

# Aprimoramentos em modelagem geológica implícita com funções distância assinaladas

Me. Roberto Mentzingen Rolo

Orientador: Prof. Dr. João Felipe Coimbra Leite Costa, PhD

Universidade Federal do Rio Grande do Sul  
Escola de Engenharia

Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais

17 de junho de 2019

## Proposta de tese para o exame de qualificação

2019-06-11

Um boa dia a todos, gostaria primeiramente de agradecer a presença de todos vocês, em especial doutor Diego, doutor Luís Eduardo, Professor João Felipe.

Meu nome é Roberto Rolo sou engenheiro de minas, fiz meu mestrado aqui na UFRGS com o professor João Felipe intitulado modelagem geológica implícita com funções distâncias assinaladas. Minha proposta de tese é uma extensão natural do trabalho desenvolvido no meu mestrado: eu proponho aprimoramentos em modelagem geológica implícita com funções distância assinaladas.

# sumário I

## Introdução

### Interpretação e modelagem geológica

#### Método tradicional

#### Incerteza do modelo geológico

#### Métodos matemáticos

#### Métodos implícitos

### Modelagem geológica implícita com funções distância assinaladas

#### O banco de dados

#### Codificando as amostras em indicadores

#### Calculando a função distância assinalada

#### Variografia das distâncias assinaladas

#### Interpolação das distâncias assinaladas

#### Visualização do modelo geológico

#### Adaptação para múltiplas categorias simultaneamente

#### Incorporação da não estacionariedade de segunda ordem

#### Refinamento iterativo

# Proposta de tese para o exame de qualificação

2019-06-11

## └ sumário

Nessa apresentação eu ser breve na introdução ao assunto modelagem geológica, vou falar sobre a metodologia tradicional, explícita, métodos automáticos e modelagem implícita, dando destaque para a modelagem implícitas com funções ditância assinaladas.

Então vou apresentar o estado da arte em modelagem geológica implícita com distâncias assinaladas usando um banco de dados tridimensional pra mostrar a mecânica e os diferentes métodos para avaliação de incerteza de modelos geológicos implícitos.

Depois disso eu vou apontar os principais problemas da metodologia e apresentar minha proposta de tese que tem como objetivo eliminar alguns desses problemas.

sumário I
Introdução
Interpretação e modelagem geológica
Método tradicional
Incerteza do modelo geológico
Métodos matemáticos
Métodos implícitos
Modelagem geológica implícita com funções distância assinaladas
O banco de dados
Codificando as amostras em indicadores
Calculando a função distância assinalada
Variografia das distâncias assinaladas
Interpolação das distâncias assinaladas
Visualização do modelo geológico
Adaptação para múltiplas categorias simultaneamente
Incorporação da não estacionariedade de segunda ordem
Refinamento iterativo

## sumário II

Incorporação de informação secundária

### Avaliação da incerteza

Avaliação heurística da incerteza

BOUNDSIM

Simulação direta das distâncias assinaladas

Simulação multi ponto

Boundary simulation

Abordagem hierárquica

Sumário dos métodos de avaliação de incerteza

### Proposta de tese

Problemas

Interpolador

Zona de incerteza

Avaliação da incerteza

Boundary simulation multi categórico

P-field

# Proposta de tese para o exame de qualificação

2019-06-11

## └ sumário

sumário II
Incorporação de informação secundária
Avaliação da incerteza
BOUNDSIM
Simulação direta das distâncias assinaladas
Simulação multi ponto
Boundary simulation
Abordagem hierárquica
Sumário dos métodos de avaliação de incerteza
Proposta de tese
Problemas
Interpolador
Zona de incerteza
Avaliação da incerteza
Boundary simulation multi categórico
P-field

## sumário III

- Simulação plurigaussiana truncada
- Validação
- Sumário
- Cronograma e atividades

## Referências bibliográficas

# Proposta de tese para o exame de qualificação

2019-06-11

## └ sumário

sumário III  
Simulação plurigaussiana truncada  
Validação  
Sumário  
Cronograma e atividades

Referências bibliográficas

# Introdução

Construir modelos numéricos de longo, médio e curto prazo para avaliação de recursos/reservas e planejamento de mina exige quatro grandes atividades:

1. Coleta e gerenciamento de dados;
  2. Interpretação e modelagem geológica;
  3. Atribuição de teores;
  4. Avaliação e gerenciamento da incerteza geológica e de teores.

posta de tese para o exame de qualificação

Introdução

## └ Introdução

mativas de recursos e reservas minerais demandam a construção de modelos numéricos de reservas:

- e longo prazo, que abrangem toda a extensão do depósito mineral e compreendem todo o  
período de vida da mina;  
Modelos de médio prazo para planejar de um a seis meses no futuro;  
Curto prazo pra balizar semanalmente ou diariamente as decisões relativas a controle de  
processos e planejamento mais detalhado da mina.

onstrução desses modelos exige quatro grandes atividades:

- coleta e gerenciamento de dados;  
interpretação e modelagem geológica;  
tribuição de teores;  
Avaliação e gerenciamento da incerteza geológica e de teores

scopo dessa proposta tese está no passo 2: interpretação e modelagem geológica

- Consultar modelos numéricos de longo, médio e curto prazo para avaliação de censores resíduos e planejamento de missões de vigilância geológica;
- 1. Coleta e gerenciamento de dados;
- 2. Interpretação e modelagem geológica;
- 3. Ativação de alertas;
- 4. Aplicação e monitoramento de inspeções geológicas e técnicas.

# Interpretação e modelagem geológica

- Identificar diferentes domínios;
- Definir os limites de cada função aleatória estacionária.

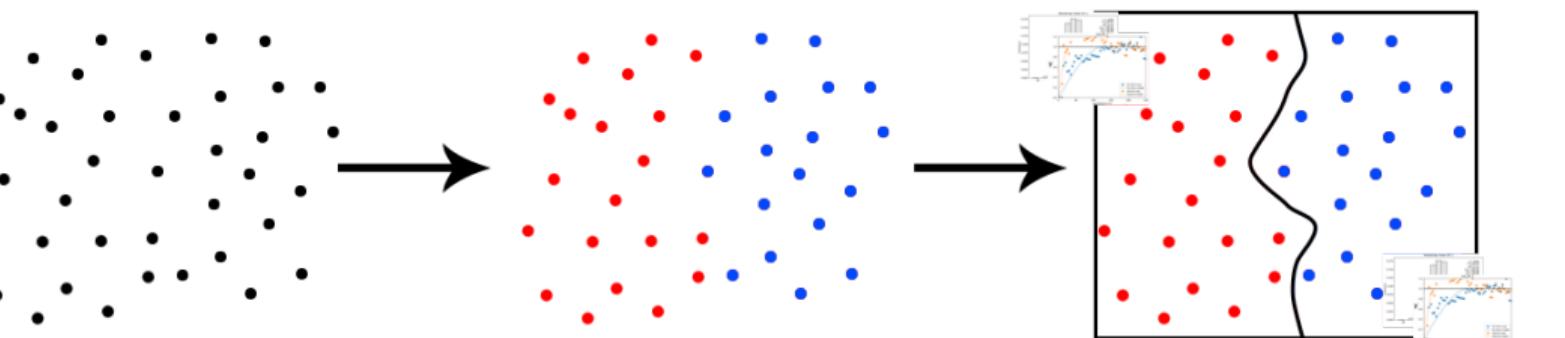


Figura 1: Interpretação e modelagem geológica.

Proposta de tese para o exame de qualificação

- Introdução
- Interpretação e modelagem geológica
  - Interpretação e modelagem geológica

2019-06-11

Após a coleta, gerenciamento e checagem dos dados, é necessário identificar diferentes domínios. A determinação dos domínios deve ser baseada no conhecimento geológico, como zonas de oxidação, diferentes litologias, alteração ou limites estruturais e deve ser suportada por uma extensiva análise estatística e variográfica e pode ser baseada na combinação de uma ou mais variáveis. A definição de diferentes domínios é necessária porque a inferência geostatística exige a decisão de estacionariedade. Os teores em cada domínio estacionários pertencem a populações estatística diferentes caracterizadas por seu modelo de distribuição de probabilidades (o histograma) e seu modelo de covariâncias (o variograma). Uma vez que os domínios tenham sido definidos um modelo tridimensional que define os limites de cada função aleatória estacionária deve ser construído. Esse é o modelo geológico: Ele define a jurisdição espacial de cada função aleatória. O modelo geológico é o alicerce para todo o trabalho de estimativa subsequente e muitas vezes é o fator de maior importância na estimativa das tonelagens mineralizadas.

Interpretação e modelagem geológica

- Identificar diferentes domínios;
- Definir os limites de cada função aleatória estacionária.

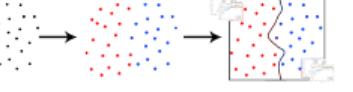


Figura 1: Interpretação e modelagem geológica.

6/75

## Metodologia tradicional

A abordagem tradicional para a criação de modelos geológicos tridimensionais é através da triangulação de polilinhas.

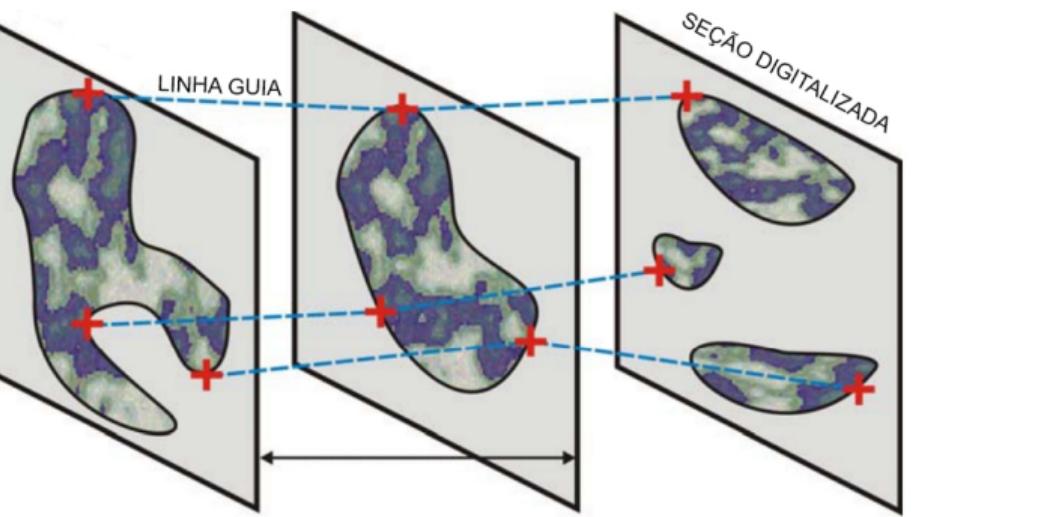
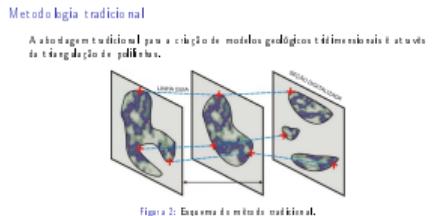


Figura 2: Esquema do método tradicional.

Proposta de tese para o exame de qualificação

- └ Introdução
- └ Método tradicional
  - └ Metodologia tradicional

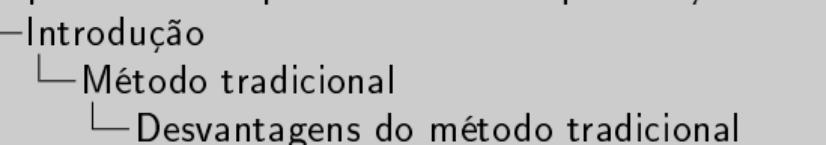
2019-06-11



Tradicionalmente os modelos geológico são construídos de forma explícita: o geomodelador digitaliza manualmente polilinhas em seções verticais e horizontais a partir dos dados de sondagem, as polilinhas são unidas por linhas guias e trianguladas gerando uma representação do sólido geológico.

# Desvantagens do método tradicional

- ▶ Tedioso e demorado;
- ▶ Exige um profissional especializado e experiente;
- ▶ Geometria dos corpos precisa ser simplificada;
- ▶ Subjetivo;
- ▶ Não replicável;
- ▶ Inflexível;
- ▶ Não avalia a incerteza.



2019-06-11

Embora a metodologia tradicional seja direta e simples e que softwares de mineração forneçam ferramentas computacionais para agilizar o processo, ainda existem uma série de desvantagens e limitações:

- \* O processo é tedioso e demorado. Em depósitos de alta complexidade não é raro o geomodelador rabalhar por três meses no modelo geológico.
- \* Digitalizar manualmente as polilinhas exige muito tempo de um profissional experiente;
- \* A geometria dos corpos geológicos muitas vezes precisa ser simplificada para que o modelo seja concebido em tempo hábil;
- \* O processo é subjetivo, Geomodeladores diferentes criam modelos geológicos diferentes a partir do mesmo banco de dados;
- \* Por esse motivo os modelos geológicos criados a partir do método tradicional não são replicáveis consequentemente não auditáveis;
- \* O método é inflexível, já que a medida que novas amostras são obtidas, a atualização do modelo demanda nova digitalização manual;
- \* Em muitas minas, talvez na grande maioria delas, apenas um modelo geológico é construído

- ▶ Tedioso e demorado;
- ▶ Exige um profissional especializado e experiente;
- ▶ Geometria dos corpos precisa ser simplificada;
- ▶ Subjetivo;
- ▶ Não replicável;
- ▶ Inflexível;
- ▶ Não avalia a incerteza.

# Incerteza do modelo geológico

Em muitos casos, a incerteza do modelo geológico pode ser uma fonte de incerteza crucial e deve ser avaliada.

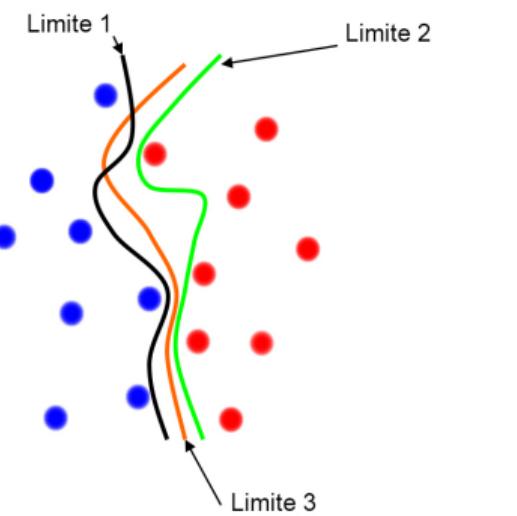


Figura 3: Incerteza do modelo geológico.

Proposta de tese para o exame de qualificação

- └ Introdução
- └ Incerteza do modelo geológico
  - └ Incerteza do modelo geológico

2019-06-11

No interior dos domínios não existe incerteza relacionada ao modelo geoógico. No esquema do slide azul é azul e vermelho é vermelho, a incerteza é associada ao limite que separa os diferentes domínios. Sua definição, em sub superfície, é um mero palpitar.

Em muitos casos a incerteza do modelo geológico pode ser uma fonte de incerteza crucial. Em depósitos de veio de ouro por exemplo, o volume mineralizado é um indicador economico vital do projeto e está diretamente ligado ao modelo geológico, ignorar a incerteza do modelo geológico pode ser devastador para o empredimento. A incerteza do modelo geológico DEVE ser avaliada.



Figura 3: Incerteza do modelo geológico.

Métodos matemáticos

Métodos determinísticos

- ▶ Vizinho mais próximo;
- ▶ Krigagem dos indicadores.

Métodos estocásticos

- ▶ Simulação sequencial dos indicadores;
- ▶ Simulação gaussiana/plurigaussiana truncada;
- ▶ Simulação multi ponto;
- ▶ Simulação baseada em objetos;

Proposta de tese para o exame de qualificação

Introdução

Métodos matemáticos

Métodos matemáticos

2019-06-11

Métodos matemáticos

Métodos determinísticos

- ▶ Vizinho mais próximo;
- ▶ Krigagem dos indicadores.

Método estocástica

- ▶ Simulação sequencial dos indicadores;
- ▶ Simulação gaussiana/plurigaussiana truncada;
- ▶ Simulação multi ponto;
- ▶ Simulação baseada em objetos;

As desvantagens do método explícito impulsionaram a criação de métodos automáticos, ou pelo menos semi automáticos, de modelagem geológica. Para a criação de modelos determinísticos, podem ser utilizados desde métodos matemáticos mais simples como o vizinho mais próximo, como geostatísticos como a krigagem dos indicadores. Já a necessidade de avaliação da incerteza impulsionou o desenvolvimento de métodos estocásticos, baseados em múltiplas realizações equiprováveis do modelo geológico. Metodologias estabelecidas da geostatística clássica, baseadas no variograma, são a simulação sequencial dos indicadores, simulação gaussiana truncada e simulação plurigaussiana truncada. Outros métodos não baseados no variograma também foram desenvolvidos: simulação multi ponto, simulação baseada em objetos são os principais expoentes.

10/75

# Métodos implícitos

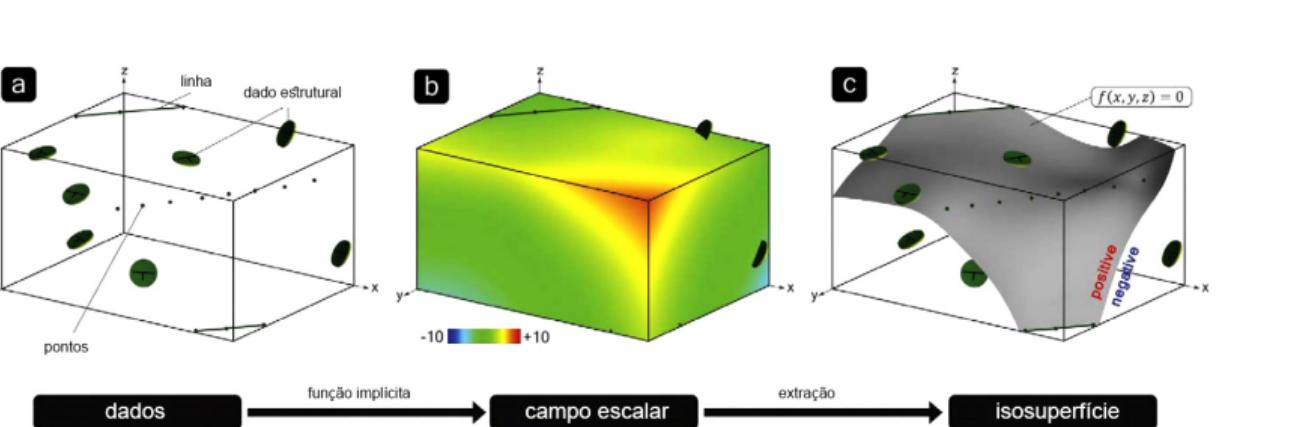


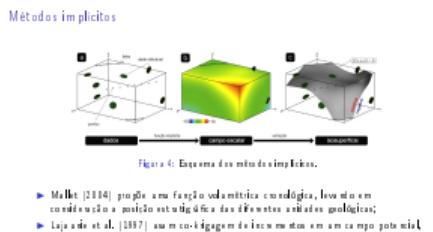
Figura 4: Esquema dos métodos implícitos.

- ▶ Mallet (2004) propõe uma função volumétrica cronológica, levando em consideração a posição estratigráfica das diferentes unidades geológicas;
- ▶ Lajaunie et al. (1997) usam co-krigagem de incrementos em um campo potencial, omitindo a função volume.

## Proposta de tese para o exame de qualificação

- └ Introdução
- └ Métodos implícitos
- └ Métodos implícitos

2019-06-11



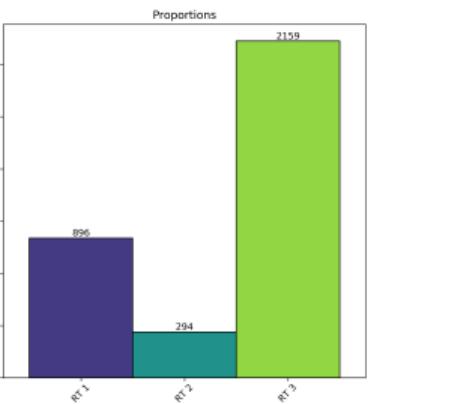
Uma outra família de métodos que impotaram técnicas da computação gráfica são os métodos implícitos. A idéia central é usar uma função implícita para demarcar regiões no espaço. Todos os métodos implícitos compartilham da mesma mecânica. A partir de dados esparsos que podem ser categorias, dados estruturais, podem ser em pontos ou linhas... Uma função implícita é derivada, o campo escalar, que têm um número infinito de isosuperfícies. Para visualizar o modelo geológico uma isosuperfície específica deve ser extraída desse modelo implícito, geralmente a isosuperfície 0.

O que difere os métodos implícitos é a função volume, que é interpolada para criação do modelo implícito. mallet propôs uma função volume cronológica, baseada na posição estratigráfica das unidades geológicas. Um outro método implícito, nascido na escola de minas da frança, usa co-krigagem de incrementos em campo potencial omitindo a função volume.

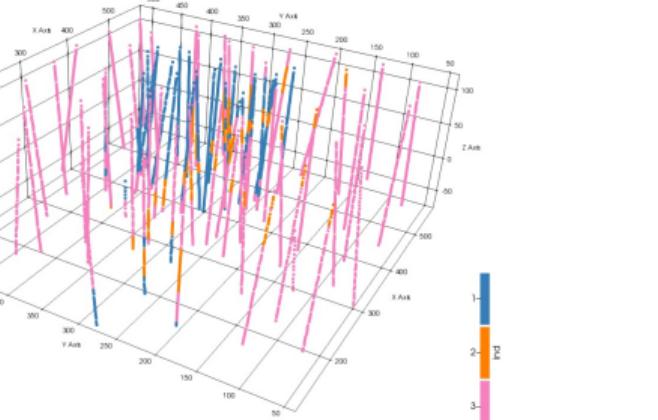
De longe a função volume mais comum é a função distância assinalada, aplicações desse método são encontradas por toda a literatura de interpolação de dados esparços. Na modelagem geológica o método é competente em capturar a geometria e extensão de corpos geológicos e tem sido aplicada com sucesso há mais de 10 anos na exploração mineral. Nos últimos anos vem

# O banco de dados

72 furos totalizando 3349 amostras distribuídas entre 3 diferentes categorias.



