas anteriormente.

Para este segundo conjunto de pontos pôde-se perceber também o comportamento assintótico dos erros dos ajustes, de modo que, com o aumento do grau do polinômio ajustado, representante da componente G, houve a redução do parâmetro estatístico SQE, conforme o que pode ser averiguado no gráfico da figura 19.

**Figura 19** – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.

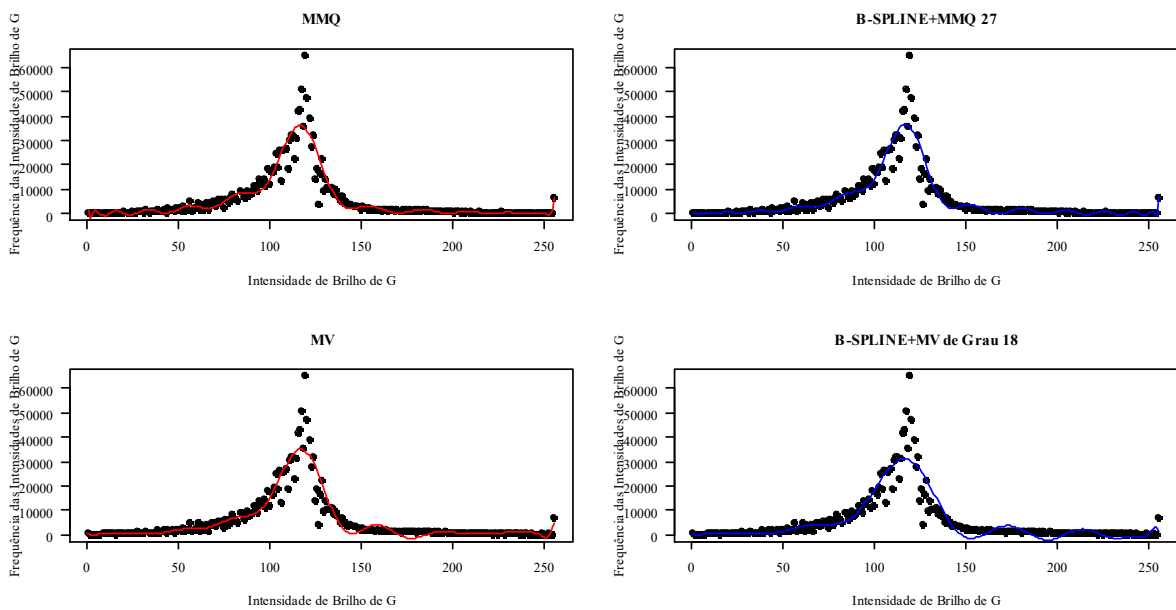


Fonte: Resultados da pesquisa.

Pode-se verificar no gráfico da Figura 19 que o menor valor de SQE, relativo ao ajustados dados da componente G, aproxima-se de  $4,00E+09$  quando o polinômio ajustado é de grau 27. Assim, assume-se que um polinômio com este grau possa ser o modelo matemático para os ajustamentos previstos na metodologia.

Nestes termos, os respectivos ajustes previstos na metodologia foram efetuados com dados provenientes da componente G, considerando como forma funcional inicial o polinômio de grau 27. Na sequência, os resultados obtidos foram comparados a fim de se identificar a metodologia que, neste caso particular, tenha alcançado melhor desempenho. As comparações/avaliações aqui apresentadas são de cunho visual (ver Figura 20) e numérico. Os gráficos da Figura 20 mostram as curvas plotadas a partir dos modelos polinomiais obtidos pelas metodologias de ajuste previstas anteriormente.

**Figura 20** – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.



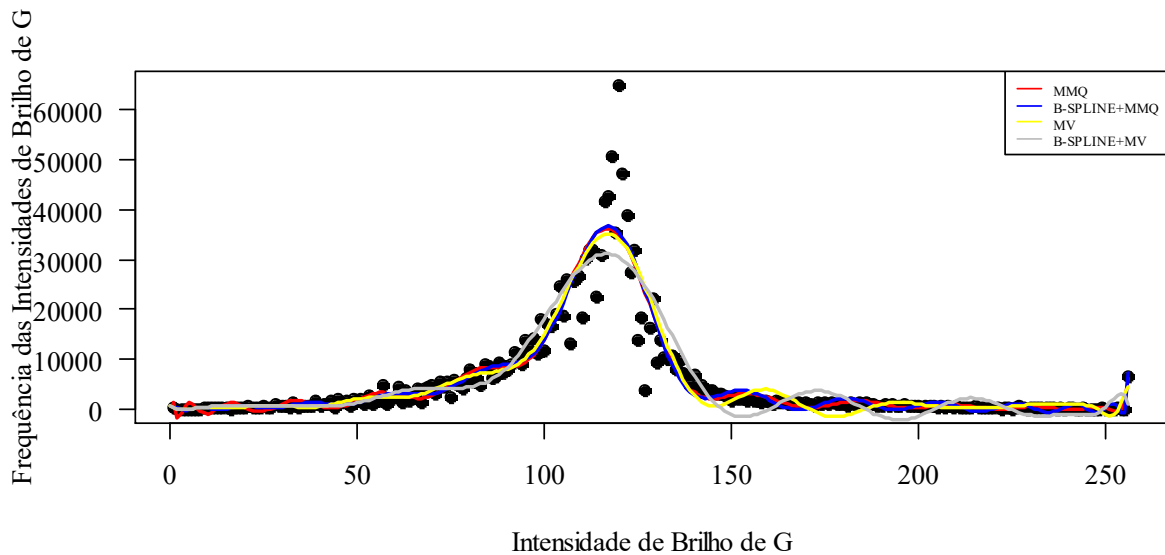
Fonte: Resultados da pesquisa.

Na Figura 20, verifica-se que o ajuste relativo ao método *B-SPLINES*+MV ocorreu tendo como modelo um polinômio de grau 18. Ou seja, o *software* RStudio não gerou resultados quando os modelos adotados foram polinômios de grau superior a 18. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada. Acredita-se que tal fato decorra, uma vez mais, de problemas associados ao *software* ou ao *hardware* utilizados. Da análise visual dos gráficos da Figura 20, observa-se algumas diferenças sutis entre os ajustes. Tais diferenças podem ser melhor



visualizadas/detectadas com a sobreposição gráfica das curvas inerentes aos modelos ajustados para componente G (Figura 21).

**Figura 21** – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente G.



Fonte: Resultados da pesquisa.

As diferenças verificadas no gráfico da Figura 21, podem ser mais bem pormenorizadas e identificadas pela análise dos coeficientes polinomiais gerados por cada ajuste. Tal análise dos coeficientes é facilitada pela especificação das respectivas significâncias estatísticas. Assim, os coeficientes e suas respectivas significâncias estatísticas são apresentadas tabela 7.

**Tabela 7** – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Componente G.

GRAU	Coefficientes -MMQ-	P- VALOR (%)	Coefficientes -B-SPLINES +MMQ-	P- VALOR (%)	Coefficientes -MV-	P- VALOR (%)	Coefficientes -B-SPLINES +MV-	P- VALOR (%)
0	5.339,61	<2,00E-16	2,64E+02	0,94	5.303,64	<2,2E-16	372	<2,2E-16
1	-17.285,40	3.38E-06	1,15E+04	0,87	-17.255,97	3,84E-04	-3.744	5,87E-04
2	-75.640,01	<2,00E-16	-2,65E+05	0,74	-74.267,03	<2,2E-16	23.378	3,09E-02
3	33.878,98	<2,00E-16	3,04E+06	0,62	35.291,69	1,34E-13	-37.260	6,57E-01
4	56.125,57	<2,00E-16	-2,33E+07	0,50	56.597,68	<2,2E-16	-309.874	5,09E-01
5	-36.517,33	<2,00E-16	1,32E+08	0,38	-38.360,15	<2,2E-16	2.818.491	1,33E-01
6	-37.141,21	<2,00E-16	-5,79E+08	0,27	-39.525,60	<2,2E-16	-12.517.795	2,36E-02

continua

continuação

7	38.755,07	<2,00E-16	2,05E+09	0,18	37.908,00	<2,2E-16	36.794.818	2,64E-03
8	28.357,07	2,01E-13	-5,95E+09	0,11	28.723,12	5,11E-11	-76.504.344	2,21E-04
9	-33.463,43	<2,00E-16	1,44E+09	0,06	-31.881,35	4,37E-14	115.324.102	2,18E-05
10	-16.643,59	7,43E-06	-2,92E+10	0,03	-16.446,33	1,58E-05	-126.918.584	4,93E-06
11	33.216,15	<2,00E-16	5,01E+10	0,02	32.538,94	5,37E-14	103.189.555	3,68E-06
12	13.171,79	0,000350	-7,29E+10	0,01	13.629,23	2,73E-03	-62.510.009	1,03E-05
13	-26.758,85	3,02E-02	9,02E+10	0,01	-25.178,98	7,00E-09	28.269.893	1,01E-04
14	-4.676,95	0,198740	-9,52E+10	0,01	-1.939,74	5,94E-01	-9.514.878	1,95E-03
15	26.038,11	1,00E-11	8,56E+10	0,01	26.753,14	2,82E-11	23.667.853	2,55E-02
16	3.510,51	0,334347	-6,56E+10	0,01	963,91	8,08E-01	-421.870	1,40E-01
17	-20.111,76	8,21E-08	4,29E+10	0,01	-23.140,61	9,71E-11	41.837	4,61E-01
18	2.396,32	0,509669	-2,38E+09	0,01	941,40	7,93E-01	570	9,37E-01
19	19.621,27	1,61E-07	1,12E+09	0,01	21.545,35	4,13E-08	-	-
20	-1.923,01	0,596660	-4,41E+09	0,01	175,48	9,61E-01	-	-
21	-14.736,91	6,71E-05	1,44E+08	0,01	-17.202,65	5,66E-06	-	-
22	5.811,53	0,110635	-3,87E+08	0,01	-208,89	9,49E-01	-	-
23	14.366,41	0,000101	8,28E+07	0,02	9.247,94	3,65E-04	-	-
24	-5.003,84	0,169251	-1,37E+07	0,03	-1.242,57	5,72E-01	-	-
25	-10.496,77	0,004189	1,68E+06	0,04	-4.359,70	9,21E-03	-	-
26	7.343,84	0,044152	-1,40E+05	0,04	-658,23	6,38E-01	-	-
27	-70,26	0,984569	6,95E+03	0,16	158,03	6,17E-01	-	-

Quantidade total de coeficientes significantes de cada estimativa

coef.sig	21	17	20	13
----------	----	----	----	----

**Obs.:** Todas as células em cinza remetem à presença de significância estatística do seu respectivo coeficiente. Os “coeficientes significantes” (**coef.sig**) são aqueles cujos valores de significância estatística são iguais ou inferiores a 10%.

Fonte: Resultados da pesquisa.

Analisando os valores presentes na tabela 7, verifica-se que os coeficientes relativos aos ajustes por MMQ e por MV possuem semelhança numérica e de sinal. E mais, os valores de suas respectivas significâncias estatísticas também se assemelham. De tais fatos, depreende-se que os modelos se assemelhem no tocante à representação do histograma da componente G. Por outro lado, averigua-se que os coeficientes gerados pelo ajuste por *B-SPLINES*+MMQ e por *B-SPLINES*+MV não se assemelham. Esta discrepância pode ser atribuída ao fato do ajuste efetuado pelo método *B-SPLINES*+MV ter sido feito com base em um modelo polinomial de grau 18, diferentemente dos outros modelos utilizados (polinômios de grau 27). Mesmo assim, optou-se por apresentar os coeficientes gerados, para fins de apreciação do leitor.

