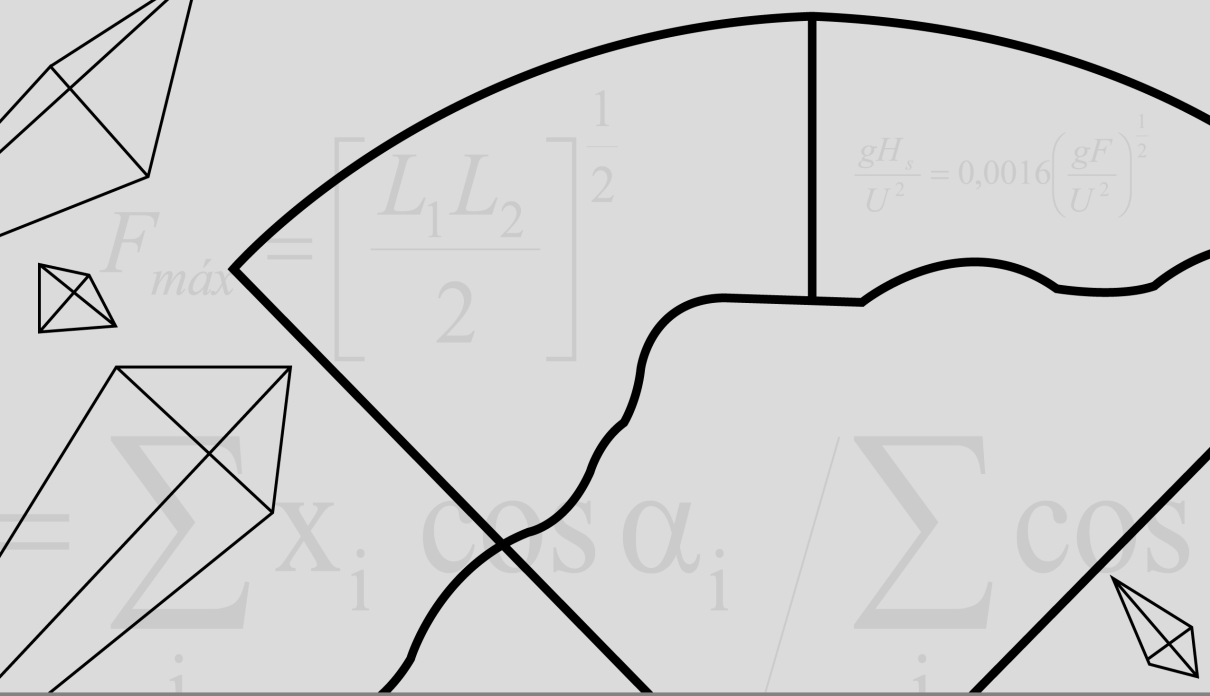


# Avanços Científicos em Matemática, Probabilidade e Estatística

Marcelo Marques  
Isabela Arantes Ferreira



# Avanços Científicos em Matemática, Probabilidade e Estatística

Marcelo Marques  
Isabela Arantes Ferreira

**Editora Chefe**

Marcia A. A. Marques

**Coordenadora Editorial**

Isabela Arantes Ferreira

**Bibliotecária**

Aline Grazielle Benitez

**Diagramação**

Marcos Antonio Ribeiro Pereira

**Arte da Capa**

Matheus Lacerra

**Imagem da Capa**

Freepik

**Revisão**

Os Autores

O conteúdo deste livro está licenciado sob uma licença de Atribuição Creative Commons. Atribuição-Não-Comercial Não Derivativos 4.0 Internacional (CC BY-NC-ND 4.0).



2022 by Bookerfield Editora

Copyright © Bookerfield Editora

Copyright do Texto © 2022 Os autores

Copyright da Edição © 2022 Bookerfield Editora

Os autores cedem à Bookerfield Editora os direitos para esta edição

Esta obra é de natureza digital (e-book). Versões impressas são permitidas, não tendo a Bookerfield Editora qualquer responsabilidade pela confecção e distribuição de exemplares físicos deste conteúdo.

Todos os manuscritos da obra passaram por rigorosa avaliação cega pelos pares, baseadas em critérios científicos e imparciais, recebendo a aprovação após atender os critérios técnicos estabelecidos pelo Conselho Editorial.

Todo o conteúdo do livro e de artigos individuais é de responsabilidade exclusiva de seus respectivos autores, não sendo a Bookerfield Editora responsável por quaisquer eventuais irregularidades.

Situações como plágio, má conduta ética/científica ou dados e resultados fraudulentos são de responsabilidade do autor, comprometendo-se a Bookerfield Editora em investigá-las rigorosamente e tomar as ações cabíveis.

O download, compartilhamento e referenciação da obra são permitidos mediante atribuição de crédito aos autores e à Editora. A comercialização desta obra é expressamente proibida.

## CONSELHO EDITORIAL

### Ciências Agrárias

Afrânio Silva Madeiro  
Alirya Magda Santos do Vale Gomes  
Ana Luiza Trovo Marques de Souza  
André Giarola Boscarato  
Carlos Eugenio Fortes Teixeira  
Daniela Kunkel  
Daniele Cristina Ficanha  
Elson Barbosa da Silva Junior  
Fabiana Schiochet  
Fernanda Beatriz Pereira Cavalcanti  
Fernanda Morcatti Coura  
Flávio José Rodrigues Cruz  
Guilherme Donadel  
Heiriane Martins Sousa  
Jairton Fraga Araujo  
João Francisco Severo Santos  
Joelma Leão Buchir  
Kleber Fernando Pereira  
Maria Cristina Bueno Coelho  
Monyck Jeane dos Santos Lopes  
Pablo Daniel Freitas Bueno  
Renato Jaqueto Goes

### Ciências Biológicas

Afrânio Silva Madeiro  
Alirya Magda Santos do Vale Gomes  
Ana Luiza Trovo Marques de Souza  
André Giarola Boscarato  
Carlos Eugenio Fortes Teixeira  
Daniela Kunkel  
Daniele Cristina Ficanha  
Elson Barbosa da Silva Junior  
Fabiana Schiochet  
Fernanda Beatriz Pereira Cavalcanti  
Fernanda Morcatti Coura

Flávio José Rodrigues Cruz  
Guilherme Donadel  
Heiriane Martins Sousa  
Jairton Fraga Araujo  
João Francisco Severo Santos  
Joelma Leão Buchir  
Kleber Fernando Pereira  
Maria Cristina Bueno Coelho  
Monyck Jeane dos Santos Lopes  
Pablo Daniel Freitas Bueno  
Renato Jaqueto Goes

### Ciências da Saúde

Alexandre Visconti Brick  
Aline Correa de Carvalho  
Ana Luiza Trovo Marques de Souza  
André de Araújo Pinto  
Andressa Ribeiro Contreira  
Bárbara de Lima Lucas  
Bianca Barros da Costa  
Carlos Vinícius Pagani Vieira Machado  
Débora Cristina Damasceno  
Deborah Margatho Ramos Gonçalves  
Diogo de Sousa Martins  
Elisângela Rodrigues Carrijo  
Emanuel Tenório Paulino  
Estélio Henrique Martin Dantas  
Eveline Fernandes Vale  
Fabiana Leticia Sbaraini  
Fabio José Antonio da Silva  
Jaqueline Rocha Borges dos Santos  
João Francisco Severo Santos  
José Aderval Aragão  
José Roberto Zaffalon Júnior

Jovino Gentilini Junior  
Juliane Campos Inácio  
Líncon Bordignon Somensi  
Luciane Cristina Arantes  
Marcela Melo dos santos  
Marcello Alberton Herdt  
Marcelo de Oliveira Pinto  
Marcos Guimarães de Souza Cunha  
Maria Patricia Costa Villela  
Nara Michelle Moura Soares  
Paulo Celso Budri Freire  
Pedro Paulo Gattai Gomes  
Raquel Ayres  
Renata Oliveira de Barcelos  
Renato Carlos Machado  
Roberson Geovani Casarin  
Rogério Wagner da Silva  
Sergio Ibañez Nunes  
Sheila Moura Amaral  
Taíza Fernanda Ramalhais  
Vivian Victoria Vivanco Valenzuela

### **Ciências Exatas e da Terra**

Andrea Sartori Jabur  
Antônio Carlos Ribeiro Araújo Júnior  
Cláudia Hitomi Watanabe Rezende  
Dalvani Fernandes  
Evandro Preuss  
Gerson Cruz Araujo  
Gisane Aparecida Michelin  
Henrique Mariano Costa do Amaral  
Henrique Pereira Oliveira d`Eça  
Neves  
Isidro ihadua  
João César Abreu de Oliveira Filho  
Lívia Sancho  
Manolo Cleiton Costa de Freitas  
Marcos Vinicius de Oliveira Peres  
Paulo Celso Budri Freire  
Sonia Tomie Tanimoto

Tânia do Carmo  
Vagner Marques de Moura  
Valdecir Alves dos Santos Júnior

### **Ciências Humanas**

Ana Margarida Theodoro Caminhas  
Ana Maria Senac Figueroa  
Anderson Dantas da Silva Brito  
Breno Henrique Ferreira Cypriano  
Bruno Cezar Silva  
Camila Bueno Grejo  
Camila de Vasconcelos Tabares  
Cássia Maria Bonifácio  
Dalvani Fernandes  
Edonilce da Rocha Barros  
Elisângela Rodrigues Carrijo  
Eulalia Fabiano  
João César Abreu de Oliveira Filho  
João Francisco Severo Santos  
Josael Jario Santos Lima  
Luciano Sérgio Ventin Bomfim  
Marcos Pereira dos Santos  
Marcos Pereira Magalhães  
Miguel Rodrigues Netto  
Rebecca Bianca de Melo Magalhães  
Roberson Geovani Casarin  
Taíza Fernanda Ramalhais  
Tatiane dos Santos Duarte

### **Ciências Sociais Aplicadas**

Ana Margarida Theodoro Caminhas  
Bruno Cezar Silva  
Camila Augusta Alves Pereira  
Camila Nathalia Padula de Godoy  
Dandara Scarlet Sousa Gomes  
Bacelar  
Daniel Nascimento e Silva  
Darline Maria Santos Bulhões  
Denise Tanaka dos Santos  
Elisângela Rodrigues Carrijo

Eulalia Fabiano  
Fabio Adriano Stürmer Kinsel  
Fabricio Lemos de Siqueira Mendes  
Gelson Mario Filho  
Hector Rodrigo Ribeiro Paes Ferraz  
Helga Midori Iwamoto  
Horácio Monteschio  
João Francisco Severo Santos  
Josael Jario Santos Lima  
Leandro Nunes Soares da Silva  
Lucas Rosas de Freitas Sá Barreto  
Miguel Rodrigues Netto  
Nagib Abrahão Duailibe Neto  
Nelson Calsavara Garcia Junior  
Patrícia Loureiro Abreu Alves  
Barbosa  
Renato Obikawa Kyosen  
Sandy Rodrigues Faidherb  
Silvia Helena Ribeiro Cruz  
Solange Kileber  
Stella Villela Florêncio  
Thiago Nery Pandolfo  
Veyzon Campos Muniz  
Ygor de Siqueira Mendes Mendonça

### **Engenharias**

Alejandro Victor Hidalgo Valdivia  
Alex Milton Albergaria Campos  
Ana Carla Fernandes Gasques  
Andrea Sartori Jabur  
Arlete Barbosa dos Reis  
Cristhiane Michiko Passos Okawa  
Daniele Cristina Ficanha  
Diego Matheus Sanches  
Elaine Patricia Arantes  
Fernando Oliveira de Andrade  
Henrique Mariano Costa do Amaral  
Jefferson Sousa Farias  
Laís Roberta Galdino de Oliveira  
Leila Cristina Nunes Ribeiro

Letícia Reis Batista Rosas  
Marcelo Henrique da Silva  
Marcelo Marques  
Marcos Guimarães de Souza Cunha  
Mileni Cristina da Silva  
Renata Jardim Martini  
Thiago Averaldo Bimestre  
Tiago Brandão Costa  
Valdecir Alves dos Santos Júnior

### **Linguística, Letras e Artes**

Adriana dos Reis Silva  
Anderson Dantas da Silva Brito  
Danuzia Marjorye Santos de Araújo  
Jane Catia Pereira Melo  
Luceni Lazara da Costa Ribeiro  
Márcia Donizete Leite-Oliveira  
Maria Christina da Silva Firmino  
Cervera  
Simone Oliveira Vieira Peres  
Thiago Blanch Pires  
Vera Regiane Brescovici Nunes

### **Multidisciplinar**

Alejandro Victor Hidalgo Valdivia  
Ana Carla Fernandes Gasques  
Ana Margarida Theodoro Caminhas  
Andrea Sartori Jabur  
Arlete Barbosa dos Reis  
Cláudia Hitomi Watanabe Rezende  
Cristhiane Michiko Passos Okawa  
Érika Alves Tavares Marques  
Fabricio Lemos de Siqueira Mendes  
Fernando Oliveira de Andrade  
Isidro ihadua  
José Amorim  
Marcelo Marques

# Avanços Científicos em Matemática, Probabilidade e Estatística

**Editora Chefe** Marcia A. A. Marques  
**Coordenadora Editorial** Isabela Arantes Ferreira  
**Bibliotecária** Aline Grazielle Benitez  
**Diagramação** Marcos Antonio Ribeiro Pereira  
**Revisão** Os Autores

## Dados Internacionais de Catalogação na Publicação (CIP) (Câmara Brasileira do Livro, SP, Brasil)

Marques, Marcelo

Avanços científicos em matemática, probabilidade e estatística [livro eletrônico] / Marcelo Marques, Isabela Arantes Ferreira. -- 1. ed. -- São Paulo : Bookerfield, 2022.

PDF.

ISBN 978-65-89929-68-0

1. Estatística 2. Matemática 3. Probabilidades  
I. Ferreira, Isabela Arantes. II. Título.

22-118062

CDD-519.5

### Índices para catálogo sistemático:

1. Estatística : Matemática 519.5

Aline Grazielle Benitez - Bibliotecária - CRB-1/3129

DOI 10.53268/BKF22080100

**Bookerfield Editora**  
São Paulo – Brasil  
Telefone: +55 (11) 98441-4444  
www.bookerfield.com  
contato@bookerfield.com

 **Bookerfield**

## DECLARAÇÃO DOS AUTORES

Os autores declaram não haver qualquer interesse comercial ou irregularidade que comprometa a integridade desta obra; declaram que participaram da elaboração e revisão da obra, atestando a confiabilidade dos dados e resultados; declaram que a obra está livre de plágio acadêmico; declaram que a publicação desta obra não fere qualquer outro contrato por eles firmados; declaram ter atendido eventuais exigências de outras partes, como instituições financiadoras, para a publicação desta obra.



## APRESENTAÇÃO

A matemática pura não possui uma aplicação objetiva a uma determinada área do conhecimento. Desta tarefa se encarrega a matemática aplicada, a qual emprega ideias maduras da matemática visando a solução de problemas que abrangem todas as áreas do conhecimento. Pela presente obra são reunidos avanços científicos em matemática, probabilidade e estatística, aplicados às áreas de meio ambiente, simulação computacional, seleção genômica e ensino de matemática.

O primeiro capítulo trata da técnica de modelagem de altura de ondas de vento recentemente desenvolvida e denominada Paramétrica Bidimensional. A técnica permite a aplicação de uma equação paramétrica de abrangência bidimensional. Possui as vantagens de ser condicionada somente por informações dinâmicas de vento (intensidade e direção) e fetch e por gerar resultados em um tempo bastante reduzido quando comparado aos modelos numéricos de base física. Um fator a ser superado, no entanto, é a variabilidade dos resultados emitidos pelas equações paramétricas atuais. Pelo presente trabalho é feita uma análise crítica a respeito da topologia dessas equações, visando a obtenção de equações paramétricas mais confiáveis de modo a aproveitar as potencialidades da técnica de modelagem Paramétrica Bidimensional em aplicações como sistemas de previsão em tempo real.

Pelo segundo capítulo é apresentado o método rápido, o qual visa determinar de modo objetivo e confiável, o comprimento da maior pista de vento, ou fetch, em um corpo de água continental de grande superfície, como lagos e reservatórios. Se constitui em um método gráfico inovador que deve ser aplicado diretamente sobre um mapa ou imagem de satélite dotado de escala gráfica. Além do caráter fundamental em permitir a determinação rápida de elementos de onda como a amplitude, o período e o comprimento, via equações paramétricas, o método permite classificar o corpo de água de acordo com a magnitude das maiores alturas de ondas geradas pelo vento. A demonstração foi realizada pela aplicação ao reservatório da Usina Hidrelétrica de Itaipu.

Pelo terceiro capítulo é realizada importante contribuição para a solução de um problema relevante e complexo na área de Geometria Computacional: determinar se há inclusão espacial entre dois sólidos com superfícies digitalizadas. Os experimentos computacionais foram realizados utilizando digitalizações de pedras brutas disponibilizadas pelo Centro Tecnológico de Pedras, Gemas e Joias do Rio Grande do Sul (CT-Pedras). Os métodos testados foram validados e se mostraram adequados para tratar o problema investigado.

Pelo quarto capítulo é aplicada a modelagem estocástica de campos

Markovianos de alcance variável para encontrar blocos independentes de SNPs, estruturando assim o genoma em regiões ilhadas de dependência. Marcadores moleculares do tipo SNP (*Single Nucleotide Polimorphism*), com alta densidade dentro do DNA humano, têm sido largamente utilizados como amostradores do genoma. Os resultados obtidos mostraram que as janelas de influência da população brasileira têm, em média, tamanho maior, evidenciando a sua história recente de miscigenação, comparativamente às demais populações mundiais. É também proposta a construção de janelas de consenso maximais em determinadas regiões genômicas, tendo como base as janelas de influência de cada população do estudo, permitindo a construção de um classificador capaz de determinar o percentual relativo de ancestralidade de um indivíduo, o qual, submetido à validação, obteve alta eficiência no acerto da população originária.

Pelo quinto capítulo são analisadas dificuldades típicas existentes no ensino-aprendizagem de Análise Combinatória para o sexto ano do ensino fundamental. Constatou-se que a aprendizagem de Análise Combinatória se mostra como um obstáculo devido à forma como o conteúdo é abordado, na quase totalidade das vezes, através de fórmulas matemáticas, resultando na perda do sentido na resolução do problema. Uma proposta de metodologia de ensino é apresentada, sendo verificada a melhora no aprendizado do conteúdo ministrado. melhorando desta maneira, o ensino e a compreensão de matemática.

Boa leitura!

# SUMÁRIO

## CAPÍTULO 1

MODELAGEM MATEMÁTICA APLICADA À TOPOLOGIA DE EQUAÇÃO ONDAS  
BASEADA NA EQUAÇÃO DO BALANÇO DE ENERGIA ..... 12

Marcelo Marques

Fernando O. de Andrade

Elaine P. Arantes

Alexandre K. Guetter

DOI: 10.53268/BKF22080101

## CAPÍTULO 2

MATEMÁTICA APLICADA À CONCEPÇÃO DE UM MÉTODO RÁPIDO PARA  
DETERMINAÇÃO DAS PISTAS DE VENTO ..... 20

Marcelo Marques

Elaine P. Arantes

Fernando O. de Andrade

Isabela A. Ferreira

DOI: 10.53268/BKF22080102

## CAPÍTULO 3

MÉTODOS COMPUTACIONAIS PARA O PROBLEMA DE IN-CLUSÃO ENTRE  
NUVEM DE PONTOS E DIGITALIZAÇÃO TRIDIMENSIONAL DE SÓLIDOS ..... 30

Vinícius Fernandes Moretti

DOI: 10.53268/BKF22080103

## CAPÍTULO 4

CARACTERIZAÇÃO DA ESTRUTURA DE DEPENDÊNCIA DO GENOMA  
HUMANO USANDO CAMPOS MARKOVIANOS: ESTUDO DE POPULAÇÕES  
MUNDIAIS E DADOS DE SNPs ..... 45

Francisco José de Almeida Fernandes

Júlia Maria Pavan Soler

DOI: 10.53268/BKF22080104

## CAPÍTULO 5

ANÁLISE COMBINATÓRIA: UMA ABORDAGEM PARA O SEXTO ANO DO  
ENSINO FUNDAMENTAL ..... 60

Juvino Pereira Dos Santos Junior

DOI: 10.53268/BKF22080105

ÍNDICE REMISSIVO ..... 74

SOBRE OS ORGANIZADORES ..... 76

## MODELAGEM MATEMÁTICA APLICADA À TOPOLOGIA DE EQUAÇÃO ONDAS BASEADA NA EQUAÇÃO DO BALANÇO DE ENERGIA

**Marcelo Marques**

Universidade Estadual de Maringá  
[lattes.cnpq.br/0855863872309532](https://lattes.cnpq.br/0855863872309532)

**Fernando O. de Andrade**

Universidade Tecnológica Federal  
do Paraná  
[lattes.cnpq.br/5175699449525205](https://lattes.cnpq.br/5175699449525205)

**Elaine P. Arantes**

Universidade Estadual de Maringá  
[lattes.cnpq.br/5673783965584352](https://lattes.cnpq.br/5673783965584352)

**Alexandre K. Guetter**

Universidade Federal do Paraná  
[lattes.cnpq.br/4438410083931646](https://lattes.cnpq.br/4438410083931646)

**RESUMO** - A técnica de modelagem recentemente desenvolvida denominada Paramétrica Bidimensional permite a aplicação de uma equação paramétrica no espaço bidimensional, com as vantagens de ser condicionada somente por informações dinâmicas de vento (intensidade e direção) e *fetch* e por gerar resultados em um tempo bastante reduzido quando comparado aos modelos numéricos de base física. Um fator a ser superado,

no entanto, é a variabilidade dos resultados emitidos pelas equações paramétricas. Pelo presente trabalho é feita uma análise crítica a respeito da topologia dessas equações, visando a obtenção de equações paramétricas mais confiáveis de modo a aproveitar as potencialidades da técnica de modelagem Paramétrica Bidimensional em aplicações como sistemas de previsão em tempo real.

**PALAVRAS-CHAVE:** Fetch; Águas continentais; Vento.

### 1. INTRODUÇÃO

Com a disseminação no uso de computadores, houve grande desenvolvimento na modelagem dos processos físicos das ondas geradas pelo vento, em detrimento do aperfeiçoamento dos modelos paramétricos. O desinteresse pela utilização de equações paramétricas é devido a ao indiscutível aumento de precisão e a possibilidade de geração de resultados bidimensionais pelos modelos numéricos de base física.