(a) Proporções.



(b) Vista das amostras.

Figura 5: O banco de dados.

Proposta de tese para o exame de qualificação

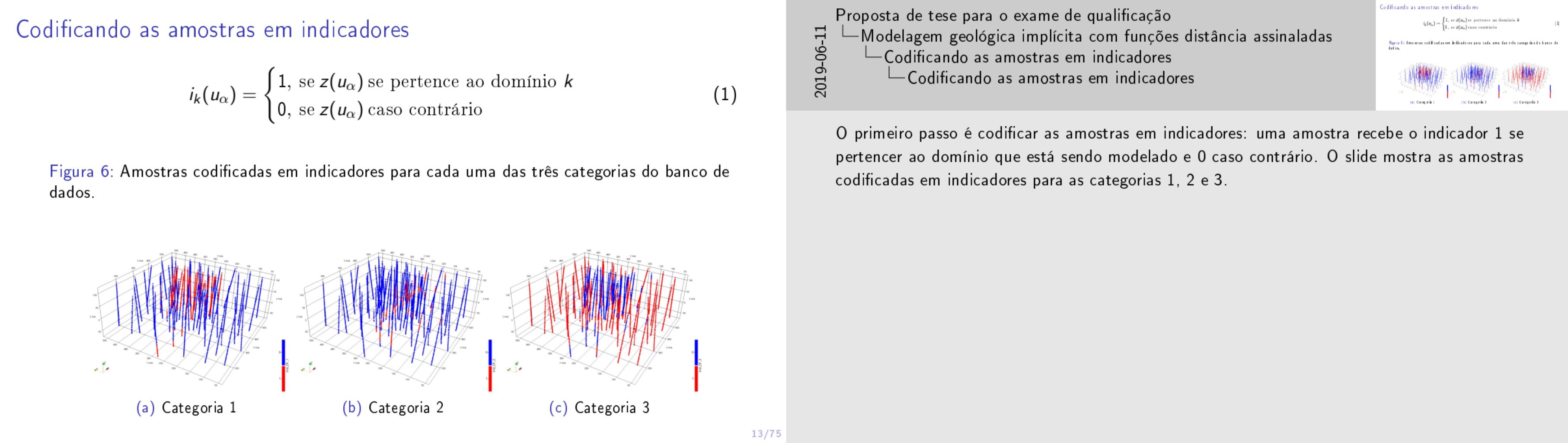
- └ Modelagem geológica implícita com funções distância assinaladas
- └ O banco de dados
- └ O banco de dados

2019-06-11



Nos próximos slides a modelagem geológica implícita com distâncias assinaladas é explicada passo a passo e os métodos para avaliação de incerteza disponíveis na literatura são apresentadas.

Um banco de dados sintético que emula um depósito de cobre pôrfiro é usado para exemplificar e comparar os métodos. São 72 furos, totalizando cerca de 3 mil amostras, distribuídas entre três categorias: uma rocha encaixante e duas intrusões: uma vertical e uma tabular com mergulho. O slide mostra as proporções de cada litologia e uma vista em perspectiva das amostras.



# Calculando a função distância assinalada

$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\|, & \text{se } u_\alpha \text{ pertence ao domínio} \\ +\|u_\alpha - u_\beta\|, & \text{se } u_\alpha \text{ não pertence ao domínio} \end{cases}$  (2)

O local  $u_\beta$  corresponde à amostra mais próxima codificada com um indicador diferente de  $u_\alpha$ .

**Figura 7:** Ilustração esquemática mostrando o cálculo das distâncias assinaladas.

Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
- └ Calculando a função distância assinalada
- └ Calculando a função distância assinalada

O segundo passo é o cálculo das distâncias assinaladas: para cada ponto amostral, a menor distância euclideana até um outro ponto amostral que pertença à um indicador oposto é computada e atribuída aquele ponto. Com o sinal negativo caso pertença ao domínio que está sendo modelado e com o sinal positivo caso contrário.

Para o esquema do slide: pra essa amostra preta aqui em cima, ela não pertence ao domínio, então foi codificada com o indicador 0, eu vou computar então a menor distância entre ela e uma outra amostra codificada com o indicador oposto, uma amostra vermelha, que pertence ao domínio. São 12 metros, então essa amostra preta vai receber o valor 12 metros com o sinal positivo, porque não pertence ao domínio. Para a amostra vermelha, vou computar a menor distância até uma amostra preta, que é essa daqui, 8 metros, então ela recebe 8 metros com o sinal negativo porque pertence ao domínio que está sendo modelado. E finalmente a ultima amostra preta está a 8 metros da vermelha, recebe 8 com o sinal positivo, porque está fora do domínio modelado.

**Figura 7:** Ilustração esquemática mostrando o cálculo das distâncias assinaladas.

**Calculando a função distância assinalada**

$$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\|, & \text{se } u_\alpha \text{ pertence ao domínio} \\ +\|u_\alpha - u_\beta\|, & \text{se } u_\alpha \text{ não pertence ao domínio} \end{cases}$$

O local  $u_\beta$  corresponde à amostra mais próxima codificada com um indicador diferente de  $u_\alpha$ .

14/75

# Calculando a função distância assinalada



(a) Categoria 1



(b) Categoria 2



(c) Categoria 3

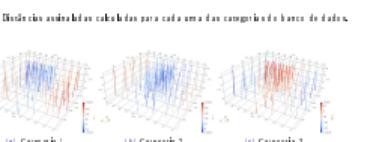
Figura 8: Distâncias assinaladas calculadas para cada uma das categorias do banco de dados.

Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
- └ Calculando a função distância assinalada
- └ Calculando a função distância assinalada

Esse slide mostra as distâncias assinaladas calculadas para as litologias 1, 2 e 3.

2019-06-11



# Variografia das distâncias assinaladas

Distâncias assinaladas não são estacionárias, o variograma não se estabiliza em um patamar. Além disso, o caráter extremamente contínuo das distâncias torna a identificação analítica das direções principais um processo embarracoso.

- ▶ Treinar o variograma usando validação cruzada;
- ▶ Tentar modelar interativamente os variogramas experimentais;
- ▶ Calcular e modelar os variogramas para as propriedades de indicadores e transformá-los em um equivalente gaussiano para as distâncias assinaladas;
- ▶ inferir um modelo de covariância plausível visualmente a partir das amostras ou de mapas delineados a mão.

## Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Variografia das distâncias assinaladas
    - └ Variografia das distâncias assinaladas

Agora as distâncias calculadas precisam ser interpoladas, alguns dos métodos de interpolação, que eu vou detalhar daqui a pouco, exigem uma função de covariância, que pode ser obtida a partir dos dados.

Aqui surge o primeiro problema do método: as distâncias assinaladas não são estacionárias, isso quer dizer que o variograma não se estabiliza em um patamar, ele cresce indefinidamente à medida que o lag aumenta, dificultando a inferência do alcance. Além disso, as distâncias apresentam um comportamento extremamente contínuo, que torna difícil a identificação de direções principais. Esse caráter contínuo das distâncias faz com que o modelo gaussiano seja o mais indicado pra ajustar os variogramas.

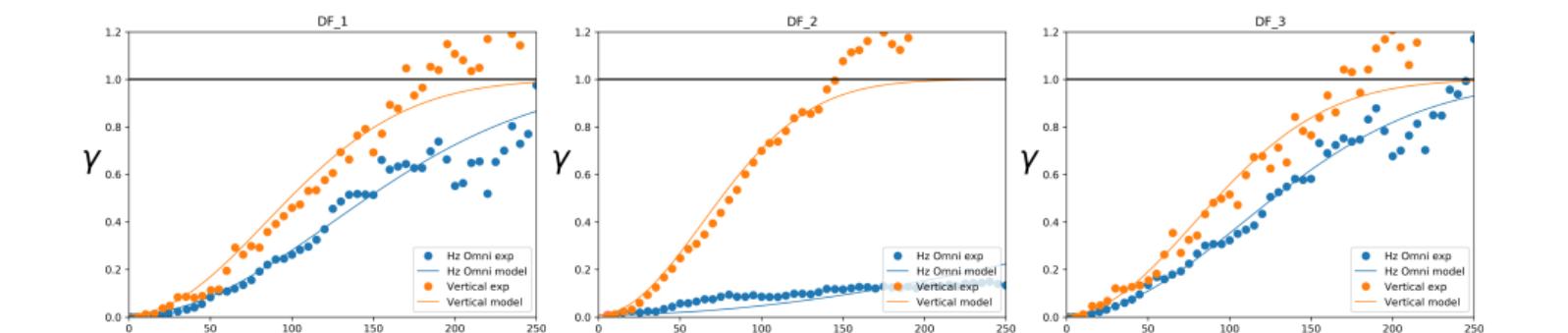
A literatura sugere alternativas:

- \* Treinar o variograma usando validação cruzada;
- \* Tentar modelar de fato os variogramas experimentais;
- \* Calcular e modelar o variograma dos indicadores e transformá-los em um equivalente gaussiano;
- \* Inferir um modelo de covariância plausível visualmente, a partir das amostras ou mapas delineados.

Distâncias assinaladas não são estacionárias, o variograma não se estabiliza em um patamar. Além disso, o caráter extremamente contínuo das distâncias torna a identificação analítica das direções principais um processo embarracoso.

- ▶ Treinar o variograma usando validação cruzada;
- ▶ Tentar modelar interativamente os variogramas experimentais;
- ▶ Calcular e modelar os variogramas para as propriedades de indicadores e transformá-los em um equivalente gaussiano para as distâncias assinaladas;
- ▶ Inferir um modelo de covariância plausível visualmente a partir das amostras ou mapas delineados a mão.

# Variogramas das distâncias assinaladas



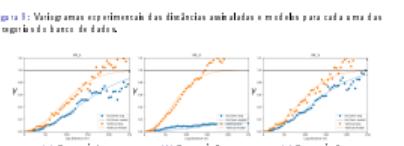
**Figura 9:** Variogramas experimentais das distâncias assinaladas e modelos para cada uma das categorias do banco de dados.

2019-06-11

Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
- └ Variografia das distâncias assinaladas
- └ Variogramas das distâncias assinaladas

Esse slide mostra os variogramas das distâncias assinaladas calculados e modelados a partir do modelo gaussiano para as três litologias. Aqui eu calculei o variograma estandarizado, Então defino o range na variância igual a um, mas note que o valor da variância continua aumentando.



# Variogramas dos indicadores

Figura 10: Variogramas experimentais dos indicadores e modelos para cada uma das categorias do banco de dados.

(a) Categoria 1      (b) Categoria 2      (c) Categoria 3

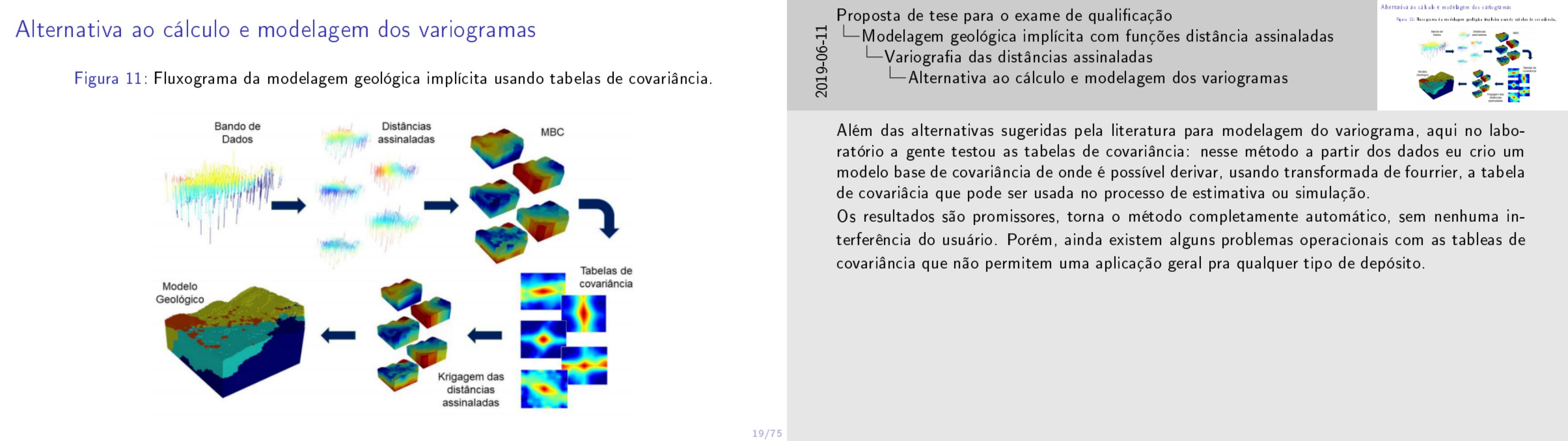
Proposta de tese para o exame de qualificação  
└ Modelagem geológica implícita com funções distância assinaladas  
  └ Variografia das distâncias assinaladas  
    └ Variogramas dos indicadores

2019-06-11

Já esse slide mostra os variograma calculados e modelados a partir do modelo esférico para os indicadores de cada categoria. Em tese, os variogramas dos indicadores são mais fáceis de modelar e são estacionários. Mas na prática, muitas vezes os variogramas dos indicadores são mais ruidosos e menos estruturados do que o das das distâncias assinaladas.

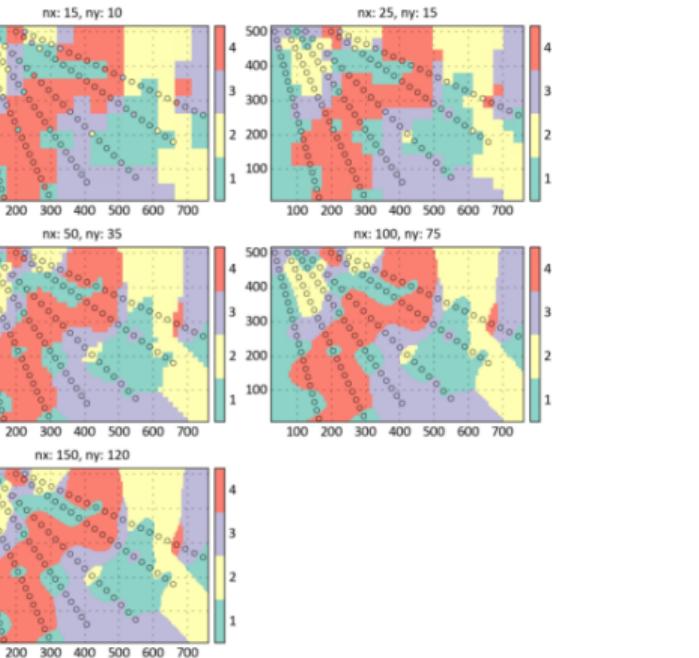
Variogramas dos indicadores

Figura 10: Variogramas experimentais dos indicadores e modelos para cada uma das categorias do banco de dados.



# Resolução do grid

Figura 12: Efeito da resolução do *grid* na reprodução de estruturas geológicas.

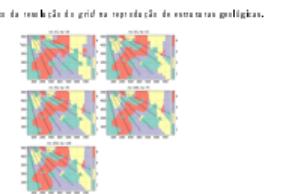


2019-06-11

Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Interpolação das distâncias assinaladas
  - └ Resolução do grid

Resolução do grid



Essas distâncias calculadas e variografadas devem ser interpoladas para todos os nós de um grid. Embora os métodos implícitos sejam gridless, que dizer, eles não dependem de um grid, os algoritmos de extração de iso superfície exigem que a propriedade exista em um grid regular. Para um número constante de amostras o tempo necessário para interpolar a função volume aumenta linearmente com o número de nós. Os parâmetros do grid, muitas vezes são determinados por parâmetros técnicos do projeto. Porém, para a definição do modelo geológico é importante que o grid possa reproduzir a menor estrutura geológica de interesse.

Essa imagem mostra um modelo geológico multi categórico em grids de diferentes resoluções: para as dimensões 15 por 10 parecem muito grosseiros e as categorias dos pontos amostrais não são reproduzidas, já a dimensão 25 por 15 eu ganho um pouco de resolução, mas as estruturas ainda não tem uma forma geológica. Pra 50 por 35 as formas o modelo se suavizam, mas ainda tem algum serrilhado, 100 por 75 as formas estão bastante suaves e as categorias das amostras são reproduzidas. Pra 150 por 120 não parece haver significativa melhora em relação ao modelo anterior.

Então é preciso encontrar um balanço entre resolução e tempo de interpolação, nesse caso algo

## Grids criados

	Grosso	Fino
nx	49	97
ny	49	98
nz	51	102
sx	10m	5m
sy	10m	5m
sz	4m	2m
num	122451	969612

- Proposta de tese para o exame de qualificação  
└ Modelagem geológica implícita com funções distância assinaladas  
  └ Interpolação das distâncias assinaladas  
    └ Grids criados

2019-06-11

Par aesse estudo de caso, dois grid foram criados, um grosso com aproximadamente 100 mil nós e um fino com aproximadamente 1 milhão de nós.

Grids criados

	Grosso	Fino
nx	49	97
ny	49	98
nz	51	102
sx	10 m	5 m
sy	10 m	5 m
sz	4 m	2 m
num	122451	969612

# Métodos de interpolação

- ▶ Hosseini and Deutsch (2007) utilizaram inverso da distância;
- ▶ Silva (2015) utilizou krigagem ordinária global;
- ▶ Rolo (2017) utilizou krigagem ordinária;
- ▶ Silva and Deutsch (2015) aplicaram *dual kriging*;
- ▶ Boisvert (2013) gerou modelos implícitos através de distâncias assinaladas com anisotropia variável local (*Locally varying anisotropy kriging - LVA*);
- ▶ Manchuck and Deutsch (2018) propuseram a utilização de mínimos quadrados móveis para incorporar interpretação manual e avaliar incerteza.

2019-06-11

## Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Interpolação das distâncias assinaladas
    - └ Métodos de interpolação

- ▶ Hosseini and Deutsch (2007) utilizaram inverso da distância;
- ▶ Silva (2015) utilizou krigagem ordinária global;
- ▶ Rolo (2017) utilizou krigagem ordinária;
- ▶ Silva and Deutsch (2015) aplicaram *dual kriging*;
- ▶ Boisvert (2013) gerou modelos implícitos através de distância assinalada com anisotropia variável local (*Locally varying anisotropy kriging - LVA*);
- ▶ Manchuck and Deutsch (2018) propuseram a utilização de mínimos quadrados móveis para incorporar interpretação manual e avaliar incerteza.

Qualquer método de interpolação pode ser usado para as distâncias assinaladas, até mesmo métodos do inverso da distância produzem resultados realistas. A krigagem e as funções de bases radiais permitem incorporar informação adicional através dos modelos de covariância, porém a literatura recomenda o uso de métodos globais, ou seja, que usa, todas as amostras em todas as estimativas.

Diferentes métodos de interpolação podem ser encontrados na literatura:

- \* housseini e deutcsch usaram iqd;
- \* SIlva usou krigagem ordinária global;
- \* Eu usei krigagem ordinária na minha dissertação;
- \* O Boisvert usa LVA kriging pra interpolar com anistropia local;
- \* e o jonh manchuck e o clayton propusera um novo método de interpolação que integra seções digitalizadas e amostras.

# Métodos de interpolação

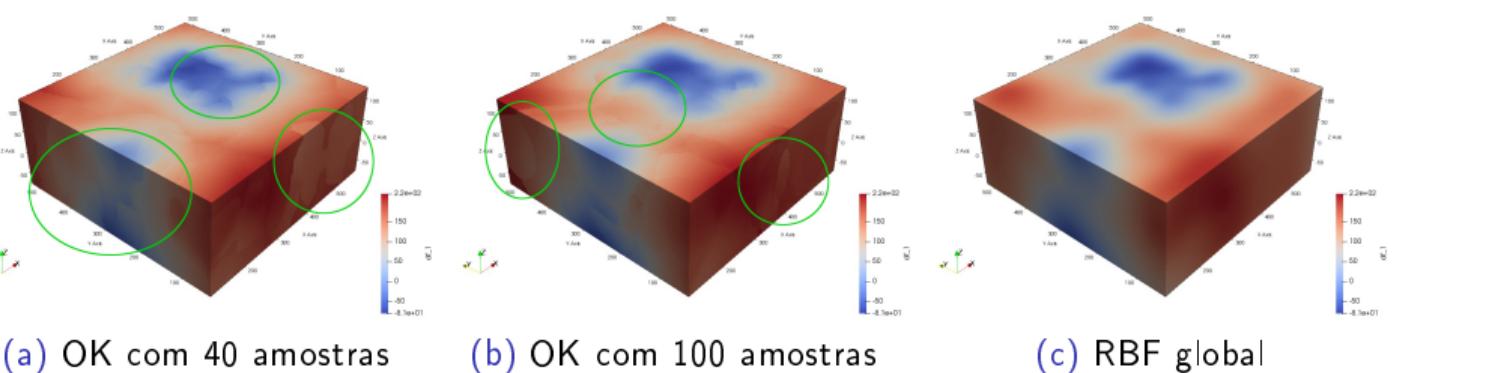


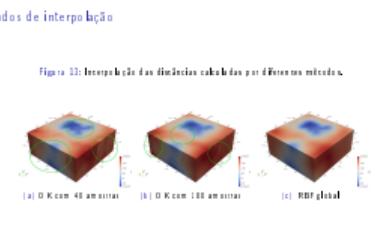
Figura 13: Interpolação das distâncias calculadas por diferentes métodos.

Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Interpolação das distâncias assinaladas
  - └ Métodos de interpolação

2019-06-11

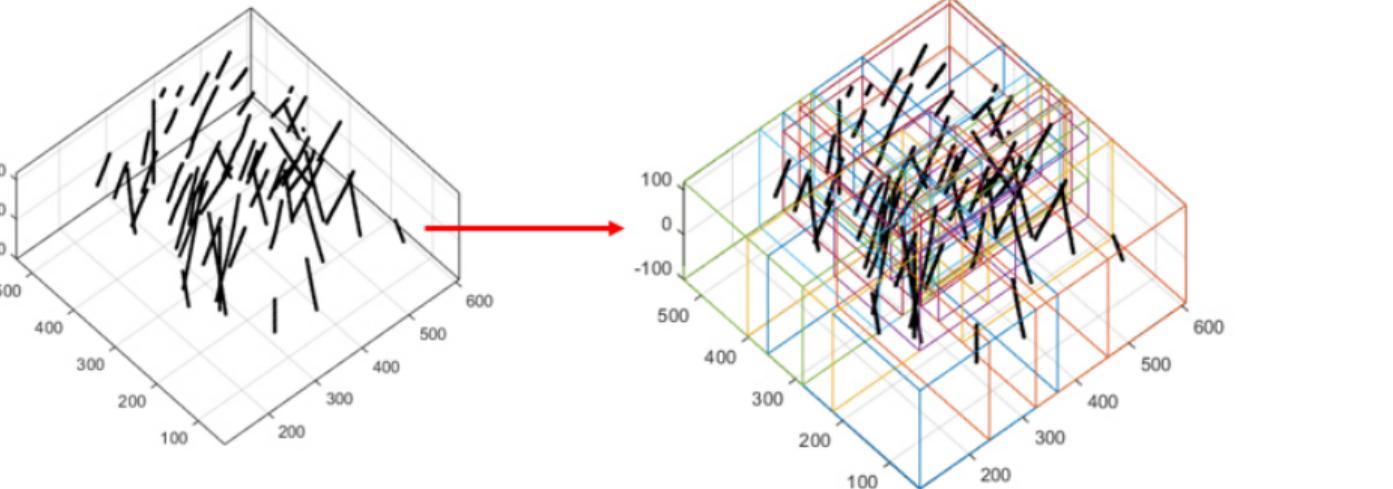
Esse slide mostra o porquê de usar métodos globais. O primeiro modelo foi interpolado por krigagem ordinária usando 40 amostras por estimativa, e fica bem evidente o surgimento de artefatos, causados pela estratégia de busca. O segundo modelo foi interpolado com 100 amostras por estimativa e mesmo assim ainda é possível observar artefatos no modelo. O último modelo foi interpolado por RBF, que é um método global e gera modelos suaves e livres de artefatos.



## Decomposição do domínio

Transforma um problema volumoso e que demanda muito esforço computacional em diversos problemas menores e eficientes que são, ao final, unidos.

Figura 14: Esquema mostrando o particionamento.



Proposta de tese para o exame de qualificação

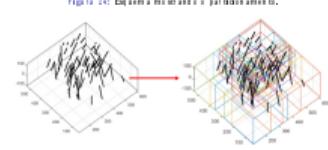
- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Interpolação das distâncias assinaladas
  - └ Decomposição do domínio

2019-06-11

Apesar de que nos métodos globais a matriz de covariâncias das amostras com elas mesmo só precise ser invertida uma vez, para bancos de dados volumosos isso ainda é uma tarefa que demanda um grande esforço computacional, então, é possível partitionar o problema em problemas menores, sobrepostos, que são mais eficientes, e unir todas as partes.

Decomposição do domínio  
Transforma um problema volumoso e que demanda muito esforço computacional em diversos problemas menores e eficientes que são, ao final, unidos.

Figura 14: Esquema mostrando o particionamento.



# Benchmark

Todos os algoritmos utilizados são da biblioteca GSLib e foram executados em um core i7 7700HQ @ 2.8 GHz com 16 Gb de RAM.

Método	Tempo grid grosso	Tempo grid fino	Classificação errônea grosso	Classificação errônea fino
krigagem global isotrópica	28min		121	
krigagem global anisotrópica	30min 34s		282	
krigagem ordinária anisotrópica (40)	1min 3s	38min	137	135
krigagem ordinária anisotrópica (100)		45min 51s		181
RBF isotrópico	21.5s		57	
RBF anisotrópico		1min 22s		38
Particionado RBF anisotrópico		1min 2s		39
Particionado RBF artefatos	16.5s			29
LVA OK			8	
LVA RBF				8
Krigagem dos indicadores		33min 27s		2

Proposta de tese para o exame de qualificação

Modelagem geológica implícita com funções distância assinaladas

Interpolação das distâncias assinaladas

Benchmark

2019-06-11

Benchmark

Todos os algoritmos utilizados são da biblioteca GSLib e foram executados em um core i7 7700HQ @ 2.8 GHz com 16 Gb de RAM.

Método	Tempo grid grosso	Tempo grid fino	Classificação errônea grosso	Classificação errônea fino
krigagem global isotrópica	28min		121	
krigagem global anisotrópica	30min 34s		282	
krigagem ordinária anisotrópica (40)	1min 3s	38min	137	135
krigagem ordinária anisotrópica (100)		45min 51s		181
RBF isotrópico	21.5s		57	
RBF anisotrópico		1min 22s		38
Particionado RBF anisotrópico		1min 2s		39
Particionado RBF artefatos	16.5s			29
LVA OK			8	
LVA RBF				8
Krigagem dos indicadores		33min 27s		2

Essa tabela mostra um benchmark dos métodos de interpolação para o banco de dados apresentado, as distâncias foram interpoladas por diferentes métodos no grid grosso que tem 100 mil nós e no grid fino com 1 milhão.

Todos os algoritmos são da biblioteca GSLib e foram rodados em um notebook comercial high end, para as três categorias.

Eu queria destacar alguns pontos: não é possível interpolar as distâncias no grid fino por krigagem global, mesmo com 16gb de ram. A krigagem ordinária no grid fino, isso para as três categorias, leva 45 minutos, enquanto o RBF leva apenas pouco mais de um minuto, ainda é possível reduzir esse tempo partionando o grid.

Não é difícil perceber porque o RBF é o método preferido para interpolação de distâncias assinaladas, além de produzir modelos suaves, desejáveis nesse contexto de modelagem geológica, é 10 vezes mais rápido que a krigagem.

A Leapfrog tem um algoritmo de rbf patenteado, chamado fast rbf, que é ainda mais rápido.

# Visualização do modelo geológico

Um bom palpite inicial para a interface que separa os domínios no espaço, seria a linha (em duas dimensões) ou superfície (em três dimensões) que corresponda ao valor zero da função distância assinalada

Figura 15: Iso superfícies para a categoria 1 extraída dos diferentes modelos implícitos.

(a) OK com 40 amostras

(b) OK com 100 amostras

(c) RBF global

Proposta de tese para o exame de qualificação  
└ Modelagem geológica implícita com funções distância assinaladas  
  └ Visualização do modelo geológico  
    └ Visualização do modelo geológico

2019-06-11

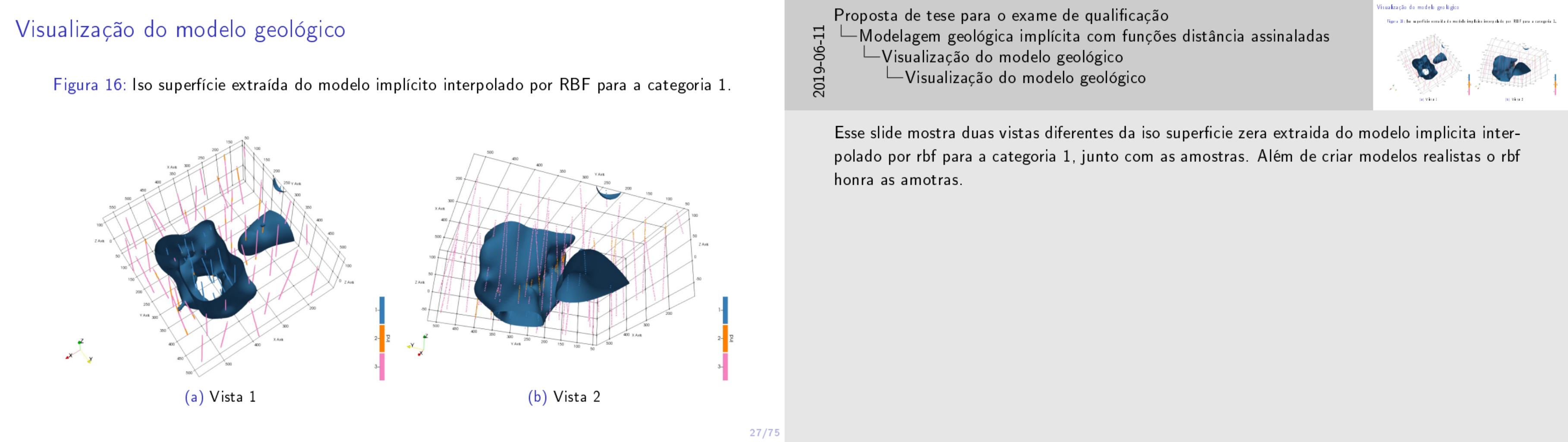
O modelo implícito, que é a função distância interpolada, tem infinitas iso superfícies, pra visualizar o modelo geológico eu preciso extrair uma dessas superfície, e o melhor palpite pra onde se localiza a interface é a iso superfície zero. esse slide mostra a isosuperfície zero extraída, usando o algoritmo de cubos marchantes no software paraview, para a categoria 1. A partir dos modelos implícitos apresentados anteriormente. krigagem ordinária com 40 amostras, 100 amostras e rbf.

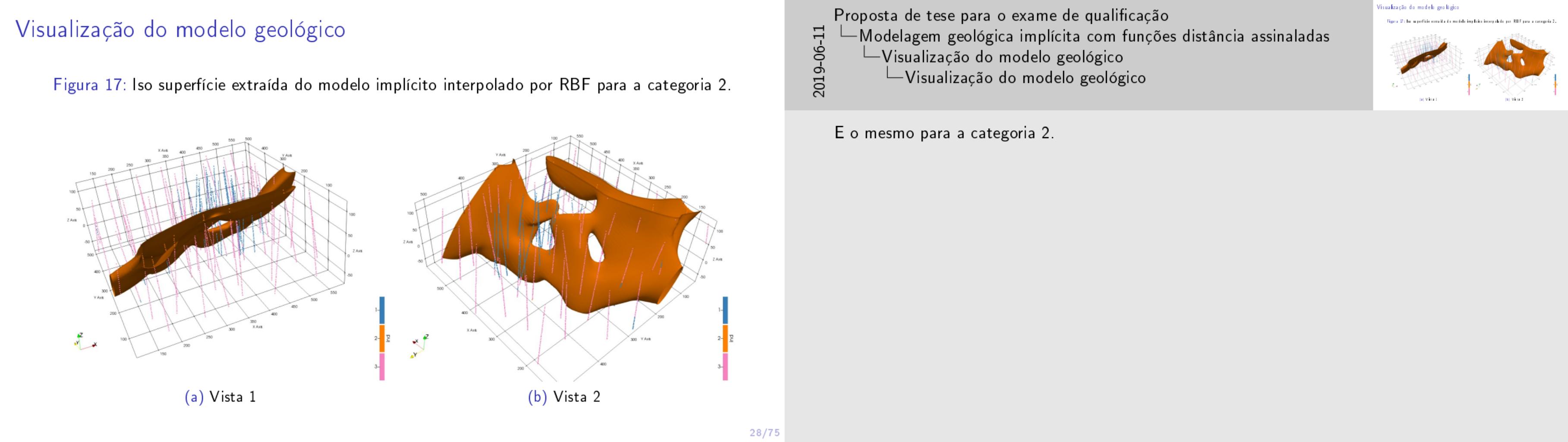
É possível observar que os artefatos dos modelos implícitos são transferidos para as isosuperfícies, e que métodos globais que geram modelos suaves, naturalmente geram iso superfícies suaves, que apresentam realismo geológico.

Visualização do modelo geológico  
Um bom palpite inicial para a interface que separa os domínios no espaço, seria a linha (em duas dimensões) ou superfície (em três dimensões) que corresponda ao valor zero da função distância assinalada

Figura 15: Iso superfícies para a categoria 1 extraída dos diferentes modelos implícitos.

(a) OK com 40 amostras (b) OK com 100 amostras (c) RBF global





## Adaptação para múltiplas categorias simultaneamente

$$\begin{aligned} i_k(u_\alpha) = & \begin{cases} 1, & \text{se } z(u_\alpha) = k \\ 0, & \text{caso contrário} \end{cases} \quad k = 1, \dots, K \end{aligned}$$

$$\begin{aligned} d_k(u_\alpha) = & \begin{cases} -\|u_\alpha - u_\beta\|, & \text{se } i_k(u_\alpha) = 1 \\ +\|u_\alpha - u_\beta\|, & \text{se } i_k(u_\alpha) = 0 \end{cases} \quad k = 1, \dots, K \end{aligned}$$

$$d_k^*(u) = \sum_{\alpha=1}^n \lambda_\alpha(u) d_k(u_\alpha) \quad k = 1, \dots, K \quad (3)$$

$$d_k^*(u) = \min_{k'=1}^K d_{k'}^*(u) \quad (4)$$

$$d_k^*(u) = \min_{k'=1}^K \lambda_{k'}(u) d_{k'}(u) \quad k = 1, \dots, K \quad (5)$$

$$i^*(u) = k' \text{ de modo que } d_{k'}^* = \min\{d_k^*(u)\}_{k=1}^K \quad (6)$$

## Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Adaptação para múltiplas categorias simultaneamente
    - └ Adaptação para múltiplas categorias simultaneamente

2019-06-11

$$\begin{aligned} i_k(u_\alpha) = & \begin{cases} 1, & \text{se } z(u_\alpha) = k \\ 0, & \text{caso contrário} \end{cases} \quad k = 1, \dots, K \end{aligned}$$

$$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\|, & \text{se } i_k(u_\alpha) = 1 \\ +\|u_\alpha - u_\beta\|, & \text{se } i_k(u_\alpha) = 0 \end{cases} \quad k = 1, \dots, K$$

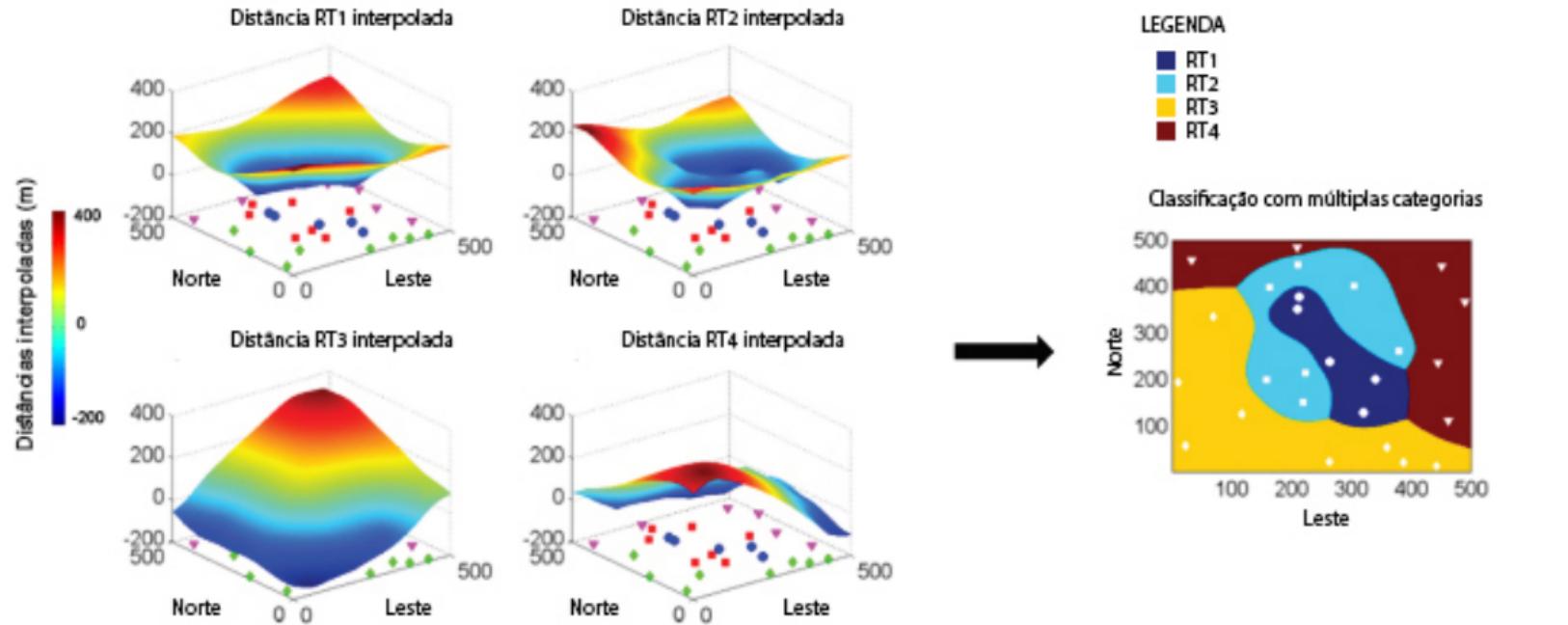
$$d_k^*(u) = \sum_{\alpha=1}^n \lambda_\alpha(u) d_k(u_\alpha) \quad k = 1, \dots, K$$

$$i^*(u) = k' \text{ de modo que } d_{k'}^* = \min\{d_k^*(u)\}_{k=1}^K$$

Silva e Deutsch propuseram uma adaptação para o método para modelar múltiplas categorias simultâneamente. Se existirem  $K$  múltiplos domínios no depósito mineral as amostras devem ser codificadas em indicadores  $K$  vezes, as distâncias assinaladas devem ser calculadas para as  $K$  categorias. As  $K$  propriedades distâncias assinaladas devem ser interpoladas para um mesmo grid, e a categoria responsável pela menor distância assinalada, a mais negativa, deve ser retida no local não amostrado.

# Adaptação para múltiplas categorias simultaneamente

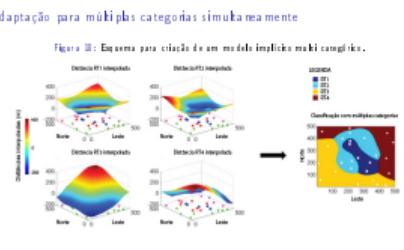
Figura 18: Esquema para criação de um modelo implícito multi categórico.



Proposta de tese para o exame de qualificação

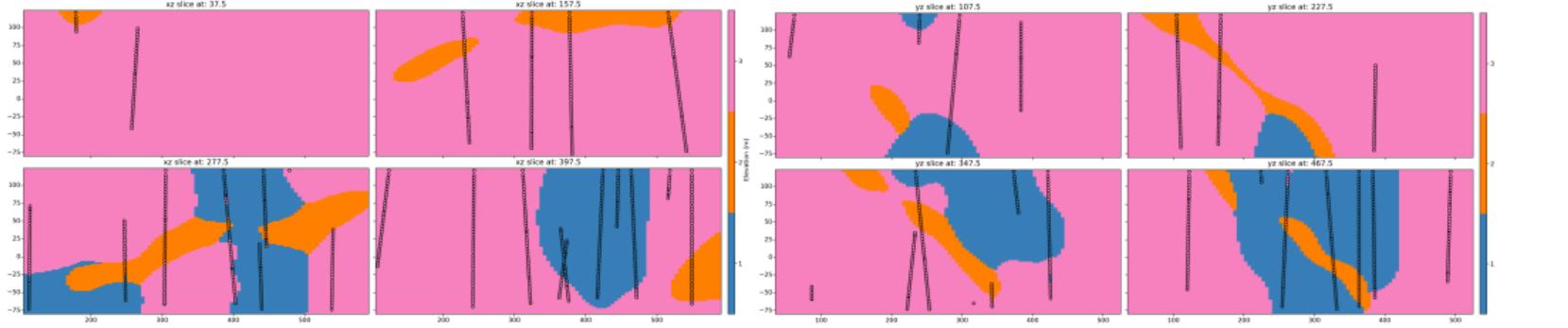
- └ Modelagem geológica implícita com funções distância assinaladas
- └ Adaptação para múltiplas categorias simultaneamente
- └ Adaptação para múltiplas categorias simultaneamente

2019-06-11



Esse slide mostra um esquema da modelagem geológica implícita para múltiplos domínios simultâneos: nas bases dos gráficos estão as amostras, no nível superior a projeção das distâncias assinaladas interpoladas, para categoria 1 as distâncias são negativas onde tem amostra da categoria 1, que é na parte central da área. mesma coisa para as categorias 2, 3 e 4. Pra cirar o modelo geológico é só reter a categoria responsável pela menor distância. na região central é a categoria 1, na parte noroeste a 3 nordeste a 4 e em volta do centro a categoria 2.

# Adaptação para múltiplas categorias simultaneamente



(a) Seções em XZ do modelo implícito gerado por RBF no *grid* fino.

(b) Seções em YZ do modelo implícito gerado por RBF no *grid* fino.

Figura 19: Modelo geológico multi categórico.