Uma vez mais, os métodos MV e MMQ atingiram as maiores quantidades de coeficientes significantes, havendo a vantagem numérica para o método MV em uma unidade. Apesar da superioridade citada acima, o método *B-SPLINES*+MMQ também teve desempenho digno de nota, gerando 17 coeficientes significantes. Assim, não se pode afirmar que o adensamento do conjunto de pontos de controle inicial por *B-Splines* tenha sido em vão.

A partir das estimativas dos coeficientes apresentadas na tabela 7, pôde-se obter as formas funcionais que aproximam a componente G, as quais são:

1) **MMQ:**

$$\hat{G}_i = 5.339,61 - 17.285,40P_i + \dots - 70,26P_i^{27} + e_{Gi} \quad (55)$$

2) ***B-SPLINES*+MMQ:**

$$\hat{G}_i = (2,64E + 02) + (1,15E + 04)P_i + \dots + (6,95E + 03)P_i^{27} + e_{Gi} \quad (56)$$

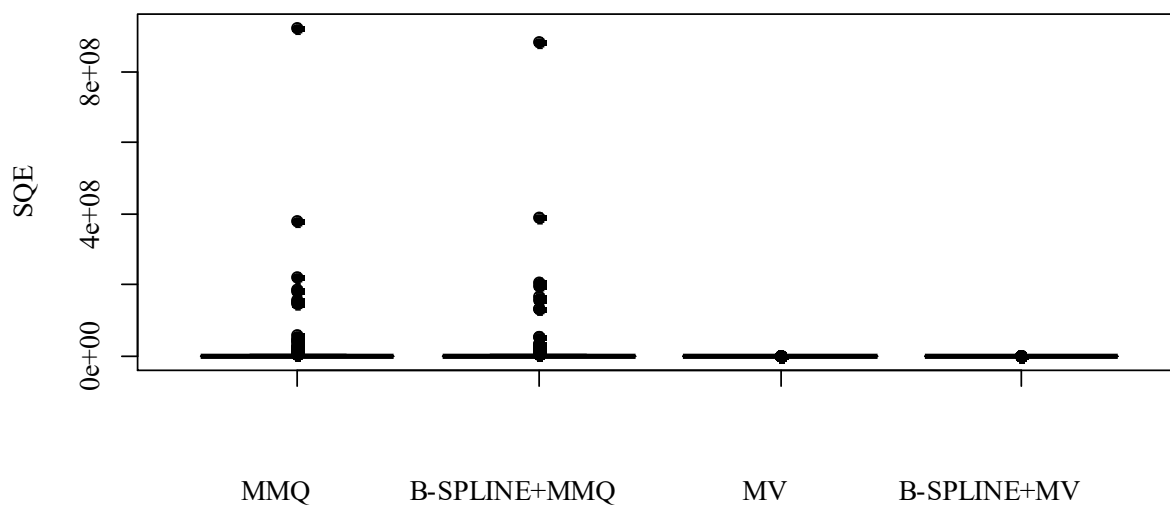
3) **MV:**

$$\hat{G}_i = 5.303,64 - 17.255,97P_i + \dots + 158,03P_i^{27} + e_{Gi} \quad (57)$$

4) ***B-SPLINES*+MV:**

$$\hat{G}_i = 372 - 3.744P_i + \dots + 570P_i^{18} + e_{Gi} \quad (58)$$

A avaliação dos resíduos ( $e_{Gi}$ ) de cada um dos modelos ajustados e a verificação de suas semelhanças são facilitadas mediante a análise da Figura 22 que apresenta o gráfico relativo ao comportamento dos erros inerentes a cada ajuste.

**Figura 22** – Gráfico *Boxplot* dos resíduos quadráticos dos ajustes para a componente G.

Mais uma vez, da comparação dos erros inerentes a cada ajuste, verifica-se a superioridade dos métodos MV e *B-SPLINES*+MV, em relação aos métodos baseados em MMQ. Porém, como o modelo utilizado para o ajuste via *B-SPLINES*+MV difere do modelo utilizado pelos demais métodos, acredita-se que a comparação entre os resultados advindos dos demais métodos e deste (*B-SPLINES*+MV) seja desaconselhável e inconsistente.

Numericamente, os métodos MV e *B-SPLINES*+MV apresentam resultados com menores valores de SQE, MÉDIA e DP, e os maiores valores de  $R^2$ . Porém, conforme o citado anteriormente, o método *B-SPLINES*+MV foi aplicado com um modelo polinomial discrepante dos demais métodos utilizados. Assim, considerando que o método MV gerou 20 coeficientes estatisticamente significativos de um total de 28 coeficientes estimados, conclui-se pela superioridade deste método em relação aos demais.

Da comparação numérica dos parâmetros estatísticos relativos aos métodos MMQ e *B-SPLINES*+MMQ (tabela 8), verifica-se que o método *B-SPLINES*+MMQ gerou resultados ligeiramente superiores. Logo, pode-se afirmar que, para a componente G, o adensamento de pontos gerado por *B-Splines* foi satisfatório.

**Tabela 8** – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente G.

<b>PARÂMETROS ESTATÍSTICOS</b>	<b>MMQ</b>	<b><i>B-SPLINES</i> +MMQ</b>	<b>MV</b>	<b><i>B-SPLINES</i> +MV</b>
<b>SQE</b>	3,00E+09	2,90E+09	4,12E-15	4,12E-15
<b>MÉDIA</b>	1,17E+07	1,31E+07	1,61E-17	1,61E-17
<b>DP</b>	6,76E+07	6,50E+07	9,91E-17	9,91E-17
<b>CV(%)</b>	5,76E+02	5,75E+02	6,16E+02	6,16E+02
<b>R<sup>2</sup>(%)</b>	8,20E+01	8,26E+01	1,00E+02	1,00E+02

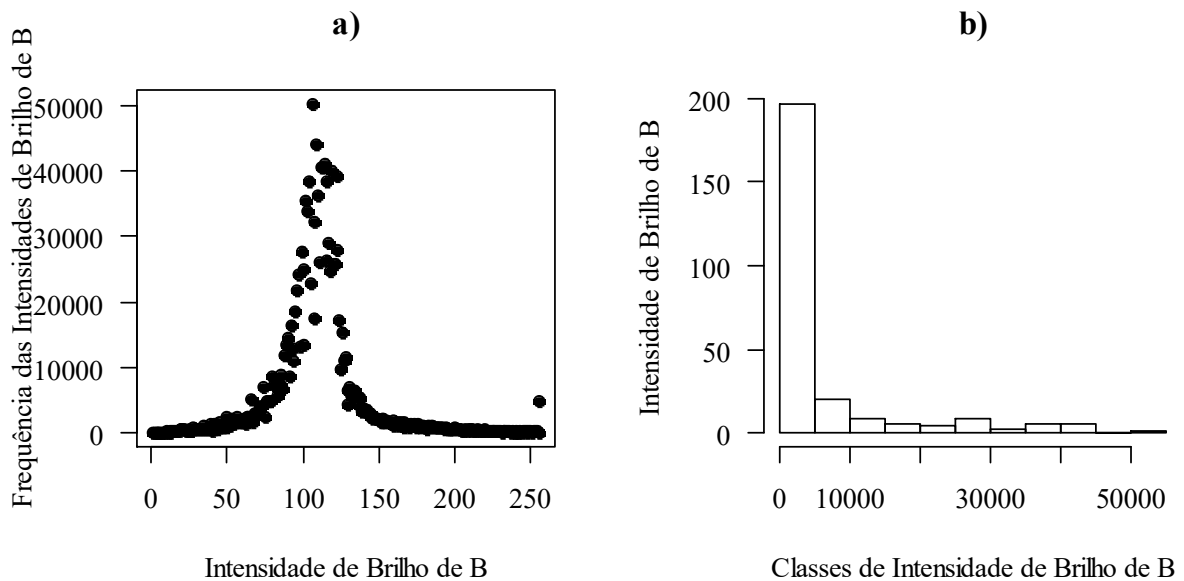
Fonte: Resultados da pesquisa.

### 4.2.3 Experimento com a componente B

De maneira análoga ao realizado nos experimentos anteriores, foram executados procedimentos semelhantes para a componente B da imagem analisada. Assim, inicialmente são apresentados os gráficos de dispersão dos valores de P contra B (Figura 23 a)) e também o histograma de B contra P (Figura 23 b)), para que se possa intuir o comportamento variacional dos dados advindos da referida componente.

Da comparação visual entre o gráfico de dispersão relativo à componente B (Figura 23 a)) e os gráficos anteriores (Figuras 11 a) e 17 a)) percebe-se alguma semelhança na distribuição das frequências, considerando a variação das intensidades de brilho. O mesmo ocorre com os respectivos histogramas. Assim, infere-se que, devido ao fato destas serem componentes de uma mesma imagem, haja correlação entre elas. De tal fato, transitivamente depreende-se que os ajustes gerados sejam semelhantes. Tal afirmação pode ser confirmada pela comparação dos coeficientes gerados pelos métodos de ajuste.

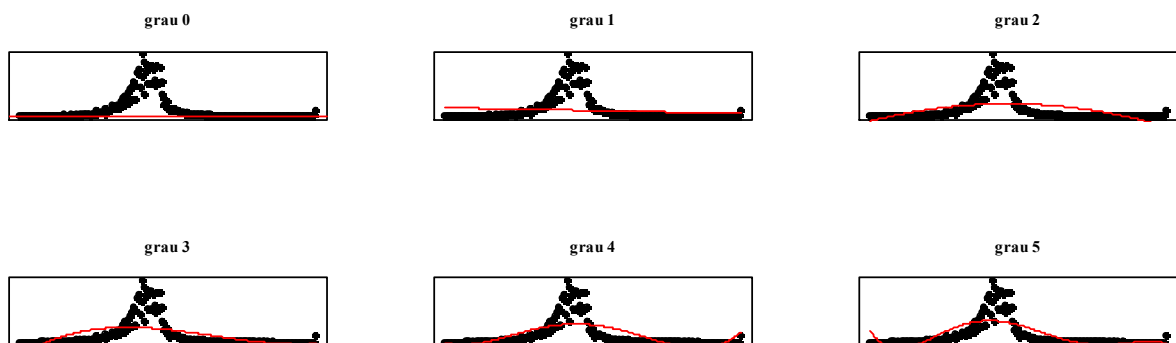
**Figura 23** – Comportamento da componente B: **a)** Gráfico de Dispersão (PxB); **b)** Histograma de B (BxP).