Dentre as equações paramétricas amplamente difundidas no meio técnico visando a simulação de altura de ondas estão as equações de: (1) Saville, proposta por Saville et al. (1962); (2) JONSWAP, conforme

Hasselmann et al. (1973); (3) SMB com coeficientes sugeridos por Sibul (1955):

$$\frac{g \cdot H_s}{U_{10}^2} = 0,0026 \cdot \left( \frac{g \cdot F}{U_{10}^2} \right)^{0,47} \quad [1]$$

$$\frac{gH_s}{U_{10}^2} = 0,0016 \left( \frac{gF}{U_{10}^2} \right)^{0,5} \quad [2]$$

$$\frac{gH_s}{U_{10}^2} = 0,0032 \left( \frac{gF}{U_{10}^2} \right)^{0,435} \quad [3]$$

Nas equações os termos  $gFU^{-1}$  e  $gHU^{-1}$  representam respectivamente os termos adimensionais de fetch ( $F^*$ ) e altura de onda ( $H^*$ ).

Comparando as equações [1] a [3] verifica-se que possuem uma estrutura comum, a qual pode ser generalizada pela equação seguinte:

$$H^* = \alpha (F^*)^\beta \quad [4]$$

Na qual os coeficientes  $\alpha$  e  $\beta$  são constantes.

Pelo presente trabalho é feita uma análise crítica a respeito da topologia das equações paramétricas, amplamente utilizadas no meio técnico. A análise visa a obtenção de equações paramétricas mais confiáveis de modo a aproveitar as vantagens da técnica de modelagem paramétrica, quais sejam, a dependência de um número reduzido de variáveis (vento e *fetch*), o reduzido tempo de processamento para os casos de utilização de recursos computacionais e a possibilidade de obtenção da altura de onda por meio de processos gráficos, seja utilizando programas do tipo CAD, seja utilizando instrumentos elementares de desenho sobre mapa impresso.

## 2. METODOLOGIA

A análise a respeito da topologia das equações paramétricas é feita partindo-se da equação do balanço de energia, utilizada pelos modelos numéricos de base física, e correlacionando-a aos termos adimensionais de energia ( $E^*$ ), altura de onda ( $H^*$ ) e fetch ( $F^*$ ). Esse desenvolvimento visa determinar a topologia de uma equação paramétrica que condicione a altura da onda somente ao *fetch* e à intensidade do vento.

## 3. DISCUSSÃO

Os modelos numéricos obtêm o campo de ondas pela propagação do espectro no espaço e no tempo. De forma geral esses modelos estão baseados na solução de uma equação de balanço de energia do seguinte

tipo:

$$\frac{\partial E}{\partial t} + \nabla \cdot (c_g E) = S = S_{in} + S_{nl} + S_{ds} \quad [5]$$

Um campo de ondas vai se desenvolvendo no espaço e no decorrer do tempo. Já o *fetch*, por depender da geometria do corpo de água, possui somente variação espacial.

Existe uma condição assumida pela quase totalidade dos métodos paramétricos de estimativa de altura de ondas, na qual o campo de ondas passa a possuir somente variação espacial. Nesta condição de campo de onda condicionado pelo *fetch*, pela invariabilidade do espectro no tempo, o termo  $\partial E/\partial t$  é nulo na equação do balanço de energia, ou seja:

$$\frac{\partial E}{\partial t} = 0 \quad [6]$$

O termo advectivo ( $\nabla \cdot c_g E$ ) representa a propagação da energia da onda no espaço. A energia ( $E$ ) e o termo adimensional de energia ( $E^*$ ), dependentes da altura da onda ( $H$ ), da massa específica da água ( $\rho$ ) e da aceleração gravitacional ( $g$ ), são definidos pelas seguintes equações:

$$E = \frac{1}{8} \rho g H^2 \quad [7]$$

$$E^* = \frac{g^2 E}{U^4} \quad [8]$$

Substituindo a equação [3] na equação [4], obtém-se uma relação entre os adimensionais de energia ( $E^*$ ) e altura de onda ( $H^*$ ), de acordo com a equação seguinte:

$$E^* = \frac{\rho g^3 H^2}{8U^4} = \frac{1}{8} \rho g (H^*)^2 \quad [9]$$

Desde os estudos de Kitaigorodskii (1962), é condição comprovada que, para campo de ondas limitado pelo *fetch*, existe uma correlação entre os adimensionais de energia da onda por unidade de área ( $g^2 E/U^4$ ) e o *fetch* adimensional ( $gF/U^2$ ) (Jinhua & Yongming, 2012; Jordi & Wang., 2012; Aiki & Greatbatch, 2012). A formulação de Kahma & Calkoen (1992) é definida como:

$$\frac{g^2 E}{U^4} = 5,4 \times 10^{-7} \left( \frac{gF}{U^2} \right)^{0,94} \quad [10]$$

Portanto, pode-se generalizar afirmando que os adimensionais de energia ( $E^*$ ) e *fetch* ( $F^*$ ) sejam associados de forma não-linear segundo a equação seguinte:

$$E^* = \alpha_1 (F^*)^{\beta_1} \tag{11}$$

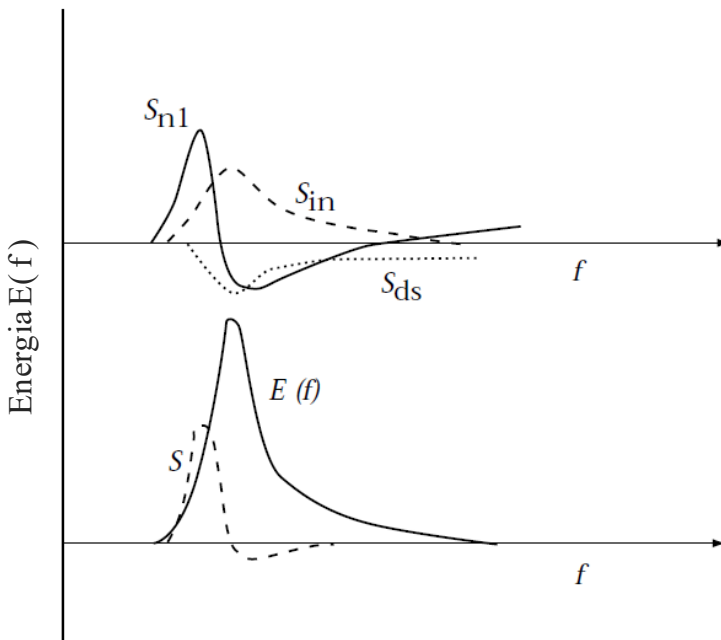
Sendo  $\alpha_1$  e  $\beta_1$  coeficientes constantes.

Como o termo advectivo da equação do balanço de energia representa a variação do espectro de energia no espaço, assume-se que esse termo pode ser representado pela derivada da energia em relação ao *fetch*. Derivando a equação [11], obtém-se:

$$\frac{\partial E^*}{\partial F^*} = \alpha_1 \beta_1 (F^*)^{\beta_1 - 1} \tag{12}$$

A respeito do termo fonte total ( $S = S_{in} + S_{nl} + S_{ds}$ ), esse é composto pelo aporte de energia pelo vento ( $S_{in}$ ), pela transferência de energia pelas interações onda-onda ( $S_{nl}$ ) e pela dissipação ( $S_{ds}$ ). A relação da energia com a frequência é mostrada na pela figura seguinte:

**Figura 1:** Espectro de energia dos termos fonte



Fonte: adaptado de Jinhua & Shen, 2012

O diagrama na parte superior da Figura 1 revela o comportamento típico de cada termo para diferentes frequências. O diagrama da parte inferior revela que o termo fonte resultante constitui-se em uma fração da energia total, permitindo admitir, em termos adimensionais, a relação seguinte:

$$S = \alpha_2 (E^*)^{\beta_2} \quad [13]$$

Onde  $\alpha_2$  e  $\beta_2$  são coeficientes constantes.

A topologia não linear adotada entre o termo-fonte resultante e a energia deve-se à verificação, pela parte inferior da Figura 1, da proporção variável entre esses termos para diferentes frequências.

### Equação paramétrica

Substituindo as equações [7], [12] e [13] na Equação [5], chega-se à equação seguinte:

$$\alpha_1 \beta_1 (F^*)^{\beta_1 - 1} = \alpha_2 (E^*)^{\beta_2} \quad [14]$$

Como o interesse é o de exprimir uma equação paramétrica em termos da altura da onda e não propriamente da energia total, substituindo a relação dada pela Equação [11] na Equação [14] e desenvolvendo-a, chega-se a:

$$H^* = \left( \frac{8}{\rho g} \left( \frac{\alpha_1 \beta_1}{\alpha_2} \right)^{\frac{1}{\beta_2}} \right)^{\frac{1}{2}} (F^*)^{\frac{\beta_1 - 1}{2\beta_2}} \quad [15]$$

A equação anterior pode ser expressa como:

$$H^* = \alpha_3 (F^*)^{\beta_3} \quad [16]$$

Trata-se de uma equação do tipo potência com coeficientes,  $\alpha_3$  e  $\beta_3$ , variáveis, o que contraria a topologia das equações paramétricas utilizadas até os dias de hoje, dotadas de coeficientes constantes, apresentada pela equação [4].

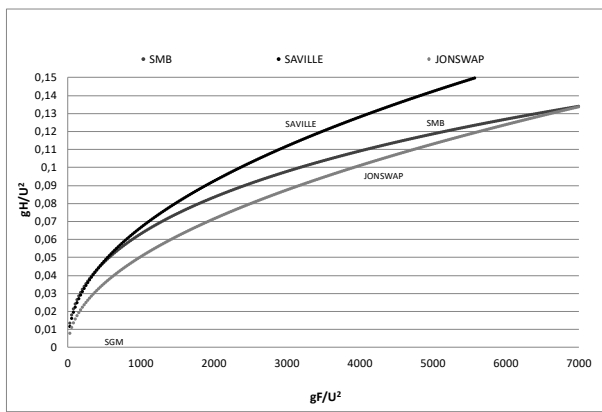
Uma equação paramétrica de coeficientes constantes gera uma curva em uma posição fixa em um diagrama adimensional de  $F^* \times H^*$ . Já uma equação paramétrica de coeficientes variáveis assume diferentes posições, dependendo do valor dos coeficientes. Isso explicaria a grande variabilidade que assumem as diferentes equações paramétricas em um diagrama



adimensional.

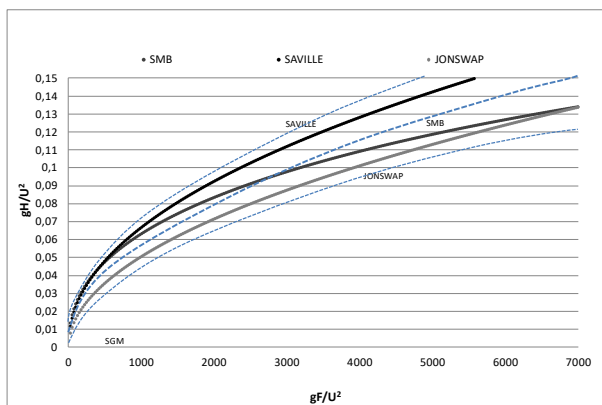
Para ilustrar essa variabilidade, observe o diagrama adimensional gerado pela Figura 2 no qual foram traçadas as curvas relacionadas às equações [1], [2] e [3]. Como as equações SMB, Saville e JONSWAP possuem coeficientes constantes, sua posição em um diagrama adimensional é invariável.

**Figura 2:** Posição das curvas que representam as equações paramétricas de coeficientes constantes no diagrama adimensional



Supondo como exemplo que a equação paramétrica possui coeficientes variáveis e dependentes, por exemplo, da intensidade do vento, uma mesma equação passaria a assumir outras posições, conforme ilustrado pela Figura 3.

**Figura 3:** Variações de posição hipotéticas que poderiam ocupar uma equação paramétrica de coeficientes variáveis



Portanto, coeficientes variáveis poderiam explicar a diversidade de equações paramétricas disponíveis na literatura, apontando para atividades de definição desses coeficientes com vistas a tornar os modelos paramétricos confiáveis.

## CONCLUSÕES

Pela abordagem adotada pelo presente trabalho foi confirmada a topologia do tipo potência como adequada para equações do tipo paramétrica visando a determinação da altura significativa da onda. Trata-se da mesma topologia adotada pela maioria das equações paramétricas apresentadas até os dias de hoje com uma sutil e relevante diferença: nestas equações os coeficientes  $\alpha$  e  $\beta$  têm sido considerados constantes. Desse modo, uma importante constatação é a possibilidade de uma equação do tipo potência composta por coeficientes variáveis poder constituir-se em uma solução promissora no aumento da precisão na utilização de equações paramétricas. Essa solução contribui para explicar a diversidade de coeficientes e conseqüentemente, a divergência de resultados de altura de onda adotados nas equações paramétricas de coeficientes constantes disponíveis na literatura.

## REFERÊNCIAS BIBLIOGRÁFICAS

**Aiki, H.; Greatbatch, R. J.** (2012) “Thickness-Weighted Mean Theory for the Effect of Surface Gravity Waves on Mean Flows in the Upper Ocean”. *Journal of Physical Oceanography* 42:5, 725-747. Online publication date: 1-May-2012.

**Hasselmann, K.; Barnett, T. P. ; Bouws, E. ; Carlson, H. ; Cartwright, D. E. ; Enke, K. ; Ewing, J. A. ; Gienapp, H. ; Hasselmann, D. E. ; Kruseman, P. ; Meerburg, A. ; Muller, P. ; Olbers, D. J. ; Richter, K. ; Sell, W. & Walden, H.** (1973). “Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP)”. *Deut. Hydrogr. Z.*, A8(12).

**Jinhua, W., Yongming, S.,** (2012) “On the development and verification of a parametric parallel unstructured-grid finite-volume wind wave model for coupling with ocean circulation models”. *Environmental Modelling & Software* 37, 179-192. Online publication date: 1-Nov-2012.

**Jordi, A.; Wang, D. P.,** (2012) “sbPOM: A parallel implementation of Princeton Ocean Model”. *Environmental Modelling & Software* 38, 59-61. Online publication date: 1-Dec-2012.

**Kitaigorodskii, S. A.,** (1962). “Applications of the theory of similarity to the analysis of wind-generated wave motion as a stochastic process”. *Izv. Geophys. Ser. Acad. Sci., USSR*, 1, 105–117.

**Saville, T. ; McClendon, E. W. & Cochran, A. L.** (1962). "Freeboard allowances for waves in inland reservoirs". *ASCE Journal of the Waterways and Harbors Division*, V. 88(WW2): 93-124.

**Sibul, O.** (1955). "Laboratory study of the generation of wind waves in shallow water". *U.S. Army Corps of Engineers Beach Erosion Board*, Technical Memo No. 72, 35 p.

### **MATHEMATICAL MODELING APPLIED TO THE WAVES EQUATION TOPOLOGY BASED ON THE ENERGY BALANCE EQUATION**

**ABSTRACT** - The newly developed modeling technique known as Two-dimensional Parametric allows the application of a parametric equation in two-dimensional space, with the advantages of being conditioned only by dynamic information of wind (direction and intensity) and fetch and generate results in a time greatly reduced when compared to numerical models of physical basis. A factor to be overcome, however, is the variability of the results issued by parametric equations. At present work is made a critical analysis regarding the topology of these equations, in order to obtain parametric equations more reliable and take advantage of the potential of two dimensional parametric modeling technique in applications such as real-time forecast systems.

**KEYWORDS:** Fetch; Continental waters; Wind.

# MATEMÁTICA APLICADA À CONCEPÇÃO DE UM MÉTODO RÁPIDO PARA DETERMINAÇÃO DAS PISTAS DE VENTO

**Marcelo Marques**

Universidade Estadual de Maringá  
[lattes.cnpq.br/0855863872309532](https://lattes.cnpq.br/0855863872309532)

**Elaine P. Arantes**

Universidade Estadual de Maringá  
[lattes.cnpq.br/5673783965584352](https://lattes.cnpq.br/5673783965584352)

**Fernando O. de Andrade**

Universidade Tecnológica Federal  
do Paraná  
[lattes.cnpq.br/5175699449525205](https://lattes.cnpq.br/5175699449525205)

**Isabela A. Ferreira**

Universidade Estadual de Maringá  
[lattes.cnpq.br/7252508396222629](https://lattes.cnpq.br/7252508396222629)

Além do caráter fundamental em permitir a determinação rápida de elementos de onda como a amplitude, o período e o comprimento, via equações paramétricas, um método rápido de determinação do fetch máximo permitiria classificar o corpo de água de acordo com a magnitude das maiores alturas de ondas geradas pelo vento. A aplicação do método independe de características geométricas dos contornos do corpo hídrico, como dendriticidade, sinuosidade e escala espacial. Pelo presente estudo o método será apresentado e validado pela aplicação ao reservatório de Itaipu.

**PALAVRAS-CHAVE:** Método Rápido; Fetch; Onda.

**RESUMO** - A finalidade do método Rápido é determinar de modo objetivo e confiável o comprimento da maior pista de vento, ou fetch, em um corpo de água continental de grande superfície, como lagos e reservatórios. Constitui-se em um método gráfico inovador que deve ser aplicado diretamente sobre um mapa ou imagem de satélite dotado de escala gráfica. O método pode ser aplicado via traçado manual ou ainda, utilizando interface computacional.

## 1. INTRODUÇÃO

A superfície do corpo hídrico em contato com o vento e que contribui na geração da onda é denominada fetch ou pista de vento.

O maior fetch em um corpo de água representa, mesmo que indiretamente, a magnitude dos eventos gerados pela ação do vento. Deste modo o fetch pode constituir-se em um importante elemento de caracterização do corpo hídrico quanto à magnitude das perturbações

geradas pelo vento.

A obtenção segura do maior valor de fetch para cada direção para qualquer corpo de água interior pode ser feita pela aplicação do conceito de campo de fetch através do modelo ONDACAD (Marques *et al.* , 2013b).

No entanto, caso o interesse se concentre apenas no maior comprimento absoluto do fetch, o processo de modelagem, apesar de preciso, pode ser bastante laborioso.

O método Rápido foi elaborado visando determinar de forma rápida e segura o maior fetch em qualquer corpo hídrico continental.

## 2. FETCH EM ÁGUAS CONTINENTAIS

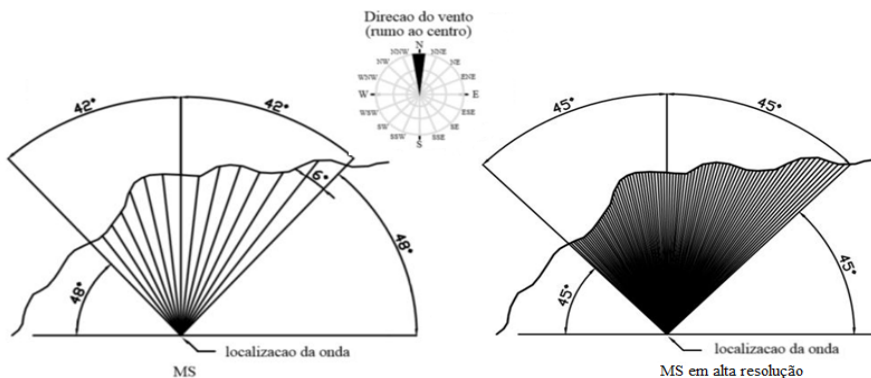
Visando ilustrar as discussões que envolvem o fetch no espaço bidimensional, decidiu-se pela utilização de mapas de distribuição do fetch obtidos pela aplicação do modelo ONDACAD desenvolvido por Marques (2013) inspirado no método de Saville (1954).

### 2.1 Método de Saville

O método apresentado por Saville (1954) e recomendado por U.S. Army Coastal Engineering Reserch Center (1966) consiste no traçado de linhas radiais com resolução angular de 6° e amplitude máxima de 42°, conforme a parte esquerda da Figura 1. De acordo com o método, o fetch (F) em cada ponto e para determinada direção é definido com base no ângulo  $\alpha_i$  entre direção do vento e a direção secundária e o comprimento na direção secundário  $x_i$ , conforme a equação 1.

$$F = \sum_i x_i \cos \alpha_i / \sum_i \cos \alpha_i \quad [1]$$

**Figura 1:** Aplicação esquemática do método de Saville (esquerda) e do método de Saville em alta resolução (direita)



Devido a aperfeiçoamentos, Marques (2013) modificou o método de Saville e passou a considerar uma resolução angular de 1° entre linhas auxiliares e amplitude angular máxima de 45° conforme a parte direita da Figura 1. Devido a este aperfeiçoamento, o fetch passou a ser determinado pela equação 2, a qual relaciona os comprimentos  $x_i$  de cada linha auxiliar ao cosseno do ângulo  $\phi$  que esta forma com a direção principal.

$$F = \frac{\int_{-45}^{45} x_i \cos \phi d\phi}{\int_{-45}^{45} \cos \phi d\phi} \quad [2]$$

Este aumento considerável da resolução angular e da quantidade de segmentos de reta no traçado, foi possível pela automação do processo via processamento computacional, permitindo obter não somente o comprimento do fetch em um ponto, mas a distribuição do fetch em um corpo hídrico qualquer. Esse novo modelo computacional se denomina ONDACAD.

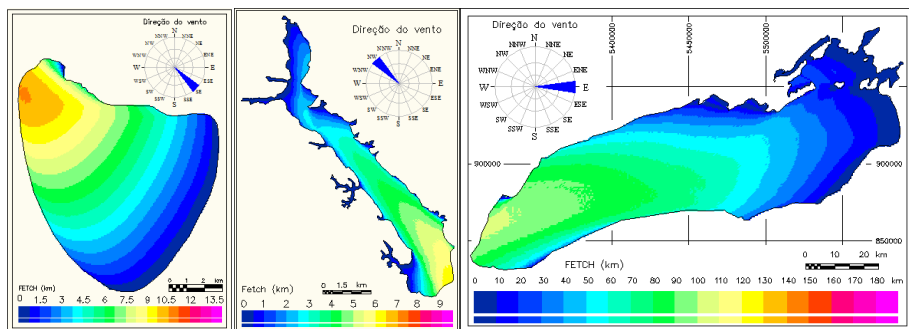
São os mapas de distribuição do fetch que fundamentam as análises que resultam na proposição do método rápido.

### 2.2 Formato do contorno do corpo hídrico

O cálculo do *fetch* como proposto por Saville, por se tratar de um processo essencialmente geométrico, raramente é influenciado por toda a extensão do reservatório.

Para locais de formatos mais ou menos regulares, há grande chance das linhas auxiliares alcançarem todo o reservatório, como mostrado pelos corpos de água da Figura 2.