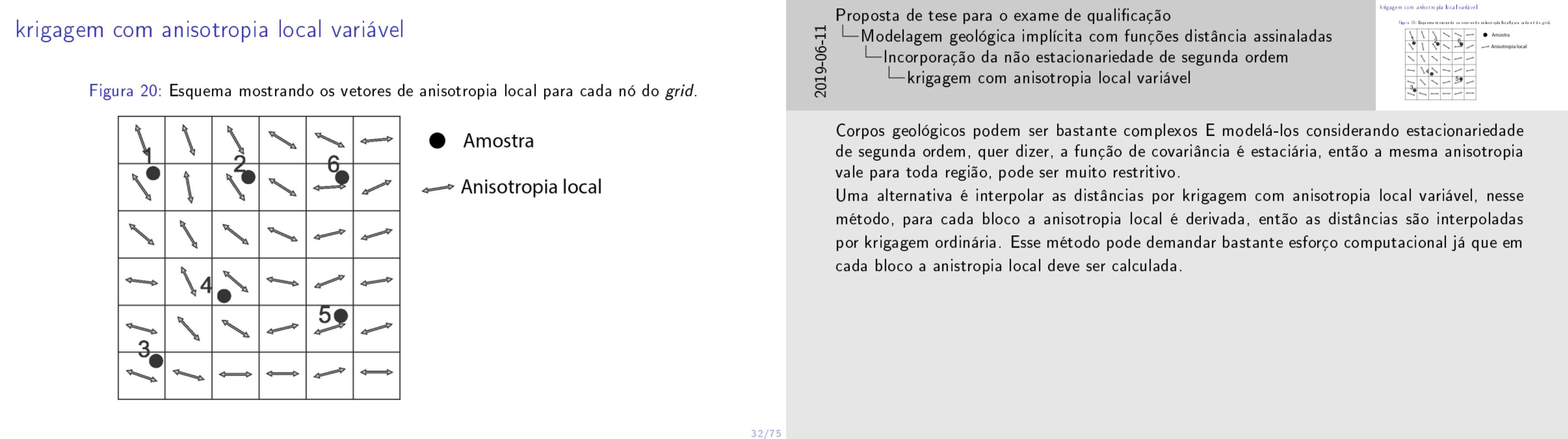
Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
- └ Adaptação para múltiplas categorias simultaneamente
- └ Adaptação para múltiplas categorias simultaneamente

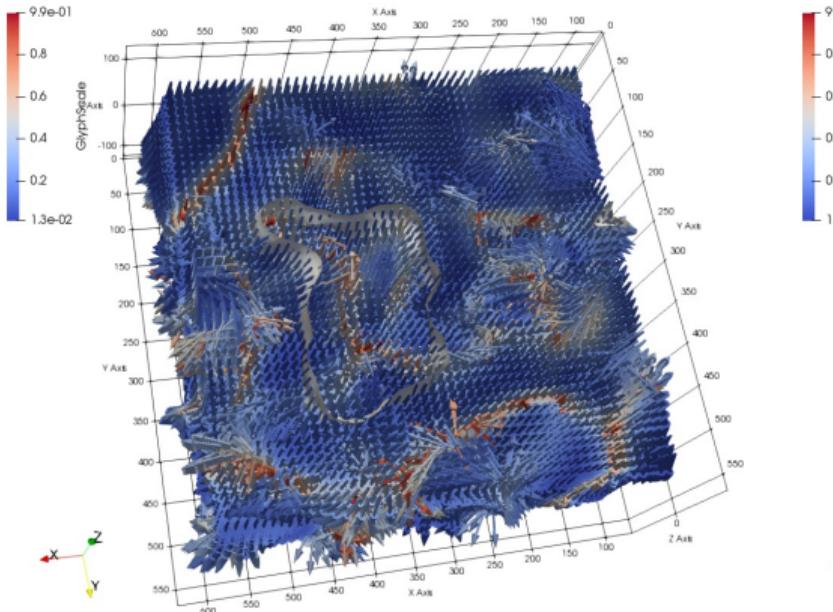
2019-06-11

Esse slide mostra 4 seções em xy e 4 seções em yz do modelo implícito multi categórico para o banco de dados do estudo de caso.

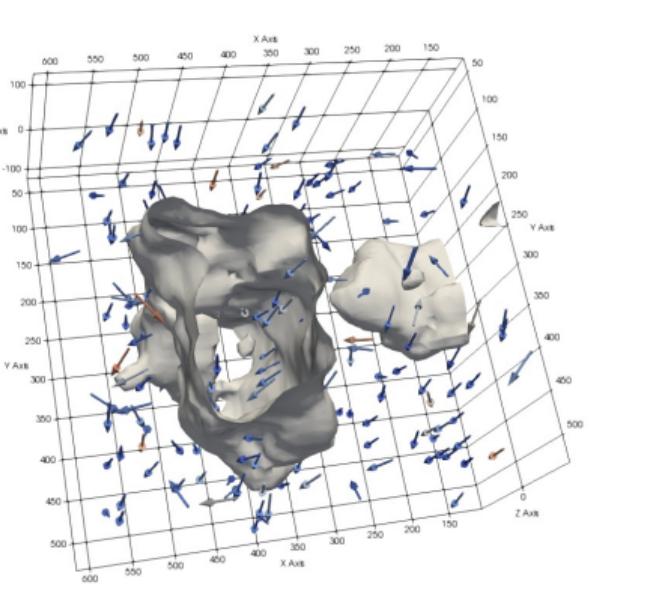




**Figura 21:** Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores.



(a) Todos os vetores de anisotropia local



(b) Um vetor a cada 100.000 blocos

## Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Incorporação da não estacionariedade de segunda ordem

2019-06-11

Esse slide mostra os vetores de anisotropia local em todos os um milhão de blocos do grid fino e a iso superfície extraída para a categoria 1, e ao lado mostrando um vetor a cada 10.000 blocos.

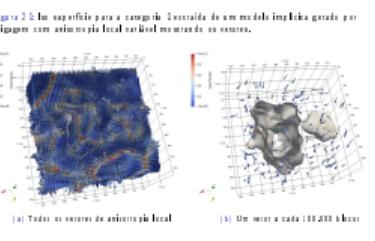
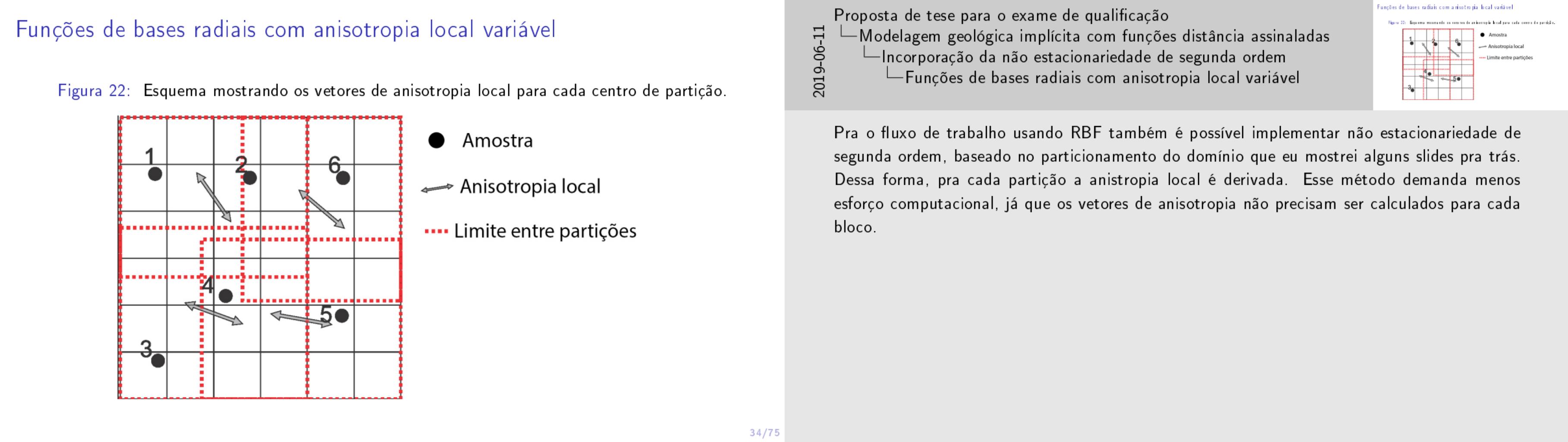
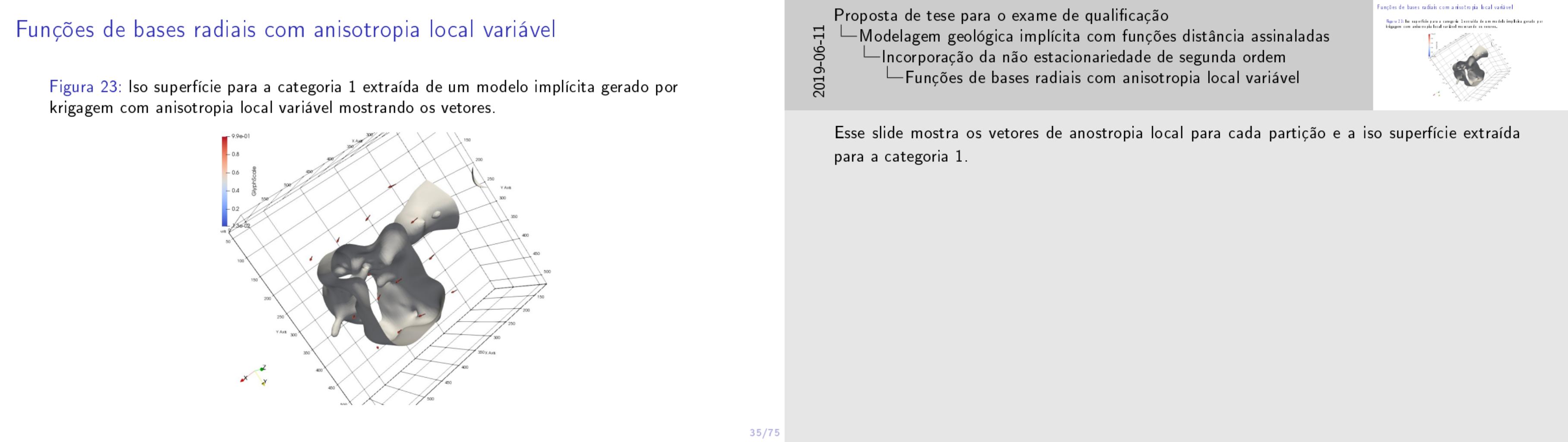


Figura 21 Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores.

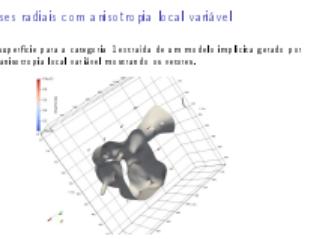




Proposta de tese para o exame de qualificação  
└ Modelagem geológica implícita com funções distância assinaladas  
  └ Incorporação da não estacionariedade de segunda ordem  
    └ Funções de bases radiais com anisotropia local variável

2019-06-11

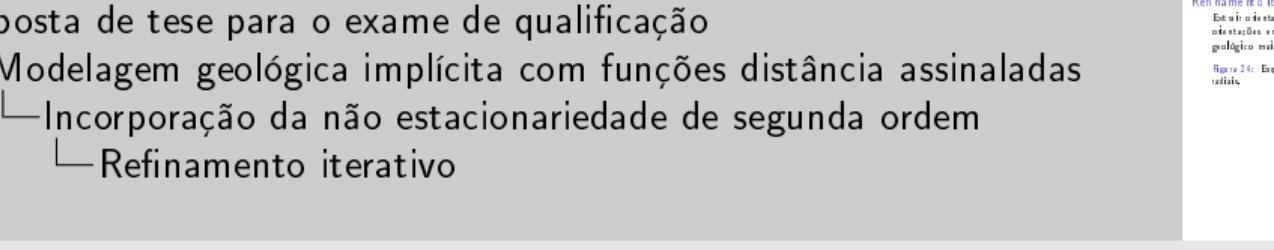
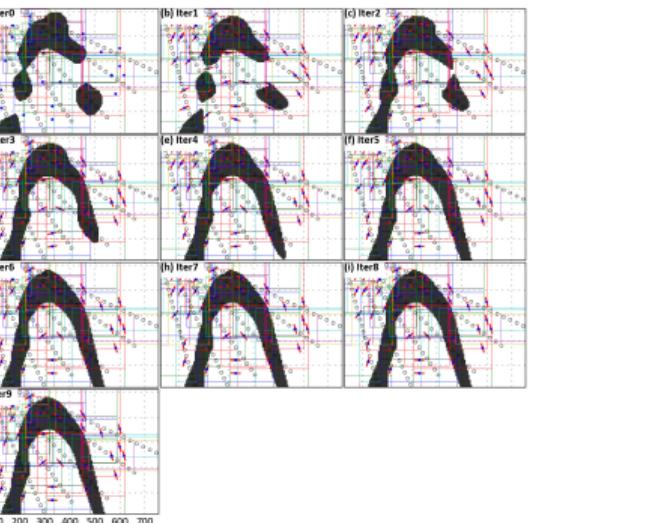
Esse slide mostra os vetores de anisotropia local para cada partição e a iso superfície extraída para a categoria 1.



## Refinamento iterativo

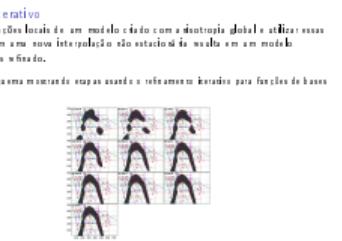
Extrair orientações locais de um modelo criado com anisotropia global e utilizar essas orientações em uma nova interpolação não estacionária resulta em um modelo geológico mais refinado.

**Figura 24:** Esquema mostrando etapas usando o refinamento iterativo para funções de bases radiais.



Pra obter resultados melhores é possível derivar os vetores de anisotropia local, interpolar as distâncias usando essa anisotropia local, então a partir das distâncias interpoladas, derivar novamente a anisotropia local e interpolar novamente, esse processo iterativo continua até que um critério de parada pré estabelecido seja atingido.

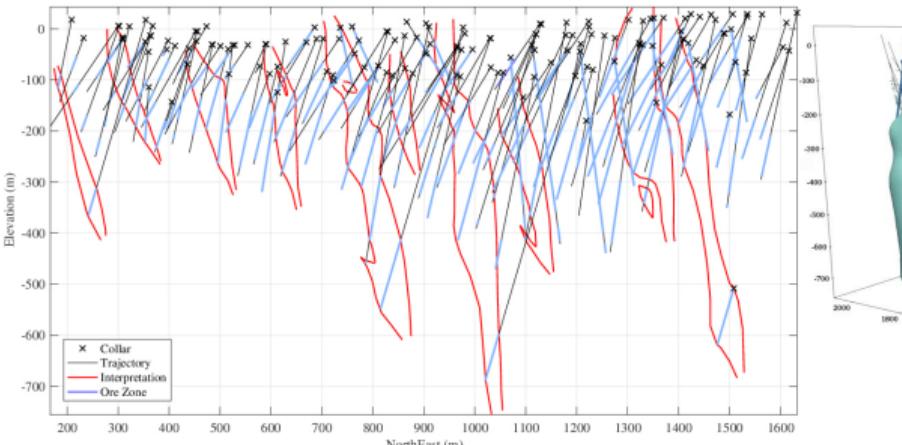
A imagem mostra uma dobra em que os flancos vão se definindo com as iterações. Apesar de ser um método iterativo, ele é bastante eficiente porque o critério de parada pode ser atingido em diferentes iterações em cada partição.



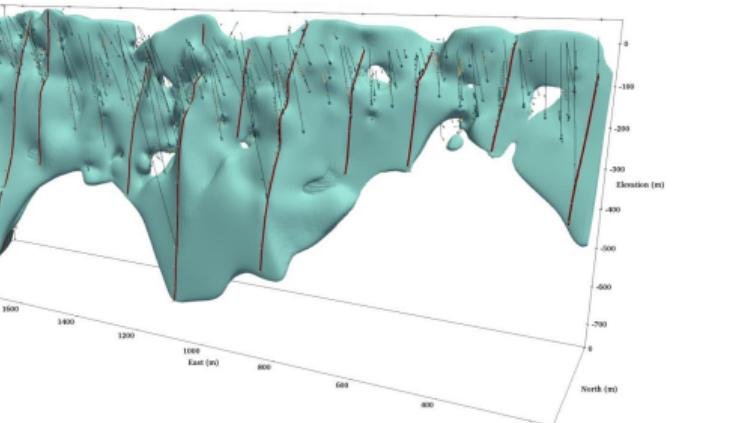
## Incorporação de informação secundária

Manchuck and Deutsch (2018) propõe o uso de uma regressão linear local para integrar modelagem geológica implícita e explícita.

Figura 25: Modelo geológico híbrido criado a partir de furos de sondagem e seções interpretadas.



(a) Seção mostrando furos e seções interpretadas.

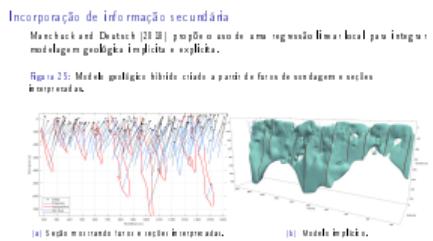


(b) Modelo implícito.

## Proposta de tese para o exame de qualificação

- └ Modelagem geológica implícita com funções distância assinaladas
  - └ Incorporação de informação secundária
  - └ Incorporação de informação secundária

2019-06-11



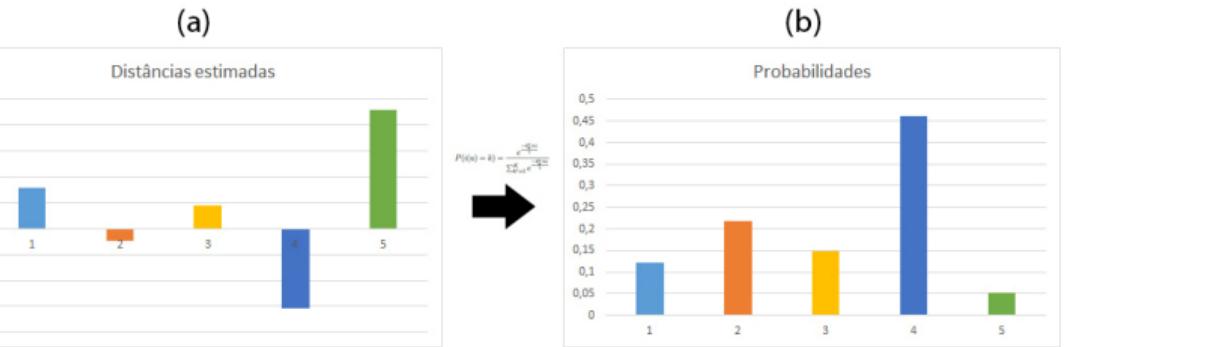
Ainda no assunto interpolação, o joh manchuck e o professor clayton desenvolveram uma metodologia para incorporar seções digitalizadas na modelagem implícita. Para isso as distancias mais seções digitalizadas devem ser interpoladas por mínimos quadrados móveis. Porém, o método ainda é recente e segundo os autores ainda apresenta alguns problemas operacionais.

# Avaliação heurística da incerteza

Transformação das distâncias em probabilidades.

$$P(i(u) = k) = \frac{e^{\frac{-d_k^*(u)}{\gamma}}}{\sum_{k'=1}^K e^{\frac{-d_{k'}^*(u)}{\gamma}}} \quad (7)$$

Figura 26: Distâncias estimadas em (a) e transformadas em probabilidades em (b) para um mesmo bloco, com cinco categorias.



Proposta de tese para o exame de qualificação

└ Avaliação da incerteza

  └ Avaliação heurística da incerteza

    └ Avaliação heurística da incerteza

Os próximos slides vão apresentar os métodos de avaliação de incerteza disponíveis na literatura e comentar um pouco sobre as características de cada um deles.

O método mais simples, é uma avaliação heurística da incerteza, porque não é baseado em múltiplas realizações. Consiste em transformar as distâncias em probabilidades. na equação para calcular a probabilidade da categoria k no local não amostrado i, o  $d^*$  é a distância estimada para a categoria k e gamma é um parâmetro que controla a incerteza. Quanto maior gamma maior a diferença entre as probabilidades calculadas.

o grafico de barras mostra as distâncias assinaladas interpoladas para um bloco em particular, e ao lado as distâncias transformadas em probabilidades, quanto menor a distância, quanto mais negativa, maior a probabilidade.

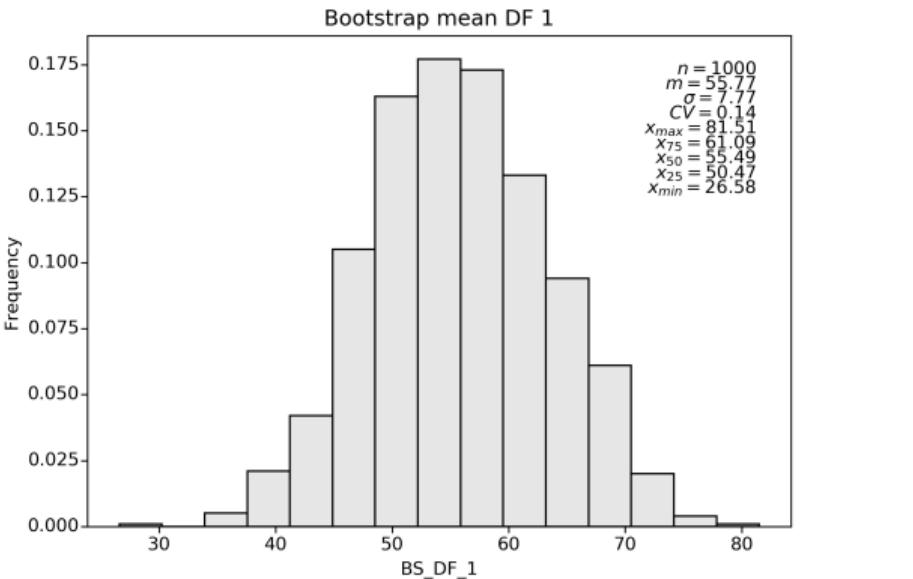
Esse é um método bem simples, rápido e direto e funciona pra multiplas categorias simultâneas. Em contrapartida, a escolha do parâmetro gamma é subjetiva, alguns autores indicam que seja a maior distância estimada entre todas as categorias, e dependendo do gamma escolhido nem blocos co locados com amostras recebem prbabilidade 100 por cento.