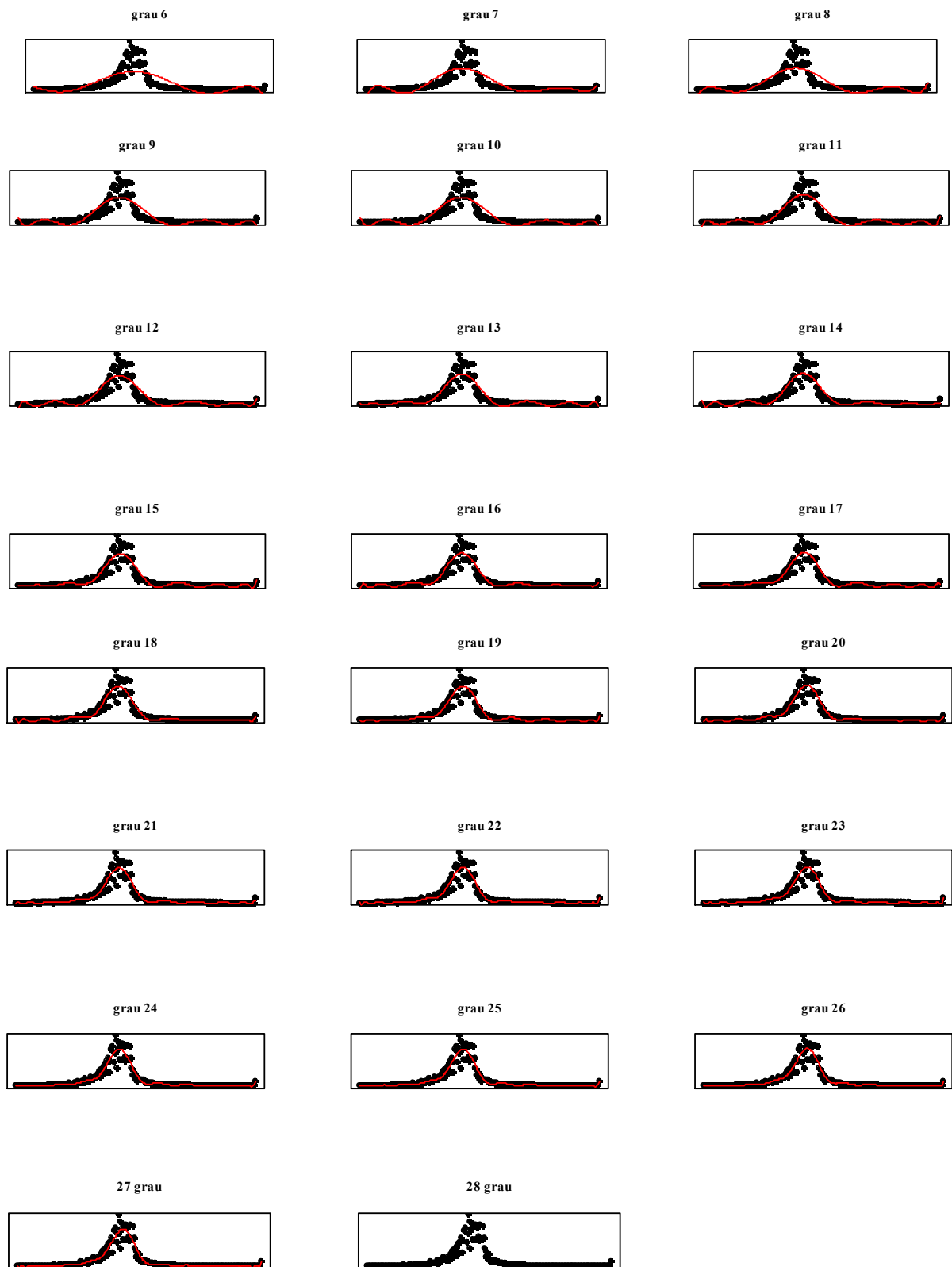


Fonte: Resultados da pesquisa.

Também para os dados provenientes da componente B foram testados diversos ajustes via MMQ, com a finalidade de se determinar o grau do polinômio a ser utilizado como modelo pelas metodologias de ajuste propostas. A Figura 24 mostra o esboço dos gráficos relativos a cada ajustamento. E mais, em cada esboço da Figura 24, os pontos pretos representam os pontos de controle iniciais e as curvas em vermelho representam os ajustes polinomiais obtidos. Dos esboços apresentados pode-se inferir aproximadamente a qualidade de cada ajustamento, no tocante à representação dos dados originais.

**Figura 24** – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio.



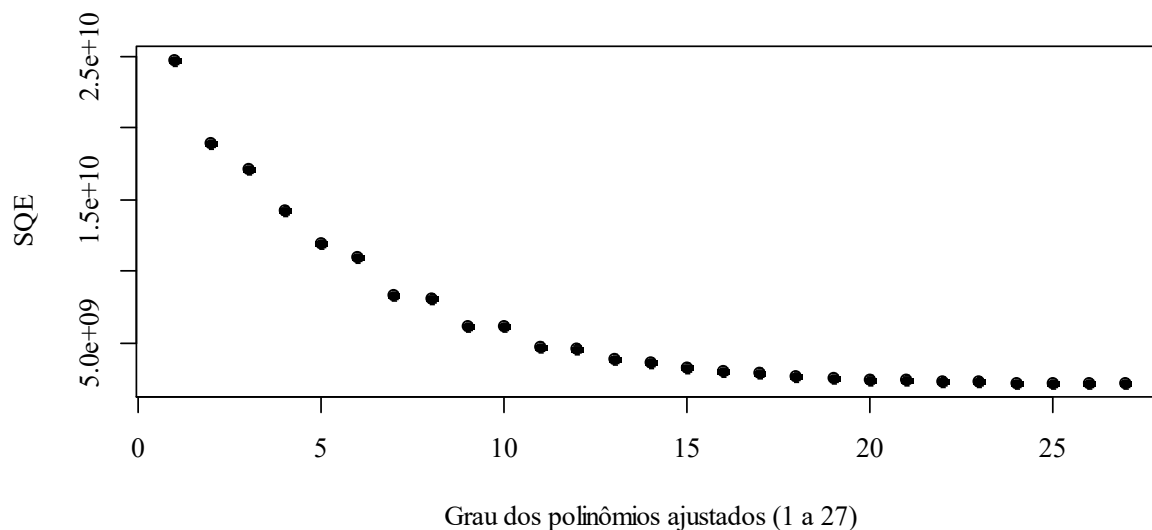


Fonte: Resultados da pesquisa.

Pode se observar, que aparentemente um modelo polinomial que se ajustou bem ao conjunto de dados originais foi o de grau 27. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada pelo RStudio. Porém, para que a escolha do polinômio de grau 27 seja consumada, é necessário se averiguar o comportamento do ajuste com relação aos erros resultantes. Para tanto, tem-se o gráfico da Figura 25, no qual o decaimento da soma do erro quadrático pode ser observado, considerando-se o aumento do grau do polinômio ajustado.

Ressalta-se que, uma vez mais o *software* RStudio não efetuou o ajustamento considerando como modelo os polinômios de grau maior que 27.

**Figura 25** – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.

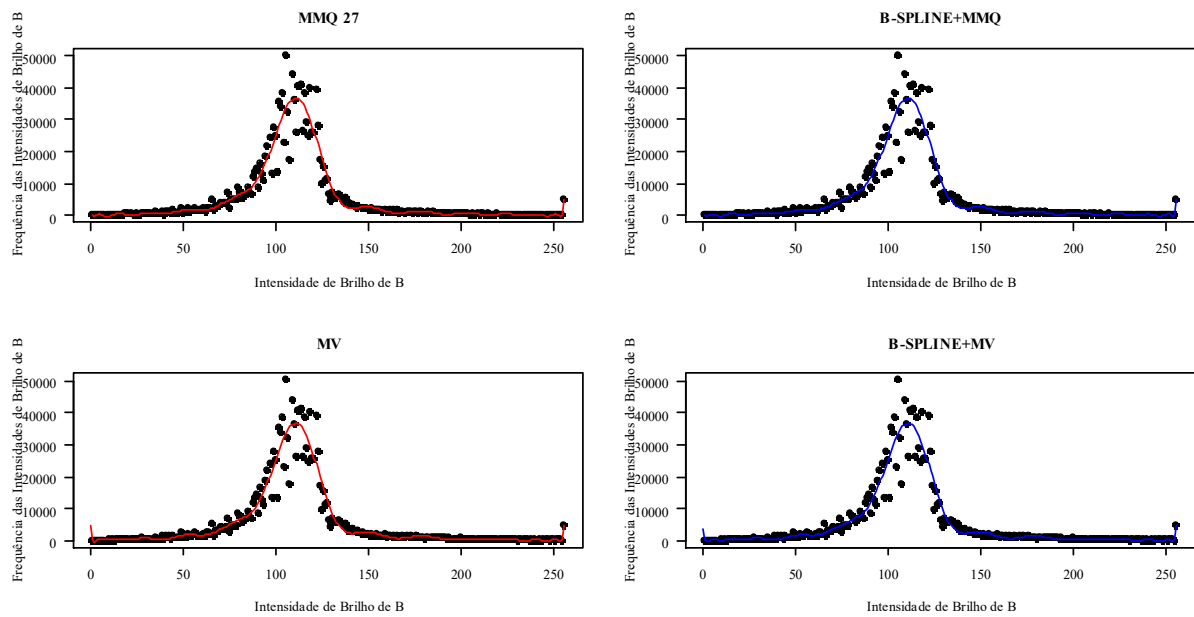


Fonte: Resultados da pesquisa.

Assim, considerando os valores da SQE utilizados para a construção do gráfico da Figura 22, verifica-se que o polinômio ajustado, que gera o menor erro em relação aos dados originais, é realmente o de grau 27. Desta maneira, foram estimados ajustes polinomiais de grau 27, representativos da componente B, segundo os métodos de ajuste previstos na metodologia. Os gráficos gerados a partir destes ajustes são apresentados na Figura 26.



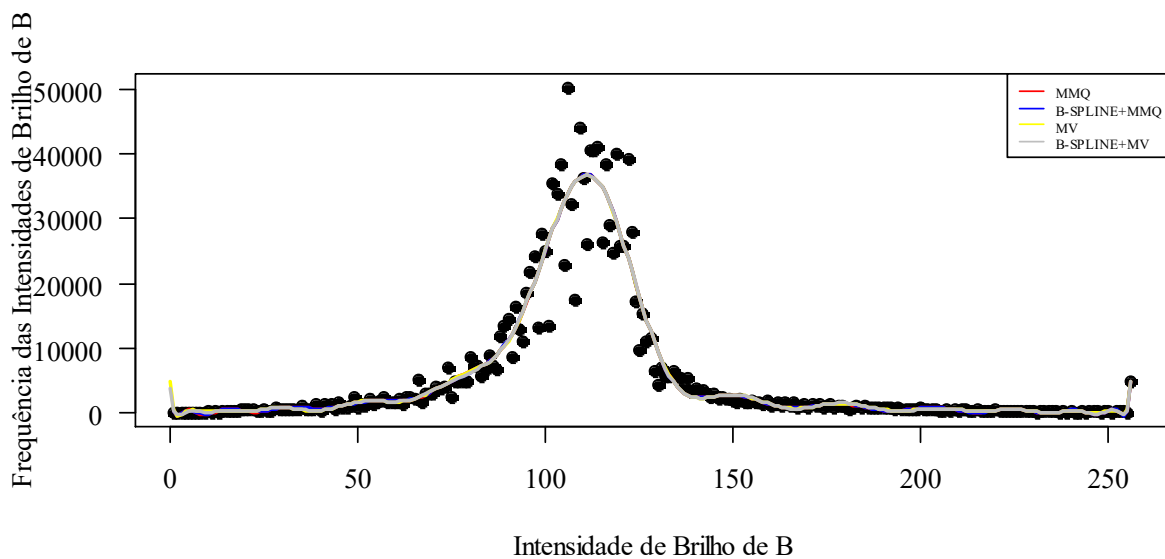
**Figura 26** – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.



Fonte: Resultados da pesquisa.

Aparentemente, as curvas representantes dos ajustes, presentes no gráfico da Figura 26, quase não apresentam diferenças. A fim de tornar as discrepâncias entre as curvas mais visíveis, tem-se o gráfico da Figura 27, no qual as curvas estão sobrepostas.

**Figura 27** – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da componente B.



Fonte: Resultados da pesquisa.

Ao se analisar o gráfico da Figura 27, novamente se tem a impressão de que as formas funcionais geradas são praticamente as mesmas. No entanto, com o intuito de complementar a

análise visual e quantificar numericamente as diferenças entre os ajustes, tem-se a tabela 9, na qual estão presentes os coeficientes estimados por cada método utilizado e os seus respectivos p-valores.