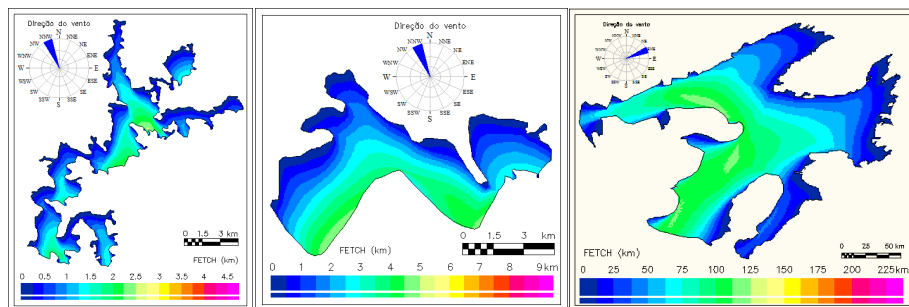
**Figura 2:** Exemplos em que o *fetch* máximo alcança toda a superfície do corpo hídrico: Lagoa Barros (esquerda), Reservatório de Jesuína (centro) e Lago Ontário (direita)



Fonte: Marques (2013)

Já para casos de formato dendrítico ou sinuoso, bem mais comuns, a determinação do *fetch* geralmente restringe-se a subáreas, invalidando métodos que considerem a totalidade da superfície do reservatório. Exemplos desse tipo são mostrados pela Figura 3.

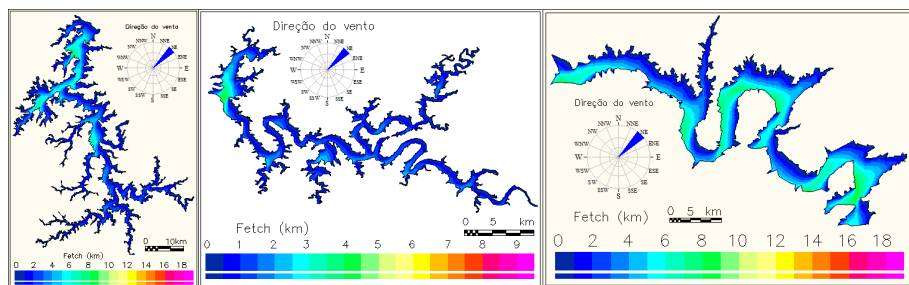
**Figura 3** - Exemplo de ocorrência do *fetch* em sub-regiões nos reservatórios de Ernestina (esquerda), Palmital (centro) e no Lago Great Bear (direita)



Fonte: Marques (2013)

A irregularidade da forma pode ser agravada pela presença de sinuosidade, como ilustra a Figura 4.

**Figura 4:** Sinuosidade nos reservatórios de Três Marias, Salto Santiago e Itaparica



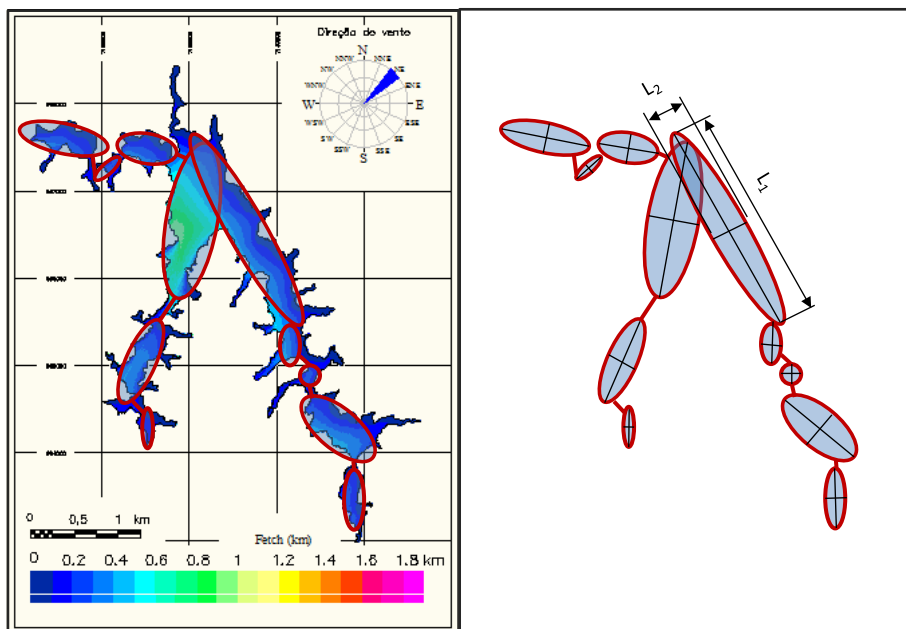
Fonte: Marques (2013)

Deste modo, uma solução geral para a determinação do *fetch* parece apontar para um método que adote a delimitação da área que contribui efetivamente para a obtenção do *fetch* máximo.

### 3. MATERIAIS E MÉTODOS

Para ilustrar a conclusão de que o maior fetch parece ocorrer de modo compartimentado no corpo hídrico, ou seja, o maior comprimento do *fetch* parece estar relacionado à maior sub-região do corpo hídrico, foi elaborado o diagrama esquemático ilustrado pela Figura 5.

**Figura 5:** Diagrama conceitual no reservatório de Vossoroca, demonstrando que o maior fetch é condicionado pelo tamanho de sub-regiões que constituem o reservatório



Através da Figura anterior é constatado que cada sub-região pode ser identificada pelos dois maiores eixos que sejam perpendiculares entre si, convencionados aqui como sendo os eixos  $L_1$  e  $L_2$ , respectivamente. Então o problema consiste em se determinar a maior superfície livre do reservatório, a qual deve corresponder a uma subárea, e relacioná-la à maior extensão do *fetch*. Considerando essa fundamentação, desenvolveu-se um método com o objetivo de determinar de modo rápido o maior *fetch* efetivo em um corpo de água qualquer conforme este procedimento:

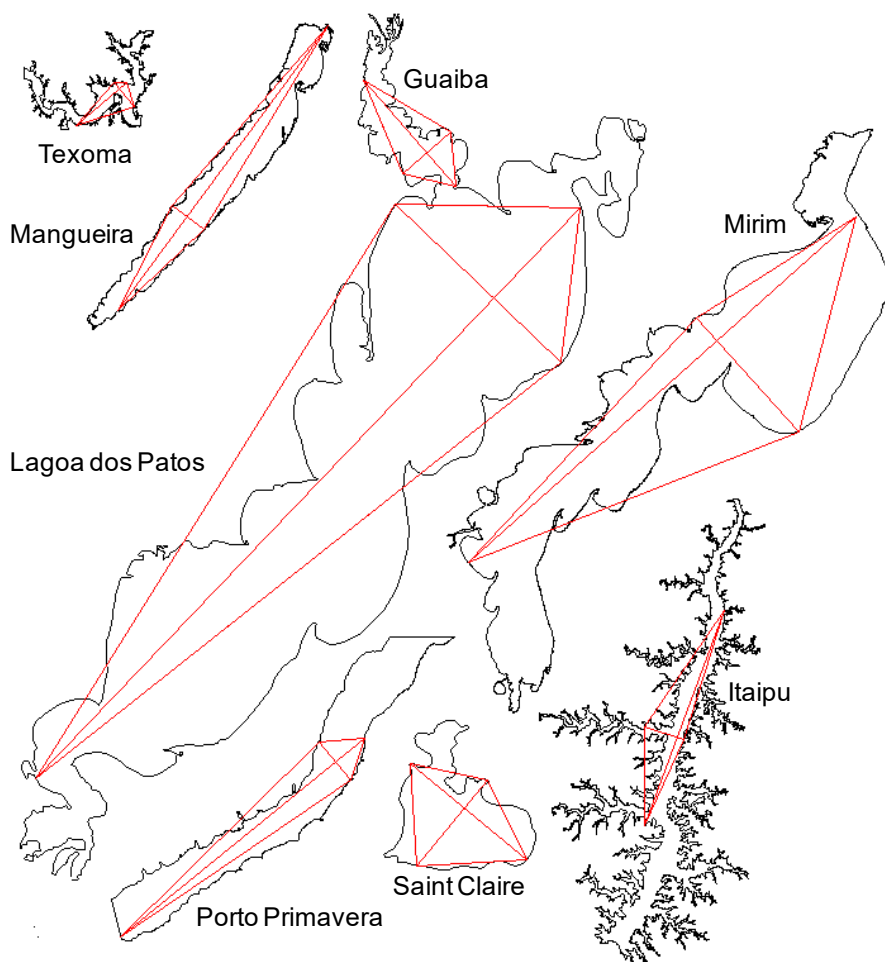
- i) traça-se a maior reta livre de comprimento  $L_1$ ;
- ii) traça-se um segundo segmento de reta, tão extensa quanto possível, perpendicular à primeira, de comprimento  $L_2$ ;



iii) a composição das linhas dá origem a um trapézio com área  $L_1L_2/2$ .

A obtenção dos comprimentos  $L_1$  e  $L_2$  é ilustrada pela Figura 6, para reservatórios com diferentes formas e tamanhos, porém, apresentados na mesma escala gráfica.

**Figura 6:** Traçado dos comprimentos  $L_1$  e  $L_2$  em diferentes corpos hídricos



A determinação da equação que correlaciona o maior *fetch* pelo método de Saville em alta resolução, ao produto dos dois comprimentos ( $L_1$  e  $L_2$ ) foi realizada por Marques (2013) precedida pela geração de 16 mapas de distribuição de *fetch* para cada um dos 68 corpos de água modelados, entre lagos e reservatórios, totalizando 1088 mapas.

Por análise de regressão, foi obtida a seguinte equação que relaciona os comprimentos  $L_1$  e  $L_2$  ao *fetch* máximo:

$$F_{\text{máx}} = \left[ \frac{L_1 L_2}{2} \right]^{\frac{1}{2}} \quad [3]$$

Portanto, pelo método aqui apresentado, a maior extensão do *fetch* efetivo para um reservatório qualquer pode ser obtida por meio de uma operação matemática elementar, com base na extração da raiz quadrada da área do trapézio gerado pelos comprimentos ( $L_1$  e  $L_2$ ) perpendiculares entre si. Observa-se que nem a direção, nem a localização do maior *fetch* são de interesse do método. Além disso, a medida de qualidade do modelo em relação à capacidade de estimar o *fetch* máximo foi feita obtendo-se um coeficiente de determinação de 98,46%.

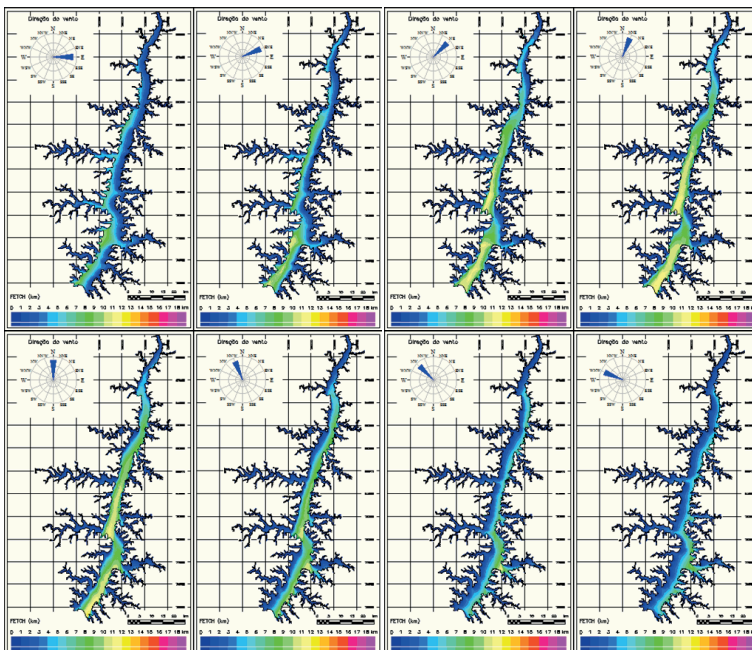
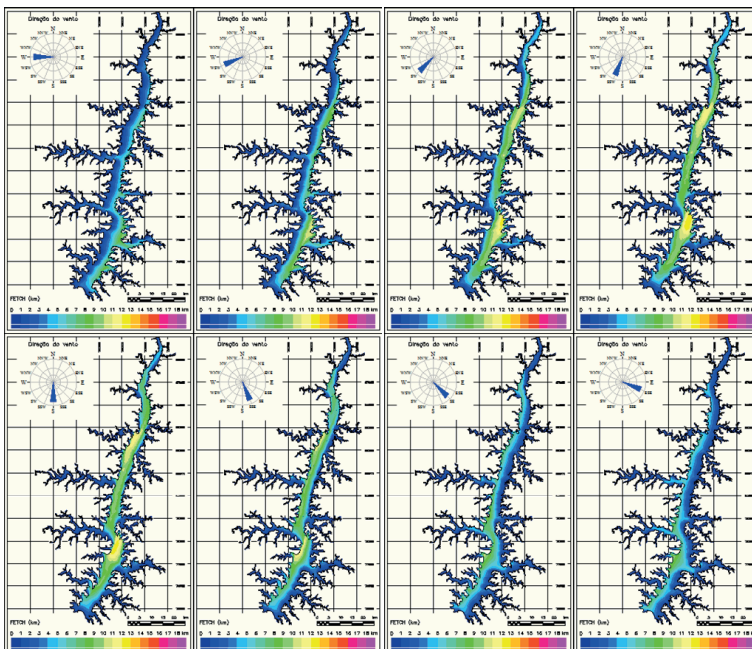
### 3.1 Modelo JONSWAP

A altura significativa da onda ( $H_s$ ), dependente da intensidade do vento ( $U$ ), da aceleração gravitacional ( $g$ ) e do comprimento do *fetch* ( $F$ ), será determinada pela equação paramétrica denominada modelo JONSWAP:

$$\frac{gH_s}{U^2} = 0,0016 \left( \frac{gF}{U^2} \right)^{\frac{1}{2}} \quad [4]$$

## 4. RESULTADOS E DISCUSSÃO

Foram gerados 16 mapas de distribuição do *fetch* para o reservatório de Itaipu, conforme Figuras 7 e 8.

**Figura 7:** Mapas de campos de fetch para as direções E a WSW

**Figura 8:** Mapas de campos de fetch para as direções W a ESE


Com base nos 16 mapas de campo de fetch gerados, conforme Figuras 6 e 7, foi obtido um fetch máximo de 12.913,0m.

Aplicando-se o método Rápido ao reservatório de Itaipu chegou-se aos comprimentos  $L_1$  e  $L_2$  de 58.950,0m e 9.115,0m. Aplicando a equação 3 chegou-se a um fetch máximo de 16.384,4 m.

Portanto o desvio devido à aplicação do método rápido foi de 21,6%. O desvio é considerado pequeno quando levado em conta a simplicidade de aplicação do método rápido em comparação à geração dos 16 mapas pelo modelo ONDACAD.

Uma importante aplicação da determinação do fetch é permitir quantificar a altura da onda gerada pelo vento em determinado local do reservatório. Desse modo, pela aplicação do método JONSWAP e adotando-se ventos de 5, 10, 15 e 20ms<sup>-1</sup> são obtidas, respectivamente, as alturas de onda máxima, em centímetros, apresentadas pela Tabela 1.

**Tabela 1:** Altura de ondas em metros

Método	Intensidade do vento (m/s)				Desvio (%)
	5	10	15	20	
Rápido	0,33	0,65	0,98	1,30	11%
JONSWAP	0,29	0,58	0,87	1,16	

Portanto, de acordo com a tabela anterior, o erro devido à determinação das alturas de ondas máximas por diferentes métodos, é de 11% e independe da intensidade do vento.

Desse modo, para o reservatório de Itaipu, a aplicação do método rápido implicaria em um erro da ordem de 20% na determinação do fetch máximo e um erro da ordem de 10% para na determinação da altura máxima da onda pelo modelo JONSWAP.

## 5. CONSIDERAÇÕES FINAIS

Nos casos em que é necessário determinar a direção do vento em que ocorre o maior fetch continua sendo imprescindível a utilização de simulação computacional.

Ao se determinar as alturas de ondas máximas no reservatório de Itaipu para ventos de diferentes intensidades, verificou-se que o desvio da ordem de 10% não é condicionado pela intensidade do vento.

Os desvios da ordem de 20% para o fetch máximo e da ordem de 10% para a altura máxima de onda foram considerados bastante satisfatórios tendo em vista a facilidade e rapidez de aplicação do método rápido.

## REFERÊNCIAS

Hasselmann, K.; Barnett, T. P.; Bouws, E.; Carlson, H. ; Cartwright, D. E. ; Enke, K. ; Ewing, J. A. ; Gienapp, H. ; Hasselmann, D. E. ; Kruseman, P. ; Meerburg, A. ; Muller, P. ; Olbers, D. J. ; Richter, K. ; Sell, W. & Walden, H. (1973). **Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP)**. Deut. Hydrogr. Z., A8(12).

Marques, M. ; Andrade, F. O. ; Arantes, E. P.; Guetter, A. K. (2013a). **Método rápido para a determinação do fetch máximo**. In: XX Congresso Brasileiro de Recursos Hídricos. Associação Brasileira de Recursos Hídricos. Bento Gonçalves, Rio Grande do Sul. Nov/2013.

Marques, M. ; Andrade, F. O. ; Guetter, A. K. (2013b). **Conceito do Campo de fetch e sua Aplicação ao reservatório de Itaipu**. Revista Brasileira de Recursos Hídricos, v. 18, p. 243-253, 2013.

Saville, T. (1954). **The effect of fetch width on wave generation**. Technical Memorandum No. 70, U.S. Army, Corps of Engineers, Beach Erosion Board, 9 pp.

U. S. Army Corps of Engineers (1966). **Coastal Engineering Manual (CEM)**, *Engineer Manual* 1110-2-1100, U.S. Army Corps of Engineers, Washington, D.C. (6 vol.).

### MATHEMATICS APPLIED TO THE DESIGN OF A QUICK METHOD FOR DETERMINING WIND TRACKS IN LAKES AND RESERVOIRS

**ABSTRACT** - The purpose of the Quick method is to determine objectively and reliably the greatest fetch length in a continental body of water. It constitutes a graphical method to be applied directly on a map or satellite image. This study is verified by the validation of the method by applying the Itaipu Reservoir.

**KEYWORDS:** Quick method; Fetch; Wind wave.

# MÉTODOS COMPUTACIONAIS PARA O PROBLEMA DE INCLUSÃO ENTRE NUVEM DE PONTOS E DIGITALIZAÇÃO TRIDIMENSIONAL DE SÓLIDOS

**Vinícius Fernandes Moretti**

UFRGS, Porto Alegre, RS

[lattes.cnpq.br/4641578411604753](https://lattes.cnpq.br/4641578411604753)

**RESUMO** - No presente trabalho são investigadas soluções eficientes para o problema de determinar se uma nuvem de pontos está contida (ou, alternativamente, invade) a digitalização tridimensional da superfície de um sólido não necessariamente convexo. Estratégias baseadas em *ray casting* (método do raio) e no Teorema da Curva de Jordan, generalizadas para o caso tridimensional, bem como outras fundamentadas no cômputo de volumes com sinal de tetraedros pela origem, foram testadas e comparadas segundo sua eficácia e eficiência computacional. Os experimentos computacionais foram realizados utilizando digitalizações de pedras brutas disponibilizadas pelo Centro Tecnológico de Pedras, Gemas e Joias do Rio Grande do Sul (CT-Pedras). Os métodos testados mostraram-se eficientes e adequados para o problema examinado. Este trabalho estabelece importante contribuição para a solução de um problema relevante e complexo na área de Geometria Computacional:

determinar se há inclusão (ou, alternativamente, invasão) espacial entre dois sólidos com superfícies digitalizadas.

**PALAVRAS-CHAVE:** Geometria Computacional; Algoritmos; Método do raio; Curva de Jordan; Lapidação.

## 1. INTRODUÇÃO

Dentre os principais objetivos das estratégias utilizadas para a otimização do corte de gemas coradas digitalizadas tridimensionalmente está a obtenção do maior volume possível para uma lapidação específica dentro de uma pedra bruta (BRUSSO *et al.*, 2010), estabelecendo matematicamente os melhores planos para se realizar os cortes. O modelo de lapidação, o qual possui representação matemática conhecida na maioria dos casos, mesmo que apenas aproximada, trata-se de um volume sólido que pode ser posicionado no espaço através de seis parâmetros: três coordenadas de centro geométrico e três ângulos de giro. A escala em relação a um protótipo de tamanho padronizado também deve ser levada em consideração, sendo um parâmetro adicional importante.

Assim, uma metodologia

matematicamente simples para resolver o problema de otimizar o aproveitamento do corte (lapidação) de pedras consiste em definir um problema de otimização naqueles seis parâmetros, com o objetivo de maximizar a escala, tendo como restrição que o modelo reposicionado e reescalado sempre permaneça interior à pedra bruta a ser trabalhada.

Entretanto, há alguns detalhes que trazem uma certa complexidade a este problema: tanto a pedra a ser trabalhada — a pedra bruta — quanto o próprio modelo padronizado de lapidação — o protótipo — são conhecidos apenas mediante as digitalizações tridimensionais das suas superfícies externas. Isso acontece, sobretudo, em situações nas quais os modelos matemáticos não conseguem aproximar satisfatoriamente a geometria pretendida. Em vista disso, podemos considerar duas situações:

- i. A restrição de que o modelo em escala deve ser interior à pedra bruta requer que uma nuvem de pontos do espaço, obtidos a partir da digitalização do modelo, mas conforme posicionamento e escala pretendidos, deve ser interior à digitalização da pedra bruta;
- ii. Uma vez que sobre o modelo de lapidação temos mais hipóteses de regularidade do que sobre a pedra bruta, é natural, do ponto de vista geométrico, estabelecer o teste dual que valida se algum dos vértices da digitalização da pedra bruta invade o sólido obtido pelo posicionamento e escala do modelo de lapidação digitalizado.

O objetivo aqui é investigar soluções eficientes para o problema de determinar se uma nuvem de pontos está inclusa (ou, alternativamente, invade) a digitalização tridimensional da superfície externa de um sólido não necessariamente convexo. A hipótese de convexidade geralmente guia a construção de muitas soluções e pode ser base para a garantia de existência e unicidade, mas os métodos utilizados durante as simulações não necessitam dessa suposição para atender ambas as situações mencionadas.

O problema de determinar se um ponto  $P$  é interior a um sólido espacial é bastante conhecido na literatura de Geometria Computacional. Por causa das tarefas cada vez mais complexas em muitas aplicações (planejamento urbano, impressão 3D, controle de trajetórias de veículos teleguiados, realidade virtual, entre outras) aplicativos computacionais vêm sendo cada vez mais usados para modelar, analisar e visualizar dados tridimensionais de maneira apropriada. Uma das tarefas mais desafiadoras é a análise em 3D do posicionamento de objetos de diferentes tipos, isto é, geometrias (CUI *et al.*, 2011). Nesse contexto, o teste de inclusão, que objetiva determinar se um ponto é interior a um poliedro dado, é uma das operações elementares que áreas como Geometria Computacional e Computação Gráfica necessitam. O teste de inclusão não é um problema complexo, o desafio maior recai em obter soluções que sejam robustas e eficientes, uma vez que tais testes são aplicados repetidamente um número muito grande de vezes (CUI *et al.*, 2011;



LIU *et al.*, 2010).