# BOUNDSIM

Krigagem simples com médias tomadas do histograma da média.

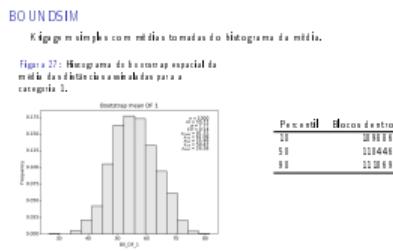
**Figura 27:** Histograma do bootstrap espacial da média das distâncias assinaladas para a categoria 1.



Proposta de tese para o exame de qualificação

- └ Avaliação da incerteza
- └ BOUNDSIM
- └ BOUNDSIM

2019-06-11



O próximo método, aumentando um pouco a complexidade, consiste, primeiramente em realizar um bootstrap espacial da média da distância assinalada, o bootstrap espacial só leva em consideração amostras não correlacionadas. Assim eu tenho um histograma da média, que é mostrado no slide, a partir desse histograma eu posso tomar valores, como o p10, p50 e p90 por exemplo, e interpolar as distâncias assinaladas por krigagem simples, usando esses valores tomados como média. Em tese, eu poderia extrair diferentes iso superfícies zero dos modelos interpolados com diferentes médias.

Porém na prática, dependendo da configuração espacial das amostras a krigagem simples não é sensível à media, para a categoria do banco de dados do estudo de caso, a diferença de volume entre as isosuperfícies extraídas do modelo p10 e p90 é menor que 1 por cento. O método não avalia adequadamente a incerteza, além disso, trabalha com uma categoria por vez, na presença de múltiplos domínios é necessário uma abordagem heirárquica.

## Simulação direta das distâncias assinaladas

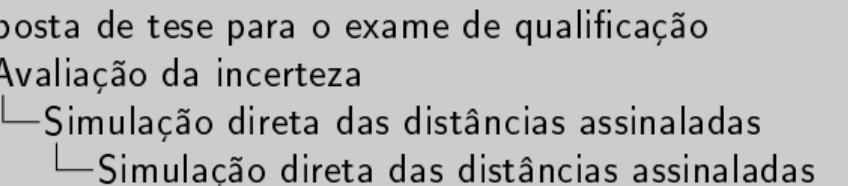
Simulação direta e classificação dos blocos baseada na menor distância simulada.  
O primeiro passo é o cálculo do coeficiente U:

$$U(u) = \frac{\max\{D_{min}\} - \min\{d_k^*(u)\}_{k=1}^K}{\max\{D_{min}\} - \min\{D_{min}\}} \quad (8)$$

Onde:

$$D_{min} = \{\min\{d_k^*(u_1)\}, \dots, \{\min\{d_k^*(u_n)\}\}_{k=1}^K\} \quad (9)$$

E  $d_k^*(u)$  é a distância estimada no local  $u$  para a categoria  $k$ .



2019-06-11

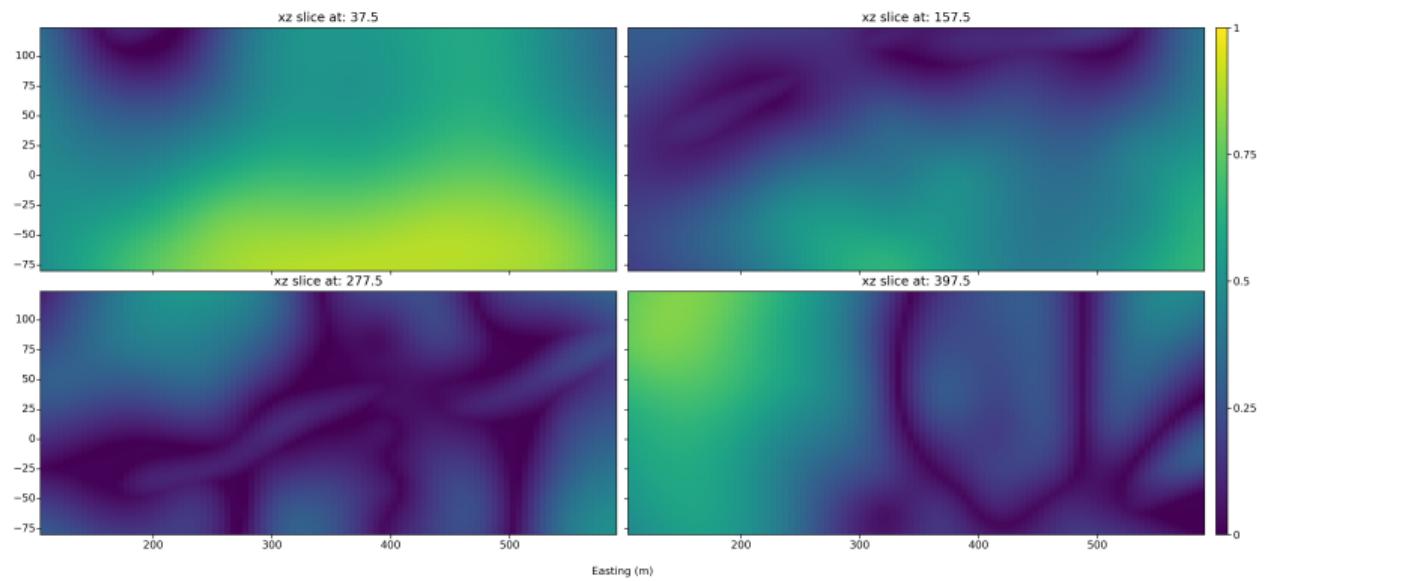
Simulação direta e classificação dos blocos baseada na menor distância simulada.  
O primeiro passo é o cálculo do coeficiente U:  
$$U(u) = \frac{\max\{D_{min}\} - \min\{d_k^*(u)\}_{k=1}^K}{\max\{D_{min}\} - \min\{D_{min}\}}$$
  
Onde:  
$$D_{min} = \{\min\{d_k^*(u_1)\}, \dots, \{\min\{d_k^*(u_n)\}\}_{k=1}^K\}$$
  
E  $d_k^*(u)$  é a distância estimada no local  $u$  para a categoria  $k$ .

O próximo método é a primeira ideia que surge em mente em relação à avaliação de incerteza de modelos geológicos implícitos: a simulação direta das distâncias.

O primeiro passo é calcular um coeficiente U de incerteza. Antes de qualquer coisa eu preciso criar uma lista com as menores distâncias interpoadas em cada bloco. Então pra cada bloco eu eu subtraio do valor maximo dessa lista a menor distância estimada naquele bloco. O denominador é pra estandardizar entre 0 e 1.

# Simulação direta das distâncias assinaladas

Figura 28: Coeficiente U calculado para todos os nós do grid.



Proposta de tese para o exame de qualificação  
└ Avaliação da incerteza  
  └ Simulação direta das distâncias assinaladas  
    └ Simulação direta das distâncias assinaladas

2019-06-11

Esse slide mostra 4 seções verticais do coeficiente U, esse coeficiente basicamente delinea os contatos, ele é mais próximo de 1 onde há mais incerteza e mais próximo de zero onde não há. Agora é necessário fazer uma truncagem entre 0 e 1 no coeficiente U pra delimitar uma zona de incerteza, quanto mais perto de 1 mais estreita é a zona de incerteza. As distâncias não precisam ser simuladas em todos os nós do grid, porque no interior dos domínios não há incerteza.

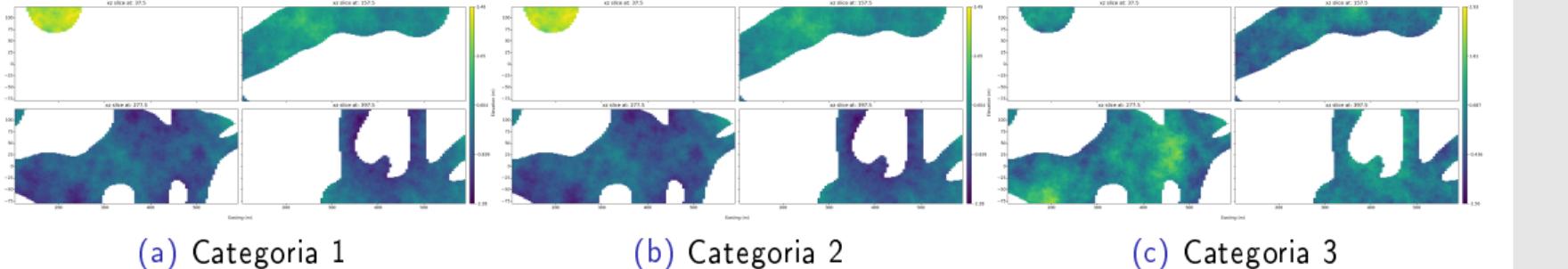
Simulação direta das distâncias assinaladas



# Simulação direta das distâncias assinaladas

Simulação das distâncias na zona de incerteza:

Figura 29: Distâncias simuladas na zona de incerteza para as categorias do banco de dados.

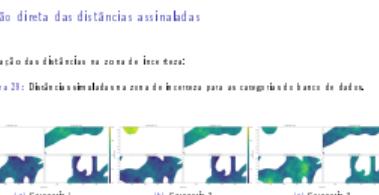


- Proposta de tese para o exame de qualificação  
└ Avaliação da incerteza  
  └ Simulação direta das distâncias assinaladas  
    └ Simulação direta das distâncias assinaladas

2019-06-11

Esse slide mostra uma realização das distâncias simuladas na zona de incerteza pra cada uma das litologias do banco de dados.

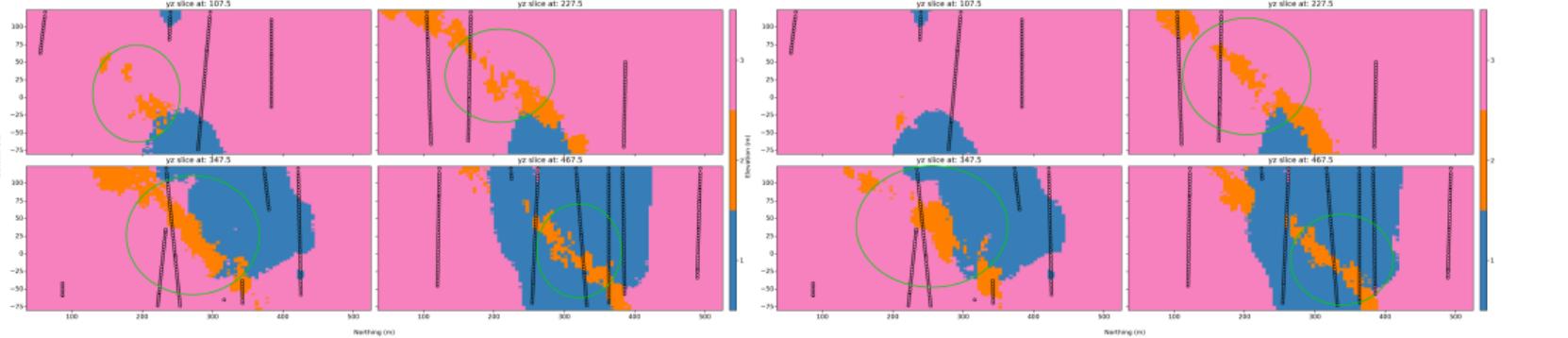
Os blocos fora da zona de incerteza ficam congelados, e recebem a categoria responsável pela menor distância estimada, enquanto os blocos na zona de incerteza recebem a categoria responsável pela menor distância simulada, em cada realização.



# Simulação direta das distâncias assinaladas

Classificação dos blocos.

Figura 30: Diferentes realizações do modelo geológico.



(a) Realização 1

(b) Realização 2

- Proposta de tese para o exame de qualificação
- └ Avaliação da incerteza
  - └ Simulação direta das distâncias assinaladas
  - └ Simulação direta das distâncias assinaladas

2019-06-11



Esse slide mostra seções verticais de duas realizações do modelo. É possível observar nas regiões destaca bastante ruído, que gera formas não realistas. Além disso, em muitos casos, não existe um bom encaixe entre os blocos congelados e os blocos simulados. A simulação é muito sensível aos parâmetros e a zona de incerteza definida só serve pra diminuir o numero de blocos simulados, os limites dos modelos dependem somente dos parâmetros do algoritmo de simulação. Porém o método trabalha com múltiplas categorias e é baseado em múltiplas realizações, o resultado final pode ser usado nas proximas etapas da avaliação de recursos e planejamento mineiro.

# Simulação direta das distâncias assinaladas

1. Cálculo das distâncias assinaladas para todas as amostras e categorias;
2. Variografia das distâncias no espaço original para todas as categorias;
3. Interpolação das distâncias;
4. criação do modelo com base na menor distância interpolada;
5. Criação da zona de incerteza;
6. Transformação Gaussiana das distâncias;
7. Variografia das distâncias no espaço gaussiano para todas as categorias;
8. Geração de múltiplos modelos baseados na menor distância simulada;
9. Validação e pós processamento das realizações.

2019-06-11

- Proposta de tese para o exame de qualificação
- └ Avaliação da incerteza
  - └ Simulação direta das distâncias assinaladas
  - └ Simulação direta das distâncias assinaladas

- Simulação direta das distâncias assinaladas
- 1. Cálculo das distâncias assinaladas para todas as amostras e categorias;
  - 2. Variografia das distâncias no espaço original para todas as categorias;
  - 3. Interpolação das distâncias;
  - 4. criação do modelo com base na menor distância interpolada;
  - 5. Criação da zona de incerteza;
  - 6. Transformação Gaussiana das distâncias;
  - 7. Variografia das distâncias no espaço gaussiano para todas as categorias;
  - 8. Geração de múltiplos modelos baseados na menor distância simulada;
  - 9. Validação e pós processamento das realizações.

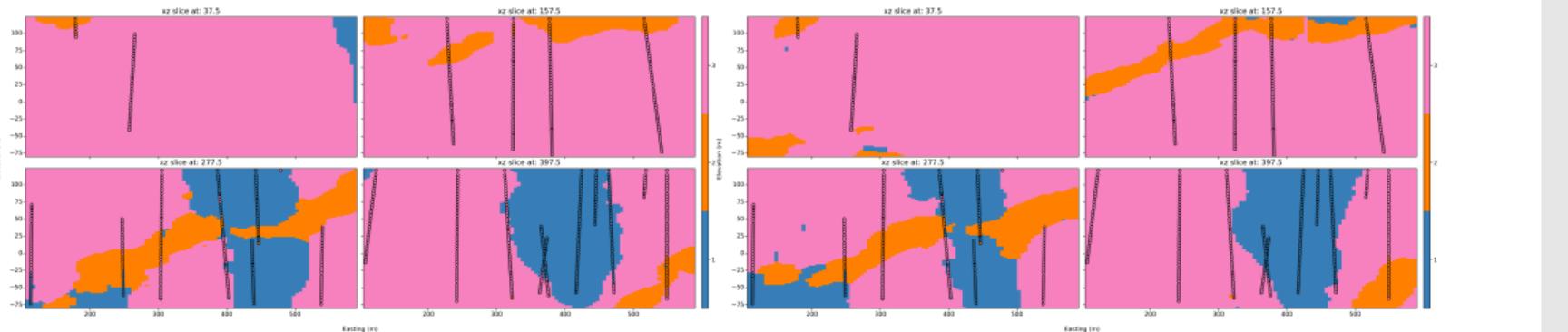
Além das desvantagens apresentadas o método possui um número execivo de passos:  
As distâncias devem ser calculadas pra cada litogia, variogradas e interpoaldas. Um modelo deterministico baseado na menor distancia deve ser criado, com base nas distâncias interpoladas posso calcular o parametro u e definir a zona de incerteza, entao as distâncias devem ser transformadas para o espaço gaussiano, variografadas no espaço gaussiano, simuladas na zona de incerteza, retro transformadas e validadas. Aí sim posso gerar as realizações do modelo geológico baseadas na menor ditância simulada.

É difícil justificar a escolha desse método.

## Simulação multi ponto

MPS em uma TI gerada pela interpolação das distâncias assinaladas.

Figura 31: Diferentes realizações do modelo geológico.



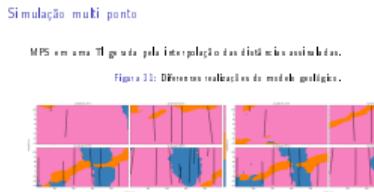
(a) Realização 1

(b) Realização 2

Proposta de tese para o exame de qualificação

- └ Avaliação da incerteza
- └ Simulação multi ponto
- └ Simulação multi ponto

2019-06-11



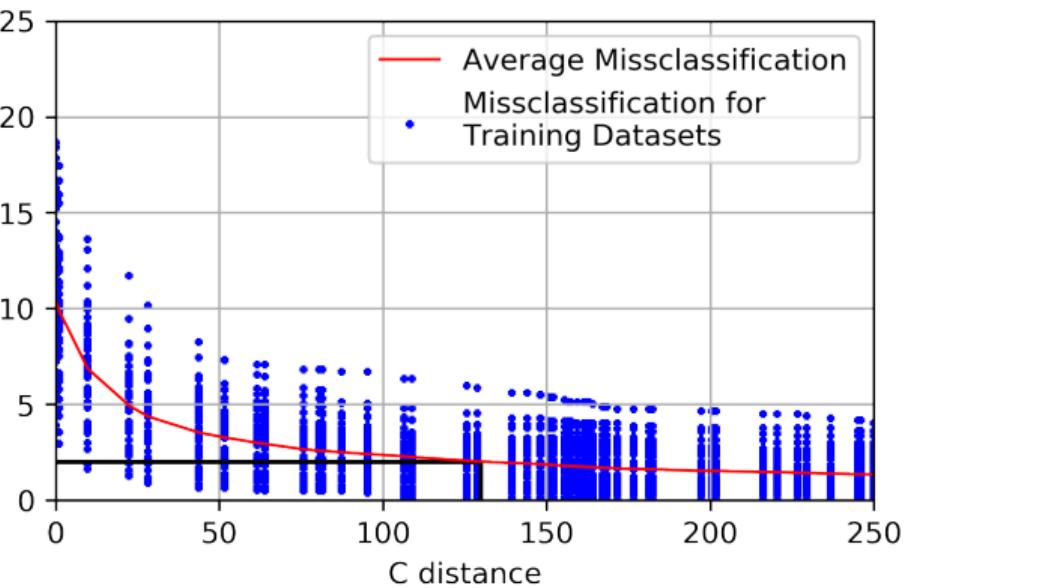
Em sua tese de doutorado Silva propôs integrar modelagem geológica implícita com simulação multi ponto. No seu trabalho ele propõe integrar múltiplas imagens de treinamento, uma metodologia pra calibrar a contribuição de cada TI e uma medida de entropia multi ponto ao longo dos furos. O esqueleto do método é criar uma TI a partir dos dados, que ele chama de data driven training image, usando funções distância assinaladas e aplicar MPS nessa TI.

O slide mostra seções verticais de uma das realizações de um modelo criado usando essa metodologia no banco de dados do estudo de caso. As formas são geológicas, os modelos não apresentam ruído excessivo, além disso, o método não depende de muitos parâmetros. Em contra partida, não é possível definir uma zona de incerteza para que os contatos variem em seu interior, tampouco controlar a natureza dos contatos.

## Boundary simulation

### Calibração do parâmetro C:

Figura 32: Calibração do parâmetro C para a categoria 1.



Proposta de tese para o exame de qualificação

- └ Avaliação da incerteza
- └ Boundary simulation
- └ Boundary simulation

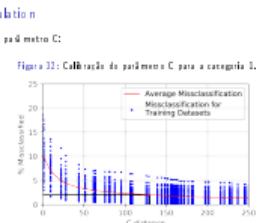
2019-06-11

O próximo e último método, é o que produz os melhores resultados e o que tem sido mais usado pra avaliação de incerteza de modelos geológicos implícitos. O primeiro passo é calibrar um parâmetro de incerteza de C. A calibração é feita por um método similar ao jackknife.

O parametro C é uma constante que vai ser somada às distâncias positivas e subtraído das distâncias negativas, isso faz com a diferença entre os valores das distâncias seja maior, é como se eu tivesse abrindo o histograma das distancias.

Eu começo a calibração com  $c = 0$ , então as distâncias não estão modificadas ainda. Eu defino uma proporção de furos do banco de dados pra ser removida, pra categoria 1 do banco de dados do estudo de caso eu escolhi 25%. Então eu removi do banco de dados, aleatoriamente, 25% dos furos, 50 vezes, gerando 50 bancos de dados diferentes. Que são os pontos verticais no  $x=0$ . Então eu vou estimar a distância assinaladas onde eu removi as amostras, e checar se eu acerto se pertence ou não ao domínio, quer dizer se eu acerto o sinal da amostra, não estou interessado no valor da distância assinalada. Aí eu marco no eixo y, o índice de amostras classificadas de forma errada, pra cada um dos 50 banco de dados.

Então eu incrementei C, nesse exemplo o c variou de 0 ate 250, e repito o processo. quanto



## Boundary simulation

*y'(u)*

$$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\| - C, & \text{se } u_\alpha \text{ pertence ao domínio} \\ +\|u_\alpha - u_\beta\| + C, & \text{se } u_\alpha \text{ não pertence ao domínio} \end{cases} \quad (10)$$

Para que a simulação seja realizada de forma uniforme entre  $-C$  e  $+C$ , o desvio padrão  $y'(u)$ , deve ser simulado e transformado pela relação:

$$df'(u) = 2 * C * G^{-1}(y'(u)) - C \quad (11)$$

onde:  $df'(u)$  é o valor da função distância simulada,  $y'(u)$  o valor normal padrão da simulação não condicional, e  $G^{-1}$  representa a determinação do valor da distribuição acumulada padrão normal correspondente a  $y'(u)$ . Para garantir que os valores pertençam a região estabelecida, os valores são multiplicados por  $2C$  e subtraídos de  $C$ .