**Tabela 9** – Valores dos coeficientes polinomiais e suas significâncias estatísticas:  
Componente B.

GRAU	Coeficientes -MMQ-	P- VALOR (%)	Coeficientes -B-SPLINES +MMQ-	P- VALOR (%)	Coeficientes -MV-	P- VALOR (%)	Coeficientes -B-SPLINES +MV-	P- VALOR (%)
0	5.339,60	<2,00E-16	3,93E+02	0,90	5.338,28	<2,2E-16	355	0,9067
1	-21.025,40	9.10E-11	-2,08E+04	0,73	-20.821,93	1,65E-13	-15.012	0,7999
2	-76.602,90	<2,00E-16	3,13E+05	0,65	-76.583,79	<2,2E-16	203.671	0,7506
3	41.905,30	<2,00E-16	-2,80E+06	0,60	41.705,99	<2,2E-16	-1.714.561	0,7101
4	53.987,60	<2,00E-16	1,73E+07	0,56	54.174,62	<2,2E-16	10.455.195	0,6669
5	-48.433,30	<2,00E-16	-7,95E+07	0,54	-48.729,55	<2,2E-16	-48.996.784	0,6283
6	-30.634,80	<2,00E-16	2,84E+08	0,53	-30.531,43	<2,2E-16	180.737.303	0,6017
7	51.034,50	<2,00E-16	-8,09E+08	0,54	51.065,66	<2,2E-16	-530.037.556	0,5933
8	15.123,60	1,90E-06	1,85E+09	0,57	14.952,00	3,51E-08	1.237.945.755	0,6077
9	-44.145,40	<2,00E-16	-3,38E+09	0,61	-43.864,86	<2,2E-16	-2.282.308.829	0,6484
10	123,10	0,968278	4,87E+09	0,68	-77,96	9,76E-01	3.222.994.828	0,7180
11	38.637,50	<2,00E-16	-5,23E+09	0,77	38.773,58	<2,2E-16	-3.177.307.025	0,8170
12	-6.335,30	0,041626	3,44E+09	0,89	-6.356,12	2,00E-02	1.342.674.199	0,9415
13	-28.105,20	<2,00E-16	4,76E+08	0,99	-28.242,93	<2,2E-16	2.160.580.364	0,9186
14	13.253,20	2,68E-05	-5,03E+09	0,85	12.951,30	9,66E-07	-5.993.283.342	0,7772
15	21.907,90	1,72E-11	8,21E+09	0,73	22.123,78	4,44E-16	8.412.517.123	0,6479
16	-12.935,40	4,10E-05	-8,84E+09	0,62	-13.057,23	2,75E-06	-8.525.718.670	0,5388
17	-13.064,60	3,46E-05	7,29E+09	0,54	-12.719,06	2,60E-07	6.779.825.593	0,4522
18	14.844,70	2,86E-06	-4,81E+09	0,47	15.014,35	8,54E-09	-4.368.622.196	0,3862
19	9.494,10	0,002397	2,60E+09	0,42	8.919,29	5,94E-04	2.312.871.067	0,3370
20	-11.836,60	0,000167	-1,15E+09	0,39	-11.415,22	2,51E-06	-1.010.679.943	0,3006
21	-3.698,30	0,232933	4,18E+08	0,36	-4.379,33	7,00E-02	363.815.416	0,2736
22	11.421,10	0,000277	-1,23E+08	0,35	12.126,53	3,65E-04	-107.045.729	0,2531
23	2.359,30	0,446251	2,91E+07	0,33	2.531,68	4,01E-06	25.381.669	0,2373
24	-7.541,90	0,015491	-5,37E+06	0,32	-7.771,79	2,68E-01	-4.741.445	0,2247
25	333,70	0,914156	7,44E+06	0,28	508,72	1,60E-03	673.403	0,2118
26	6.583,10	0,034330	-7,38E+04	0,21	5.430,20	1,31E-02	-69.240	0,1739
27	-62,90	0,983789	4,51E+03	0,29	31,22	9,92E-01	4.484	0,2961
<b>Quantidade total de coeficientes significantes de cada estimativa</b>								
<b>coef.sig</b>	<b>23</b>		<b>0</b>		<b>26</b>		<b>0</b>	

**Obs.:** Todas as células em cinza remetem à presença de significância estatística do seu respectivo coeficiente. Os “coeficientes significantes” (**coef.sig**) são aqueles cujos valores de significância estatística são iguais ou inferiores a 10%.

Fonte: Resultados da pesquisa.

Da análise dos valores presentes na tabela 8, infere-se que os métodos MMQ e MV geram resultados semelhantes. Ou seja, os coeficiente homólogos possuem valores numéricos semelhante e sinais idênticos. E mais, as quantidades de coeficientes significantes gerados por ambos os método também são semelhantes, havendo uma pequena vantagem para o método MV. De forma análoga, os ajustes gerados pelos métodos *B-SPLINES*+MMQ e *B-SPLINES*+MV também se assemelham.

As estimativas dos coeficientes mostradas na tabela 8 integram as seguintes formas funcionais polinomiais:

1) **MMQ:**

$$\hat{B}_i = 5.339,60 - 21.025,40P_i + \dots - 62,90P_i^{27} + e_{Bi} \quad (59)$$

2) ***B-SPLINES*+MMQ:**

$$\hat{B}_i = (3,93E + 02) - (2,08E + 04)P_i + \dots + (4,51E + 03)P_i^{27} + e_{Bi} \quad (60)$$

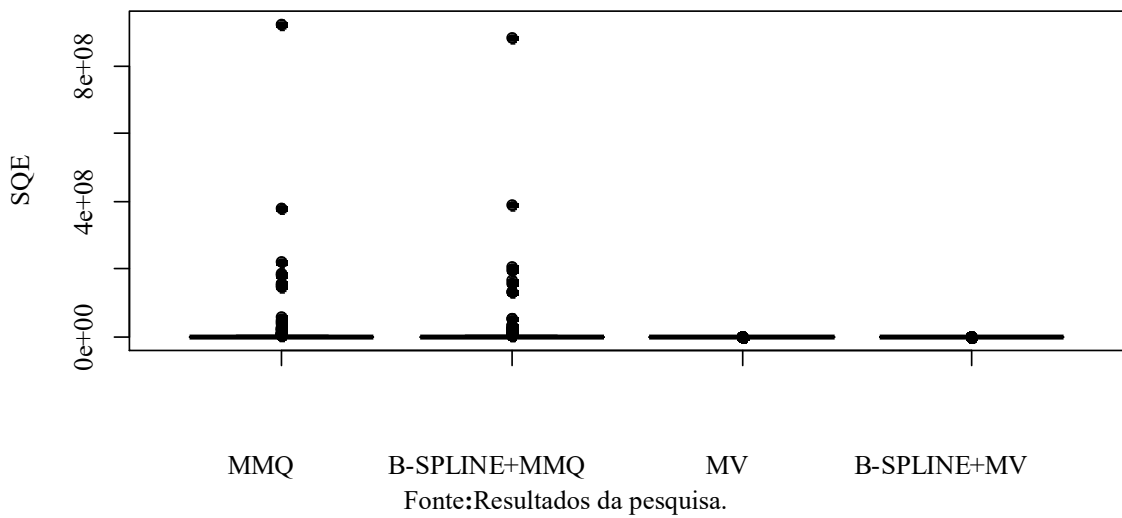
3) **MV:**

$$\hat{B}_i = 5.338,28 - 20.821,93P_i + \dots + 31,22P_i^{27} + e_{Bi} \quad (61)$$

4) ***B-SPLINES*+MV:**

$$\hat{B}_i = 355 - 15.012P_i + \dots + 4.484P_i^{27} + e_{Bi} \quad (62)$$

Ainda com a intenção de capturar semelhanças e diferenças entre os ajuste, examinou-se o comportamento dos resíduos ( $e_{Bi}$ ) para cada modelo. Por meio da análise do gráfico da Figura 28, pode-se depreender o comportamento dos resíduos para cada ajuste efetuado.

**Figura 28** - Gráfico *Boxplot* dos resíduos quadráticos dos ajustes para a componente B.

Assim como o que se verificou para as componentes R e G, os ajustes gerados pelos métodos MV e *B-SPLINES*+MV encontram-se em vantagem, pois suas respectivas formas funcionais geradas apresentam o menor erro na representação da componente B

Ainda com intuito de descrever de forma mais precisa e detalhada o comportamento dos resíduos advindos dos ajustes, construiu-se a tabela 10, na qual os parâmetros estatísticos baseados no erro quadrático são apresentados.

**Tabela 10** – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a componente B.

PARÂMETROS ESTATÍSTICOS	MMQ	<i>B-SPLINES</i> + MMQ	MV	<i>B-SPLINES</i> + MV
SQE	3,00E+09	2,90E+09	11056,33	11056,33
MÉDIA	1,17E+07	1,13E+07	43,19	43,19
DP	6,75E+07	6,50E+07	314,01	314,01
CV(%)	5,76E+02	5,75E+02	727,07	727,07
R <sup>2</sup> (%)	8,20E+01	8,26E+01	99,99	99,99

Fonte: Resultados da pesquisa.

Observa-se nesta análise dos resíduos que os resultados obtidos pelos métodos MV e *B-SPLINES*+MV apresentam menores valores de SQE, MÉDIA e DP, e também os maiores valores de R<sup>2</sup>. Assim, depreende-se que tais métodos gerem os melhores ajustes. Porém, considerando ainda que o método MV gerou 26 coeficientes estatisticamente significativos de um total de 28 coeficientes estimados, conclui-se que este método obteve desempenho

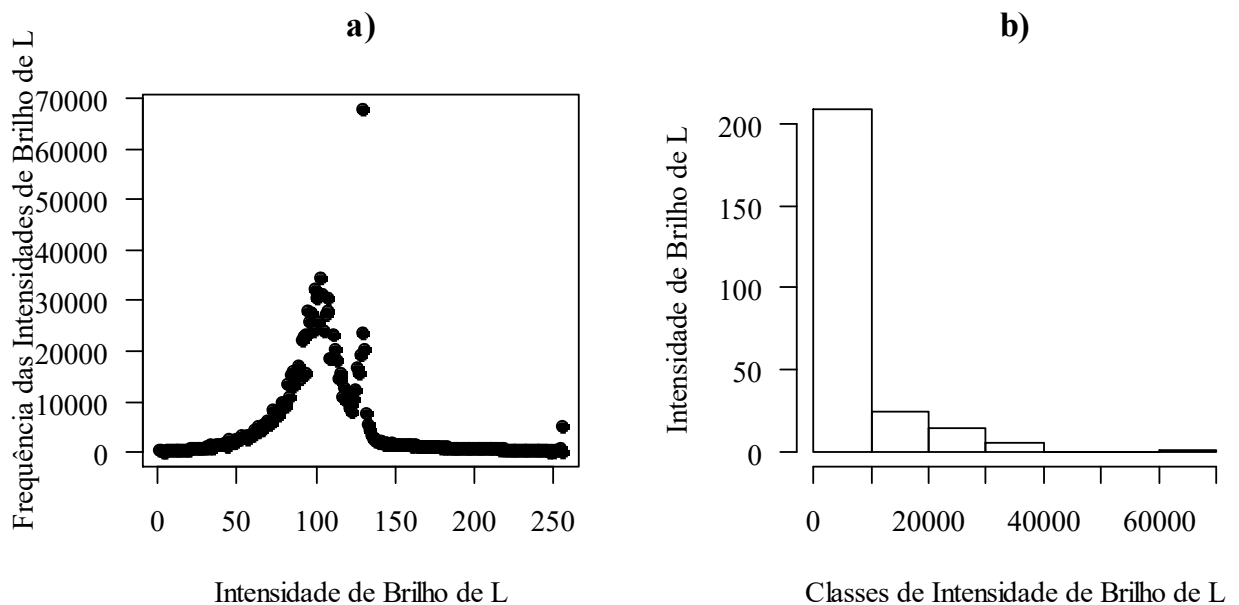
superior aos demais métodos utilizados. Cabe esclarecer que, apesar do método *B-SPLINES+MV* gerar resultados que resultam em parâmetros estatísticos significativos, este método não gerou coeficientes estatisticamente relevantes. Daí a superioridade do método MV.

Assim como o que ocorreu para a componente G, o adensamento gerado por *B-Splines* e que antecedeu ambos os métodos (*B-SPLINES+MMQ* e *B-SPLINES+MV*), ocasionou ganho de qualidade apenas nos resultados advindos do método *B-SPLINES+MMQ*. Tal afirmação pode ser confirmada pela comparação dos valores numéricos presentes na tabela 10.