A estratégia mais conhecida na literatura é o método baseado em *ray casting* (ou Método do Raio), que basicamente conta a paridade do número de intersecções entre um raio (semirreta), originado em um ponto  $P$ , e a fronteira do sólido, baseando-se no Teorema da Curva de Jordan. No caso particular de um sólido cuja fronteira é uma estrutura (malha) de faces poligonais adjacentes, oriunda do processo de digitalização, encontros singulares do raio com arestas ou vértices da discretização geram intersecções simultâneas com um número muitas vezes incerto de faces contíguas, o que normalmente dificulta tal contagem. Além disso, situações onde o raio é paralelo ou coplanar ao plano de alguma das faces, ou ainda quando existem faces quase-degeneradas, podem acarretar sérias dificuldades numéricas (LIU *et al.*, 2010). Para aumentar a eficiência, o Método do Raio pode ser aplicado repetidamente ou então usando uma direção aleatória (O'ROURKE, 1998). Normalmente, melhor performance é obtida com uma combinação do Método do Raio e estruturas de dados hierárquicas do tipo *octree* ou *grid* (OGAYAR; SEGURA; FEITO, 2005), mas a possíveis grandes custos de armazenamento em memória (LIU *et al.*, 2010). Um outro método muito conhecido é baseado em estruturas espaciais chamadas de *BSP* (*Binary Space Partitioning*), que têm como desvantagem um custo de processamento muito grande quando a geometria/topologia dos modelos é irregular (CUI *et al.*, 2011; HAINES, 1994). Para a versão bidimensional do problema de inclusão, quando é necessário determinar se um ponto é interior a uma curva plana, inúmeras soluções têm sido propostas para o caso de polígonos com número qualquer de arestas, convexos ou não (BADOUEL, 1990; FEITO; TORRES, 1997; HAINES, 1994; MÖLLER; TRUMBORE, 1997).

De maneira alternativa, alguns trabalhos propõem procedimentos alicerçados na contabilização de volumes com sinal de tetraedros formados por faces da digitalização e o ponto  $P$  a ser testado (FEITO; TORRES, 1997; FEITO; TORRES; URENA, 1995; JIMÉNEZ; SEGURA; FEITO, 2010; OGAYAR; SEGURA; FEITO, 2005), o que tem se mostrado simples e seguro, contudo, ineficiente para discretizações com número muito grande de faces (LIU *et al.*, 2010). LIU *et al.*, (2010) apresenta um novo método de solução que usa pré-processamento e determinação de triângulos espaciais no interior de um poliedro, e que generaliza o método de Horn (HORN; TAYLOR, 1989). Existem também estratégias bastante conhecidas que resolvem o problema de inclusão bidimensional que se baseiam no cálculo do chamado *winding number* — ou índice do ponto — conforme (HAINES, 1994; O'ROURKE, 1998). Tais estratégias têm sido preteridas por requererem avaliações de funções trigonométricas, que sabidamente trazem ineficiência por implicarem em um maior número de operações básicas de aritmética de máquina e por trazerem embutidos erros de ponto flutuante, que podem sofrer acúmulo catastrófico.

O presente trabalho apresenta sucintamente as ideias fundamentais



de alguns algoritmos bastante utilizados na literatura e na prática para o teste de inclusão de uma nuvem pontos em sólidos digitalizados tridimensionalmente em malha discreta, bem como resultados de testes de comparação no contexto de pedras brutas. O conjunto de malhas utilizado foi fornecido pelo Centro Tecnológico de Pedras, Gemas e Joias do Rio Grande do Sul (CT-Pedras). Uma discussão e pormenorização mais detalhada desses métodos e resultados é feita em MORETTI (2015).

## 2. ALGORITMOS UTILIZADOS

São apresentados aqui, de maneira breve, as principais ideias dos métodos computacionais utilizados nos testes realizados.

### 2.1. Método de Badouel

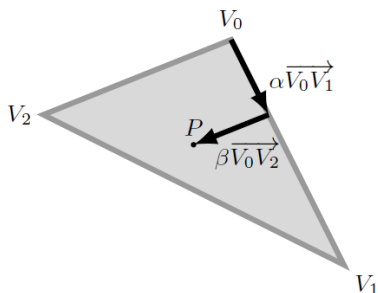
Esse algoritmo (BADOUEL, 1990) determina primeiramente se o cruzamento entre um segmento de reta e um triângulo ocorre e, posteriormente, as coordenadas desse ponto de intersecção com respeito aos vértices do triângulo. Para tal, essa intersecção e o triângulo são projetados nos planos  $xy$ ,  $yz$  ou  $xz$ , com o objetivo de resolver o problema em duas dimensões usando as suas coordenadas baricêntricas. Esses parâmetros podem ser usados também para o cálculo do vetor normal que tem como origem o ponto de intersecção, caso ele seja de interesse.

Se  $V_0V_1V_2$  é um triângulo, então a posição de um ponto  $P$  coplanar ao plano definido pelo triângulo pode ser descrita na forma vetorial da seguinte maneira:

$$\overrightarrow{V_0P} = \alpha \overrightarrow{V_0V_1} + \beta \overrightarrow{V_0V_2} \tag{1}$$

Podemos visualizar essa representação na figura (1) a seguir:

**Figura 1:** Parametrização do ponto  $P$  segundo suas coordenadas baricêntricas



Para que o ponto  $P$  seja interior ao triângulo devemos ter  $\alpha \geq 0$ ,  $\beta \geq 0$

e também  $\alpha + \beta \leq 1$ .

### 2.2. Método de Möller

Esse procedimento (MÖLLER; TRUMBORE, 1997) é construído por meio de uma transformação da face triangular e da origem do raio  $R(t) = O + tD$ , sendo  $D$  uma direção escolhida, não necessariamente normalizada, e  $O$  o seu ponto de origem. Tal transformação produz a distância  $t$  de  $O$  até o ponto de intersecção com a face e as coordenadas baricêntricas deste ponto. O funcionamento do método se baseia na translação do triângulo até a origem do sistema de coordenadas e numa troca de escala a fim de que o triângulo seja unitário no plano  $yz$  e que o raio esteja alinhado com eixo  $x$ .

Um ponto  $P$  em um triângulo  $V_0V_1V_2$  pode ser parametrizado por

$$P(\alpha, \beta) = (1 - \alpha - \beta)V_0 + \alpha V_1 + \beta V_2 \tag{2}$$

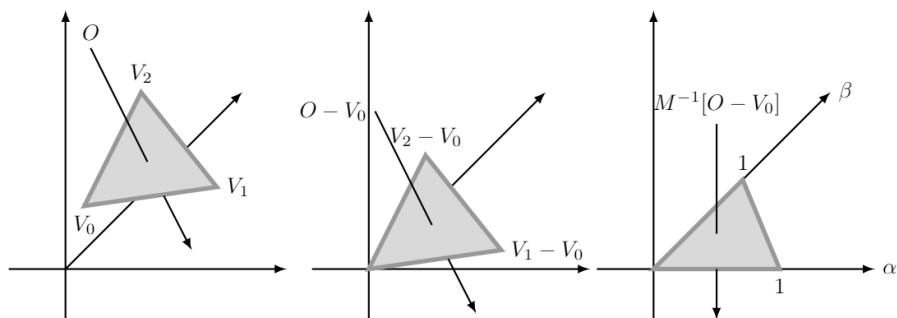
onde  $\alpha$  e  $\beta$  são as coordenadas baricêntricas de  $P$  com respeito ao triângulo dado.

Para a intersecção entre o raio e o triângulo, temos  $R(t) = P(\alpha, \beta)$ , ou seja,

$$\begin{bmatrix} -D & V_1 - V_0 & V_2 - V_0 \end{bmatrix} \begin{bmatrix} t \\ \alpha \\ \beta \end{bmatrix} = O - V_0 \tag{3}$$

A matriz  $M = [-D \ V_1 - V_0 \ V_2 - V_0]$  realiza as transformações desejadas com o raio e os vértices do triângulo. Essa transformação pode ser visualizada na figura (2).

**Figura 2:** Transformações no triângulo e no raio realizadas pela matriz  $M$



Desse modo, resolvendo o sistema (3) obtemos os valores da coordenadas baricêntricas  $\alpha$  e  $\beta$  e podemos decidir se ponto é interno ao

triângulo ou não.

### 2.3. Método de Segura

Esse algoritmo (SEGURA *et al.*, 2005) é baseado no estudo do volume com sinal de diferentes tetraedros formados com os vértices do triângulo e dois pontos pertencentes ao raio a ser testado  $R(t) = O + tD$ .

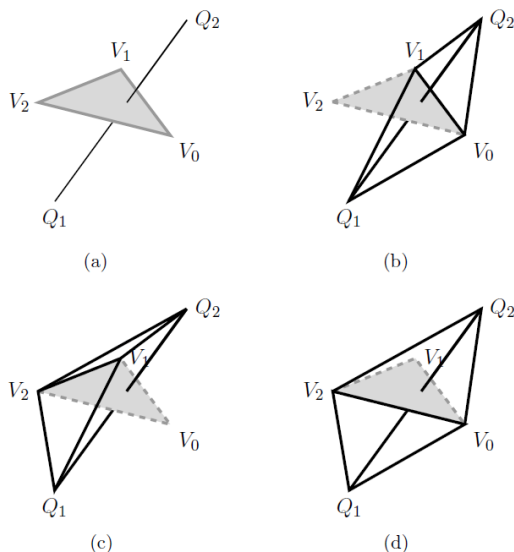
Sejam  $V_0V_1V_2$  um triângulo e  $Q_1$  e  $Q_2$  as extremidades do segmento  $Q_1Q_2 \in R(t)$ , os quais estão em lados opostos em relação ao plano que contém o triângulo. A orientação dos vértices é escolhida de tal forma que o tetraedro  $Q_1V_0V_1V_2$  tenha orientação positiva (isto é, o seu volume com sinal  $[Q_1V_0V_1V_2]$  seja positivo). Assim, o segmento  $Q_1Q_2$  intercepta o triângulo  $V_0V_1V_2$  se, e somente se,

$$\text{sgn}([Q_2V_0V_1Q_1]) \geq 0 \text{ e } \text{sgn}([Q_2V_2V_1Q_1]) \geq 0 \text{ e } \text{sgn}([Q_2V_0V_2Q_1]) \geq 0 \quad (4)$$

Esse método indica se a intersecção de um segmento com um triângulo existe ou não, mas originalmente não calcula as coordenadas baricêntricas do ponto de intersecção. Os casos possíveis para a interceptação são mostrados na figura (3). Caso o triângulo  $V_0V_1V_2$  tenha orientação negativa, a inclusão de P nele fica estabelecida se, e somente se,

$$\text{sgn}([Q_2V_0V_1Q_1]) \leq 0 \text{ e } \text{sgn}([Q_2V_2V_1Q_1]) \leq 0 \text{ e } \text{sgn}([Q_2V_0V_2Q_1]) \leq 0 \quad (5)$$

**Figura 3:** Visualização do algoritmo de Segura: (a) problema de intersecção; (b) tetraedro  $Q_2V_0V_1Q_1$ ; (c) tetraedro  $Q_2V_2V_1Q_1$ ; (d) tetraedro  $Q_2V_0V_2Q_1$



## 2.4. Um novo método de Segura

Um novo algoritmo foi proposto por Segura (JIMÉNEZ; SEGURA; FEITO, 2010) para determinar a intersecção entre um segmento e um triângulo. Dados um triângulo  $V_0V_1V_2$ , um raio  $R(t) = O + tD$  e um segmento  $Q_1Q_2 \in R(t)$ , a estratégia consiste encontrar as coordenadas baricêntricas de um dos extremos do segmento  $Q_1Q_2$ , digamos  $Q_2$ , com respeito ao tetraedro  $Q_1V_0V_1V_2$  e verificar os seus sinais.

Seja  $Q_2 = \alpha Q_1 + \beta V_0 + \gamma V_1 + \delta V_2$ , com  $\alpha, \beta, \gamma, \delta \in \mathbb{R}$  a equação que determina a posição de  $Q_2$  com respeito ao tetraedro  $T = Q_1V_0V_1V_2$ . Desse modo, temos  $\delta = 1 - \alpha - \beta - \gamma$ . Para o cálculo dessas coordenadas, é usado o volume com sinal do tetraedro  $T$ . O sinal depende da ordem em que os pontos são listados: se os triângulos do tetraedro têm orientação positiva (sentido anti-horário), o sinal é positivo; caso contrário, é negativo. Ainda é possível o volume ser nulo caso os quatro pontos sejam coplanares.

Dessa forma, o ponto  $Q_1$  assume o papel de vértice de origem do tetraedro  $T$ . A orientação será positiva se o triângulo  $V_0V_1V_2$  tiver orientação positiva em relação ao vértice de origem e os triângulos  $Q_1V_1V_0$ ,  $Q_1V_2V_1$  e  $Q_1V_0V_2$  tiverem orientação positiva com relação aos demais vértices. Assim, as coordenadas baricêntricas de  $Q_2$  em relação ao tetraedro  $T$  são calculadas como:

$$\begin{aligned} \alpha &= \frac{[Q_2V_0V_1V_2]}{[Q_1V_0V_1V_2]} \\ \beta &= \frac{[Q_2Q_1V_2V_1]}{[Q_1V_0V_1V_2]} \\ \gamma &= \frac{[Q_2Q_1V_0V_2]}{[Q_1V_0V_1V_2]} \\ \delta &= \frac{[Q_2Q_1V_1V_0]}{[Q_1V_0V_1V_2]} \end{aligned} \tag{6}$$

De posse das coordenadas baricêntricas de  $Q_2$  em relação ao tetraedro  $T = Q_1V_0V_1V_2$ , é possível decidir se um ponto é interno ou externo a ele. Para isso, basta verificar se  $\alpha + \beta + \gamma + \delta = 1$  e se  $\alpha, \beta, \gamma, \delta \in [0,1]$ .

## 2.5. Método de Feito-Torres

Esse método (FEITO; TORRES, 1997) difere-se bastante dos demais descritos anteriormente no seu funcionamento e notação. Dada a discretização em malha triangular de um sólido, são considerados tetraedros determinados pela origem, um vértice fixo de cada face e cada uma das demais arestas da mesma face que não sejam incidentes a esse vértice fixado. São definidos os seguintes conjuntos:

**i.  $V_+(Q)$ :** conjunto dos vértices que determinam arestas pela origem que contêm  $Q$  e ou que pertençam a tetraedros positivos ou que são pontos iniciais ou finais de faces positivas, de acordo com a ordem em que aparecem na descrição da face.  $\#V_+(Q)$  denota a cardinalidade deste conjunto;

**ii.  $V_-(Q)$ :** conjunto dos vértices que determinam arestas pela origem que contêm  $Q$  e ou que pertençam a tetraedros negativos ou que são pontos iniciais ou finais de faces negativas, de acordo com a ordem em que aparecem na descrição da face.  $\#V_-(Q)$  indica a cardinalidade deste conjunto;

**iii.  $T_{or}(OV_0V_1V_2)$ :** conjunto de faces de tetraedros pela origem determinados pela origem  $O$  e os pontos  $V_0, V_1$  e  $V_2$ .

O método se baseia no seguinte teorema, enunciado aqui sem demonstração:

**Teorema:** Seja  $P = F_1 F_2 \dots F_n$  um poliedro qualquer com faces  $F_1, F_2, \dots, F_n$ , onde  $F_i = V_{i,1} V_{i,2} \dots V_{i,e_i}$ . Então, um ponto  $Q \in \mathbb{R}^3$ , não-contido na fronteira de  $P$ , é interior ao poliedro se, e somente se,

$$\sum_{i=1}^n \sum_{j=1}^{e_i} \lambda_{ij} + \#V_+(Q) - \#V_-(Q) = 1 \tag{7}$$

onde  $1 < j < e_i$ , e

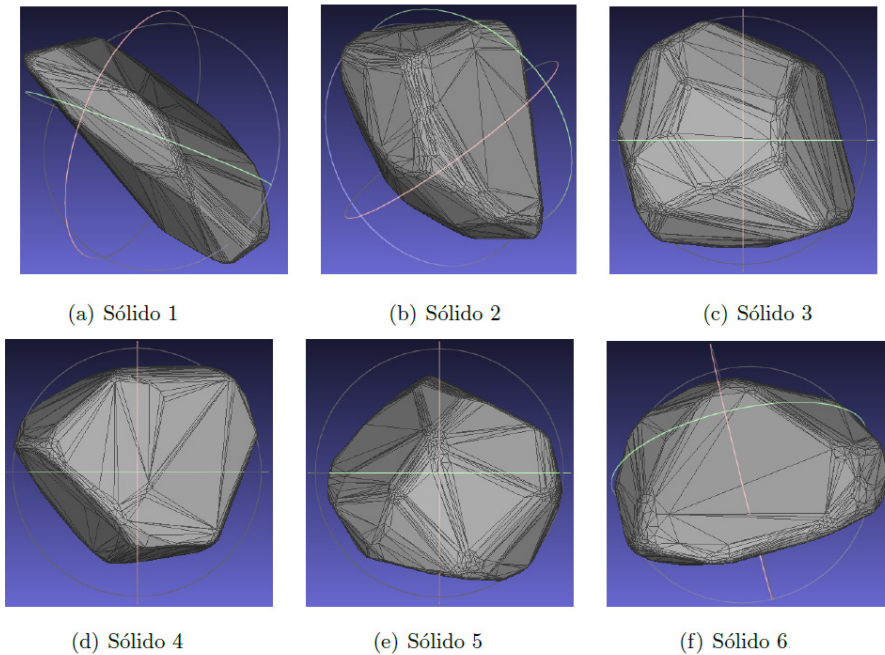
$$\lambda_{ij} = \begin{cases} \text{sgn}([OV_{i,1}V_{ij}V_{i,j+1}]), & \text{se } Q \in \text{int}(OV_{i,1}V_{ij}V_{i,j+1}) \\ 0,5 \cdot \text{sgn}([OV_{i,1}V_{ij}V_{i,j+1}]), & \text{se } Q \in \text{int}(T_k), \text{ com } T_k \in T_{or}(OV_{i,1}V_{ij}V_{i,j+1}) \\ 0, & \text{caso contrário} \end{cases} \tag{8}$$

O cerne do método é a avaliação da equação (7) para a decisão se o ponto é interior ao sólido ou não.

### 3. TESTES COMPUTACIONAIS

Para os testes de comparação entre os métodos apresentados foram utilizadas malhas de pedras brutas fornecidas pelo Centro Tecnológico de Pedras, Gemas e Joias do Rio Grande do Sul (CT-Pedras). O conjunto de sólidos utilizados utilizado é apresentado a seguir na figura (4).

Figura 4: Sólidos testados



Os métodos apresentados na seção anterior foram implementados em linguagem C por meio do compilador DevC++ 5.11, utilizando diretivas padrão de compilação. Os testes foram realizados em um processador Intel Core i5- 7400, 3 GHz, 4 GB RAM, em sistema Windows 10, versão 64 bits.

Para a realização dos testes foram gerados três lotes de 5000 pontos aleatórios: um com pontos fora do sólido, outro com pontos dentro e outro sem restrições de localização. Para cada lote de pontos utilizou-se um dos métodos 10 vezes seguidas, sendo tomada a média do tempo de execução para cada algoritmo. Ao total, foram testados seis métodos diferentes, baseados nas apresentações anteriores: duas versões o algoritmo de Badouel, uma se utilizando do um vetor normal a cada face pré-processado e outra versão com o método típico; uma versão do método de Möller; uma versão para o método de Segura; uma versão para o novo método de Segura; e uma versão para o método de Feito-Torres (esse é o único dos métodos testados que não utiliza o Teorema da Curva de Jordan). O tempo para o pré-processamento dos vetores normais (onde necessário) e também para a geração da nuvem de pontos aleatória não foi contabilizado.

Após essas baterias de testes, temos os seguintes resultados apresentados nas tabelas (1), (2) e (3), normalizados segundo o método de Badouel, versão sem vetor normal:

**Tabela 1:** Teste 1: tempo relativo para 5000 pontos gerados dentro do sólido

S	V/F	MÉTODO					
		Badouel		Möller	Segura		Feito-Torres
		s/N	c/N		Antigo	Novo	
1	838/1672	1	0,7141	0,7263	0,9169	0,8158	3,1711
2	990/1976	1	0,7212	0,7336	0,9251	0,8233	3,1817
3	1086/2168	1	0,7263	0,7382	0,9301	0,8284	3,1883
4	1142/2280	1	0,7285	0,7409	0,9329	0,8314	3,1918
5	1192/2380	1	0,7306	0,7435	0,9352	0,8341	3,1952
6	1199/2394	1	0,7311	0,7439	0,9355	0,8346	3,1958

**Tabela 2:** Teste 2: tempo relativo para 5000 pontos gerados fora do sólido

S	V/F	Pontos D/F	MÉTODO					
			Badouel		Möller	Segura		Fe
			s/N	c/N		Antigo	Novo	
1	838/1672	2441/2559	1	0,7150	0,7296	0,9147	0,7950	
2	990/1976	2497/2503	1	0,7219	0,7364	0,9229	0,8021	
3	1086/2168	2486/2514	1	0,7268	0,7414	0,9280	0,8071	
4	1142/2280	2424/2576	1	0,7289	0,7443	0,9310	0,8097	
5	1192/2380	2423/2577	1	0,7314	0,7471	0,9334	0,8124	
6	1199/2394	2489/2511	1	0,7321	0,7475	0,9339	0,8128	

**Tabela 3:** Teste 3: tempo relativo para 5000 pontos gerados dentro e fora do sólido

S	V/F	Pontos D/F	MÉTODO					
			Badouel		Möller	Segura		Feito-Torres
			s/N	c/N		Antigo	Novo	
1	838/1672	2441/2559	1	0,7150	0,7296	0,9147	0,7950	3,1706
2	990/1976	2497/2503	1	0,7219	0,7364	0,9229	0,8021	3,1811
3	1086/2168	2486/2514	1	0,7268	0,7414	0,9280	0,8071	3,1882
4	1142/2280	2424/2576	1	0,7289	0,7443	0,9310	0,8097	3,1915
5	1192/2380	2423/2577	1	0,7314	0,7471	0,9334	0,8124	3,1950
6	1199/2394	2489/2511	1	0,7321	0,7475	0,9339	0,8128	3,1955

Nessas tabelas, são indicados na coluna V/F o número de vértices e de faces de cada sólido, respectivamente. Também, na coluna D/F, da tabela (3), são indicadas as quantidades de pontos gerados dentro e fora dos sólidos, respectivamente, para o teste dos métodos.