## Proposta de tese para o exame de qualificação

- └ Avaliação da incerteza
- └ Boundary simulation
- └ Boundary simulation

2019-06-11

Boundary simulation

$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\| - C, & \text{se } u_\alpha \text{ pertence no domínio} \\ +\|u_\alpha - u_\beta\| + C, & \text{se } u_\alpha \text{ não pertence no domínio} \end{cases}$  [10]  
Para que a simulação seja realizada de forma uniforme entre  $-C$  e  $+C$ , o desvio padrão  $y'(u)$ , deve ser simulado e transformado pela relação:  
 $df'(u) = 2 * C * G^{-1}(y'(u)) - C$  [11]  
onde:  $df'(u)$  é o valor da função distância simulada,  $y'(u)$  o valor normal padrão da simulação não condicional, e  $G^{-1}$  representa a determinação do valor da distribuição acumulada padrão normal correspondente a  $y'(u)$ . Para garantir que os valores pertençam a região estabelecida, os valores são multiplicados por  $2C$  e subtraídos de  $C$ .

A partir do valor escolhido para  $C$ , eu modifico as distâncias calculadas, somando  $C$  às distâncias positivas e subtraindo  $C$  das distâncias negativas.

Então essas distâncias modificadas são interpoladas para todos os nós do grid, uma truncagem no modelo interpolado deve ser feita entre  $-c$  e  $+c$  pra definir a zona de incerteza.

Na zona de incerteza eu realizo uma simulação gaussiana não condicional, o variograma pode ser o mesmo usado para interpolar as distâncias assinaladas. Para tornar a simulação gaussiana entre  $+c$  e  $-c$  deve ser tomado seu valor acumulado, multiplicado por  $2c$  e subtraído de  $C$ .

## Boundary simulation

Classificação dos blocos.

Figura 33: Classificação dos locais comparando valores estimados e simulados.

The graph plots 'Função distância' (Distance Function) against 'u'. It features two curves: a green line labeled 'dist modificada interpolação' and a red line labeled 'valor simulado'. The horizontal axis is divided into three regions: 'Fora' (red), 'Dentro' (black), and 'Fora' (red). A vertical line marks the 'Limite simulado'. The green curve starts at a positive value, crosses zero at the boundary, and then oscillates. The red curve starts below zero, crosses zero at the boundary, and then oscillates. The intersection point between the two curves defines the 'Limite simulado'.

Proposta de tese para o exame de qualificação

2019-06-11

- └ Avaliação da incerteza
- └ Boundary simulation
- └ Boundary simulation

Boundary simulation

Classificação dos blocos.

Figura 33: Classificação dos locais comparando valores estimados e simulados.

A small version of the graph from Figure 33, showing the distance function versus position  $u$ . The x-axis is labeled 'u' and has markers for 'Fora', 'Dentro', 'Limite simulado', and 'Fora'. The y-axis is labeled 'Função distância'. The graph shows two curves: a green line labeled 'dist modificada interpolação' and a red line labeled 'valor simulado'. The intersection point between the two curves is labeled 'Limite simulado'.

Agora finalmente eu posso classificar os blocos comparando as distâncias assinaladas modificadas com a simulação não condicional, blocos onde a simulação é menor que a interpolação não pertencem ao domínio enquanto blocos em que a simulação é maior que a interpolação, pertencem ao domínio.

# Boundary simulation

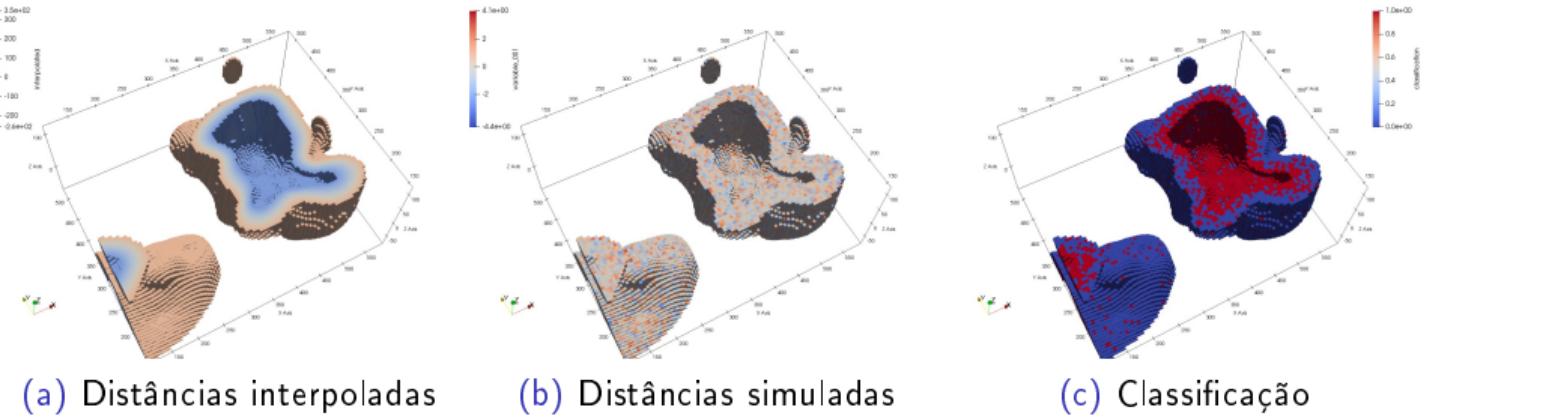
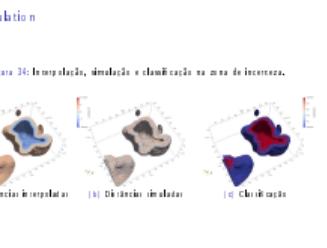


Figura 34: Interpolação, simulação e classificação na zona de incerteza.

Proposta de tese para o exame de qualificação  
└ Avaliação da incerteza  
  └ Boundary simulation  
    └ Boundary simulation

2019-06-11

Esse slide mostra as distâncias interpoladas na zona de incerteza para a categoria 1, as distâncias simuladas na zona de incerteza, e os blocos classificados comparando distâncias intepoladas e simuladas.



## Boundary simulation

Blocos classificados.

**Figura 35:** Seções verticais de uma realização para a categoria 1.

(a) Seção em XY

(b) Seção em YZ

Proposta de tese para o exame de qualificação

2019-06-11

- └ Avaliação da incerteza
- └ Boundary simulation
- └ Boundary simulation

Boundary simulation

Blocos classificados.

Figura 35: Seções verticais de uma realização para a categoria 1.

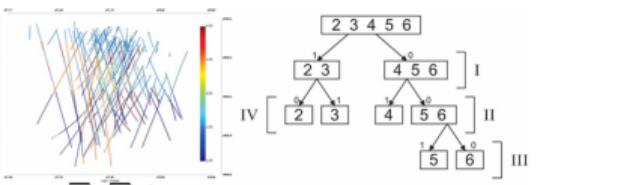
O próximo slide mostra seções verticais em xy e xz de uma das realizações para a categoria 1. É possível observar que não há presença de ruído, as formas são suaves e realistas, além disso, ao contrário da simulação direta das distâncias, aqui o contato varia dentro da zona de incerteza, se a zona for larga a posição do contato e consequente o volume do sólido vai ter grande variação, se for estrito terá pouca variação, o variograma controla a natureza do contato, um range maior produz contatos suaves enquanto ranges menores e efeito pepita produzem contatos rugosos. Apesar dessas vantagens a escolha do C pode ser subjetiva, e o método trabalha de forma binária, uma litologia por vez.

50/75

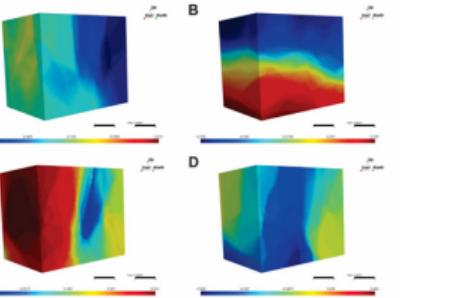
# Abordagem hierárquica

Figura 36: Esquema do método mostrando os passos necessários.

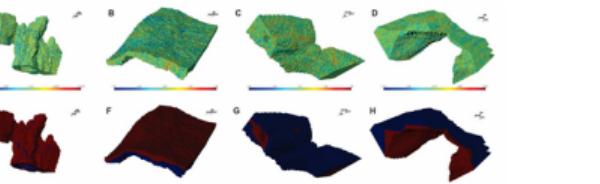
## 1-Definição dos grupos



## 2-Interpolação das distâncias assinaladas modificadas para cada grupo



## 3-Definição da zona de incerteza para cada grupo e simulação das distâncias



## 4-Definição do modelo geológico aplicando a mesma regra do passo 1 em cada realização

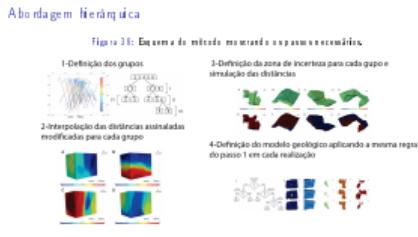


## Proposta de tese para o exame de qualificação

- └ Avaliação da incerteza
- └ Boundary simulation
- └ Abordagem hierárquica

2019-06-11

Aqui no laboratório nós desenvolvemos uma metodologia pra trabalhar com múltiplas categorias simultâneas de forma hierárquica, porém a escolha dos grupos é subjetiva e depende de conhecimento a respeito da gênese do depósito.



# Sumário dos métodos de avaliação de incerteza

Proposta de tese para o exame de qualificação

- Avaliação da incerteza
- Sumário dos métodos de avaliação de incerteza
- Sumário dos métodos de avaliação de incerteza

Sumário dos métodos de avaliação de incerteza

Método	Simplicidade	Velocidade	Multi categorico	Realismo geológico	Controle da incerteza	Controle do tipo de contato
Heurístico	simples	rápido	sim	não	sim	não
Boundsim	simples	rápido	não	sim	não	não
Simulação direta	complexo	demorado	sim	não	sim	não
MPS	simples	rápido	sim	sim	não	não
Boundary simulation	simples	rápido	não	sim	sim	sim

2019-06-11

Proposta de tese para o exame de qualificação

- Avaliação da incerteza
- Sumário dos métodos de avaliação de incerteza
- Sumário dos métodos de avaliação de incerteza

Sumário dos métodos de avaliação de incerteza

Essa tabela mostra um sumário dos métodos de avaliação de incerteza apresentados, o método heuristico é simple's, rápido, multi categorico, porem nao produz modelo com realismo geologico, mas controla a incerteza por meio do parametro gamma mas nao o tipo de contato.

o metodo boundsim é simples, rapido, não trabalha com diferentes categorias, não apresenta realismo geologico, em tese seria possivel controlar a incerteza tomando diferentes valores para a media, mas isso nao acontece na pratica, também nao é possivel controlar o tipo de contato.

A simulacao direta das distancias é um metodo complexo e demorado que trabalha com multiphas categorias simultaneamente porem nao apresenta realismo geologico nem permite o controle do tipo de contato, mas permite controlar a incerteza por meio dos parametros da simulacao.

O metodo que mescla mps com modelagem implicita é simples rapido, multicategorico apresenta realismo geologico mas nao permite controlar a incerteza nem o tipo do contato.

Finalmente o ultimo método, boundary simulation, é simples, rapido ja que a simulacao nao é condicional, nao é multi categorico mas apresenta realismo geologico, permite controle da incerteza e do tipo de contato. é o método que marca mais checkboxes disponivel na literatura.

Isso encerra o estado da arte e os próximos slides tratam da minha proposta de tese

## Problemas I

- ▶ Não estacionariedade da função distância assinalada torna a modelagem dos variogramas arbitrária e questionável;
- ▶ É preciso encontrar um balanço entre número de nós e resolução necessária. A resolução do *grid* influencia diretamente a avaliação de incertezas;
- ▶ A escolha do interpolador muitas vezes é subjetiva e confusa;
- ▶ Na presença de múltiplos domínios, principalmente em ambientes geológicos complexos, é necessário a aplicação de uma lógica de precedência de estruturas ao invés de simplesmente tomar a menor distância assinalada para a criação de modelos realistas;
- ▶ Em alguns métodos a definição da zona de incerteza é subjetiva e não segue nenhuma regra matemática ou geológica, em outros a definição da zona de incerteza é extremamente laboriosa e complicada;
- ▶ Estruturas geológicas específicas, como lentes ou diques, podem desaparecer, ou não serem bem reproduzidas nos modelos implícitos;

## Proposta de tese para o exame de qualificação

- └ Proposta de tese
- └ Problemas
- └ Problemas

2019-06-11

- Problemas I
- ▶ Não estacionariedade da função distância assinalada torna a modelagem dos variogramas arbitrária e questionável;
  - ▶ É preciso encontrar um balanço entre número de nós e resolução necessária. A resolução do *grid* influencia diretamente a avaliação de incertezas;
  - ▶ A escolha do interpolador muitas vezes é subjetiva e confusa;
  - ▶ Na presença de múltiplos domínios, principalmente em ambientes geológicos complexos, é necessário a aplicação de uma lógica de precedência de estruturas ao invés de simplesmente tomar a menor distância assinalada para a criação de modelos realistas;
  - ▶ Em alguns métodos a definição da zona de incerteza é subjetiva e não segue nenhuma regra matemática ou geológica, em outros a definição da zona de incerteza é extreamente laboriosa e complicada;
  - ▶ Estruturas geológicas específicas, como lentes ou diques, podem desaparecer, ou não serem bem reproduzidas nos modelos implícitos;

Embora a modelagem geoética implícita seja um método consagrado que vem sendo usado com sucesso há mais de uma década, a metodologia apresenta alguns problemas pontuais:

- \* o ponto mais critico, que todos os revisores de artigo questionam e é uma dificuldade prática pra quem usa método também é a não estacionariedade das distâncias, que torna a modelgaem do variograma questionável;
- \* além disso, a escolha dos parametros do grid pode ser subjetiva;
- \* Assim como a escolha do interpolaor;
- \* Na presença de multiplos dominios, para geologias complexas, simplesemnte tomar a menor distância não produz bons resultados, é necessário uma abordagem hierárquica baseada na gênese do depósito;
- \* Nos métodos de avaliação de incerteza que usam algum tipo de zona de incerteza, muitas vezes, essa zona não é de fato incerta, existem amotras dentro das zona;
- \* Estructuras geológicas específicas, como lentes, diques, falhas, dobradas, podem não ser reproduzida ou desaparecerem dos modelos;
- \* é necessário checar se os modelos implícitos honram a geologia do depósito e serão úteis nas

## Problemas II

Proposta de tese para o exame de qualificação

- └ Proposta de tese
- └ Problemas
- └ Problemas

2019-06-11

► É necessário checar se os modelos implícitos honram a geologia do depósito e serão úteis para o processo de avaliação de recursos/reservas.

Problemas II

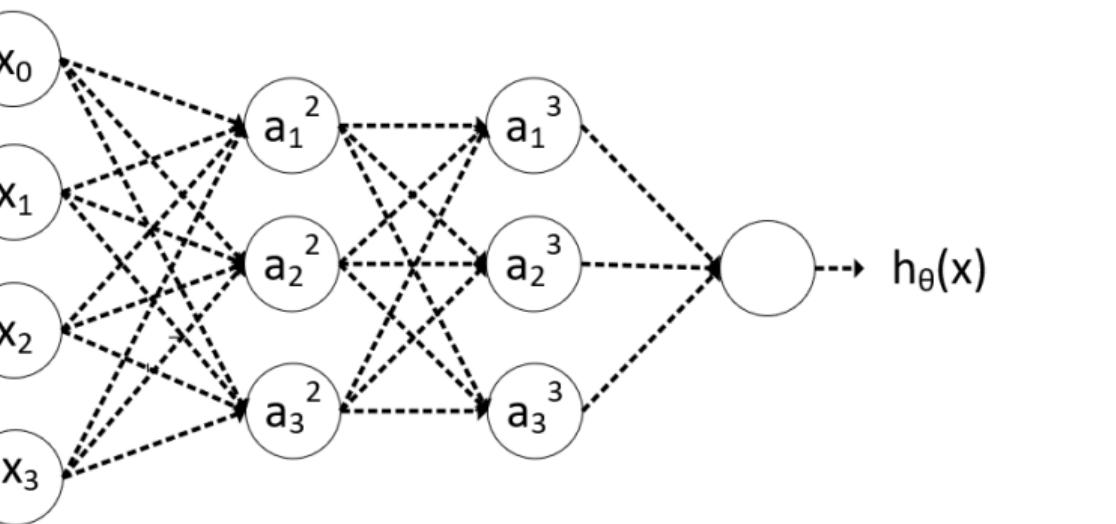
► É necessário checar se os modelos implícitos honram a geologia do depósito e serão úteis para o processo de avaliação de recursos/reservas.

- É necessário checar se os modelos implícitos honram a geologia do depósito e serão úteis para o processo de avaliação de recursos/reservas.

# Interpolador

Redes neurais trabalham em um sistema binário de zeros e uns e por isso podem se basear nas propriedades da função sigmóide.

Figura 37: Rede neural.



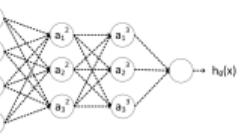
Proposta de tese para o exame de qualificação  
└ Proposta de tese  
  └ Interpolador  
    └ Interpolador

2019-06-11

O primeiro aprimoramento que essa tese propõe é um novo interpolador para as distâncias assinaladas.

Redes neurais são baseadas em neurônios, e usadas tipicamente em problemas de identificação. Esse slide esquematiza uma rede neural. Os nós  $x$  na camada um da rede representam os inputs, os nós  $a$  na camada dois e três representam as camadas escondidas e a camada 4 é o nó de output e representa a hipótese.

As redes neurais trabalham em um sistema binário de zeros e uns e por isso podem se basear nas propriedades da função sigmoide, que tem seu valor em  $y$  entre zero e um para qualquer valor de  $x$ . Em uma rede neural uma hipótese é gerada para cada nó e é determinado se seu valor será zero ou um, um valor binário é passado para a próxima camada.



# Interpolador

Samson and Deutsch (2018) desenvolveram um algoritmo, usando TensorFlow.

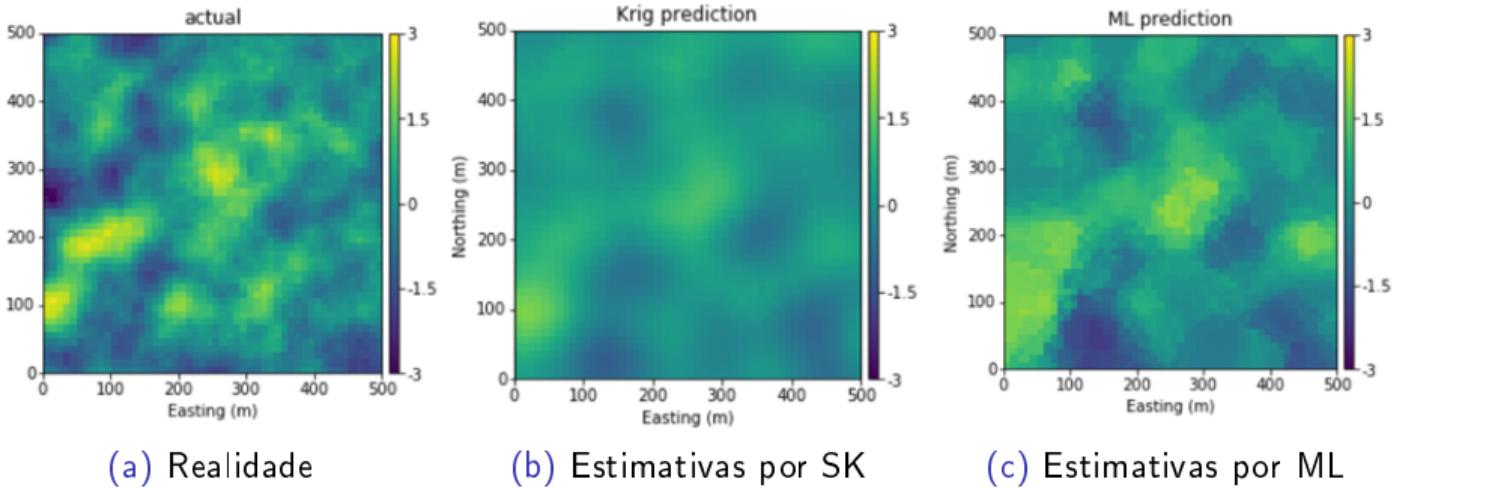
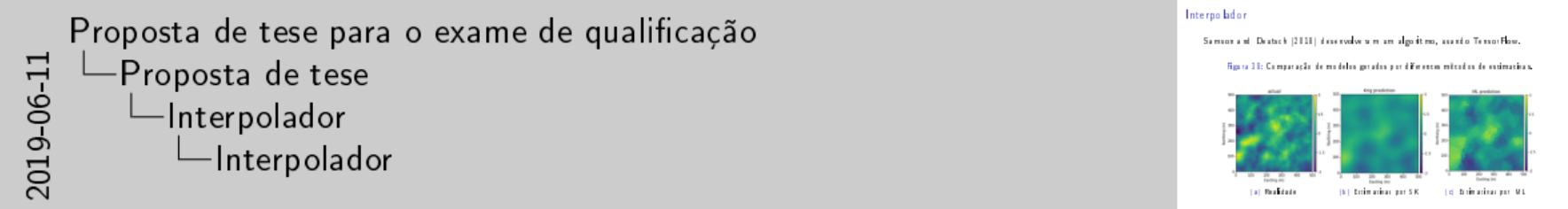


Figura 38: Comparação de modelos gerados por diferentes métodos de estimativas.