#### 4.2.4 Experimento com a imagem de Luminância (L)

Por meio da análise do gráfico de dispersão (PxL), presente na Figura 29a), e do histograma (LxP), presente na Figura 29b), pode-se intuir o comportamento variacional da imagem de luminância.

**Figura 29** - Comportamento da imagem de luminância L: **a)** Gráfico de Dispersão (PxL); **b)** Histograma de L (LxP).

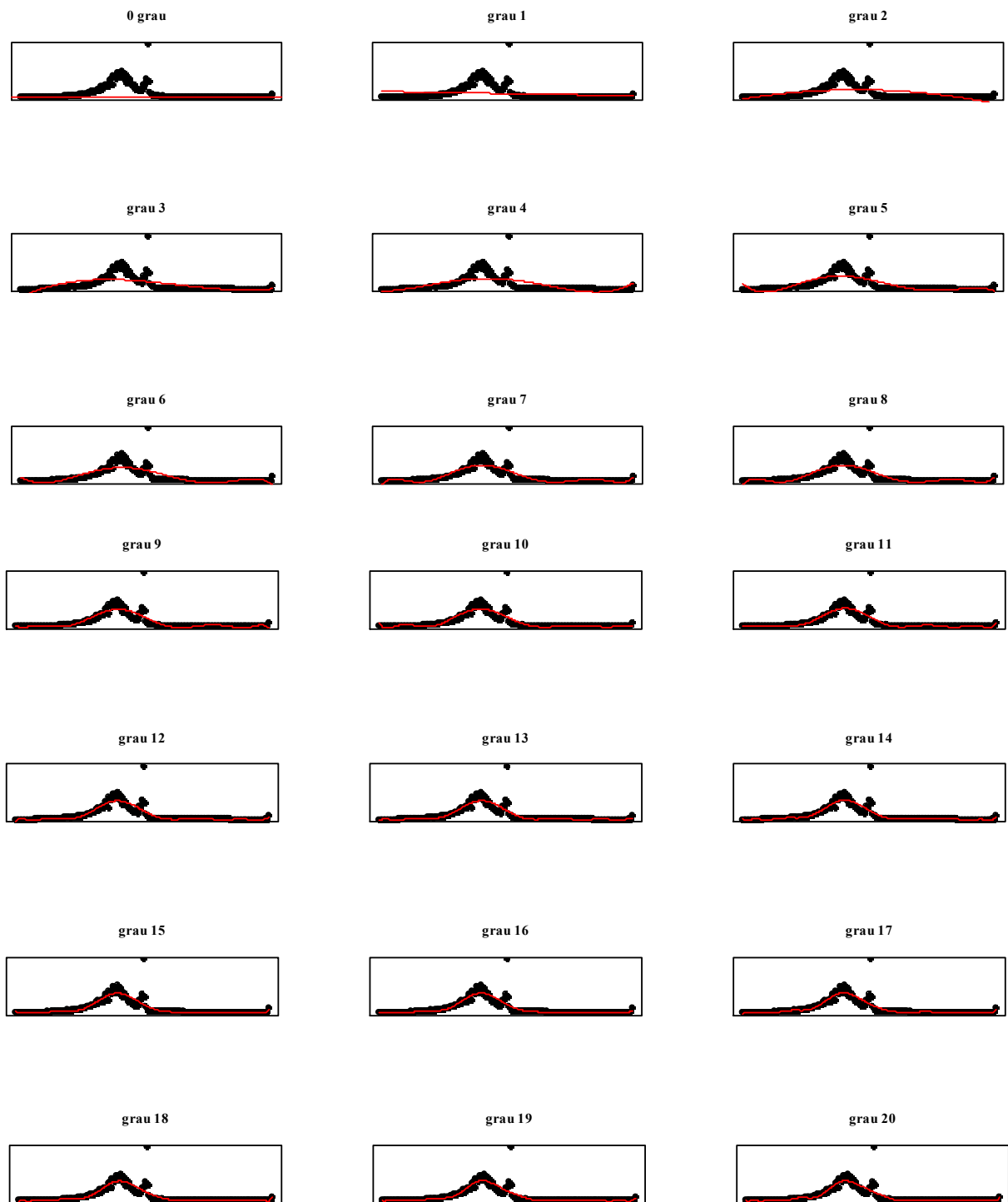


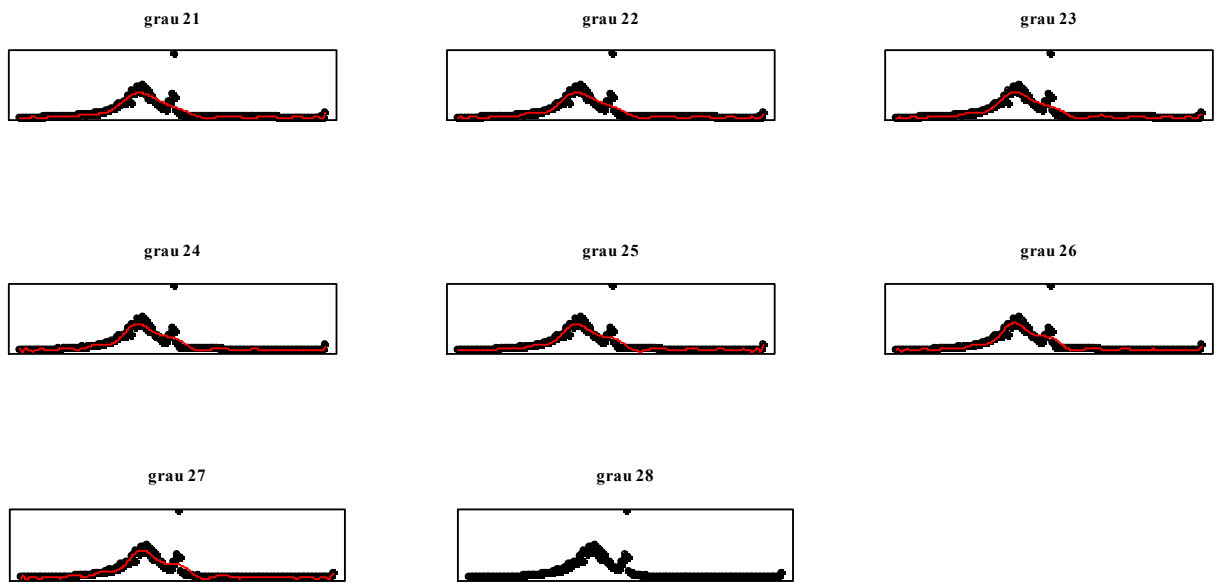
Fonte: Resultados da pesquisa.

Por meio do MMQ buscou-se estimar o grau de polinômio a ser utilizado como modelo pelas metodologias de ajuste propostas. O esboço dos gráficos relativos a cada ajustamento são apresentados na Figura 30. Nos esboços, os pontos pretos representam os pontos de controle iniciais e as curvas em vermelho representam os ajustes polinomiais obtidos. A qualidade de

cada ajustamento, no tocante à representação dos dados originais, pode ser intuída a partir dos referidos esboços.

**Figura 30** – Esboços dos gráficos relativos aos sucessivos ajustes via MMQ para a determinação de grau do polinômio.

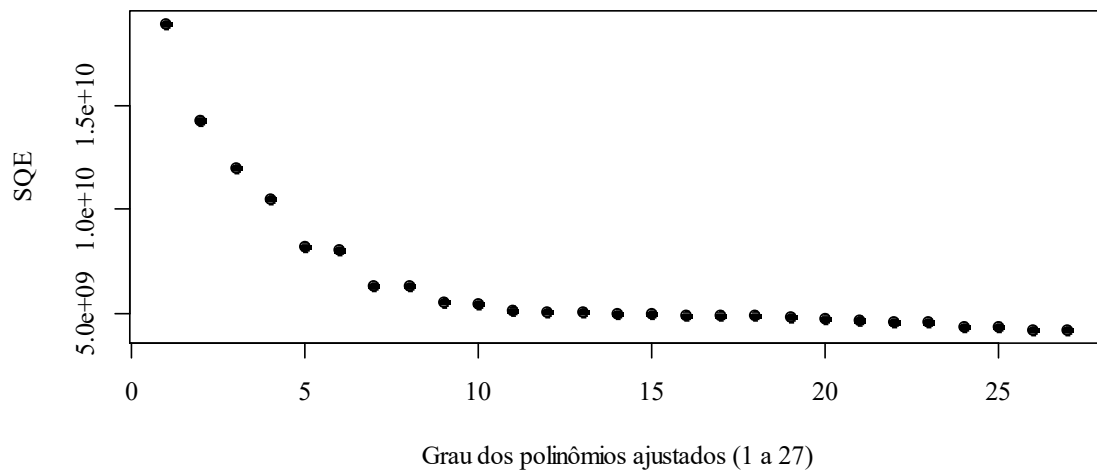




Fonte: Resultados da pesquisa.

A partir dos cálculos efetuados via MMQ, chegou-se à conclusão preliminar que o modelo a ser utilizado nos ajustamentos previstos é um polinômio de grau 27. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada pelo RStudio. Porém, antes da decisão definitiva pelo polinômio de grau 27 é coerente que o comportamento do erro quadrático seja analisado. Assim, para facilitar a referida análise tem-se o gráfico da Figura 31, no qual o comportamento do erro pode ser observado, mediante a evolução do grau do polinômio ajustado via MMQ.

**Figura 31** – Comportamento da soma dos resíduos quadráticos de cada polinômio ajustado.

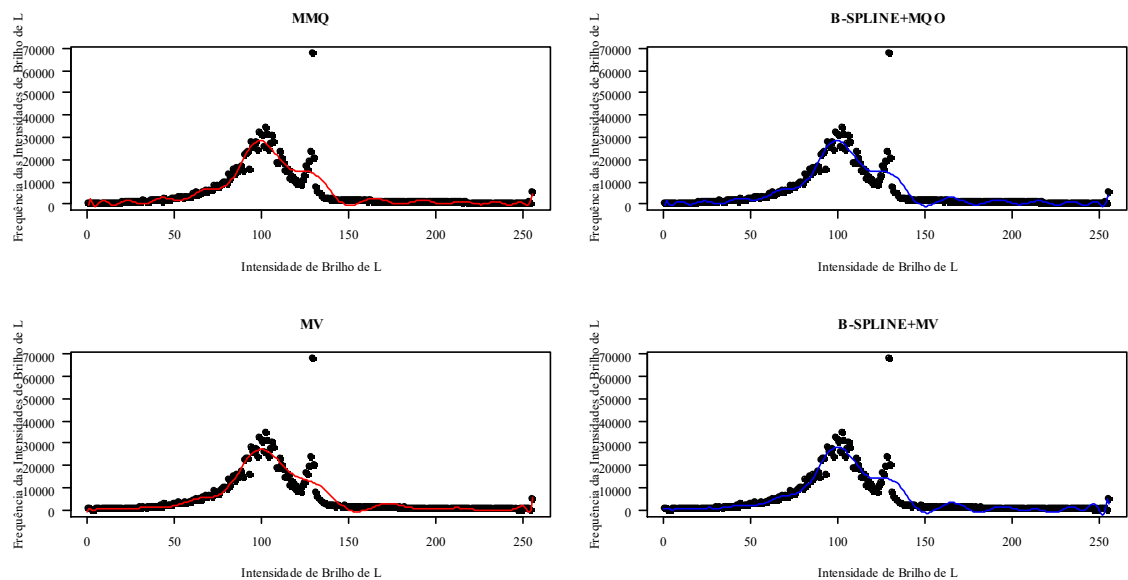


Fonte: Resultados da pesquisa.