A tabela (4) a seguir mostra, para fins de completude nas comparações, os tempos absolutos, em segundos, dos testes para o método de Badouel, versão sem vetor normal pré-processado, já que esse tempo foi o utilizado

como referência nas tabelas anteriores.

**Tabela 4:** Tempo absoluto (segundos) para o algoritmo de Badouel (sem vetor normal)

S	V/F	Teste 1	Teste 2	Teste 3
1	838/1672	4,6525	4,6509	4,6533
2	990/1976	4,8947	4,8958	4,8953
3	1086/2168	5,0577	5,0579	5,0569
4	1142/2280	5,1392	5,1389	5,1377
5	1192/2380	5,2205	5,22	5,2184
6	1199/2394	5,2319	5,231	5,2297

#### 4. RESULTADOS

Como dito em (JIMÉNEZ; SEGURA; FEITO, 2010; LÖFSTEDT; AKENINE-MÖLLER, 2005), não é uma tarefa simples desenvolver um conjunto de testes computacionais que certifiquem que um dado método é mais rápido que outros, a performance depende de diversos fatores distintos, tais como: o compilador, as especificações de hardware, o tipo de aplicação e de dados disponíveis e também detalhes na construção dos programas. Contudo, podemos observar indícios para estabelecer algumas comparações baseadas no conjunto de testes propostos nesse trabalho.

Analisando as tabelas (1) (2) e (3), podemos observar que o algoritmo de Möller é a mais eficiente das estratégias implementadas dentre as que não usam informações obtidas por pré-processamento — isto é, desconsiderando os métodos Segura Novo e Feito-Torres. Este algoritmo obteve resultados superiores que oscilam entre 25% a 27%, aproximadamente, quando comparados ao algoritmo de Badouel, sem normais. A performance do método de Segura (Segura Antigo) é também superior ao de Badouel, mas não mais do que 9%, aproximadamente, nesses testes. SEGURA e FEITO (1998, 2001) atestam superioridade desse algoritmo perante o método de Möller, o que não foi observado em nenhum momento nos testes realizados no presente trabalho. Conforme apontado em (JIMÉNEZ; SEGURA; FEITO, 2010), essa melhor performance foi obtida nos trabalhos citados porque nesses estudos comparativos foram utilizadas informações pré-calculadas, como por exemplo, o volume com sinal dos tetraedros e, naquele contexto, nenhuma informação pré-processada foi utilizada para o método de Möller.

O novo método proposto por Segura (Segura Novo) tem resultados mais satisfatórios quando comparados ao método de Badouel, variando de 17% a 20%, aproximadamente, a sua melhoria no tempo de execução para o conjunto de malhas testadas. No entanto, quando a sua performance é comparada com a do algoritmo de Möller ela é inferior. Em (JIMÉNEZ; SEGURA;



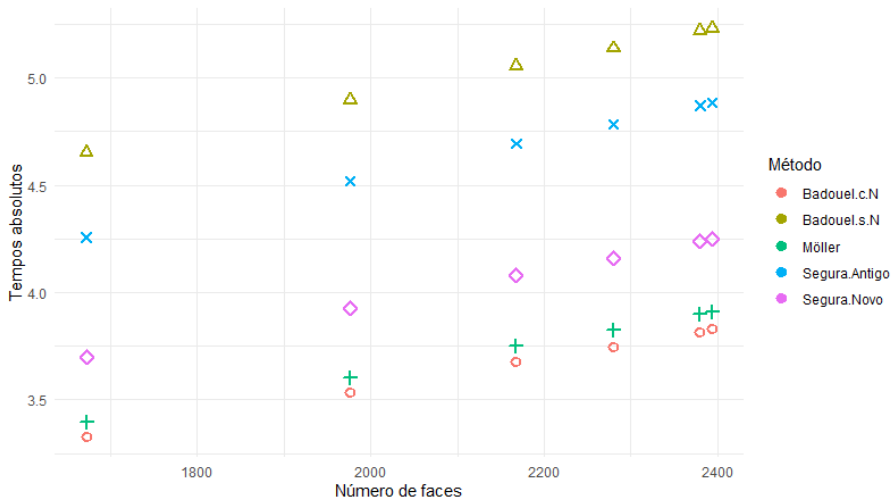
FEITO, 2010) foram feitos testes em que o método obtém resultados melhores que o de Möller, no contexto de simulação de interferência entre objetos numa mesma cena, que é o principal objetivo do artigo. É mencionado, contudo, que o método de Möller é mais rápido quando são aplicados testes de *ray casting* similares aos executados neste trabalho. Isso sugere investigações adicionais futuras de possíveis melhorias no método, como utilização de pré-processamento de algumas informações que possam ser armazenadas na própria estrutura dos triângulos e compartilhadas entre eles, mas que não foram levadas em conta nem em (JIMÉNEZ; SEGURA; FEITO, 2010) nem no presente trabalho.

A implementação do método proposto por Feito-Torres (FEITO; TORRES, 1997) obteve resultados mais desfavoráveis para o conjunto de malhas testadas. (SEGURA *et al.*, 2005) alegam que esse método é mais eficiente do que qualquer método baseado no Teorema da Curva de Jordan; entretanto, foi utilizada uma versão modificada do algoritmo, a qual não foi totalmente especificada, apenas foi apontado que estruturas de divisão espacial (*octrees*, *BSP* e *voxels*) foram empregadas nos testes. Nos testes realizados aqui, a performance deste algoritmo ficou muito aquém de todos os outros métodos utilizados, o tempo de execução sendo até 220% mais lento do que o método de Badouel. De acordo com (HAINES; AKENINE-MÖLLER, 2002), os testes baseados na Curva de Jordan são os mais rápidos para se testar a inclusão, sem pré-processamento, de pontos em malhas, o que foi de fato verificado no presente trabalho.

Quando o algoritmo de Badouel foi melhorado com o uso dos vetores normais às faces (obtidos por pré-processamento) os seus resultados foram superiores, inclusive ao método de Möller. Segundo (JIMÉNEZ; SEGURA; FEITO, 2010), isso realmente se verifica quando há informações adicionais pré-calculadas, como os vetores normais ou então os planos de projeção das faces. Comparando sua performance a do caso sem pré-processamento, os resultados variaram entre 27% e 29% de ganho de tempo, aproximadamente.

Ademais, segundo os resultados obtidos, podemos observar que há indicativo do esforço computacional ser linear no número de faces dos sólidos. Tal evidência é reforçada pela análise da figura (5) a seguir, que mostra os tempos absolutos para o teste 3 (o método de Feito-Torres não estão presentes por questões de escala).

**Figura 5:** Tempos absolutos para o teste 3.



## 5. COMENTÁRIOS FINAIS

Com base nos resultados obtidos por este trabalho, contextualizados para inclusão/invasão de uma nuvem de pontos e digitalizações em malhas triangulares de pedras brutas, fica evidente que métodos conceitualmente muito simples como os de Badouel e Möller, baseados no Teorema da Curva de Jordan, são competitivos frente a estratégias mais complexas encontradas na literatura, considerando algoritmos que não fazem uso de estruturas pré-processadas além de normais de face. Mais especificamente, se os normais exteriores das faces da digitalização não são conhecidos — por exemplo, arquivos com extensão *.OFF* — obtém-se melhor performance com o método de Möller. Por outro lado, quando os vetores normais exteriores das faces são conhecidos — arquivos do tipo *.STL*, por exemplo — estão disponíveis, o método de Badouel melhorado consegue um melhor desempenho.

Como já mencionado anteriormente, não foram utilizadas estruturas de hierarquização espacial e nem pré-processamentos prévios para os testes; sugerem-se aprimoramentos podem a serem incorporadas em novas versões dos algoritmos a fim de melhorar suas performances, tanto a nível de informações pré-implementadas (envelopes convexos, volumes com sinal de tetraedros ou técnicas de coerência, por exemplo) quanto a nível de código nos métodos.

## REFERÊNCIAS

BADOUEL, D. **An efficient ray-polygon intersection**. Academic Press Professional, Inc., p. 390–393, 1990.

BRUSSO, M. J. *et al.* **Tecnologia 3DGemas**: otimização do aproveitamento de gemas coradas digitalizadas tridimensionalmente. *In*: (L. A. Hartmann & J. T. da Silva, Org.) *Tecnologias para o setor de gemas, joias e mineração*. Porto Alegre: IGEO/UFRGS, p. 40–52, 2010.

CUI, S. *et al.* **Point-in-polyhedra test with direct handling of degeneracies**. *Geo-Spatial Information Science*, v. 14, n. 2, p. 91–97, 2011.

FEITO, F. R.; TORRES, J. C. **Inclusion test for general polyhedra**. *Computers & Graphics*, v. 21, n. 1, p. 23–30, 1997.

FEITO, F.; TORRES, J. C.; URENA, A. **Orientation, simplicity, and inclusion test for planar polygons**. *Computers & Graphics*, v. 19, n. 4, p. 595–600, 1995.

HAINES, E. **Point in polygon strategies**. *Graphics Gems IV*, p. 24–46, 1994.

HAINES, E.; AKENINE-MÖLLER, T. **Real-time rendering**. 2. ed. AK Peters, Ltd., 2002.

HORN, W. P.; TAYLOR, D. L. **A theorem to determine the spatial containment of a point in a planar polyhedron**. *Computer Vision, Graphics and Image Processing*, v. 45, n. 1, p. 106–116, 1989.

JIMÉNEZ, J. J.; SEGURA, R. J.; FEITO, F. R. **A robust segment/triangle intersection algorithm for interference tests**. *Efficiency study*. *Computational Geometry*, v. 43, n. 5, p. 474–492, 2010.

LIU, J. *et al.* **A new point containment test algorithm based on preprocessing and determining triangles**. *Computer-Aided Design*, v. 42, n. 12, p. 1143–1150, 2010.

LÖFSTEDT, M.; AKENINE-MÖLLER, T. **An evaluation framework for ray-triangle intersection algorithms**. *Journal of Graphics, GPU and Game Tools*, v. 10, n. 2, p. 13–26, 2005.

MÖLLER, T.; TRUMBORE, B. **Fast, minimum storage ray-triangle intersection**. *Journal of Graphics Tools*, v. 2, n. 1, p. 21–28, 1997.

MORETTI, V. F. **Inclusão entre nuvem de pontos e digitalização 3D**: estratégias e implementação. 2015. 88 f. Dissertação de Mestrado (UFRGS), Porto Alegre, 2015. Disponível em: <https://lume.ufrgs.br/handle/10183/132245>.

OGAYAR, C. J.; SEGURA, R. J.; FEITO, F. R. **Point in solid strategies**. *Computers & Graphics*, v. 29, n. 4, p. 616–624, 2005.

O'ROURKE, J. **Computational geometry in C**. Cambridge University Press, 1998.

SEGURA, R. *et al.* **An efficient point classification algorithm for triangle meshes**. *Journal of Graphics, GPU, and Game Tools*, v. 10, n. 3, p. 27–35, 2005.

SEGURA, R. J.; FEITO, F. R. **Algorithms to test ray-triangle intersection.** Comparative study. WSCG (Short Papers), 2001. p. 76–81.

SEGURA, R. J.; FEITO, F. R. **An algorithm for determining intersection segment-polygon in 3D.** Computers & Graphics, v. 22, n. 5, p. 587–592, 1998.

### COMPUTATIONAL METHODS FOR THE PROBLEM OF INCLUSION BETWEEN POINT CLOUD AND THREE-DIMENSIONAL DIGITIZED SOLIDS

**ABSTRACT** - The present work investigates efficient solutions to the problem of determining whether a point cloud is contained (or, alternatively, invades) the three-dimensional digitization of the surface of a not necessarily convex solid. Strategies based on ray casting and on the Jordan Curve Theorem, generalized for the three-dimensional case, as well as others based on the computation of volumes with sign of tetrahedrons by the origin, were tested and compared according to their computational effectiveness and efficiency. The computational experiments were carried out using digitization of unpolished stones made available by Centro Tecnológico de Pedras, Gemas e Joias do Rio Grande do Sul (CT-Pedras). The tested methods proved to be efficient and adequate for the examined problem. This work establishes an important contribution to the solution of a relevant and complex problem in the field of Computational Geometry: determining whether there is spatial inclusion (or, alternatively, invasion) between two solids with digitized surfaces.

**KEYWORDS:** Computational Geometry; Algorithms; Ray casting; Jordan curve; Stone cutting.

## CARACTERIZAÇÃO DA ESTRUTURA DE DEPENDÊNCIA DO GENOMA HUMANO USANDO CAMPOS MARKOVIANOS ESTUDO DE POPULAÇÕES MUNDIAIS E DADOS DE SNPs

**Francisco José de Almeida  
Fernandes**

Instituto de Matemática e Estatística  
– USP – SP

[lattes.cnpq.br/9912184132196127](https://lattes.cnpq.br/9912184132196127)

**Júlia Maria Pavan Soler**

Instituto de Matemática e Estatística  
– USP – SP

[lattes.cnpq.br/2947693375237766](https://lattes.cnpq.br/2947693375237766)

**RESUMO** - Marcadores moleculares do tipo SNP (*Single Nucleotide Polimorphism*), com alta densidade dentro do DNA humano, têm sido largamente utilizados como amostradores do genoma. Este trabalho faz uso de uma modelagem estocástica de campos Markovianos de alcance variável para encontrar blocos independentes de SNPs, estruturando assim o genoma em regiões ilhadas de dependência. Foram utilizados dados públicos de SNPs de diferentes populações mundiais (projeto HapMap), além de uma amostra da população brasileira. Os blocos de SNPs configuram regiões de dependência ou janelas de influência as quais foram usadas para caracterizar as diferentes populações de acordo com sua ancestralidade e

os resultados obtidos mostraram que as janelas da população brasileira têm, em média, tamanho maior, evidenciando a sua história recente de miscigenação, comparativamente às demais populações mundiais. É também proposta a construção de janelas de consenso maximais em determinadas regiões genômicas, tendo como base as janelas de influência de cada população do estudo. Identificando as janelas de consenso mais heterogêneas, a Análise de Correspondência e a teoria Biplots foram aplicadas na construção de um classificador capaz de determinar o percentual relativo de ancestralidade de um indivíduo, o qual, submetido à validação, obteve alta eficiência no acerto da população originária.

**PALAVRAS-CHAVE:** Campos Markovianos de alcance variável; HapMap; Biplots; blocos de SNPs; Classificador de ancestralidade.

### 1. INTRODUÇÃO

Polimorfismos são variações genéticas entre indivíduos, para um mesmo local do genoma. Uma vez que o DNA é uma molécula longa, com duas fitas ligadas entre si por pares de bases, A (adenina) com T

(timina) e C (citosina) com G (guanina), denomina-se SNP (do inglês *Single Nucleotides Polimorphisms*) ao polimorfismo resultante da alteração de um único par dessas bases. Os SNPs estão distribuídos ao longo de todo o DNA, são originados de mutações e dada sua alta densidade podem ser utilizados para amostrar o genoma com uma cobertura representativa de sua variabilidade.

O projeto HapMap (2005) foi criado a partir de um consórcio formado por vários países com o objetivo de identificar e catalogar as diferenças e similaridades entre o genoma de seres humanos através do levantamento do mapa dos SNPs em várias populações, distribuídas em regiões geográficas distintas e com diferentes ancestralidades<sup>1</sup>. A teoria mais aceita atualmente no que diz respeito à origem do *Homo sapiens* dá conta que sua migração teve início a partir da África, mais especificamente, na atual Etiópia (SCHLESINGER, 2010). Entre as evidências científicas que sustentam essa teoria, está o fato de que, sendo as populações africanas mais antigas, o tempo permitiu um maior número de recombinações gênicas (mistura entre os genes dos pais), levando a regiões, que são transmitidas em conjunto para seus descendentes, menores do que aquelas encontradas em populações mais recentes como os europeus. Neste contexto, a população brasileira tem uma história recente de miscigenação, com cerca de 500 anos ou, equivalentemente, cerca de 20 gerações.

Neste trabalho, essas regiões foram inferidas utilizando-se uma abordagem de campos Markovianos para levantamento das janelas (ou blocos) de SNPs que possuem dependência entre si, em cada população do projeto HapMap, bem como em uma amostra da população brasileira. O Quadro 1 contém uma descrição das populações utilizadas neste estudo<sup>2</sup>. Denotaremos estes blocos de SNPs como as “*janelas populacionais*” específicas de cada população (Jp). Os resultados obtidos foram utilizados para caracterizar as diferentes populações e, conseqüentemente, suas ancestralidades. Estendendo o conceito de influência entre SNPs para a unificação das janelas (Jp) de cada população, propõe-se a construção de janelas de consenso maximais (Jc), cada uma contendo as diferentes janelas que se sobrepõem nas diferentes populações. As janelas populacionais que compõem uma determinada janela de consenso, podem ser similares ou não, assim, para medir a dissimilaridade entre as diferentes Jp's de cada janela de consenso, foi elaborado um Índice de Heterogeneidade (IH). Finalmente, utilizando as janelas de consenso mais heterogêneas é proposto um classificador da população originária de um indivíduo.

**Quadro 1:** Populações utilizadas neste trabalho

Rótulo	Origem da População	Ancestralidade
ASW	população do sudoeste dos EUA com ancestralidade africana	Africana
CEU	residentes de Utah com ancestralidade do norte/ocidente da Europa	Europeia
CHB	chineses em Beijing, China, com ancestralidade da dinastia Han	Asiática
CHD	chineses residentes na área metropolitana de Denver, Colorado	Asiática
GIH	índios Gujarati em Houston, Texas	Ameríndia
JPT	japoneses em Tokyo, Japão	Asiática
LWK	residentes em Webuye, Quênia com ancestralidade Luhya	Africana
MEX	residentes em Los Angeles, com ancestralidade mexicana	Ameríndia
MKK	residentes em Kinyawa, Quênia com ancestralidade Maasai	Africana
TSI	residentes próximos a Florência, com ancestralidade toscana	Europeia
YRI	residentes em Ibadan, Nigéria com ancestralidade Yoruba	Africana
BRZ	brasileiros residentes na região urbana de São Paulo	Brasileira

## 2. MODELAGEM

### 2.1. Notação

SNPs são marcadores moleculares dialélicos, assim, sem perda de generalidade, seja o par  $(x, y)$  o genótipo de um SNP, tal que, em um loco cromossômico de um determinado indivíduo teremos  $(x, y) \in \{A, C, T, G\}^2$ , onde  $A = \{A, C, T, G\}$  é o alfabeto e  $|A| = 4$  a sua cardinalidade.

A codificação dos genótipos de um SNP pode ser feita atribuindo o valor 1 para bases complementares de referência, por exemplo, atribuindo 1 às bases A e T e o valor 0 para as bases C e G. Cada SNP configura, então, uma variável aleatória cuja realização pertence ao conjunto  $\{0, 1, 2\}$ . Assim, para um determinado SNP, em uma posição específica  $j$ , temos  $X_j$  a variável aleatória definida como  $X_j = x_j | x_j \in \{0, 1, 2\}$ , em que:

$$\begin{cases} P(X_j = 0) = P(\text{genótipoSNP} \in \{CC, CG, GC, GG\}) \\ P(X_j = 1) = P(\text{genótipoSNP} \in \{AC, AG, TC, TG, CA, GA, CT, GT\}) \\ P(X_j = 2) = P(\text{genótipoSNP} \in \{AA, AT, TA, TT\}) \end{cases}$$

Um conjunto  $S$ , de uma sequência de  $s$  SNPs, será então uma sequência de variáveis aleatórias,  $S = \{X_1, X_2, X_3, \dots, X_s\}$ , sendo,  $|S| = s$  a cardinalidade de  $S$ . Estendendo a notação para  $n$  indivíduos independentes<sup>3</sup>, um determinado SNP na posição  $j$  passa a ter uma amostra de tamanho  $n$  com  $X_j$  tendo distribuição trinomial,  $X_j \sim \text{Trinomial}(n, \phi_j)$ ,  $X_j = x_j^i \in \{0, 1, 2\}$ , em que  $i$  é o indivíduo ( $1 \leq i \leq n$ ) e  $\phi_j$  é o vetor de parâmetros da trinomial, contendo as probabilidades de ocorrência dos valores 0, 1 e 2, restritas a somarem 1.

Uma amostra dos genótipos de  $s$  SNPs para  $n$  indivíduos será então uma matriz  $D_{(n \times s)}$ . Considerando a amostra estratificada em diferentes populações, denotaremos por  $D^g = D_{(n_g \times s)}^g$  como a matriz dos genótipos dos  $n_g$  indivíduos da população  $g$ , onde  $1 \leq g \leq G$ , sendo  $G$  o número total de populações, e  $n = n_1 + \dots + n_G$ .

## 2.2 Campos Markovianos - Conceito de Vizinhança

Cadeias de Markov são um tipo de modelagem bastante utilizada para dados com dependência. Como o interesse é encontrar blocos de SNPs que guardam entre si alguma relação de dependência e, considerando cada realização da variável aleatória  $X_j$  como um estado com distribuição Trinomial, pode-se utilizar a abordagem de cadeias de Markov de alcance variável, com a diferença importante, no caso do problema aqui colocado, que deve-se inspecionar ambos os lados de um determinado estado genotípico de um loco fixado (SNP). Em outras palavras, o objetivo é encontrar a memória da cadeia de Markov em ambas as direções de um determinado estado. O problema assim proposto, e considerando dados de uma única população, foi abordado por Leonardi (2006, 2010) utilizando campos Markovianos, verificando a vizinhança de cada posição da cadeia e encontrando quantas posições antes e depois são necessárias para determinação da distribuição de probabilidade dos estados da posição em estudo.