2019-06-11

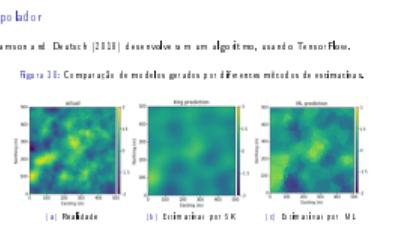
Samson e deutsh propuseram um algoritmo usando tensor flow, que é uma biblioteca da google de código aberto para aprendizado de máquina, para implementar a arquitetura da rede neural. O método é baseado em aprendizado não supervisionado, com o algoritmo k means e supervisionado com uma rede neural de bases radiais.

Redes neurais de bases radiais apresentam uma arquitetura semelhante as redes neurais artificiais, a principal diferença é que nas redes neurais de bases radiais só há uma camada escondida e a função de ativação é uma função de base radial ao invés de uma função sigmoide.

Em seu trabalho os autores dividiram o banco de dados em 2/3 para treinamento e 1/3 para teste. Usaram o algoritmo k-means pra determinar o numero de unidades de bases radiais na rede neural. A função de ativação alimentou o algoritmo com coordenadas x, y e z e teores das 5 amostras mais próximas do local a ser estimado.

As imagens mostram mostram a realidade estimativas por krigagem simples e estimativas pelo metodo proposto, que reprodu as caracteristicas visuais da imagem de referencia bem como o variograma e histograma.

Esse algoritmo resolve o problema do interpolador e do variograma ja que essa tecnica nao exige calculo e modelagem da covariânci. A pesquisa de samsom e deutsch ainda está em seus estágios



## Zona de incerteza

Gerar múltiplas realizações com a finalidade de avaliar a incerteza de modelos geológicos em todos os nós do *grid* é desperdício de tempo e poder computacional, já que no interior dos domínios não há incerteza.

**Figura 39:** Zonas de incerteza e contatos.

The figure consists of two side-by-side geological cross-sections labeled (a) and (b). Both diagrams show a red base layer labeled 'Litologia 2' and a blue top layer labeled 'Litologia 1'. A green wavy line represents a contact between the two layers. In diagram (a), the contact is relatively straight and horizontal at the top, with a narrow green band labeled 'Zona de incerteza' extending from it. In diagram (b), the contact is more complex and undulating, with a wider green band labeled 'Zona de incerteza' extending from it. Labels 'contato1' and 'contato2' are placed near the contacts in both diagrams.

(a) Zona de incerteza 1

(b) Zona de incerteza 2

Proposta de tese para o exame de qualificação

- Proposta de tese
- Zona de incerteza
- Zona de incerteza

2019-06-11

Um outro objetivo dessa tese é desenvolver uma metodologia objetiva e direta para a definição da zona de incerteza. As metodologias disponíveis na literatura podem gerar zonas de incerteza em interseção com amostras. É necessário definir essas zonas de incerteza porque eu não preciso simular o grid inteiro, reduzindo o tempo e o esforço computacional e além disso, o método de avaliação de incerteza proposto nessa tese depende diretamente da largura da zona de incerteza, os contatos variam em seu interior, zonas de incerteza maiores gera maior variação de volume dos corpos geológicos.

Zona de incerteza

Gerar múltiplas realizações com a finalidade de avaliar a incerteza de modelos geológicos em todos os nós do *grid* é desperdício de tempo e poder computacional, já que no interior dos domínios não há incerteza.

**Figura 39:** Zonas de incerteza e contatos.

The figure consists of two side-by-side geological cross-sections labeled (a) and (b). Both diagrams show a red base layer labeled 'Litologia 2' and a blue top layer labeled 'Litologia 1'. A green wavy line represents a contact between the two layers. In diagram (a), the contact is relatively straight and horizontal at the top, with a narrow green band labeled 'Zona de incerteza' extending from it. In diagram (b), the contact is more complex and undulating, with a wider green band labeled 'Zona de incerteza' extending from it. Labels 'contato1' and 'contato2' are placed near the contacts in both diagrams.

(a) Zona de incerteza 1

(b) Zona de incerteza 2

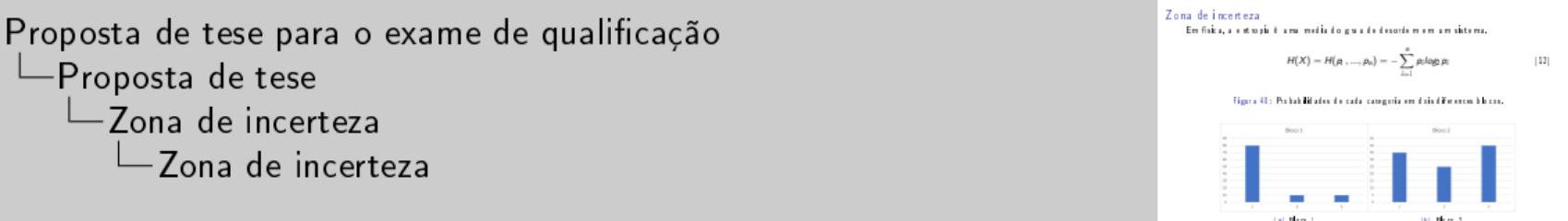
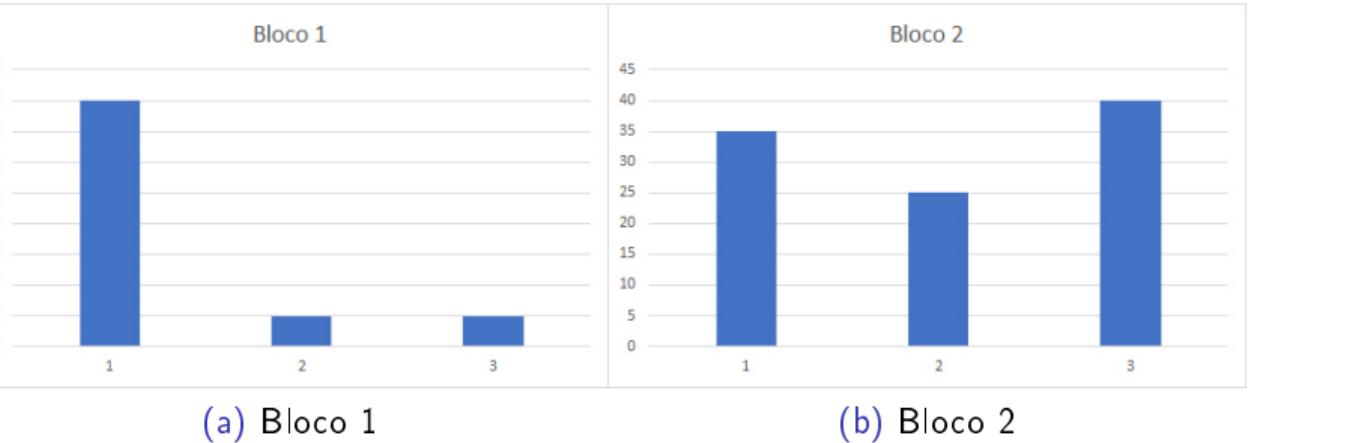
57 / 75

## Zona de incerteza

Em física, a entropia é uma media do grau de desordem em um sistema.

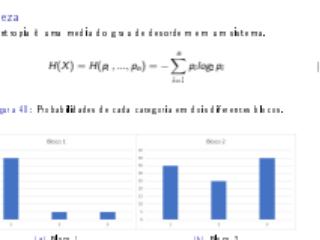
$$H(X) = H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log_2 p_i \quad (12)$$

Figura 40: Probabilidades de cada categoria em dois diferentes blocos.



A ideia central da metodologia proposta é a entropia, que é calculada como: menos o somatório do produto da probabilidade de cada de cada litologia pelo logaritmo da probabilidade. Dessa forma, a entropia pode ser vista como uma medida da incerteza.

Suponha dois blocos e as probabilidades deles pertencerem a cada uma de três litologias, as probabilidades podem ser facilmente calculadas transformando distâncias estimadas em probabilidades, como mostrado no método heurístico de avaliação de incerteza. A entropia do primeiro bloco é baixa, já que a categoria 1 tem probabilidade 80% enquanto a entropia do bloco 2 é alta, porque todas as categorias têm probabilidades similares.



## Zona de incerteza

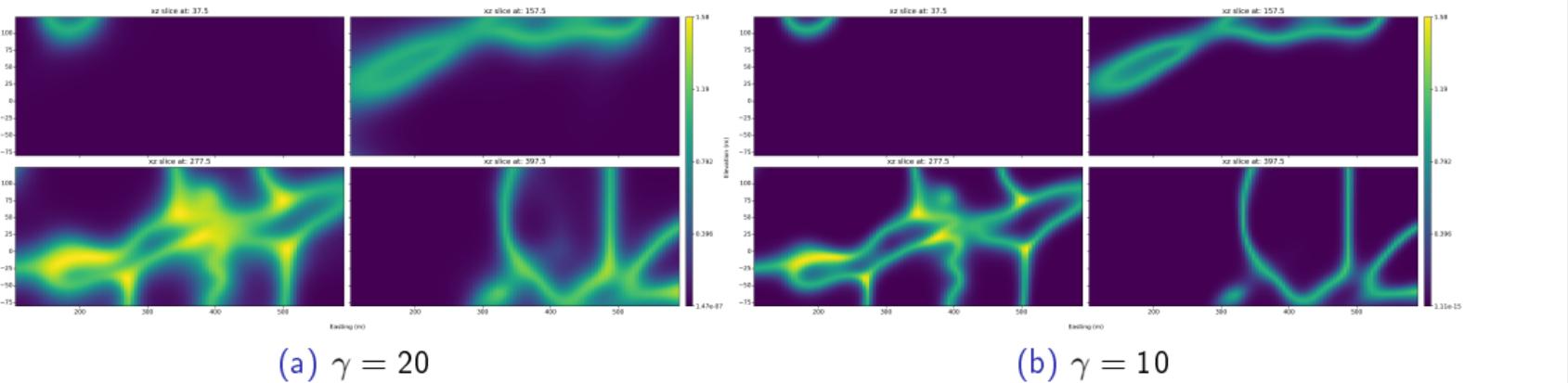


Figura 41: Entropias calculadas para diferentes valores de  $\gamma$ .

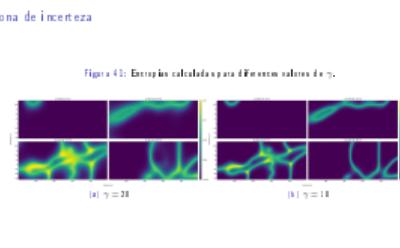
Proposta de tese para o exame de qualificação  
└ Proposta de tese  
  └ Zona de incerteza  
    └ Zona de incerteza

2019-06-11

A transformação das distâncias em probabilidades envolve um paâmetro gamma, que controla a incerteza, esse slide mostra duas seções verticais da entropia calculada para o banco de dados do estudo de caso, com dois valores de gamma diferentes.

O desafio é desenvolver uma metodologia para calibrar o parametro gamma, de forma que onde existam amostras a entropia seja máxima.

depois de calcular a entropia, uma truncagem deve ser realizada para definição da zona de incerteza. Essa truncagem pode ser baseada no tipo de depósito, depositos com menor entropia de formação, como de bauxita, por exemplo, terão uma zona de incerteza mais estreita. enquanto depositos com alta entropia de formação, como os de ouro, terão zonas de incertezas mais largas.



Zona de incerteza

Figura 42: Entropias calculadas para diferentes valores de  $\gamma$ .

# Boundary simulation multi categórico

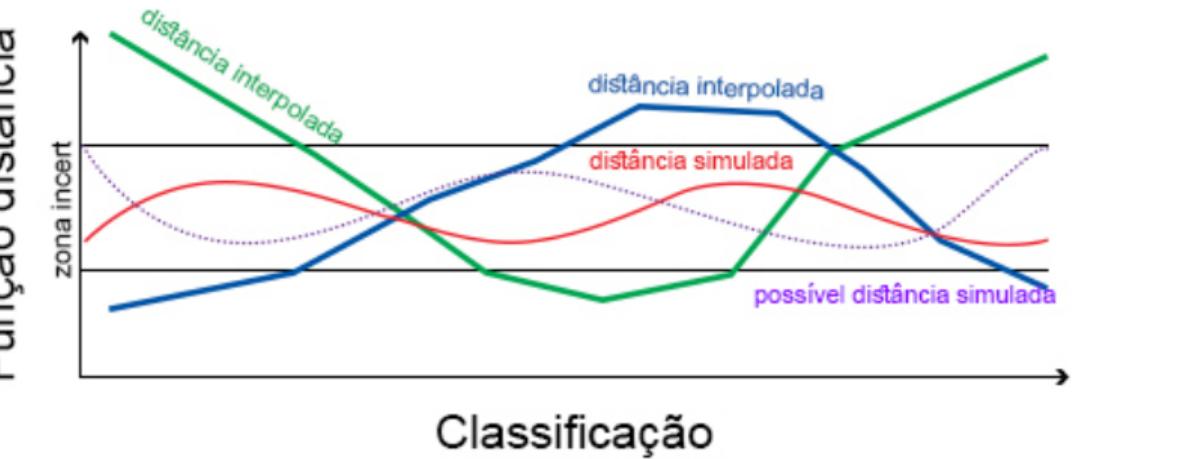


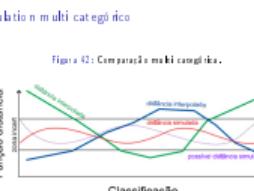
Figura 42: Comparação multi categórica.

- Proposta de tese para o exame de qualificação
- └ Proposta de tese
  - └ Avaliação da incerteza
  - └ Boundary simulation multi categórico

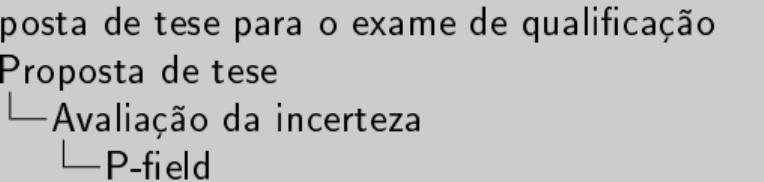
2019-06-11

Essa tese também propõe investigar diferentes técnicas para a avaliação da incerteza em modelos geológico multi categóricos. Para que os modelos gerados sejam livres de ruído, tenham contatos suaves e contínuos e para que o que método seja rápido e não exija muito esforço computacional é preciso que esses métodos sejam baseados em uma variável auxiliar contínua, gerada por simulação não condicional.

O primeiro método que eu quero investigar é baseado no boundary simulation, que mostrei alguns slides atrás. A diferença aqui é que pra múltiplas categorias preciso de múltiplas distâncias simuladas que serão comparadas às estimadas ou será necessário estandardizar as distâncias interpoladas para que possam ser comparadas com uma única variável simulada.



Uma alternativa à simulação não condicional das distâncias na zona de incerteza são os campos de probabilidade (*P-field*). A ideia central desse método de simulação é dissociar a tarefa de estimar distribuições de probabilidades locais para geração de múltiplas realizações equiprováveis Froidevaux (1993). Uma premissa é de que as distribuições locais são conhecidas, o que é uma premissa razoável, já que as distribuições locais podem ser calculadas transformando distâncias em probabilidades.



Uma alternativa à simulação não condicional das distâncias na zona de incerteza são os campos de probabilidade (*P-field*). A ideia central desse método de simulação é dissociar a tarefa de estimar distribuições de probabilidades locais para geração de múltiplas realizações equiprováveis Froidevaux (1993). Uma premissa é de que as distribuições locais são conhecidas, o que é uma premissa razoável, já que as distribuições locais podem ser calculadas transformando distâncias em probabilidades.

Uma outra avenida que eu acredito que vale a pena ser investigada é o o p-field ou campos de probabilidade. Um campo gaussiano é gerado na zona de incerteza, a partir dos dos valor acumulado, eu posso amostrar uma distribuição local de probabilidades, bloco a bloco, que foi criada a partir da transformação das distâncias em probabilidades.

## Simulação plurigaussiana truncada

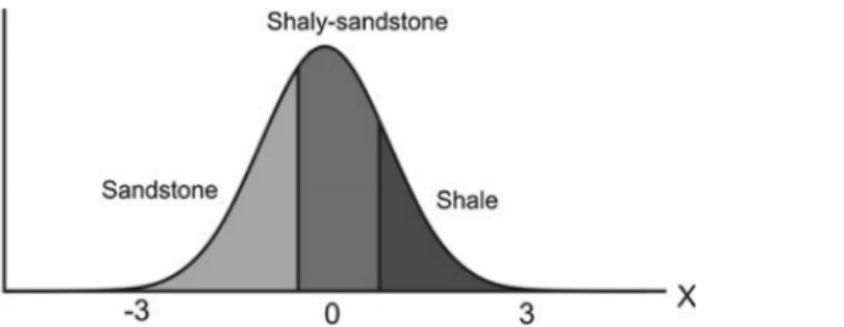


Figura 43: Esquema da SGT. Note que as litologias sandstone e shale não aparecerão juntas nas realizações.

- 2019-06-11
- Proposta de tese para o exame de qualificação
    - Proposta de tese
    - Avaliação da incerteza
    - Simulação plurigaussiana truncada



Ainda outra abordagem para a avaliação de incerteza que eu pretendo investigar é baseada em simulação plurigaussiana truncada.

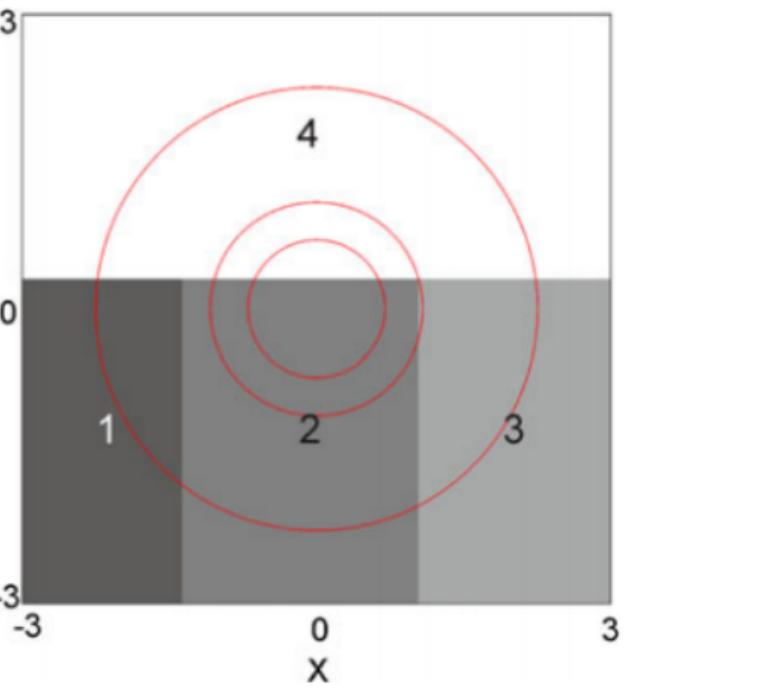
Na simulação gaussiana truncada, uma variável continua gaussiana é simulada, e o histograma de cada realização é truncado em limites pré definidos para gerar uma variável categórica.

A imagem mostra a representação de um histograma de uma variável gaussiana, entao toda vez que o valor simulado para um bloco cair nessa regiao da esquerda ela é classificada como arenito, arenito-xisto na regiao central e xisto na regiao à direita.

O uso de apenas uma variável gaussina pra derivar a variável categórica, em muitos casos é restritivo demais, porque limita a técnica ambientes estratificados, onde algumas das litologias não se tocam.

## Simulação plurigaussiana truncada

Figura 44: Esquema da SPT. Note que as litologias 1 e 3 não aparecerão juntas nas realizações.



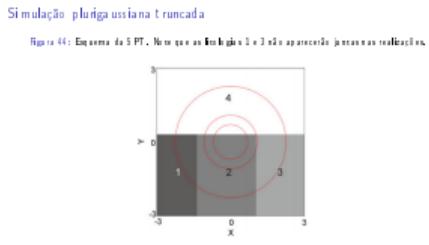
Proposta de tese para o exame de qualificação  
└ Proposta de tese  
└ Avaliação da incerteza  
└ Simulação plurigaussiana truncada

2019-06-11

Uma evolução natural da simulação gaussiana truncada, é a simulação pluri gaussiana truncada, que usa mais de uma variável gaussiana para derivar a variável categórica, permitindo a reprodução de estruturas mais complexas. Embora seja possível usar quantas variáveis gaussianas se desejar, o uso geralmente é limitado à apenas duas, pela dificuldade em se criar regras de truncagem em mais de duas dimensões.

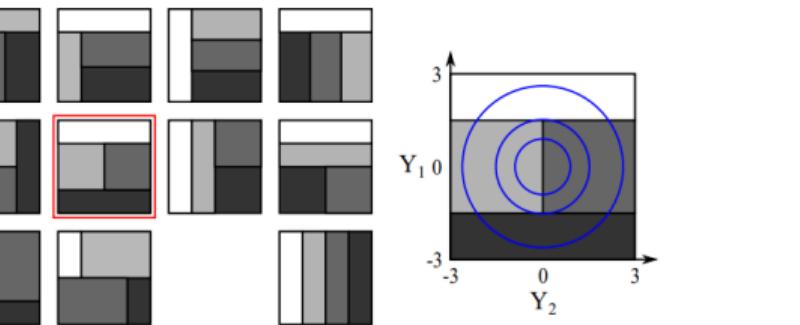
Regra de truncagem é uma parte importante da metodologia, já que ela controla os contatos entre as categorias, suas transições e proporções. A regra de truncagem, tradicionalmente, é baseada nas probabilidades de transição calculadas a partir dos dados.