Pela análise visual do gráfico da Figura 31, tem-se a noção do decaimento assintótico do erro quadrático mediante o aumento do grau do polinômio ajustado via MMQ, representante dos dados provenientes da imagem de luminância (L). Verifica-se que o mínimo resíduo está atrelado ao ajuste relativo ao modelo polinomial de grau 27. Desta maneira, escolheu-se um polinômio de grau 27 como modelo para os ajustes.

Os gráficos relativos aos ajustes efetuados pelos métodos previstos na metodologia são apresentados na Figura 32.

**Figura 32** – Gráficos das curvas obtidas mediante a execução dos métodos de ajuste propostos.

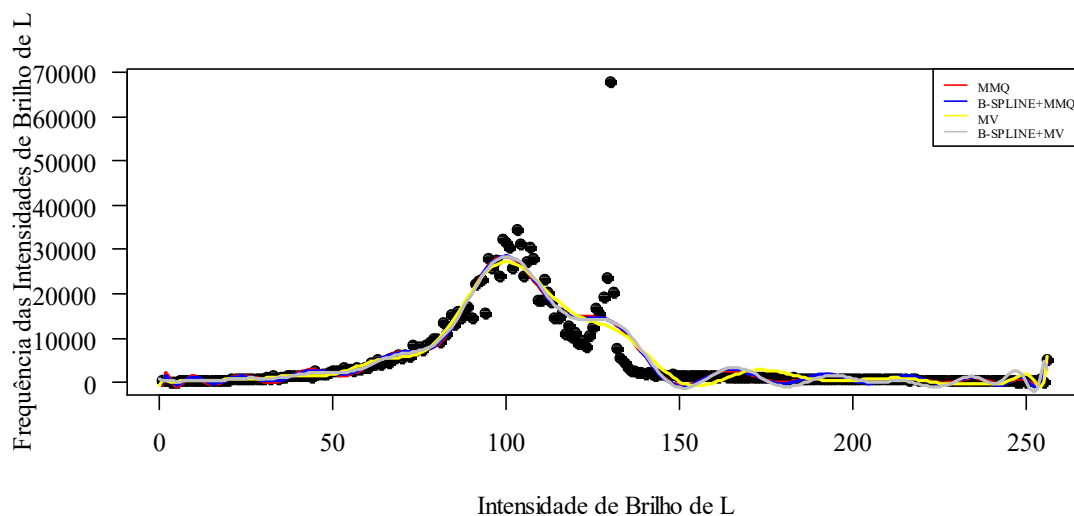


Fonte: Resultados da pesquisa.

Observa-se que as curvas presentes na Figura 32 são bastante semelhantes, sendo que a que é mais discrepante das demais é a curva relativa ao ajuste efetuado via MV. Afim de tentar melhorar a identificação visual das discrepâncias entre as curvas, estas são apresentadas conjuntamente no gráfico da Figura 33.



**Figura 33** – Gráfico das curvas sobrepostas relativas aos polinômios representativos do comportamento da imagem de luminância L.



Fonte: Resultados da pesquisa.

Considerando as curvas presentes no gráfico da Figura 33, verifica-se que a curva relativa ao ajuste efetuado via MV realmente é a mais discrepante das demais. Porém, afim de se poder avaliar mais precisamente a qualidade dos ajustes tem-se a tabela 11, na qual os coeficientes, resultantes de cada método de ajustamento, são apresentados, juntamente com os respectivos coeficientes de significância.

**Tabela 11** – Valores dos coeficientes polinomiais e suas significâncias estatísticas: Imagem de luminância L.

GRAU	Coefficiente s -MMQ-	P- VALOR (%)	Coefficientes -B-SPLINES +MMQ-	P- VALOR (%)	Coefficiente s -MV-	P- VALOR (%)	Coefficientes -B-SPLINES +MV-	P- VALOR (%)
0	5.339,60	<2,00E-16	-3,44E+02	0,94	5.342,32	<2,2E-16	254	2,08E-06
1	-25.786,00	7.88E-11	4,32E+04	0,61	-26.086,20	3,78E-06	-1.072	6,30E-01
2	-68.216,30	<2,00E-16	-6,50E+05	0,50	-69.048,50	<2,2E-16	29.189	4,83E-01
3	47.464,60	<2,00E-16	5,58E+06	0,45	46.816,76	6,66E-16	-480.986	3,03E-01
4	38.617,90	<2,00E-16	-3,24E+07	0,44	39.093,28	6,74E-13	4.994.259	1,71E-01
5	-47.966,00	<2,00E-16	1,38E+08	0,45	-46.226,99	<2,2E-16	-35.740.782	8,91E-02
6	-12.512,10	3,98E-03	-4,43E+08	0,49	-10.853,72	4,56E-02	190.130.533	4,26E-02
7	41.914,50	<2,00E-16	1,09E+09	0,56	42.559,79	4,91E-14	-786.773.155	1,88E-02
8	-283,40	0,94752	-1,98E+09	0,66	-1.205,19	8,00E-01	2.607.681.331	7,78E-03
9	-27.245,40	1,26E-09	2,39E+09	0,80	-28.736,74	1,51E-09	-7.056.821.937	3,11E-03
10	10.091,9	0,01982	-7,07E+08	0,97	9.502,50	4,34E-02	15.796.182.062	1,26E-03
11	17.226,20	8,39E-05	-4,85E+09	0,85	16.644,74	7,09E-04	-29.507.112.792	5,41E-04
12	-9.717,30	0,02481	1,46E+10	0,66	-10.246,67	4,31E-02	46.281.013.608	2,61E-04

continua

continuação

13	-4.516,10	0,29484	-2,61E+10	0,49	-5.993,72	1,99E-01	-61.218.979.079	1,49E-04
14	9.710,70	0,02491	3,47E+10	0,36	8.207,63	4,29E-02	68.513.797.969	1,02E-04
15	-1.292,50	0,76407	-3,67E+10	0,27	-329,69	9,39E-01	-65.019.420.622	8,54E-05
16	-3.047,10	0,47939	3,18E+10	0,20	-772,26	8,62E-01	52.382.893.551	8,60E-05
17	7.602,30	0,07848	-2,28E+10	0,17	10.547,33	5,95E-03	-35.825.064.545	1,03E-02
18	-448,10	0,91712	1,37E+10	0,14	358,00	9,32E-01	20.763.976.731	5,78E-02
19	-7.974,30	0,065503	-6,69E+09	0,14	-9.854,17	2,86E-02	-10.159.689.370	2,02E-01
20	6.774,30	0,11664	2,72E+09	0,14	6.383,64	1,21E-01	4.167.262.108	5,42E-01
21	9.536,60	0,02759	-8,98E+08	0,16	12.013,47	3,78E-03	-1.416.321.246	7,56E-01
22	-8.791,90	0,04209	2,36E+08	0,19	-3.395,64	3,54E-01	391.454.045	0,2531
23	-6.946,30	0,10769	-4,76E+07	0,25	-4.159,31	1,77E-01	-85.345.556	0,2373
24	12.579,60	0,00380	6,91E+06	0,35	6.031,88	5,66E-02	13.909.906	0,2247
25	6500,50	0,13208	-6,13E+05	0,52	2.583,19	2,94E-01	-1.514.121	0,2118
26	-12.663,20	0,00357	1,35E+04	0,87	-2.102,70	2,05E-01	-78.663	0,1739
27	599,70	0,88923	4,31E+03	0,47	75,78	8,32E-01	2.637	0,2961

---

**Quantidade total de coeficientes significantes de cada estimativa**

coef.sig	19	0	17	15
----------	----	---	----	----

**Obs.:** Todas as células em cinza remetem à presença de significância estatística do seu respectivo coeficiente. Os “coeficientes significantes” (**coef.sig**) são aqueles cujos valores de significância estatística são iguais ou inferiores a 10%.

Fonte: Resultados da pesquisa.

Dos valores presentes na tabela 11, verifica-se que as metodologias que geraram o maior número de coeficientes validados em suas formas funcionais foram o MMQ e o MV, com ligeira vantagem para o MMQ. Pode-se observar também que o método *B-SPLINES*+MV também foi eficiente em gerar coeficientes significantes, porém em menor quantidade.

As estimativas dos coeficientes mostradas na tabela 11 dizem respeito às seguintes formas funcionais polinomiais:

1) **MMQ:**

$$\hat{L}_i = 5.339,60 - 25.786,00P_i + \dots - 599,70P_i^{27} + e_{Li} \quad (63)$$

2) ***B-SPLINES*+MMQ:**

$$\hat{L}_i = -(3,44E + 02) + (4,32E + 04)P_i + \dots + (4,31E + 03)P_i^{27} + e_{Li} \quad (64)$$

3) **MV:**

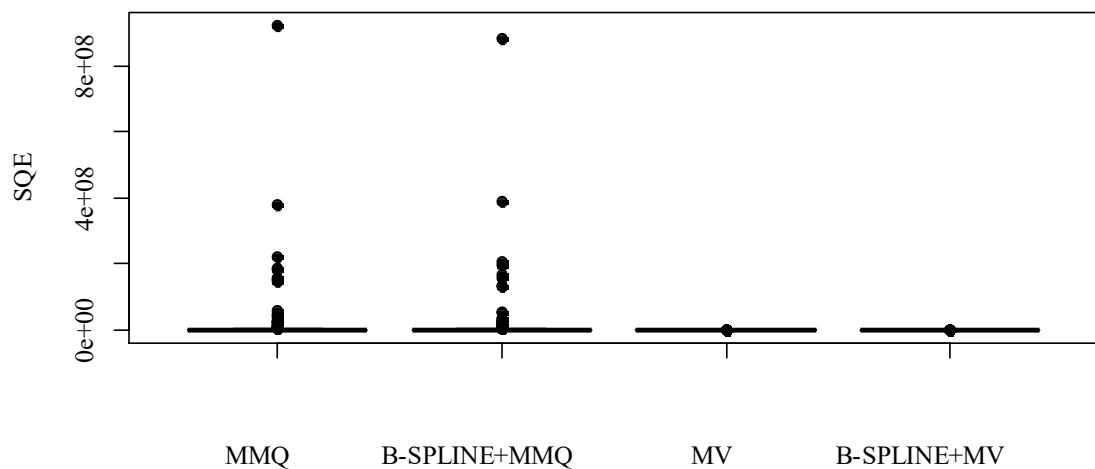
$$\hat{L}_i = 5.342,32 - 26.086,00P_i + \dots + 75,78P_i^{27} + e_{Li} \quad (65)$$

4) **B-SPLINES+MV:**

$$\hat{L}_i = 254 - 1.072P_i + \dots + 2.637P_i^{27} + e_{Li} \quad (66)$$

Com o objetivo de complementar a análise dos resultados, o comportamento dos erros/resíduos ( $e_{Li}$ ) quadráticos, relativo aos resultados provenientes de cada método de ajustamento, pode ser depreendido da análise do gráfico presente na Figura 34.

**Figura 34** – Gráfico *Boxplot* dos resíduos quadráticos dos ajustes para a componente L.



Fonte: Resultados da pesquisa.

Pela análise do gráfico da Figura 34, pode-se exarar que os métodos MV e *B-SPLINES+MV* são aqueles que geram resultados com menor erro. Em adição a esta informação, a tabela 12 revela mais informações sobre os resíduos decorrentes de cada ajuste.

**Tabela 12** – Parâmetros estatísticos baseados nos resíduos relativos aos ajustes efetuados para a imagem de luminância L.

<b>PARÂMETROS ESTATÍSTICOS</b>	<b>MMQ</b>	<b><i>B-PLINES+</i> MMQ</b>	<b>MV</b>	<b><i>B-SPLINES+</i> MV</b>
<b>SQE</b>	3,00E+09	2,90E+09	1.657,45	1.657,45
<b>MÉDIA</b>	1,17E+07	1,13E+07	6,47	6,47
<b>DP</b>	6,75E+07	6,50E+07	51,46	51,46
<b>CV(%)</b>	5,76E+02	5,75E+02	794,79	794,79
<b>R<sup>2</sup>(%)</b>	8,20E+01	8,26E+01	99,99	99,99

Fonte: Resultados da pesquisa.