Dado um alfabeto  $A$  pode-se definir uma palavra  $w$  de tamanho  $m$  como sendo uma sequência de letras desse alfabeto ( $w \in A^m$ ). Fixada uma posição  $j$  dentro da palavra  $w$ ,  $\omega$  com  $|\omega| = l$  será prefixo de  $j$  e  $\tau$  com  $|\tau| = r$  será sufixo de  $j$  se  $w = \omega j \tau$ . Denominamos  $\omega$  a vizinhança à esquerda da posição  $j$  em  $w$  e  $\tau$  a vizinhança à direita da posição  $j$  em  $w$ . À concatenação  $(w.\tau)$  denominamos simplesmente *vizinhança de  $j$  dentro de  $w$* . Usando esse conceito para a sequência de SNPs da amostra, dado um SNP na posição  $j$  sua vizinhança será o conjunto:

$$V_j^{l,r} = \{k \in \mathbb{N} | j - l \leq k \leq j + r, k \neq j, j - l \geq 1, j + r \leq s\}.$$



### 2.3 Função de Pseudo-Verossimilhança

Dada a matriz  $D$  da amostra, considerando os indivíduos independentes e dado um indivíduo  $i$ , cada SNP é assumido independente condicionalmente ao restante dos SNPs. Assim, pode-se utilizar a definição de vizinhança da seção anterior, com a abordagem introduzida por Besag (1975) e escrever a função de pseudo-verossimilhança:

$$\tilde{L}(\varphi_j|D) = \prod_{i=1}^n \prod_{j=1}^s P(X_j = x_j^i | X_k = x_k^i, \forall k \in V_j^{l_j r_j}). \quad (1)$$

Assim, assume-se que a distribuição condicional da posição  $j$  é especificada em termos de um vetor  $\varphi_j$  dos parâmetros  $(l_j, r_j)$  desconhecidos, sendo  $l_j$  a vizinhança à esquerda e  $r_j$  a vizinhança à direita da posição  $j$ . Como a amostra é composta por várias realizações independentes, correspondentes aos diferentes indivíduos na amostra, o processo é denominado não homogêneo, com tamanhos diferentes de vizinhança, dependendo da posição específica  $j$ . Aplicando então o logaritmo à expressão (1), teremos:

$$\ell(l_j, r_j|D) = \sum_{j=1}^s \sum_{i=1}^n \log \left( P(X_j = x_j^i | X_k = x_k^i, \forall k \in V_j^{l_j r_j}) \right). \quad (2)$$

Reescrevendo a probabilidade em termos de prefixos e sufixos, temos:

$$P(X_j = a | X_k = x_k^i, \forall k \in V_j^{l_j r_j}) =$$

$$P(X_j = a | X_{j-l} = \omega_1, \dots, X_{j-1} = \omega_l, X_{j+1} = \tau_1, \dots, X_{j+r} = \tau_r),$$

a qual denotaremos simplesmente por  $P_j(a|\omega, \tau)$ . Considerando  $N_j^D(\omega, a, \tau)$  como sendo o número de vezes que o símbolo  $a$  aparece na posição  $j$ , com a vizinhança  $\omega$  à esquerda e  $\tau$  à direita, dentro da amostra  $D$ , pode-se reescrever a equação (2) como:

$$\ell(l_j, r_j|D) = \sum_{j=1}^s \sum_{\omega} \sum_{\tau} \sum_a \left( \log(P_j(a|\omega, \tau))^{N_j^D(\omega, a, \tau)} \right). \quad (3)$$

Maximizar a verossimilhança em (3), representa maximizar cada parcela da soma em  $j$ , isto é, considerando cada SNP individualmente. A  $j$ -ésima componente da pseudo-verossimilhança em função da vizinhança  $(\omega, \tau)$  pode ser definida em termos da distribuição multinomial. Assim, pode-se escrever (3) na forma:

$$\ell(l_j, r_j|D) = \sum_{j=1}^s \sum_{\omega} \sum_{\tau} \sum_a \left( N_j^D(\omega, a, \tau) \log \left( \frac{N_j^D(\omega, a, \tau)}{N_j^D(\omega, \cdot, \tau)} \right) \right), \quad (4)$$

em que  $N_j^D(\omega, \cdot, \tau)$  corresponde ao número de vezes que a vizinhança  $(\omega, \tau)$  aparece na amostra no entorno da posição  $j$ , para qualquer valor de  $X_j$ . Da forma como está escrita na equação (4), a pseudo-verossimilhança sempre aumenta com o aumento da vizinhança, o que levaria a considerar todos os outros SNPs para maximizar a função de uma determinada posição. É

necessário, portanto, introduzir um termo de penalização proporcional ao tamanho da vizinhança. A penalização utilizada é a proposta por Csiszar e Talata (2006), considerando uma constante  $c$  dada por  $\frac{(|A|-1)}{2}$ , onde  $A$  é o alfabeto, no nosso caso,  $A = \{0,1,2\}$  e um termo proporcional ao tamanho da vizinhança  $t(\omega, \tau) = |A|^{|\omega\tau|}$ . Finalmente, considerando a estratificação da amostra nas diferentes populações, a expressão final da pseudo-verossimilhança será:

$$\ell_j^g(l_j^g, r_j^g | D^g) = \sum_{\omega} \sum_{\tau} \sum_a \left( N_j^{D^g}(\omega, a, \tau) \log \left( \frac{N_j^{D^g}(\omega, a, \tau)}{N_j^{D^g}(\omega, \cdot, \tau)} \right) \right) - \frac{(|A|-1)}{2} |A|^{|\omega\tau|} \log(n_g). \quad (5)$$

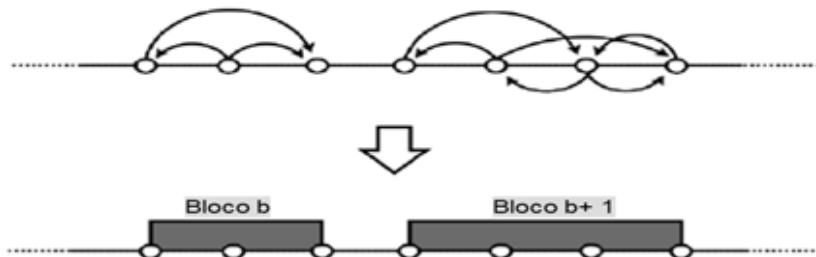
### 2.4 Janelas Populacionais (Jp) e de Consenso (Jc)

A vizinhança  $V_j$  estimada é uma medida da dependência do SNP na posição  $j$  em relação aos seus vizinhos à esquerda e à direita. Os parâmetros  $l_j$  e  $r_j$  estimados para cada variável  $X_j$  definem a janela de dependência do SNP na posição  $j$ . Analogamente, os pares  $(l_{j-1}, r_{j-1})$  e  $(l_{j+1}, r_{j+1})$  definem as janelas de dependências dos SNPs vizinhos, nas posições  $j - 1$  e  $j + 1$ , respectivamente. Avaliando a sobreposição das janelas de dependência dos SNPs adjacentes, pode-se definir uma janela de influência populacional  $J_p$  como sendo a sequência de SNPs cujas vizinhanças (janelas de dependência) estão contidas em  $J_p$ . Assim, cada janela populacional  $J_p$  é um par de valores tais que:

$$J_p \in \{(a, b) | a, b \in \{1, \dots, s\}, a < b\} \text{ e } \forall j \in J_p \Rightarrow V_j \subseteq J_p$$

A Figura 1 ilustra a montagem das janelas populacionais; na metade superior estão representados os SNPs adjacentes na forma de pequenos círculos e suas vizinhanças estimadas na forma de arcos; na parte inferior da figura estão representadas as janelas populacionais no formato de blocos, formando uma região de SNPs adjacentes que possuem suas vizinhanças sobrepostas.

**Figura 1:** Visão esquemática da construção de janelas populacionais de SNPs



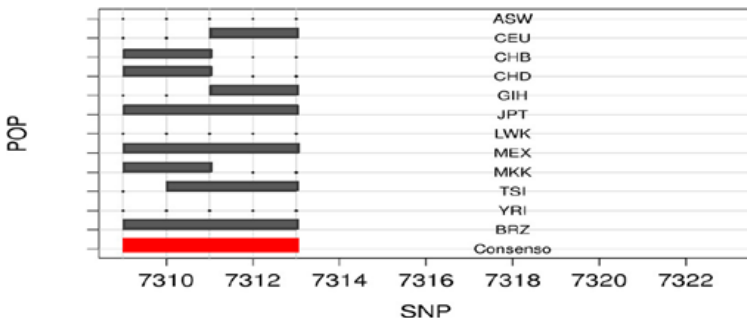
Unificando as estimativas das várias populações, propomos a formação de uma nova janela, a qual definiremos como janela de consenso, denotada por  $J_c$ , generalizando o mesmo conceito utilizado para encontrar as janelas de influência dos SNPs, desta vez, sobrepondo as janelas de influência de cada população, em cada região específica do genoma. Cada janela de influência estimada em uma população representa um conjunto de SNPs que guardam entre si uma relação de dependência. Para cada intervalo de SNPs adjacentes é possível verificar se existem ou não janelas de influência parcial ou totalmente inseridas nesse intervalo, em cada uma das populações em estudo. Assim, pode-se definir a janela de consenso  $J_c$  como sendo uma sequência de SNPs na qual existem janelas populacionais ( $J_p$ ) que estão contidas em  $J_c$ . Portanto, cada janela de consenso  $J_c$  é um par de valores tais que:

$$J_c \in \{(c, d) | c, d \in \{1, \dots, s\}, c < d\} \text{ e } \forall J_p \in J_c \Rightarrow J_p \subseteq J_c$$

mais especificamente, se houverem  $K$  janelas de influência contidas em uma janela de consenso  $J_c$ , teremos:  $J_p = (a_p, b_p), 1 \leq p \leq K$ , se  $J_p \subseteq J_c \Rightarrow (c \leq a_p \text{ e } d \geq b_p), \forall p, 1 \leq p \leq K, 1 \leq K \leq G$ , com  $G$  sendo o número de populações sob estudo.

A Figura 2 mostra esquematicamente a montagem de uma janela de consenso maximal, em uma região específica do genoma, neste caso, um intervalo de 5 SNPs: cada linha representa um estrato, isto é, uma população, os pontos são SNPs independentes, os blocos pretos as janelas populacionais ( $J_p$ ) e a janela de consenso ( $J_c$ ), está evidenciada na cor vermelha.

**Figura 2:** Visão esquemática da construção de uma janela de consenso maximal



### 2.5 Índice de Heterogeneidade

A janela de consenso,  $J_c$ , pode servir, portanto, como uma medida da heterogeneidade das regiões de dependência entre as populações, uma vez que pode ser resultado de regiões do genoma bastante diferentes em termos das janelas de influência de cada população ( $J_p$ ) ou, ao contrário,

de segmentos do genoma nos quais as  $J_p$  são muito parecidas ou até mesmo iguais entre as populações. Para avaliar a heterogeneidade de uma determinada janela de consenso é preciso quantificar a distância entre as janelas  $J_p$  formadoras da mesma. Para isso utilizou-se o conceito de medida de distância/similaridade entre variáveis categóricas (GREENACRE, 2007).

Para cada janela de consenso  $J_c$ , o trecho respectivo foi codificado, para cada população, em termos binários, considerando 1 para os intervalos entre SNPs que pertencem a uma mesma  $J_p$ , ou 0 caso contrário. O cálculo do Índice de Heterogeneidade (IH) foi feito através de uma matriz de distância entre as populações, somando-se os elementos da triangular superior, ou seja, a soma de todas as dissimilaridades e normalizando pelo número de comparações pareadas entre as populações. O IH calculado para uma janela de consenso  $l$  qualquer ( $J_{c_l}$ ), em  $G$  populações será,

$$IH_{J_{c_l}} = \sum_{g,g'=1;g<g'}^G \frac{d_{gg'}}{m_l} \text{ sendo, } m_l = \frac{G(G-1)}{2}.$$

## 2.6 Classificador - Mapa Biplot

Se janelas heterogêneas acabam por refletir áreas de dependência entre os SNPs que são discriminantes das diferentes populações, estas podem, então, ser utilizadas para identificar, a partir dos genótipos dos SNPs correspondentes, a população originária de um determinado indivíduo, ou ainda, a proporcionalidade de contribuições ancestrais de cada população, configurando assim um classificador. Para cada janela heterogênea existem diferentes genótipos com frequências que variam entre as populações do estudo. Uma tabela de contingência construída a partir das populações (linhas) e dos diferentes genótipos (colunas) pode sintetizar essas informações. O problema assim colocado passa a ser multivariado no sentido de relacionar as linhas e colunas da tabela com a distribuição de genótipos entre as populações sob estudo. Uma técnica capaz de lidar com esse tipo de abordagem em tabelas de contingência é a Análise de Correspondência (AC). Dentre os diversos mapas de representação da AC, optou-se pelo mapa Biplot, por ser uma alternativa mais acurada para juntar linhas e colunas numa mesma representação (GREENACRE, 2007). As colunas (genótipos) são representadas como vetores de resposta, sendo de interesse a direção desses vetores, pois é nela que os perfis linha (populações trinomiais) serão projetados. Neste caso, a representação Biplot equivale à representação Simplex de populações trinomiais.

Em cada janela de consenso, dado o genótipo de um determinado indivíduo, calcula-se a distância de cada população ao vetor correspondente no mapa biplot desse genótipo. O indivíduo é, então, para essa janela, classificado na população mais próxima do vetor. Finalmente, depois de classificar o indivíduo em todas as janelas mais heterogêneas, a classificação final, segundo um critério adotado, será a população na qual ele foi classificado

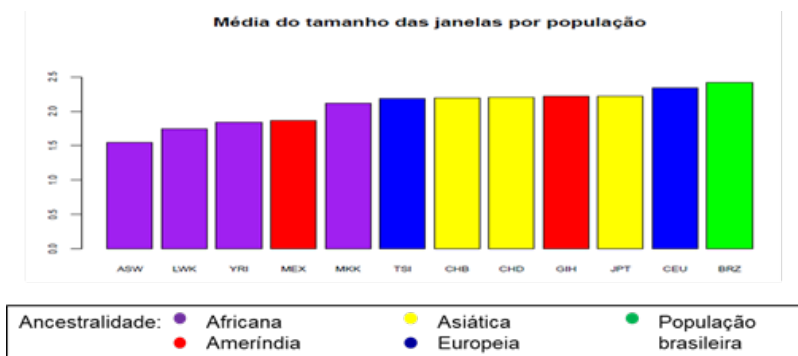
mais vezes.

### 3. RESULTADOS

Este trabalho fez uso dos dados utilizados por Giolo *et al.* (2012), resultado da fusão dos arquivos do projeto HapMap (disponível em <http://www.sanger.ac.uk/humgen/hapmap3>) e de uma amostra da população brasileira constituída por 138 indivíduos não relacionados residentes na cidade de São Paulo (Quadro 1). A combinação dos dados de todas as 11 populações HapMap e da Brasileira resultou em 365.116 marcadores SNPs genotipados em 1.129 indivíduos (138 brasileiros e 991 das demais populações). Considerando os dados genômicos (SNPs) codificados de cada população, foram estimadas as janelas de influência ( $J_p$ ) por meio da aplicação da metodologia descrita na Seção 2.4.

Quando analisamos o tamanho médio dessas janelas, definido, para cada população, como a média aritmética do número de SNPs presentes nas janelas encontradas, as populações africanas são aquelas que possuem as menores médias, enquanto que a população brasileira apresenta o maior tamanho médio entre todas as populações estudadas (ver Figura 3). As janelas das populações africanas, mais antigas na história da humanidade, são menores, de acordo, portanto, com o que é esperado, uma vez que, devido ao maior número de gerações há uma redução progressiva no tamanho dos blocos do genoma que são transmitidos aos descendentes, consequência das recombinações gênicas ocorridas. A população brasileira, ao contrário, possui janelas maiores em média, representativas de blocos maiores do genoma ainda presentes (conservados), evidenciando sua história recente de miscigenação.

**Figura 3:** Tamanho médio das janelas por população

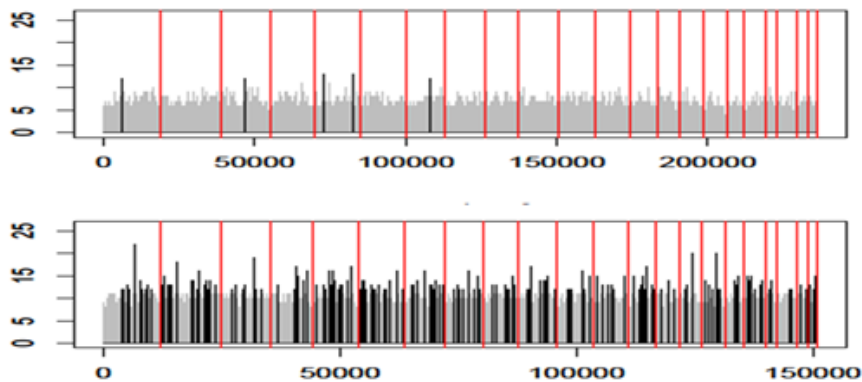


As populações envolvidas neste estudo acabaram por se diferenciar não apenas no tamanho médio, mas também no número de janelas com

tamanhos considerados 'grandes'. Neste sentido, foi adotado o tamanho 12 como ponto de corte entre janelas 'grandes' e janelas 'pequenas' e analisada a distribuição dessas janelas ao longo do genoma, em cada população. Nota-se também que o número dessas janelas (maiores ou iguais a 12 SNPs) é superior nas populações mais recentes, como a brasileira.

A Figura 4 ilustra bem essa afirmação, comparando a ocorrência de janelas 'grandes' nas populações ASW e BRZ, as populações com menor e maior tamanho médio de janelas, respectivamente. Nesta figura, todas as janelas encontradas são representadas por linhas verticais sendo, na cor cinza as janelas 'pequenas' e na cor preta, em destaque, as janelas 'grandes'. As linhas verticais que tocam o topo da escala, são na verdade, representações dos limites de cada cromossomo. Reflete-se bem, dessa forma, a diferença no tamanho dos blocos do genoma como consequência, muito provavelmente, do número de cruzamentos aos quais cada população foi exposta ao longo dos tempos. Esse mesmo comportamento foi observado para as demais populações, ou seja, as mais antigas possuem poucas ocorrências de janelas 'grandes' e, à medida que caminhamos para populações mais recentes, o número dessas janelas aumenta.

**Figura 4:** Distribuição de janelas 'grandes' nas populações ASW e BRZ

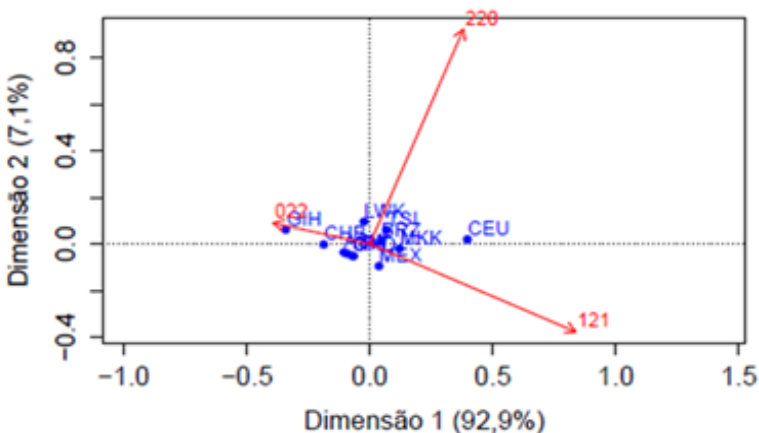


De posse das janelas de influência de cada população do estudo, dadas suas respectivas posições em termos do número relativo do SNP e do número do cromossomo, foram determinadas as janelas de consenso maximais (denominadas  $J_c$ ), conforme discutido e proposto na Seção 2.4. Quanto mais heterogênea uma determinada janela de consenso, mais discriminante ela deve ser para as populações. Adotou-se como ponto de corte para seleção das janelas o índice de heterogeneidade igual ou maior a 0,50, equivalente ao percentil 98% (considere que, independentemente do cromossomo, o terceiro quartil do IH foi igual a 0,34), obtendo assim 1.646 janelas, distribuídas ao longo dos 22 cromossomos, as quais foram utilizadas

no classificador.

A Figura 5 mostra o mapa biplot obtido para a janela de consenso 48 presente no cromossomo 13, a qual tem apenas 3 genótipos distintos, o que permite que o mapa biplot possa ser representado em duas dimensões, sem perda de informação. A análise dessa figura permite concluir que o genótipo '121' e o genótipo '220' estão mais associados com a população 'CEU', enquanto que o genótipo '022' associa-se com a população 'GIH'. Foi retirada uma amostra aleatória de tamanho 100 do conjunto total de 1.124 indivíduos. Para cada uma das 1.646 janelas heterogêneas selecionadas, foi atribuída uma população originária à cada indivíduo, de acordo com o classificador proposto. Finalmente, depois de todas as janelas analisadas, o indivíduo é classificado na população escolhida no maior número de janelas.

**Figura 5:** Biplot produzido pela ACS na janela 48 no cromossomo 13



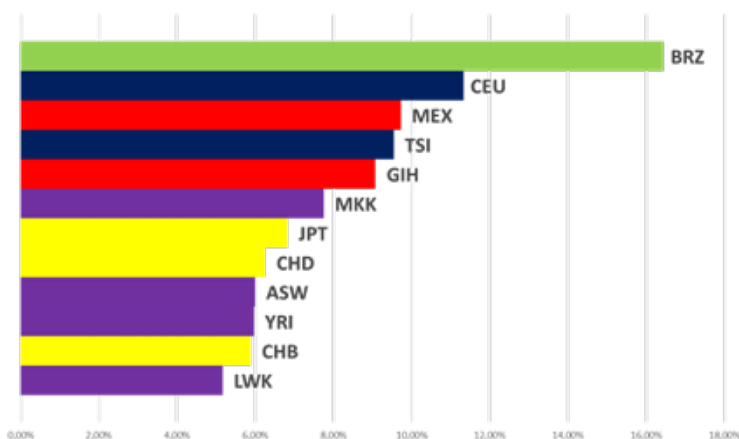
Para avaliar a eficiência do classificador, foram executadas 10 realizações do procedimento, utilizando diferentes amostras de 100 indivíduos e, em todos os casos, o acerto foi sempre superior a 90% da população originária. Adicionalmente, o mesmo classificador foi utilizado, porém, usando um conjunto de 1.646 janelas consideradas homogêneas (neste caso, adotou-se  $IH \leq 0,05$  como ponto de corte para janelas homogêneas), obtendo resultados finais sempre abaixo dos índices atingidos quando utilizadas as janelas heterogêneas, evidenciando o alto poder discriminante das últimas. Os resultados detalhados de cada uma das realizações não estão mostrados, contudo, a Tabela 1 retrata a comparação da taxa de acerto do classificador para cada uma das realizações.

**Tabela 1:** Tabela resumo dos resultados obtidos para o classificador (em %)

Realização	1	2	3	4	5	6	7	8	9	10
Jc Heterogêneas	96	97	96	96	95	96	95	96	91	92
Jc Homogêneas	75	66	72	79	67	82	80	71	64	63

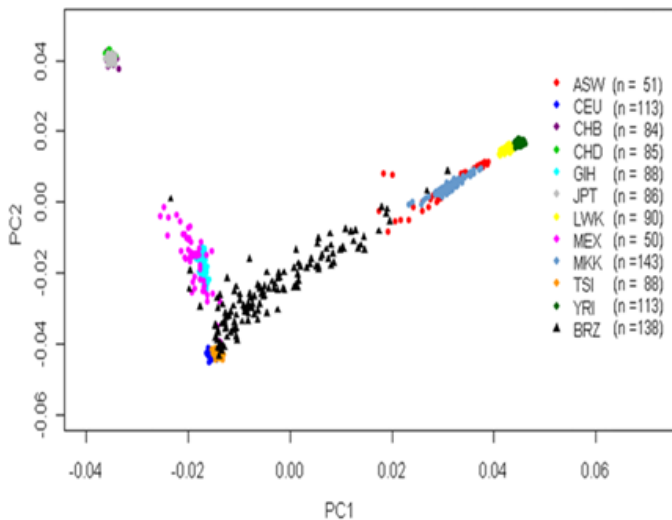
Adicionalmente, utilizando-se o classificador aqui proposto, pode-se analisar a proporção de janelas classificadas em cada população, para uma determinada amostra de indivíduos. Em particular, nos interessa compreender como essa proporção acontece na população brasileira. A Figura 6 ilustra o resultado dessa análise. Claramente o maior número de janelas são representativas da própria população brasileira, como era de se esperar. Na sequência, destacam-se como as 5 populações mais encontradas, CEU (européia), MEX(ameríndia), TSI(européia), GIH(ameríndia) e MKK(africana).

**Figura 6:** Distribuição de janelas 'grandes' nas populações ASW e BRZ



Esse resultado está em concordância com o obtido por Giolo *et al.* (2010) através do estudo de componentes principais dessas mesmas populações. A Figura 7 mostra o gráfico das duas primeiras componentes obtidas no estudo. Nota-se claramente que a população brasileira (em preto), situa-se entre as populações GIH, MEX, TSI e CEU de um lado e MKK do outro, evidenciando a forte ancestralidade européia, ameríndia e africana que compõe a miscigenação da população brasileira.



**Figura 7:** Distribuição de janelas 'grandes' nas populações ASW e BRZ

Fonte: Gráfico reproduzido a partir dos dados de Giolo *et al.* (2010)

#### 4. CONCLUSÃO

Este trabalho teve foco em identificar blocos do genoma a partir de marcadores genéticos conhecidos como SNPs (variações em um único nucleotídeo) amostrados em diversas populações mundiais, incluindo a brasileira, com ancestralidades distintas. Tendo em vista que variabilidade genotípica deve ter percorrido caminhos distintos em cada população de estudo, os blocos conservados do genoma devem ser da mesma forma diferentes. Estes blocos, sob o ponto de vista estatístico são, então, regiões de dependência dentro do genoma, transmitidas e modificadas ao longo do tempo.

A análise descritiva das janelas de influência para cada população mostrou que na população brasileira o tamanho médio dessas janelas é maior do que nas outras populações, de acordo, portanto, com a história recente de miscigenação desta população. No outro extremo, as janelas de influência obtidas para as populações de origem mais remota, como as africanas, possuem tamanho médio menor. Não só o tamanho médio, mas também a frequência das janelas de influência maiores, é relativamente diferente entre as populações e indicativa da idade evolutiva de cada uma em termos do número de gerações que as distanciam de um marco ancestral.

Estendendo o conceito de janelas de influência populacional ( $J_p$ ) para janelas de consenso ( $J_c$ ) foi possível caracterizá-las em termos de sua heterogeneidade, sendo as mais heterogêneas, utilizadas em um classificador de indivíduos à sua população de origem, o qual obteve um acerto da população originária em torno de 90%, indicando que essas são

regiões com dependência dentro do genoma e com grande diversidade entre populações, podendo revelar as diferenças no que diz respeito ao registro genético de suas ancestralidades.

O conjunto de resultados obtidos neste trabalho, mostram que a obtenção de blocos dependentes no genoma através do processo estocástico utilizando campos Markovianos é bastante promissor e merece aprofundamento.

## AGRADECIMENTOS

Os autores agradecem à doutora Suely Giolo pelo suporte na elaboração do arquivo de dados, matéria prima deste trabalho, à doutora Florencia Leonardi por compartilhar a primeira versão do algoritmo de construção das janelas de influência, bem como ao Laboratório de Cardiologia e Genética Molecular (InCor/USP) por disponibilizar os dados do Projeto Trios de São Paulo.

## REFERÊNCIAS

BESAG, J. **Statistical analysis of non-lattice data**, The statistician, 1975. 179-195p .

CSISZÁR, I.; TALATA, Z. **Context tree estimation for not necessarily finite memory processes, via BIC and MDL**, Information Theory, IEEE Transactions on, 52/3, 1007-1016, 2006 .

GIOLO, S. R.; SOLER, J. M. P. *et al.* **Brazilian urban population genetic structure reveals a high degree of admixture** European Journal of Human Genetics, 1, 111–116, 2012

GIOLO, S. R. *et al.* **Brazilian genetic population structure along with other worldwide human populations**. European J of Human Genetics, 2010.

GREENACRE, M. **Correspondence Analysis in Practice**, 2007.

LEONARDI, F. G. **A generalization of the PST algorithm: modeling the sparse nature of protein sequences** Bioinformatics, 22/11, 1302-1307, 2006.

LEONARDI, F. G. *et al.* **Some upper bounds for the rate of convergence of penalized likelihood context tree estimators**. Brazilian Journal of Probability and Statistics, 24(2):321-336, 2010.

SCHLESINGER, D. **Ancestralidade da população de São Paulo e correlação com alterações neuropatológicas no idoso**. Tese de Doutorado - Universidade de São Paulo, 2010.

THE INTERNATIONAL HAPMAP CONSORTIUM. **A haplotype map of the human genome** Nature, 437/27, 1299-1320, 2005.

## CHARACTERIZATION OF THE HUMAN GENOME DEPENDENCE STRUCTURE USING MARKOV RANDOM FIELDS POPULATION WORLDWIDE STUDY AND SNP DATA

**ABSTRACT** - Molecular markers such as SNP (Single Nucleotide Polymorphisms), with high density inside the human DNA, have been widely used as samples of our genome. This work makes use of stochastic modeling of Markov random fields, to find SNPs blocks, independent of each other, thus structuring the genome in stranded regions of dependency. Public data from different worldwide populations were used (HapMap project), beyond a Brazilian population. The SNPs blocks configure dependence regions or windows of influence which were used to characterize the different populations according to their ancestry and the results showed that the Brazilian population's windows have, on average, a bigger size, showing their recent history of admixture, compared to other world populations. It is also proposed the maximal windows of consensus, based on the windows of influence of each population within the study, in a specific genomic region. By identifying the most heterogeneous windows of consensus, the Correspondence Analysis and Biplot theory were applied to build a classifier able to determine the percentual of the ancestry of a given individual, which, submitted for validation, achieved high efficiency in hitting the original population.

**KEYWORDS:** Variable range Markov random fields; HapMap; Biplot; SNP blocks; Ancestry classifier.

# ANÁLISE COMBINATÓRIA: UMA ABORDAGEM PARA O SEXTO ANO DO ENSINO FUNDAMENTAL

**Juvino Pereira Dos Santos Junior**  
[lattes.cnpq.br/4960715718766728](http://lattes.cnpq.br/4960715718766728)

Erem José Do Patrocínio Mota  
São Bento Do Una – Pernambuco

**RESUMO** - Ao observar as dificuldades que existem no ensino-aprendizagem de matemática, propomos neste trabalho uma abordagem à Análise Combinatória para o 6° (sexto) ano do ensino fundamental de forma intuitiva, indutiva e, em um primeiro momento, sem formalidades, favorecendo o desenvolvimento cognitivo do aluno. Percebemos ao longo do nosso trabalho docente que a aprendizagem de Análise Combinatória se mostra como um obstáculo aos alunos devido à forma como esse conteúdo é abordado, na quase totalidade das vezes através de fórmulas matemáticas, o que faz perder o sentido da resolução de problemas. Nossa proposta é utilizar a resolução de problemas e as técnicas do princípio fundamental de contagem, que normalmente é citado apenas no início do estudo de Análise Combinatória no segundo ano do Ensino Médio para em seguida enunciar as fórmulas de arranjos, combinações e permutações, para

o desenvolvimento do raciocínio combinatório em alunos do sexto ano do ensino fundamental, melhorando desta maneira, o ensino e a compreensão de matemática. Tivemos como principais fontes de pesquisa, livros, os Parâmetros Curriculares Nacionais, artigos em revistas, dissertações e teses que tratam do tema.

**PALAVRAS-CHAVE:** Análise combinatória; Raciocínio combinatório; Resolução de problemas.

### 1. INTRODUÇÃO

É tendo em vista que o papel do professor de matemática é fazer com que a aprendizagem dessa disciplina se torne algo fundamental para a formação da cidadania, de forma a preparar o sujeito para enfrentar os problemas do seu cotidiano, que apresentamos este trabalho sobre Análise Combinatória. Tema este, trabalhado por nós professores, de forma geral, no segundo ano do ensino médio, mas que aqui o abordaremos no sexto ano do ensino fundamental.

Nos dias atuais, podemos perceber que nossa sociedade requer pessoas que tenham a capacidade de analisar e propor

soluções para problemas de forma prática e rápida. Para tal, seria necessário que nossas instituições de ensino tivessem uma matriz curricular voltada para a resolução de problemas. Contudo, não é isso que observamos ao nos depararmos com uma educação voltada para a memorização de fórmulas e o ensino de algoritmos, sequências de instruções bem definidas que podem ser executadas mecanicamente. Identificamos nas escolas uma educação matemática voltada para a resolução de questões sem aplicabilidade e sem significado, o que impossibilita ao aluno um real aprendizado. Ao não identificar a contextualização dos conteúdos e das respectivas questões em seu cotidiano, o aluno simplesmente “aprende” para ser aprovado no final do ano letivo, isso quando consegue assimilar os conteúdos estudados dessa forma.

Segundo CARRAHER (1986):

“Os problemas de Matemática, dos quais o aluno tem que utilizar precisamente as fórmulas que acabou de estudar, não são verdadeiros problemas que exijam reflexão, mas sim, exercícios que exigem apenas memória; não lhe é exigida compreensão dos conceitos matemáticos, nem que faça relações entre o que já aprendeu e a possível solução do problema. Nestes casos, os problemas são tratados mecanicamente, sem que, muitas vezes, o aluno compreenda o que está fazendo. Esta abordagem não exige estímulo do raciocínio do aluno”.

É concordando com CARRAHER em relação ao uso de fórmulas como ponto de partida na resolução de problemas de Análise Combinatória que STURM (1999) expõe também em sua pesquisa uma abordagem alternativa, na qual as fórmulas apresentam-se como decorrência da experiência do aluno, inseridas em um processo de resolução de problemas de contagem.

Portanto, segundo STURM (1999):

“[...] o ensino de Análise Combinatória deve se dar através de situações-problemas. As fórmulas devem aparecer em decorrência das experiências dos alunos na resolução de problemas, devem ser construídas e não ser o elemento de partida para o ensino de cada tema: Arranjo, Permutação e Combinação”.

A dicotomia teoria/prática, ensino/aprendizado, muitas das vezes se dá devido à falta de preparo e atualização profissional da equipe docente e pedagógica nas escolas, onde se faz necessária a busca de um novo paradigma para que, dessa forma, seja substituído o debilitado processo ensino-aprendizagem baseado numa educação tradicional que, apesar de

ter funcionado por um bom tempo, não se aplica nos dias atuais.

Acreditamos que a Análise Combinatória seja uma das mais importantes ferramentas de resolução de problemas, parte importante do estudo das Probabilidades e que desenvolve o raciocínio lógico-matemático de forma plena e eficaz, fazendo com que o aluno consiga desenvolver diversas outras capacidades de resolução de problemas.

Portanto, é ciente das dificuldades encontradas não somente pelos alunos, mas também por muitos colegas professores, que pretendemos mostrar uma abordagem diferenciada para o ensino da Análise Combinatória, tirando do foco a metodologia da fórmula-aplicação e incentivando o uso das técnicas de contagem para a resolução de problemas. Sabemos que não existe um caminho único e correto para o ensino desse conteúdo, mas é necessário aprimorar técnicas pedagógicas para que sejam desencadeadas mudanças notáveis no modo como o aluno aprende, para que seja percebida a real construção do conhecimento feito pelo próprio discente, aprendendo a matemática de uma forma contextualizada e significativa, desenvolvendo um raciocínio dedutivo e lógico que será útil para a resolução dos mais diversos problemas que podem surgir ao longo da vida.

Entendemos que os Problemas de Contagem (conceitos básicos de Análise Combinatória) tratam de estabelecer procedimentos e estratégias próprios de raciocínio que envolve ideias da operação de multiplicação de números naturais – o raciocínio combinatório – para a contagem de casos presentes em situações-problema de agrupamentos de objetos que contenham determinado número de elementos.

Nos PCN's (1997, 1998, 1999) foi criado um Bloco de Conteúdos denominado Tratamento da Informação, que contempla-os dando ênfase nos problemas de contagem – envolvendo o Princípio Aditivo e o Princípio Multiplicativo – e neles estão prescritas claras recomendações quanto ao desenvolvimento de atividades que contemplem tais conteúdos, como se pode constatar na citação a seguir:

Um olhar mais atento para nossa sociedade mostra a necessidades de acrescentar a esses conteúdos aqueles que permitam ao cidadão “tratar” as informações que recebe cotidianamente, aprendendo a lidar com dados estatísticos, tabelas, gráficos, a raciocinar utilizando ideias relativas à probabilidade e à combinatória (BRASIL, 1997, p.53).

Já nos anos iniciais do Ensino Fundamental os problemas de contagem – parte integrante da combinatória – deveriam ser mais desenvolvidos com os alunos, valorizando o raciocínio combinatório.

Para tal, aplicamos algumas atividades em sala de aula que tinham como objetivos, por um lado, proporcionar condições aos alunos para apropriarem-se de conhecimentos, procedimentos, estratégias e habilidades relacionados às noções básicas de Combinatória com a construção e a compreensão dos dados presentes em uma representação gráfica, que

permitem a contagem de possibilidades para a obtenção da solução a um dado problema de contagem proposto, e, por outro lado, sem a necessidade de explorar aspectos formais e algorítmicos no encaminhamento da solução. Nessas atividades foi utilizada uma linguagem clara nos enunciados de tal maneira que diminuíssem as dificuldades apresentadas por parte dos alunos para compreender o que estava sendo solicitado na resolução de cada situação-problema.

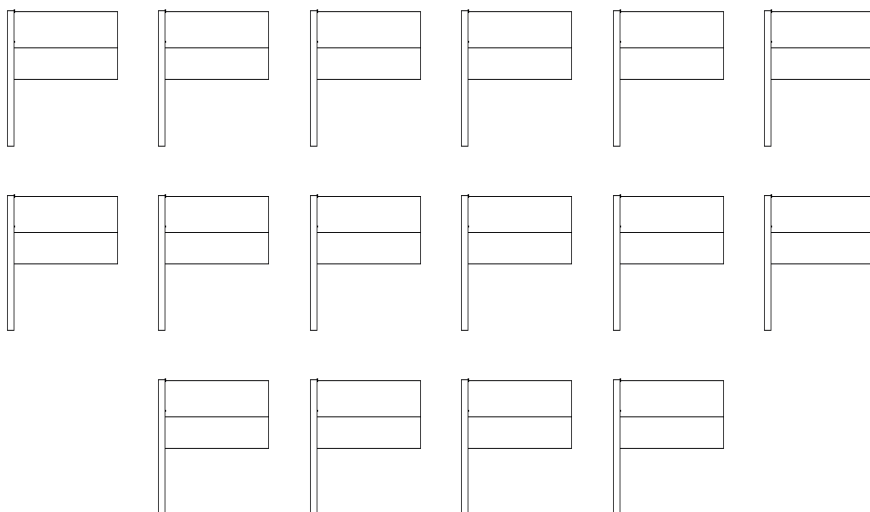
Por fim, a ideia deste trabalho tem como objetivo mostrar como o Princípio Fundamental da contagem e as técnicas de resolução de problemas, quando trabalhados de forma planejada e correta, atingem resultados surpreendentes relacionados aos aspectos cognitivos dos alunos nessa área da matemática.

## 2. PROPOSTAS DE ATIVIDADES A SEREM REALIZADAS

### 2.1. Atividade I

São dadas 16 bandeiras com mastros, contendo duas listras horizontais cada uma. Dispondo de três bastões de giz de cera nas cores azul, vermelho e verde, faça o que se pede:

- Apresente todas as diferentes maneiras de se colorir as listras horizontais de cada bandeira, de modo que se obtenha o maior número de bandeiras pintadas diferentemente.

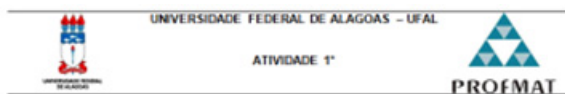


Responda o que se pede a seguir:

- Todas as bandeiras que estão desenhadas acima foram

pintadas? Quantas dessas bandeiras não foram pintadas?

- c) Quantas bandeiras foram pintadas utilizando cores diferentes?
- d) Acrescentando um quarto bastão de giz de cera numa cor diferente das apresentadas no enunciado, apresente todas as possibilidades que identificam as diferentes pinturas de bandeiras entre si, considerando que nenhuma listra poderá deixar de ser pintada em cada bandeira considerada. Pinte as diferentes bandeiras.
- e) Quantas possibilidades diferentes de pinturas das bandeiras você obteve quando dispõe de quatro cores de giz de cera disponíveis para a pintura, se nenhuma listra pode ficar sem ser pintada?
- f) Quantas possibilidades diferentes de pinturas das bandeiras você obteve quando dispõe de quatro cores de giz de cera disponíveis para a pintura, se nenhuma listra pode ficar sem ser pintada e as listras pintadas devem ter cores distintas (diferentes)?
- g) Se você não tivesse as bandeiras desenhadas e também não quisesse ter o trabalho de desenhá-las, escreva como você procederia para mostrar para um amigo quantas e quais são as bandeiras com duas listras que podem ser pintadas de maneiras diferentes entre si, quando você dispõe de giz de cera nas quatro cores indicadas para fazer as pinturas?



**TABELA DE DUPLA ENTRADA**

Cor da 2ª listra Cor da 1ª listra	Verde	Vermelho	Azul	Amarelo
Verde	( _____ )	( _____ )	( _____ )	( _____ )
Vermelho	( _____ )	( _____ )	( _____ )	( _____ )
Azul	( _____ )	( _____ )	( _____ )	( _____ )
Amarelo	( _____ )	( _____ )	( _____ )	( _____ )



UNIVERSIDADE FEDERAL DE ALAGOAS – UFAL

ATIVIDADE 1

PROJMAT

PRODUTO CARTESIANO

COR DA 1ª LISTRA

COR DA 2ª LISTRA

$$\begin{matrix} \text{AXB} = \text{E} \\ \left\{ \begin{matrix} \text{V} \\ \text{A} \\ \text{V} \\ \text{A} \end{matrix} \right\} \times \left\{ \begin{matrix} \text{V} \\ \text{A} \\ \text{V} \\ \text{A} \end{matrix} \right\} = \left\{ \begin{matrix} \text{VV} \\ \text{VA} \\ \text{VAV} \\ \text{VAA} \\ \text{AV} \\ \text{AA} \\ \text{AVV} \\ \text{AVA} \\ \text{VV} \\ \text{VA} \\ \text{VAV} \\ \text{VAA} \\ \text{AV} \\ \text{AA} \\ \text{AVV} \\ \text{AVA} \end{matrix} \right\} \end{matrix}$$

Para a atividade 1 procuramos explorar situações que envolvessem o conceito de Princípio Multiplicativo e a ideia de arranjos simples.

Propomos a pintura de algumas bandeiras, sabendo que cada bandeira é composta de duas listras. Essa primeira situação-problema foi elaborada com o propósito de inicializar o aluno com as primeiras ideias relacionadas com o raciocínio combinatório enquanto compreende os procedimentos necessários para a construção de diversas representações gráficas, as quais mostram todas as possibilidades que atendem à solução (agrupamentos) para um mesmo problema de contagem. As maneiras como as cores para a pintura das listras de uma bandeira podem ser escolhidas e combinadas determina, igualmente, a totalidade de diferentes bandeiras que podem ser pintadas, nas quais as “combinações de escolhas cores para a pintura das listras de bandeira” favorecem o desenvolvimento do raciocínio combinatório enquanto são obtidas todas as diferentes possibilidades em que as pinturas de uma bandeira podem ser feitas.

Para esta atividade temos como objetivos: Apresentar diferentes representações gráficas que podem ser utilizadas para resolver problemas de contagem, de modo que o aluno conheça as características que permitem a construção de cada uma delas; Oferecer procedimentos que favoreça ao aluno identificar os diferentes agrupamentos de objetos envolvidos em cada situação-problema; Criar condições para que o aluno enumere todos os agrupamentos de objetos envolvidos na situação-problema, após ter construído qualquer uma representação e, a partir daí, efetue a contagem direta desses agrupamentos para concretizar a resolução do problema de contagem proposto; Permitir que quando da construção de uma árvore de possibilidades, por exemplo, o aluno identifique em todas as suas “folhas”, todos os diferentes agrupamentos de objetos e as características que os

diferem entre si, certificando-se de que estabeleceu a enumeração de todas as possibilidades.

A resolução da primeira ficha de atividades pode ser realizada de quatro maneiras: Enumeração das possibilidades (de seus agrupamentos), Tabela de dupla entrada, Produto Cartesiano e Árvore de Possibilidades.

Na primeira maneira de resolver a atividade proposta utilizamos a Enumeração das possibilidades. Quando queremos determinar a quantidade total de possibilidades diferentes de se fazer algo dizemos que estamos diante de um problema de contagem, como é o caso da atividade apresentada. O objetivo de um Problema de Contagem é contabilizar a totalidade das diferentes possibilidades que se pode considerar para a situação-problema que está sendo proposta resolver.

Uma segunda maneira de resolver o problema proposto é utilizando uma tabela de dupla entrada. Uma tabela de dupla entrada também pode ser utilizada quando não se tem todas as bandeiras desenhadas e disponíveis para colorir e quando tratar-se de bandeiras que têm duas listras e se deseja conhecer o total de possibilidades de fazer pinturas distintas, não importando o total de cores disponíveis para se utilizar na pintura. A tabela de dupla entrada também é útil quando se deseja apresentar todas as possibilidades para colorir as diferentes bandeiras que têm duas listras e se dispõe de quatro cores para a pintura, ou então nos casos em que se dispõe de uma só cor, duas ou três cores, e é possível deixar de pintar uma listra ou as duas listras, e considerar “pinturas diferentes” entre as bandeiras. Nesse caso, deixar de pintar uma listra corresponde a usar uma segunda cor na pintura, que indicamos como “cor branca”.