Da análise dos parâmetros estatísticos presentes na tabela 12, depreende-se a melhor qualidade dos resultados advindos dos métodos MV e *B-SPLINES*+MV, que apresentam menores valores de SQE, MÉDIA e DP, também os maiores valores de R<sup>2</sup>. Considerando a igualdade numérica dos parâmetros estatísticos para ambos os métodos, conclui-se pela superioridade do método MV, pois, de acordo com a tabela 11, o método MV dá origem a 17 coeficientes estatisticamente significantes, ao passo que, o método *B-SPLINES*+MV gera apenas 15 coeficientes validados estatisticamente. Daí a justificativa de se eleger o método MV como o melhor neste caso particular.

Ainda da análise dos valores da tabela 12, verifica-se que o adensamento de pontos gerado por meio de *B-Splines* contribuiu positivamente somente com os resultados advindos do MMQ, não refletindo em ganho de qualidade junto ao método MV.

## 5 CONCLUSÕES E RECOMENDAÇÕES

As seções que seguem se destinam à apresentação das principais conclusões advindas da execução da pesquisa aqui relatada. Além disso, as principais recomendações associadas ao tema estudado e à pesquisa efetuada são listadas e se discorre a respeito dessas recomendações.

### 5.1 CONCLUSÕES

A pesquisa aqui relatada teve como objetivo o estudo de métodos de aproximação/interpolação (*B-Splines*, método dos Mínimos Quadrados Ordinários e método da Máxima Verossimilhança) a fim de, a partir de funções de domínio discreto, estabelecidas com base em histogramas dos canais vermelho, verde e azul de imagens digitais e em imagens de luminância, se obter fórmulas matemáticas que representassem, mesmo que de modo aproximado, tais funções (histogramas). As formulações foram obtidas por meio dos referidos métodos de aproximação/interpolação e foram estatisticamente comparadas.

Para tanto, inicialmente foi feito um estudo dos métodos de aproximação/interpolação objetivando abstrair os conceitos matemáticos e algorítmicos relativos aos métodos. Na sequência, construiu-se um programa, em linguagem de programação C++, que decompunha imagens coloridas em componentes R, G e B e estabelecia a frequência de cada intensidade de brilho relativa a cada uma das componentes. E mais, tal programa ainda é capaz de calcular a imagem de luminância da respectiva imagem de entrada e gerar o histograma relativo à esta imagem calculada.

De posse dos conteúdos pertinentes aos métodos de interpolação/aproximação e dos dados de entrada, efetuou-se, por meio do *software* RStudio, o processamento dos dados relativos aos histogramas gerados, considerando o modelo polinomial e os métodos MMQ, *B-Splines* acoplado aos MMQ, MV e *B-Splines* acoplado ao MV. Cabe explicar que, o método de aproximação *B-Splines* foi introduzido na pesquisa com o objetivo de gerar um adensamento nos conjuntos de pontos originais, a fim de que se pudesse verificar os efeitos que tais adensamentos poderiam causar aos resultados advindos dos métodos MMQ e MV.

Após a geração dos resultados, estes foram sujeitos à avaliação/comparação estatística a fim de se determinar a qualidade dos resultados e quantificar o desempenho de cada método utilizado.

Nestes termos, após análises visuais dos gráficos e parâmetros estatísticos pôde-se concluir o seguinte:

### **1. Teste relativo à componente R:**

O teste com os dados relativos à componente R apontou para uma semelhança gráfica das curvas obtidas a partir dos polinômios gerados. Neste caso, as formas funcionais ajustadas foram polinômios de grau 27. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada pelo RStudio. A escolha do grau do polinômio usado como modelo para os métodos de ajustamento se baseou também na análise dos erros (via MMQ).

Da análise visual, supõem-se que se possa escolher qualquer dos métodos testados para a geração de curvas polinomiais. Tal escolha dependerá do tipo de problema e das características particulares inerentes ao problema. No entanto, numericamente os resultados advindos do método MV foram aqueles que continham o menor erro. E mais, o polinômio ajustado via MV foi o que possuiu a maior quantidade de coeficientes significativos (18 coeficientes de um total 28).

Neste caso em particular, o adensamento de pontos dos conjuntos de dados de entrada, efetuado via *B-Splines*, não gerou acréscimo de qualidade em nenhum dos resultados gerados.

### **2. Teste relativo à componente G:**

De modo similar ao que ocorreu para componente R, verificou-se também uma semelhança gráfica das curvas obtidas a partir dos polinômios ajustados. Como no caso anterior, as formas funcionais ajustadas foram polinômios de grau 27. A escolha do grau do polinômio usado como modelo para os métodos de ajustamento se deu com base nos mesmos critérios utilizados anteriormente.

Durante o processamento via *B-SPLINES+MV*, o *software* RStudio gerou resultado apenas quando o modelo utilizado foi um polinômio de grau 18. Ou seja, para a quantidade de graus de liberdade disponíveis, um polinômio de grau 27 foi a especificação máxima suportada. Acredita-se que tal fato seja proveniente de limitações associadas ao *hardware* ou inconsistências inerentes ao *software*.

Numericamente, os resultados que continham o menor erro foram, uma vez mais, aqueles advindos do método MV. E mais, novamente o polinômio ajustado via MV foi o que possuiu a maior quantidade de coeficientes significativos (20 coeficientes de um total 28).

Quanto ao adensamento do conjunto de pontos via *B-Splines*, este impactou positivamente apenas nos resultados advindos do MMQ. Quanto ao método MV, o adensamento de pontos não gerou ganho de qualidade nos seus resultados.

### **3. Teste relativo à componente B:**

A semelhança gráfica das curvas obtidas a partir dos polinômios ajustados também foi verificada neste caso. E mais, igualmente aos casos anteriores, o modelo utilizado para o ajustamento foi o polinômio de grau 27.

Numericamente, o método MV foi tido como o de melhor desempenho, uma vez que, os seus resultados foram os continham o menor erro. Em adição a isso, novamente o polinômio ajustado via MV foi o que possuiu a maior quantidade de coeficientes significativos (26 coeficientes de um total 28).

Neste caso, pôde-se perceber que o adensamento do conjunto de pontos via *B-Splines* impactou positivamente apenas nos resultados advindos do MMQ. Quanto ao método MV, o adensamento de pontos não gerou ganho de qualidade nos resultados.

### **4. Teste relativo à imagem de luminância (L):**

A análise das curvas geradas a partir das formas funcionais ajustadas apontou para a semelhança dos ajuste, os quais foram realizados tendo como modelo um polinômio de grau 27.

O método MV foi o que gerou os melhores resultados, visto que estes foram o que apresentaram os melhores valores de parâmetros estatísticos associados ao erro. No entanto, neste caso, o MMQ gerou uma maior quantidade de coeficientes polinomiais significativos (19 coeficientes de um total de 28).

Assim como nos dois casos anteriores, apenas os resultados advindos do ajuste via MMQ tiveram acréscimo de qualidade em decorrência do adensamento do conjunto de pontos de controle.

Nestes termos, considerando os testes e análises efetuadas, conclui-se que o método que gerou melhores resultados para todos os conjuntos de dados utilizados foi a Máxima Verossimilhança. Tal fato justifica o crescente uso desta metodologia, conforme o verificado na literatura pertinente. E mais, o adensamento do conjunto de pontos de controle efetuado via *B-Splines* foi eficaz em gerar resultados de melhor qualidade apenas quando acoplado ao MMQ. Por meio da experimentação, pôde-se verificar também que o aumento do grau do polinômio utilizado como modelo para os métodos de ajustamento implicava na redução do erro resultante.

Por fim, cabe expressar que o ganho conceitual, algorítmico e experimental advindo da execução da pesquisa não é quantificável e é o que de melhor se pôde abstrair.

## 5.2 RECOMENDAÇÕES

Em virtude do ganho de experiência gerado pela execução da pesquisa recomenda-se o seguinte:

- **Criação de métodos híbridos:** a partir das bases teóricas e experimentais estabelecidas neste trabalho, vislumbra-se a aplicação do estado da arte pela geração de teorias híbridas que abarquem elementos dos métodos estudados e, se conveniente, de novos e distintos métodos;
- **Uso de MMQ e do método MV nas suas formas ponderadas:** recomenda-se reeditar a pesquisa considerando os métodos MMQ e MV na sua forma ponderada;
- **Uso de MMQ e do método MV com modelos não lineares:** considerando que o modelo utilizado nesta pesquisa era linear, poder-se-ia reeditar a pesquisa considerando situações-problema que envolvessem modelos não-lineares;
- **Implementação de *software* que execute de modo autônomo a metodologia aqui apresentada:** considerando a abrangência do tema e a necessidade de se desenvolver novas habilidades e ferramentas, poder-se-ia implementar computacionalmente a metodologia aqui apresentada afim de se ter um *software* dedicado que executasse todo o processo aqui descrito. Dessa forma, poder-se-ia



testar exaustivamente e com inúmeras/diferentes entradas o *software* construído para fins de validação de seus resultados. Em adição a isso, seria possível comparar os resultados advindos do *software* com aqueles gerados no trabalho de pesquisa aqui descrito, também para fins de validação da implementação.

- **Estudo de outros métodos de aproximação/interpolação:** de forma similar ao proposto e executado na pesquisa aqui descrita, seria possível buscar por outros métodos de aproximação/interpolação e se verificar de modo comparativo o seu desempenho. Para tanto, tais métodos poderiam ser utilizados em sua forma original, acoplados a outras metodologias ou matematicamente alterados.

## REFERÊNCIAS

- ANTON, H.; RORRES, C. *Álgebra linear com aplicações*. Porto Alegre: Bookmann, 2001.
- BRASIL, Parâmetros Curriculares Nacionais; SEMTEC, *Parâmetros Curriculares Nacionais. ensino médio*. Brasília: Ministério da Educação, 1999.
- BRASIL, Secretaria da educação Básica. *Orientações Curriculares para o Ensino Médio: Ciências da Natureza, Matemática e suas Tecnologias*. Brasília, MEC, 2006.
- De BOOR, C. *Package for calculating with B-Splines*. Journal of Numerical Analysis, 1977, 14:441–472.
- De BOOR, C. *A practical guide to Splines. Applied Mathematical Sciences*. New York: SpringerVerlag, 2001.
- DE SOUZA, Sandro Amorim. *ESTUDO DO PRODUTO MATRICIAL POR MEIO DO MÉTODO DOS MÍNIMOS QUADRADOS: UMA ABORDAGEM DESTINADA AO ENSINO MÉDIO*. 2014.
- DRAPER, N.; SMITH, H. *Applied regression analysis*. 3 ed. New York: J. Wiley, 1998.
- GELADI, Paul; GRAHN, Hans F. *Multivariate image analysis*. John Wiley & Sons, Ltd, 1996.
- CRÓSTA, A. P. *Processamento digital de imagens de sensoriamento remoto*. Campinas: IG/UNICAMP, 1992.
- GOMES, J.; VELHO, L. *Computação Gráfica: Imagem*. IMPA-SBM, 1994.
- GONZALEZ, Rafael C.; WOODS, Richard E. *Processamento de imagens digitais*. Edgard Blucher, 2000.
- GRAYBIL, F.A. *Theory and application of the linear model*. Duxbury, North State, Massachusetts, 1976, \_\_p.
- HARTLEY, H.O. ; RAO, J.N.K. *Maximum likelihood estimation for the mixed analysis of variance model*. Biometrika, 54, p. 93-108, 1967.
- HASTIE, T, J; TIBSHIRANI, R, J. *Generalized Additive Models*. 1 ed. New York: Chapman and Hall, 1990.
- MENEGUETTE JR, M., ARTERO, A. O BARBOSA, R. L. *Introdução à visão computacional*. Semana dos cursos de Matemática, Estatística e Computação – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, 2003.