Na terceira forma de resolver o problema proposto utilizaremos o Produto Cartesiano. Em um produto cartesiano fazemos a indicação das possibilidades de pintura “associando” as possibilidades de colorir as listras da bandeira, por meio de um tracejado. Para que se tenha um Produto Cartesiano é preciso que cada “elemento” que está no conjunto A (o conjunto das cores que serão utilizadas na pintura da 1ª listra) se “relacione” com todos os “elementos” do conjunto B (o conjunto das cores que serão utilizadas na pintura da 2ª listra), de maneira a identificar todas as diferentes bandeiras que podem ser pintadas.

Após fazer a associação de cada elemento do conjunto A com todos os elementos do conjunto B, conta-se o total de “pares associados” que utilizou para essas associações e tem o total de possibilidades de colorir essas diferentes bandeiras, quando dispõe de quatro cores. O mesmo se aplica quando se tem mais de quatro cores e bandeiras com duas listras.

O Produto Cartesiano entre os conjuntos A e B é indicado como  $A \times B$ . O conjunto  $A \times B$  é formado por “pares ordenados”, onde o primeiro elemento de cada par ordenado pertence ao conjunto A e o segundo elemento de cada par ordenado pertence ao conjunto B. Observemos que o total de elementos

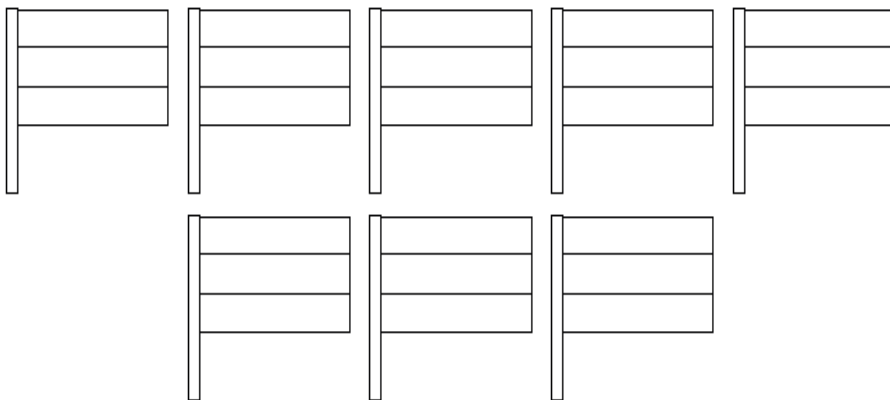
$A \times B$ , ou seja, o total de pares ordenados presentes no produto cartesiano, são todas as possibilidades diferentes que podemos pintar uma bandeira com duas listras quando dispomos de quatro cores diferentes e todas as listras devem ser pintadas.

O quarto modo de se resolver o problema apresentado é dado pela árvore das possibilidades. Ela recebe esse nome porque sua representação gráfica se assemelha ao de uma árvore em que é possível mostrar cada um de seus “galhos” e as suas “folhas”, que seriam as partes terminais da árvore. Em cada “folha” indica-se a cor que foi utilizada para a pintura da última listra de uma bandeira e, por essa razão, a construção de uma “folha” e a indicação de sua cor coincide com a escolha da cor que será usada na pintura de cada uma das bandeiras.

## 2.2. Atividade II

São dadas 8 bandeiras com mastros, contendo três listras horizontais cada uma. Dispondo de dois bastões de giz de cera nas cores azul e vermelho, faça o que se pede:

- a) Apresente todas as diferentes maneiras de se colorir as listras horizontais de cada bandeira, de modo que se obtenha o maior número de bandeiras pintadas diferentemente entre si, considerando que nenhuma listra poderá deixar de ser pintada em cada bandeira considerada.



- b) Todas as bandeiras que estão desenhadas acima foram pintadas por você?
- c) Quantas possibilidades diferentes de pinturas das bandeiras que têm três listras horizontais você obteve, quando dispõe de duas cores de giz de cera para a pintura e nenhuma listra pode ficar sem ser pintada? \_\_\_\_\_

### Tabela de Dupla Entrada

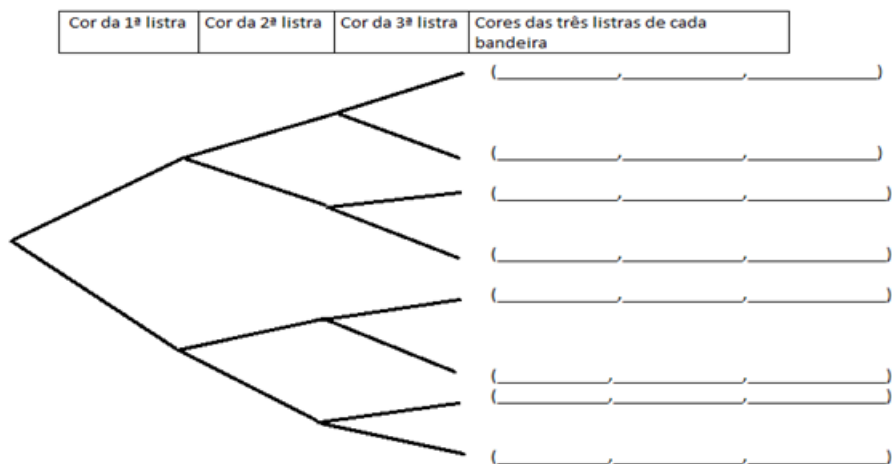
Abaixo apresentamos uma Tabela de Dupla Entrada. Você deve completar a tabela colocando nos espaços em branco a combinação das cores em que as duas primeiras listras das bandeiras poderão ser pintadas de maneiras diferentes.

	cor da 2ª listra		
cor da 1ª listra		cor azul	cor vermelho
cor azul	( _____ , _____ )	( _____ , _____ )	
cor vermelho	( _____ , _____ )	( _____ , _____ )	

Precisamos de uma nova Tabela de Dupla Entrada para indicar as cores que foram utilizadas para pintar as três listras de cada uma das bandeiras. Você completar a tabela abaixo colocando nos espaços em branco a combinação das cores que já foram feitas por você para as duas primeiras listras das bandeiras, as quais estão indicadas em uma das entradas da tabela de dupla entrada. A outra cor refere-se àquela que se destina a pintar a terceira listra da bandeira de modo que se obtenha todas as possibilidades diferentes que se têm de pintar uma bandeira com três listras horizontais, quando se dispõe de duas cores: azul e vermelho.

	Cor da 3ª listra		
Cores das 1ª e 2ª listras da bandeira		cor Azul	cor Vermelho
( _____ )	( _____ )	( _____ )	( _____ )
( _____ )	( _____ )	( _____ )	( _____ )
( _____ )	( _____ )	( _____ )	( _____ )
( _____ )	( _____ )	( _____ )	( _____ )

Preencha a árvore de possibilidades desenhada abaixo de modo a conhecer as diferentes possibilidades que se tem de pintar uma bandeira com três listras horizontais, quando se dispõe de duas cores: azul e vermelho.



Concluindo: se uma bandeira tem três listras e queremos determinar todas as possibilidades de pintar essa bandeira quando se dispõe de duas distintas cores de giz de cera, o total de diferentes possibilidades para se pintar essa bandeira é de quantas maneiras?

Na segunda atividade exploramos o Princípio Multiplicativo. Propomos a pintura de bandeiras com três listras dispondo de dois bastões de giz de cera. Apresentamos ainda duas tabelas de dupla entrada, a primeira tabela para o preenchimento das duas primeiras listras e a segunda tabela relaciona o resultado da primeira tabela com as possibilidades de cores para a pintura da terceira listra. É proposto ainda o preenchimento de uma árvore de possibilidades de modo a conhecer as diferentes possibilidades para a pintura de uma bandeira com três listras horizontais.

Os objetivos da atividade 2 são o de identificar se os alunos apresentam algum padrão sistemático para efetuar as pinturas das bandeiras e construir duas tabelas de dupla entrada, uma vez que há três ações a serem feitas, quais sejam as escolhas das cores das três listras das bandeiras.

A utilização de duas Tabelas de dupla entrada e uma árvore de possibilidades, é dada de tal maneira que os alunos possam comparar vantagens e desvantagens na construção entre as duas possibilidades do uso de uma representação gráfica para determinar todas as possibilidades de pintura de uma bandeira com três listras horizontais.

A partir daí os alunos poderão certificar-se de que a escolha da Tabela de dupla entrada como representação gráfica para apresentar todas as possibilidades de solução para esta situação-problema obriga-os à construção de duas Tabelas de dupla entrada, enquanto a opção pela construção de uma árvore de possibilidades é o bastante para determinar a solução.

A construção da árvore de possibilidades também pode indicar ao aluno alguma padronização de pintura das listras na qual ele poderia ter se baseado enquanto efetuava as pinturas das listras da bandeira, uma vez que essa padronização guarda relação direta entre as ações utilizadas quando da construção dos “galhos” da árvore de possibilidades e o estabelecimento de suas “folhas”.

A árvore de possibilidades permite que os alunos reflitam sobre as informações que ali estão contidas e acerca da viabilidade de utilizá-la quando da resolução de problemas de contagem, podendo vir a adotá-la com frequência.

Pode ocorrer que o aluno efetue as pinturas das listras das bandeiras sem que ele tenha estabelecido algum padrão sistemático de pintura. Mas, enquanto constrói a árvore de possibilidades e ao identificar todos os agrupamentos (que são as possibilidades de pintura das faixas da bandeira) ele pode vir a fazer correlações entre a sequência das ações estabelecidas quando da pintura das listras e a ordenação dos elementos, presentes nos agrupamentos.

No momento da pintura de listras de bandeira com três listras, com a condição de que listras adjacentes não possam ser pintadas com a mesma cor o quantitativo de bandeiras, a maior exige que o aluno se utilize de alguma estratégia que permite identificar esse fato, razão por que salientamos a proposição de situações semelhantes, como uma das razões para o desenvolvimento do raciocínio combinatório.

Durante os momentos de construção de árvore de possibilidades pelos alunos ele está desenvolvendo o raciocínio combinatório enquanto decide acerca das ações necessárias e o quantitativo destas para o estabelecimento dos agrupamentos de objetos (sequência das cores das listras das bandeiras) que devem ser consideradas.

### 2.3. Atividade III

Analise as situações-problema indicadas a seguir:

**Situação-problema 1:** A professora do 5º ano A foi a um Mercado popular de sua cidade para comprar presentes para seus alunos por ocasião do “Dia das Crianças”. Ela decidiu que gastaria, no máximo, R\$ 5,00 pelo presente de cada uma das crianças. Ela foi anotando os preços dos presentes que gostaria de selecionar para distribuir aos seus alunos e registrou-os na tabela a seguir:

Brinquedos que vai poder comprar	Preço unitário
Bolas de gude	R\$ 1,50
Carrinho plástico	R\$ 2,00
Jogo de damas	R\$ 3,00

Jogo de dominó	R\$ 2,50
Jogo de varetas	R\$ 1,50
Bola plástica	R\$ 2,50
“iô-iô”	R\$ 2,00
Boneca plástica	R\$ 3,50

- Quantos presentes diferentes (contendo dois brinquedos cada um) ela poderá formar para gastar os R\$ 5,00 de que dispõe para cada aluno?
- Represente por meio de uma árvore de possibilidades todas as diferentes maneiras que a professora dispõe para “montar” presentes contendo dois brinquedos diferentes em cada um deles, gastando R\$ 5,00 em cada um deles.
- Represente por meio de uma árvore de possibilidades todas as diferentes maneiras que a professora dispõe para “montar” presentes contendo três brinquedos diferentes em cada um deles, gastando R\$ 5,00 em cada um deles.
- Represente por meio de uma árvore de possibilidades todas as diferentes maneiras que a professora dispõe para “montar” presentes contendo um, dois ou até três brinquedos diferentes em cada um deles, gastando valores até o máximo de R\$ 5,00 em cada um dos presentes possíveis.

**Situação-problema 2:** Considere que você dispõe dos algarismos 1, 2, 3, 4 e 7 para formar números de dois algarismos.

- Construa uma tabela de dupla entrada mostrando todos os números com dois algarismos que podemos formar, dispondo desses algarismos: 1, 2, 3, 4 e 7.
- Quantos números com dois algarismos podem ser formados atendendo às condições do item (a)?
- Construa uma árvore de possibilidades mostrando todos os números pares com dois algarismos que podemos formar quando se dispõe dos algarismos 1, 2, 3, 4 e 7.
- Quantos números com dois algarismos podem ser formados atendendo às condições do item (c)?
- Construa uma árvore de possibilidades mostrando todos os números ímpares com dois algarismos diferentes que podemos formar, quando se dispõe dos algarismos 1, 2, 3, 4 e 7.
- Quantos números com dois algarismos podem ser formados atendendo às condições do item (e)?

- g) Quantos números com dois algarismos, maiores que 40, podem ser formados quando se dispõe dos algarismos 1, 3, 4, 6 e 9?

**Situação-problema 3:** Carlos constrói aviõezinhos de brinquedo em madeira e, depois de pintados, eles estão prontos para serem vendidos na feira. Carlos dispõe de latas de tinta nas cores azul, vermelho, verde e amarelo. Carlos vai pintar a cabine do avião, as asas direita e esquerda e a cauda do avião, todos em cores diferentes.

- a) Quantos “aviõezinhos de brinquedo em madeira” diferentes, combinando suas cores como quiser podem ser pintados, dispondo de tintas nas cores azul, vermelho, amarelo e verde?
- b) Desenhe uma árvore de possibilidades que contenha todos os diferentes “aviõezinhos de brinquedo em madeira” conforme as condições estabelecidas no item (a).
- c) Quantos “aviõezinhos de brinquedo em madeira” diferentes podem ser pintados, combinando suas cores como quiser, mas tendo as duas asas do aviõezinho pintadas com a mesma cor, dispondo de tintas nas cores azul, vermelho, amarelo e verde?

Para esta atividade dispomos de três situações-problema. Para a primeira situação-problema, temos como propósito levar o aluno à leitura atenta dos enunciados das situações, a construção de árvores de possibilidades com o desenvolvimento do raciocínio combinatório. Para a segunda situação-problema temos como objetivo discutir com o nosso estudante as possibilidades de obter diferentes números a partir de certo quantitativo de algarismos disponíveis e que o total de possibilidades está associado ao fato de que um mesmo algarismo assume diferentes valores em função da posição que ele ocupe na formação do número, ou seja, está ligado ao uso do princípio do valor posicional na obtenção de todos os diferentes números a partir de um dado quantitativo de algarismos disponíveis. Para a terceira situação-problema temos como objetivo sistematizar aspectos relacionados com a apropriação do raciocínio combinatório, combinando com a construção de diferentes tipos de árvores de possibilidades e a aplicação do Princípio Multiplicativo. A contagem final das possibilidades depende de duas ou mais etapas de escolha, configurando situações-problema associadas à problema de contagem, as quais exigem formas sistematizadas para obtê-la, em cada uma das diferentes situações propostas.

Utilizamos, na resolução dessa atividade, os conceitos de Princípio Multiplicativo, Arranjos Simples e Combinação Simples.



## REFERÊNCIAS

BRASIL. Ministério da Educação. **Parâmetros curriculares nacionais: matemática**. 3. ed. Brasília: Secretaria da Educação Fundamental. 2001.

CARRAHER, David. Educação Tradicional e Educação Moderna. Em: CARRAHER, D.; NUNES, T. (Org.). **Aprender pensando: contribuições da psicologia cognitiva para a educação**. Petrópolis: Vozes, 1986.

DANTE, Luiz Roberto. **Matemática: contexto e aplicações**. 1. ed. São Paulo: Ática, 2010.

HAZZAN, Samuel. **Fundamentos de Matemática Elementar**, 5. 7ª ed. São Paulo: Atual, 1993.

PAIVA, Manoel. **Matemática**. 1. ed. São Paulo: Moderna, 2009.

STURM, Wilton. **As possibilidades de um ensino de análise combinatória sob uma abordagem alternativa**. Campinas, 1999, 132 p. Dissertação (Mestrado em educação) – Faculdade de Educação, Universidade de Campinas, 1999.

TEIXEIRA, Paulo Jorge Magalhães. **Resolvendo Problemas de Análise Combinatória nos Anos Iniciais do Ensino Fundamental**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2014.

### COMBINATORY ANALYSIS: AN APPROACH FOR THE SIXTH YEAR OF ELEMENTARY EDUCATION

**ABSTRACT** - Noting the difficulties that exist in the teaching and learning of mathematics, we propose in this paper an approach to Combinatorial Analysis to the sixth grade of elementary school, inductive and intuitively, in a first moment, without formality, favoring the cognitive development of the student. We noticed along our learning staff who work Combinatorics shown as an obstacle to students due to the way that content is discussed in almost all of the time through mathematical formulas, which makes losing the sense of problem solving. Our proposal is to use the resolution of problems and techniques of fundamental principle of counting, which is usually in the second year of high school is quoted only at the beginning of the study of Combinatorics to then list the formulas of arrangements, combination and permutation, combinatorial reasoning development in students of the sixth grade of elementary school, improving this way the teaching and understanding of mathematics. We had as main sources of research, books, the National curricular parameters, magazine articles, dissertations and theses that deal with the subject.

**KEYWORDS:** Combinatorics; Combinatorial Reasoning; Troubleshooting.

## ÍNDICE REMISSIVO

### A

Abordagem 11, 18, 46, 48, 49, 52, 60-62, 73  
Adimensional 14, 16, 17  
Alcance variável 10, 45  
Algoritmo 30, 33, 35, 36, 38, 40-42, 58, 61  
Altura de onda 13, 14, 18  
Análise Combinatória 10, 60-62, 73  
Análise de Correspondência 45, 52  
Ancestralidade 45-47, 56-58  
Árvore de possibilidades 68-72

### B

Bidimensional 9, 12, 21, 32  
Biplots 45

### C

CAD 21  
Cadeia de Markov 48  
Campo de ondas 13, 14  
Campos Markovianos 45, 46, 48, 58  
Classificador 10, 45, 45, 46, 52, 52, 55-57  
Coerência 42  
Compilação 38  
Comprimento 9, 20-22, 24-26, 28  
Computação Gráfica 31  
Convexidade 31  
Coordenadas 30, 33-36  
Corpo hídrico 20-22  
Cromossomo 54, 55

Curva de Jordan 41

### D

Dependência 10, 11, 13, 45, 46, 48, 50-52, 57, 58  
Desvio 28  
Diagrama 16, 17, 24  
Digitalização tridimensional 11, 30  
Discretização 32, 36  
Dissipação 15  
Distribuição 21, 22, 22, 26, 48, 49, 52, 54, 56, 57

### E

Energia 11-16  
Ensino fundamental 11, 60, 62  
Ensino-aprendizagem 10, 61  
Enumeração das possibilidades 66  
Equação paramétrica 9, 12, 13, 16, 17, 26  
Espectro 13-15

### F

Fetch 9, 12-15, 19-24, 26-29  
Fonte total 15

### G

Genótipos 47, 48, 52, 55  
Geometria Computacional 31

### H

HapMap 45, 46, 53, 58, 59

**I**

Índice de Heterogeneidade 51, 52, 54  
Índice do ponto 32  
Itaipu 9, 20, 25, 26, 28, 29

**J**

Janela de consenso 51, 52, 55  
Janelas populacionais 50, 51  
JONSWAP 12, 17, 26, 28

**L**

Lapidação 30, 31

**M**

Magnitude 9, 20  
Malhas 33, 37, 40-42  
Método de Feito-Torres 36, 38, 41  
Método de Möller 34, 41  
Método de Segura 35, 36, 40  
Método do raio 32  
Método rápido 11, 20, 21, 28, 29  
Modelagem 9, 11-13, 21, 45, 47, 48

**N**

Notação 47, 48  
Nuvem de pontos 11, 30, 31, 38, 42

**P**

Parametrização 33  
PCN 62  
Polimorfismo 45, 46  
População 10, 45-48, 51-58  
Pré-processamento 32, 38, 40-42

Probabilidade 7, 9, 48, 48, 49, 62  
Problemas de Contagem 62  
Produto cartesiano 66  
Protótipo 30, 31

**R**

Raciocínio 60-62, 65, 70, 72  
Reservatório 9, 20, 22-26, 28, 29  
Resolução de problemas 60, 61, 70

**S**

Saville 12, 17, 19, 21, 22, 26, 29  
Simulação computacional  
SNPs 10, 11, 45-54, 57, 59  
Sólidos 9, 11, 30, 33, 37-39, 41

**T**

Tabela de Dupla Entrada 66, 68-69, 71  
Teorema 30, 32, 37, 38, 41, 42  
Tetraedro 30, 32, 35-37, 40, 42  
Topologia 9, 11-13, 16, 18  
Transferência de energia 15  
Transformação 34

**V**

Variabilidade 9, 12, 16, 46, 57  
Vento 9, 11-13, 15, 17, 20, 21, 26, 28  
Verossimilhança 49  
Vértices 31-33, 35-37, 39  
Vetor 33, 38, 39, 41, 42, 48, 49, 52  
Vizinhança 48-50

## **SOBRE OS ORGANIZADORES**

### **Marcelo Marques**

Professor da Universidade Estadual de Maringá (UEM) desde 2005, doutor em Engenharia de Recursos Hídricos e Ambiental pela Universidade Federal do Paraná (2013) e mestre em Engenharia de Recursos Hídricos e Tecnologias Ambientais pela UNESP (2005) realizando o estágio de docência e parte das atividades de mestrado na COPPE-UFRJ (2004). Especialista em obras hidráulicas pela Universidad Internacional de Andalucia (UNIA), Espanha (2004) e especialista em Geoprocessamento pela UEM (2006). Graduado em Engenharia Civil pela UEM (1999). Coordena o Núcleo de Pesquisa em Engenharia de Recursos Hídricos e Ambiental (NUPEHIDRO). Atua nas grandes áreas de Engenharia, Ciências Exatas e da Terra e Multidisciplinar.

### **Isabela Arantes Ferreira**

Engenheira Civil pela Universidade Estadual de Maringá (UEM) e mestranda do Programa de Pós-Graduação em Engenharia Química da Universidade Estadual de Maringá (PEQ-UEM). Atua nas áreas de Gestão Ambiental e Recursos Hídricos.



# Avanços Científicos em Matemática, Probabilidade e Estatística

[www.bookerfield.com](http://www.bookerfield.com) 

[contato@bookerfield.com](mailto:contato@bookerfield.com) 

[@bookerfield](https://www.instagram.com/bookerfield) 

Bookerfield Editora 



# Avanços Científicos em Matemática, Probabilidade e Estatística

[www.bookerfield.com](http://www.bookerfield.com) 

[contato@bookerfield.com](mailto:contato@bookerfield.com) 

[@bookerfield](https://www.instagram.com/bookerfield) 

Bookerfield Editora 

ISBN 978-658992968-0