- MENEZES, Isabella Barbosa; DUARTE, Marco Aparecido Queiroz; SILVA, Willian Carvalho; KIKUTI, Hamilton; NOGUEIRA, Nadia Rodrigues. *Interpolação Spline usada no cálculo da altura da cana-de-açúcar*. Revista Omnia Exatas, v.3, n.1, p.17-25, 2010.
- MENEZES, G. R. O; TORRES, R.A; SARMENTO, J. L; RODRIGUES, M.T; BRITTO, L.F; LOPES, P. S; SILVA, F. G. *Modelos de regressão na avaliação da produção de leite em cabras da raça Saanen*. Revista Brasileira de Zootecnia, v. 40, p.1526-1532, 2012.
- MEYER, K. *Random regression analyses using B-splines to model growth of Australian Angus cattle*. Genetics, Selection, Evolution, v. 37, p.473-500, 2005.
- NUNHEZ, J. R. *Método Spline modificado: acoplamento do Método da Máxima Verossimilhança ao Método Spline*. 1990. 117 f. Dissertação (Mestrado em Engenharia Química). Faculdade de Engenharia de Campinas, UNICAMP – Universidade de Campinas, Campinas, 1990.
- OLIVEIRA, Diana Campos de et al. *Funções Splines para estudo de curvas de crescimento em ovinos cruzados*. 2011.
- PEREIRA, L. R.; CARVALHO, C. M. C.. *Método de ajuste de curvas usando B-Spline com peso*. In: ABCM SYMPOSIUM SERIES IN MECHATRONICS - Vol. 6. 2014, Rio de Janeiro.
- PIEGL, L.; TILLER, W.. *The NURBS Book*. 2. Ed. NY: Springer, 1997. 641 p.
- RAWLINGS, John O.; PANTULA, Sastry G.; DICKEY, David A. *Applied regression analysis: a research tool*. 1998. Wadsworth and Brooks.
- RIBEIRO, Renato de Assis et al. *Análise temporal através da interpolação e do ajuste de curvas pelo método dos quadrados mínimos*. 2014.
- RESENDE, M. D. V; SILVA, F. F; AZEVEDO, C. F. *Estatística Matemática, Biométrica e Computacional: Modelos Mistos, Multivariados, Categóricos e Generalizados (REML/BLUP), Inferência Bayesiana, Regressão Aleatória, Seleção Genômica, QTL-GWAS, Estatística Espacial e Temporal, Competição e Sobrevivência*. Visoça, MG: Suprema, 2014.
- RUGGIERO, M. e LOPES, V., *Cálculo Numérico: Aspectos Teóricos e Computacionais*. Makron Books, 1996.
- SÁNCHEZ, J.P; MISZTAL, I; AGUILAR, I; BERTRAND, J.K. *Genetic evaluation of growth in a multibreed beef cattle population using random regression-linear spline models*. Journal de la Société Française de Statistique, v. 143, p.155-165, 2008.
- SEARLE, S.R. *Linear models*. New York, John Wiley & Sons, 1971. 532p.
- SILVA, Talita F.. *Análise de comparação dos detectores de borda Canny e Canny-EDP*. Monografia (graduação). Universidade do Estado de Mato Grosso – UNEMAT, Campus Universitário do Vale do Teles Pires. Colider, Mato Grosso. 2010.

VALE, Giovane Maia do et al. *Uso de técnicas de otimização e agrupamento hierárquico para a complementação automática da malha viária em regiões rurais a partir de imagens aéreas de média e alta resolução*. 2007.

WELLAM, S.J; CULLIS, B. R; KENWARD, M; THOMPSON, R. *The analysis of longitudinal data using mixed model L-splines*. Biometrics, v.62, p.392-401, 2005.

WELLAM, S.J; CULLIS, B. R; KENWARD, M; THOMPSON, R. *A comparasion of mixed model splines for curve fitting*. Australian and New Zealand. Journal of Statistics, v. 49, p. 1-23, 2007.

## APÊNDICE I

Quadro contendo os histogramas dos canais R, G e B da imagem digital de entrada e o histograma da imagem de luminância (L).

Intensidade de Brilho (P)	Frequência de R	Frequência de G	Frequência de B	Frequência de L
0	5018	376	146	250
1	624	134	34	268
2	529	117	63	178
3	721	155	30	189
4	12	130	20	212
5	699	159	40	263
6	745	184	149	285
7	801	154	71	364
8	94	235	138	393
9	911	290	191	461
10	897	224	141	269
11	937	418	219	415
12	76	320	150	464
13	1115	332	193	517
14	1164	290	263	429
15	1261	75	279	485
16	93	508	361	542
17	1410	449	280	496
18	1443	346	491	519
19	1452	246	456	532
20	276	664	448	633
21	1704	422	443	618
22	1754	497	396	585
23	2010	374	369	577
24	331	601	510	862
25	1964	477	283	663
26	2484	436	387	779
27	2517	940	713	832
28	372	606	628	798
29	2540	760	539	784
30	5849	721	501	969
31	509	515	517	709
32	3618	1467	551	1397
33	267	475	552	806
34	6819	828	1131	1147
35	820	746	618	977
36	4628	884	691	1414
37	423	664	562	1312
38	7938	1689	1440	1434
39	1609	617	304	1511
40	6075	1184	1355	1597

41	1105	1052	1036	1321
42	8659	1661	1250	1850
43	4090	406	457	1285
44	7913	1955	1647	2415
45	2319	1328	707	1833
46	11250	1869	894	1984
47	7686	878	1205	1872
48	11430	2132	2313	2332
49	3055	1854	879	2720
50	13496	2184	1233	2719
51	9573	1057	1174	2517
52	16332	2669	2048	3426
53	4018	1164	1265	2775
54	15750	1664	1574	2985
55	10909	1921	2017	2451
56	22962	4755	2363	3279
57	7678	1133	1629	2805
58	20360	2203	1720	3469
59	20367	2510	1520	3367
60	22035	4702	2112	4304
61	14151	1564	1341	4151
62	24006	3895	2350	5003
63	29533	1890	2364	3985
64	12234	2323	2041	4630
65	22283	4133	5119	4412
66	32640	1580	1717	5573
67	30966	4312	2868	5323
68	16620	4380	2884	6318
69	22410	4843	2911	5389
70	35095	3321	4136	6309
71	34850	5668	3791	5762
72	15337	4442	3944	8245
73	23877	5422	7049	7274
74	32580	2311	2342	7948
75	32777	6078	4819	7322
76	15543	4853	4853	8526
77	24708	4052	4915	9012
78	26380	6348	4724	9926
79	26332	7985	8571	9802
80	14978	5593	7309	9154
81	20871	6381	7210	13559
82	20212	4837	5750	11047
83	20853	8896	6866	15209
84	13450	8789	9015	13204
85	14269	6155	7243	15843
86	14793	6704	6862	14715
87	14472	9371	11724	16132
88	11065	7352	13345	17200

89	10474	8056	14423	14699
90	11640	9154	8667	21979
91	9308	11537	16474	23029
92	7370	10976	13029	23175
93	7162	8905	11013	15558
94	8196	13918	18678	28013
95	6684	12266	21681	25658
96	5780	14414	24111	27480
97	5356	11034	13083	23787
98	5778	17972	27620	32443
99	4991	11840	25001	31483
100	4808	16940	13413	30373
101	4687	16536	35493	25875
102	3953	19222	33829	34645
103	5131	24693	38280	31325
104	5001	18715	22862	23832
105	3813	26032	50300	27161
106	3139	13355	32300	30608
107	3415	25678	17477	27924
108	2934	26914	44180	18689
109	2768	18515	36267	18698
110	2761	30327	26161	23206
111	3020	32037	40693	20206
112	2288	31958	40629	18322
113	1722	22726	41044	14366
114	2159	31083	26206	14484
115	1619	41511	38526	15577
116	2784	42714	29090	10879
117	2428	50867	24695	12791
118	2251	35477	40003	10207
119	2311	65023	25688	11229
120	2516	47288	25792	8636
121	2450	38774	39109	9574
122	2244	27534	27967	7951
123	2054	32019	17322	10560
124	2446	13825	9727	12411
125	2208	18432	15332	16566
126	1988	3884	10906	15796
127	1813	16313	11644	19288
128	2305	22182	6364	23755
129	1974	9511	4446	67957
130	1990	14045	6869	20180
131	1947	10480	5695	7508
132	1895	10221	5558	5466
133	2090	10649	6491	4446
134	2009	7935	5790	3141
135	2100	9726	5385	2751
136	2171	7312	4051	2404

137	2517	6961	5429	2164
138	2331	4942	3222	2062
139	2964	6884	3704	1960
140	2233	4606	3339	2117
141	4735	4880	3435	1852
142	7723	3650	3009	1621
143	7400	3817	2552	1740
144	2197	3814	3039	1919
145	5286	3119	2113	1640
146	4560	3270	2478	1553
147	5649	2909	2284	1752
148	4088	2477	2194	1627
149	4300	1927	1828	1486
150	3532	2780	1515	1519
151	6879	1967	2254	1548
152	8432	2159	1504	1576
153	12418	1707	1520	1391
154	7335	1981	1903	1515
155	11097	1878	1513	1655
156	53125	1867	1764	1470
157	19960	1506	1405	1424
158	4557	1773	1889	1616
159	1984	1430	1179	1480
160	2901	1226	1307	1281
161	1243	1811	1676	1536
162	1347	1488	1097	1454
163	1317	1710	1595	1277
164	1195	1016	1125	1224
165	978	1582	1655	1286
166	1028	1426	889	1223
167	987	1469	1201	1145
168	1097	913	1370	1203
169	861	1437	895	1251
170	1031	1139	1375	1084
171	1019	753	1067	1153
172	1092	1407	1291	1195
173	861	956	757	1042
174	1026	1364	1158	1117
175	910	962	1245	1050
176	950	1292	921	1038
177	851	915	1158	1067
178	1077	1144	1107	840
179	957	1265	950	1066
180	821	802	592	876
181	920	1482	950	929
182	848	1024	1029	941
183	1021	1272	829	852
184	926	920	1047	884



185	816	1264	735	776
186	921	1115	878	889
187	839	990	712	922
188	1005	745	695	773
189	767	801	820	784
190	916	958	702	848
191	781	657	789	850
192	993	1097	620	911
193	709	882	811	853
194	1152	887	490	903
195	945	1097	621	867
196	752	769	583	793
197	552	596	457	795
198	901	830	517	804
199	861	708	694	792
200	968	503	524	774
201	510	694	625	705
202	857	680	563	869
203	943	772	524	806
204	840	526	342	684
205	503	721	393	683
206	862	540	553	666
207	961	710	367	616
208	801	423	524	632
209	499	709	408	775
210	1307	525	446	772
211	1069	417	474	707
212	235	437	294	596
213	387	449	399	645
214	891	597	496	677
215	937	417	253	515
216	761	525	373	582
217	357	266	400	575
218	881	533	309	531
219	915	362	497	444
220	815	492	250	498
221	266	152	372	451
222	797	479	272	445
223	827	489	183	422
224	804	135	344	512
225	202	510	330	433
226	1382	706	343	513
227	862	324	202	486
228	139	192	218	343
229	168	441	309	417
230	774	317	117	421
231	753	310	238	384
232	780	234	292	430

233	61	344	204	435
234	728	323	207	393
235	648	267	164	411
236	737	199	128	309
237	43	233	209	366
238	700	201	115	299
239	650	216	209	293
240	657	297	83	241
241	28	192	156	319
242	1164	199	160	238
243	562	329	84	264
244	51	146	162	300
245	18	133	203	293
246	584	193	201	190
247	539	112	7	210
248	515	215	185	228
249	1	153	191	185
250	540	143	3	214
251	464	16	169	396
252	509	204	201	467
253	0	182	162	759
254	0	0	0	0
255	13500	6772	4898	5152