Essa figura mostra duas variáveis gaussianas  $x$  e  $y$ , e uma regra de truncagem definida para quatro categorias, diferentemente da simulação gaussiana truncada, eu preciso observar o valor das duas gaussianas simultaneamente, e a partir da regra de truncagem, classificar o bloco simulado. Então se na gaussiana  $y$  o valor simulado caiu nessa região, e na gaussiana  $x$ , caiu nessa, por exemplo, o bloco é classificado como litologia 3.



# Simulação plurigaussiana truncada

Figura 45: Diferentes templates, para o caso de duas variáveis latentes e quatro categorias.



2019-06-11

Proposta de tese para o exame de qualificação  
└ Proposta de tese  
└ Avaliação da incerteza  
└ Simulação plurigaussiana truncada

Simulação plurigaussiana truncada

Figura 45: Diferentes templates, para o caso de duas variáveis latentes e quatro categorias.

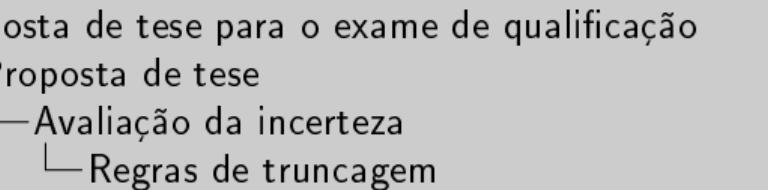
A small diagram in the top right corner shows a 4x4 grid with a central circle. The axes are labeled  $Y_1$  and  $Y_2$ .

Diferentes regras de truncagem pode ser definidas, geralmente o espaço gaussiano multivariado é particionado em paralelepípedos, como os mostrados no slide, mas diferentes tecnicas para a regra de truncagem foram propostas:

# Regras de truncagem I

- ▶ Allard et al. (2012) introduziram o *assination diagram*, que automaticamente constrói a regra de truncagem para o caso bivariado usando regressão baseada em kernel em variáveis auxiliares;
- ▶ Sadeghi and Boisvert (2012) usaram *simulated annealing* para otimizar a regra de truncagem. A função objetivo é a minimização da classificação errônea entre as probabilidades de transição calculadas de realizações e das probabilidades de transição calculadas a partir dos dados.
- ▶ Deutsch and Deutsch (2014) usaram *multidimensional scaling* (MDS) para definir regras de truncagem complexas com o foco em reproduzir as probabilidades de transição. Essa metodologia pode ser aplicada para qualquer número de variáveis gaussianas.
- ▶ Astrakova et al. (2015) propuseram uma metodologia semelhante a de Deutsch and Deutsch (2014) usando entropia bayesiana máxima em conjunto com *simulated annealing* para otimizar a regra de truncagem bigaussian;

2019-06-11



- \* Allard constroi automaticamente a regra de truncagem usando regressao baseada em kernel
- \* sadeghi e boisvert usaram sikulated annealing para otimizar a regra de truncagem
- \* deutsh pai e dutsh filho usaram MDS pra definir regras de truncage complexas
- \* astrakiva propos uma metodoloia semelhante a do deutsh
- \* e finalmente o emery e o silva propuseram uma abordagem hierárquica pra definir a regra de truncagem

Regras de truncagem I

- ▶ Allard et al. (2012) introduziram o *assination diagram*, que automaticamente constrói a regra de truncagem para o caso bivariado usando regressão baseada em kernel em variáveis auxiliares;
- ▶ Sadeghi and Boisvert (2012) usaram *simulated annealing* para otimizar a regra de truncagem. A função objetivo é a minimização da classificação errônea entre as probabilidades de transição calculadas de realizações e das probabilidades de transição calculadas a partir dos dados.
- ▶ Deutsch and Deutsch (2014) usaram *multidimensional scaling* (MDS) para definir regras de truncagem complexas com o foco em reproduzir as probabilidades de transição. Essa metodologia pode ser aplicada para qualquer número de variáveis gaussianas.
- ▶ Astrakova et al. (2015) propuseram uma metodologia semelhante a de Deutsch and Deutsch (2014) usando entropia bayesiana máxima em conjunto com *simulated annealing* para otimizar a regra de truncagem bigaussian;

## Regras de truncagem II

- ▶ Madani and Emery (2015) e Silva and Deutsch (2016) propuseram uma abordagem hierárquica para definir a regra de truncagem em espaços multi dimensionais.

2019-06-11

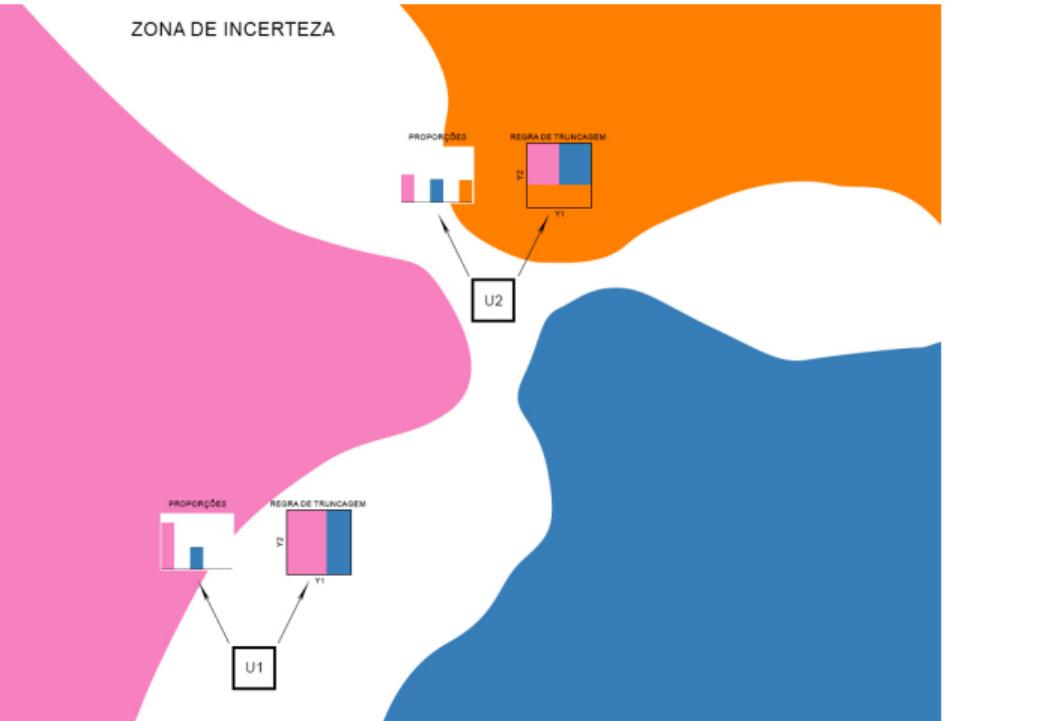
Proposta de tese para o exame de qualificação  
└ Proposta de tese  
  └ Avaliação da incerteza  
    └ Regras de truncagem

Regras de truncagem II

► Madani and Emery (2015) e Silva and Deutsch (2016) propuseram uma abordagem hierárquica para definir a regra de truncagem em espaços multi dimensionais.

## Regra dinâmica de truncagem

Figura 46: Diferentes regras de truncagem definidas para cada bloco a partir das probabilidades calibradas e calculadas.



- Proposta de tese para o exame de qualificação  
└ Proposta de tese  
└ Avaliação da incerteza  
└ Regra dinâmica de truncagem

2019-06-11

Geralmente essas regras de truncagem são definidas globalmente e usadas para todo o depósito, mas em muitos casos as proporções não são estacionárias matheron propos em 1987 curvas de proporções verticais, nesse caso, a regra de truncagem variava com a profundidade.

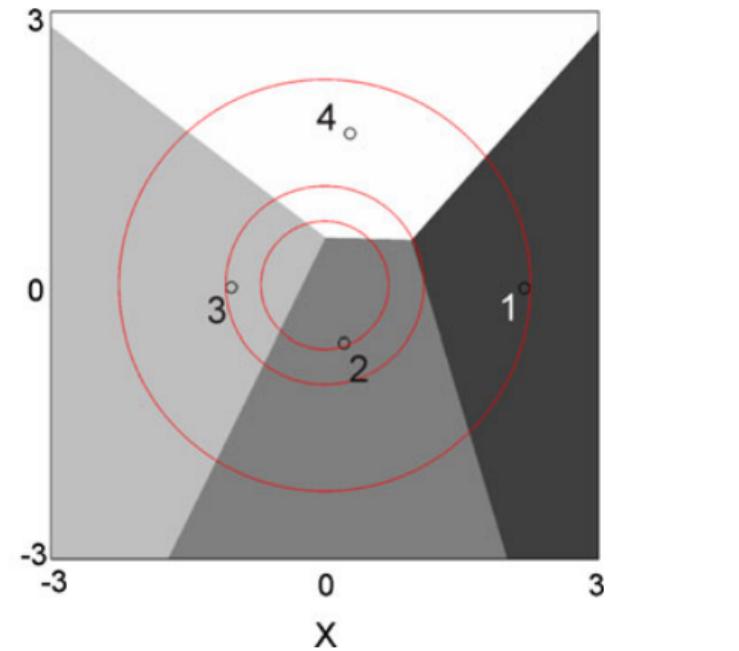
Essa tese propõe a definição dinâmica da regra de truncagem, bivariada, pra cada bloco mas somente dentro da zona de incerteza. A regra de truncagem é construída a partir das probabilidades locais, que foram derivadas a partir das distâncias estimadas.

O slide mostra dois blocos em duas regiões diferentes do depósito, dentro da zona de incerteza. Então pra esse bloco debaixo existe uma probabilidade alta dele pertencer a categoria rosa, e uma probabilidade mais baixa de pertencer a azul, já o bloco de cima, existe a probabilidade de pertencer as três categorias, a partir dessas probabilidades eu posso construir uma regra de truncagem, simular duas gaussianas não condicionais, comparar o valor das duas com a regra de truncagem e classificar o bloco em cada realização.



## Regra de truncagem via MDS

Figura 47: Esquema do uso de MDS para definição da regra de truncagem.



Proposta de tese para o exame de qualificação  
└ Proposta de tese  
└ Avaliação da incerteza  
└ Regra de truncagem via MDS

2019-06-11

Essa regra de truncagem deve ser definida de forma automática pra cada bloco, uma das técnicas que eu acredito ser bastante promissora é a proposta por deutshc pai e deutsh filho, que constrói a regra de truncagem a partir de MDS e decomposição de voronoi.

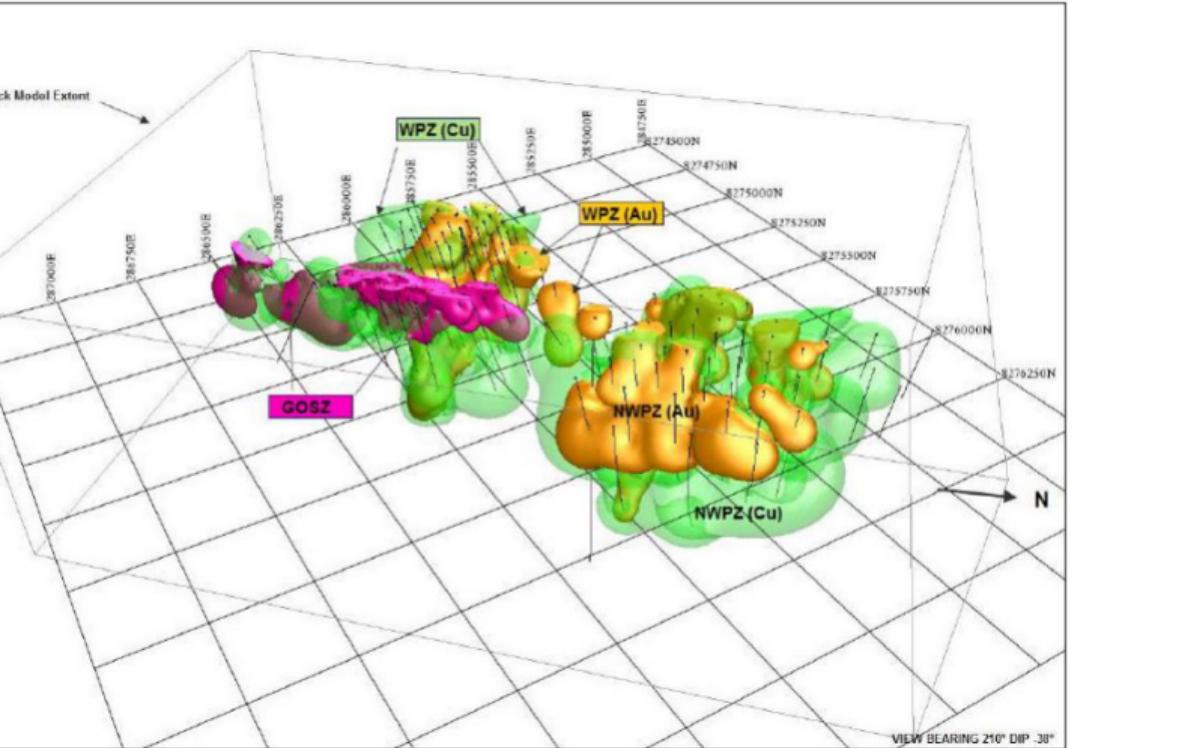
Um outro desafio é definir os variograms das variáveis gaussianas que controlam a continuidade dos domínios.

Regra de truncagem via MDS



# Validação

Figura 48: Modelo implícito ruim, com a presença de estruturas e formas indesejadas ("blobs").



Proposta de tese para o exame de qualificação  
└ Proposta de tese  
└ Validação  
└ Validação

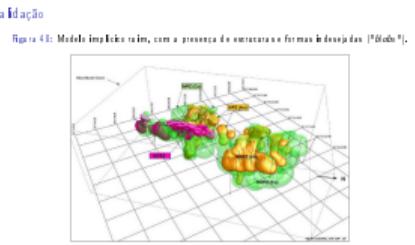
2019-06-11

Finalmente, o ultimo aprimoramento proposto nessa tese é um método de validação de modelos geológicos implícitos.

O estatístico box diz que todos os modelos estão errados mas alguns são úteis. O Jun Cowan que é um dos precursores da modelagem implícita, hoje trabalha na leapfrog, é um grande crítico da modelagem implícita, ele diz que a modelagem implícita cria modelos de forma fácil e rápida mas cria modelos ruins de forma fácil e rápida.

A modelagem implícita tende a gerar estruturas esféricas ou em forma de salsicha ao redor dos furos, como essas mostradas no slide, que não são estruturas geológicas.

Minha proposta é desenvolver um algoritmo para checagem visual automática de modelos geológicos implícitos, esse algoritmo deve reconhecer padrões indesejados, que serão corrigidas com a alteração de parâmetros ou utilizando outros métodos.



# Sumário

Em suma, essa tese propõe desenvolver e investigar:

- ▶ Um interpolador baseado em redes neurais independente de variograma para a variável distância assinalada;
- ▶ Uma metodologia para calibração e definição de uma zona de incerteza multi categórica estatisticamente justa, livre de viés e que seja de fato incerta;
  - ▶ Calibração do parâmetro  $\gamma$ ;
  - ▶ Truncagem no campo de entropia.
- ▶ Uma metodologia para avaliação da incerteza que gere múltiplas realizações de modelos geológicos multi categóricos dentro da zona de incerteza, sem ruído e com a possibilidade de controlar a natureza dos contatos.
  - ▶ Definição de uma regra de truncagem local para cada bloco.
  - ▶ Definição de um modelo de covariância para cada variável latente, que seja representativo.
- ▶ Uma metodologia para validação de modelos geológicos implícitos.
  - ▶ Identificação automática de estruturas indesejadas.

## Proposta de tese para o exame de qualificação

- └ Proposta de tese
- └ Sumário
- └ Sumário

2019-06-11

Sumário

- Em suma, essa tese propõe desenvolver e investigar:
- ▶ Um interpolador baseado em redes neurais independente de variograma para a variável distância assinalada;
  - ▶ Uma metodologia para calibração e definição de uma zona de incerteza multi categórica estatisticamente justa, livre de viés e que seja de fato incerta;
    - ▶ Calibração do parâmetro  $\gamma$ ;
    - ▶ Truncagem no campo de entropia.
  - ▶ Uma metodologia para avaliação da incerteza que gere múltiplas realizações de modelos geológicos multi categóricos dentro da zona de incerteza, sem ruído e com possibilidade de controlar a natureza dos contatos.
    - ▶ Definição de uma regra de truncagem local para cada bloco.
    - ▶ Definição de um modelo de covariância para cada variável latente, que seja representativo.
  - ▶ Uma metodologia para validação de modelos geológicos implícitos.
    - ▶ Identificação automática de estruturas indesejadas.

Em suma essa tese propõe:

- \* O desenvolvimento de um interpolador baseado em redes neurais que não depende de um variograma;
- \* O desenvolvimento de uma metodologia para calibração de uma zona de incerteza justa e livre de viés, pra isso é preciso calibrar o parametro gamma, usado na transformação das distâncias em probabilidades e truncar o campo de entropia, com base no tipo de depósito;
- \* uma metodologia para avaliação de incerteza que gere modelos continuos, livres de ruido, pra isso é preciso definir uma regra de truncagem local em cada bloco e definir um modelo de covariância pra cada variavel latente, ou variavel gaussiana auxiliar.
- \* e finalmente uma metodologia pra validar modelos implícitos, que identifique de forma automatica padroes indesejáveis.

# Progresso de desenvolvimento

- ▶ Cálculo das distâncias assinaladas ✓
- ▶ Validação cruzada ✓
- ▶ Interpolação tradicional por RBF ✓
- ▶ Interpolação baseada em redes neurais
- ▶ Calibração do parâmetro  $\gamma$  a partir das amostras ✓
- ▶ Calibração do parâmetro  $\gamma$  (nova abordagem)
- ▶ Cálculo das probabilidades ✓
- ▶ Cálculo da entropia ✓
- ▶ Definição do *cutt-off* no campo de entropia
- ▶ Simulação da variável contínua auxiliar
- ▶ Definição da regra de truncagem local
- ▶ Algoritmo de validação visual

2019-06-11

Proposta de tese para o exame de qualificação

- └ Proposta de tese
- └ Cronograma e atividades
- └ Progresso de desenvolvimento

Progresso de desenvolvimento

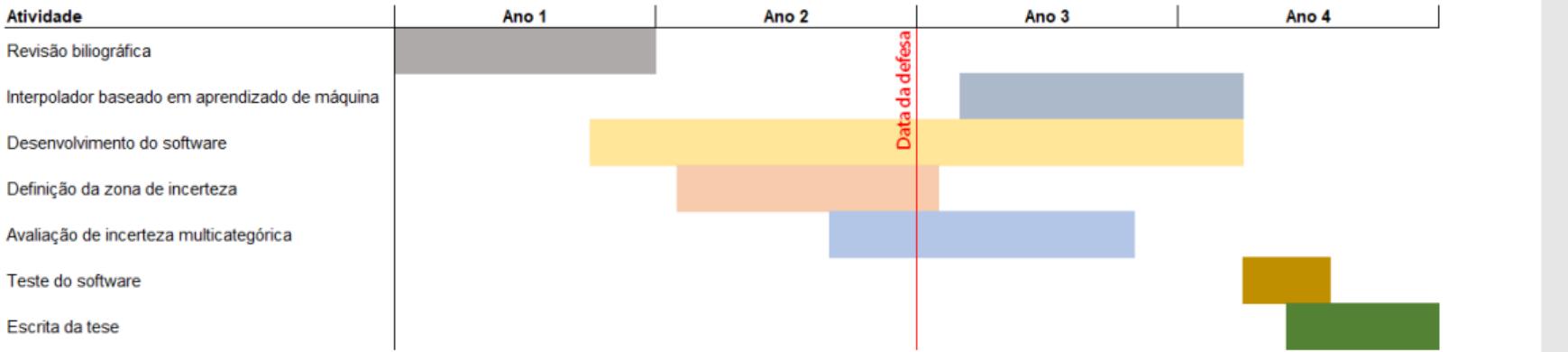
- ▶ Cálculo das distâncias assinaladas ✓
- ▶ Validação cruzada ✓
- ▶ Interpolação tradicional por RBF ✓
- ▶ Interpolação baseada em redes neurais
- ▶ Calibração do parâmetro  $\gamma$  a partir das amostras ✓
- ▶ Calibração do parâmetro  $\gamma$  (nova abordagem)
- ▶ Cálculo das probabilidades ✓
- ▶ Cálculo da entropia ✓
- ▶ Definição do *cutt-off* no campo de entropia
- ▶ Simulação da variável contínua auxiliar
- ▶ Definição da regra de truncagem local
- ▶ Algoritmo de validação visual

Eu já venho desenvolvendo uma biblioteca de modelagem geológica implícita em python, tá no github, convido todos os colegas a testarem e editarem o código. Por enquanto ala calcula as distâncias, faz validação cruzada baseadas em k-folds, interpola por rbf, eu trabalhei numa primeira tentativa de calibrar o gamma a partir das moastras, mas não deu muito certo, então estou trabalhando em uma nova abordagem. A biblioteca ainda calcula as probabilidades, e calcula a entropia, por enquanto.

71/75

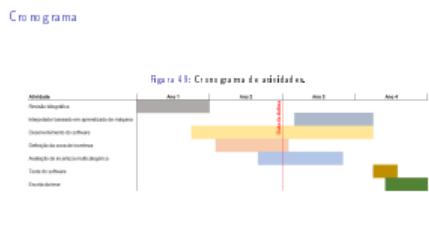
# Cronograma

Figura 49: Cronograma de atividades.



- Proposta de tese para o exame de qualificação
- └ Proposta de tese
  - └ Cronograma e atividades
  - └ Cronograma

2019-06-11



Estou entrando no segundo ano de doutorado, a revisão biliográfica já foi feita, eu comecei a trabalhar no desenvolvimento de software, no momento estou trabalhando na definição da zona de incerteza, e dando os primeiros passos nas metodologias de avaliação de incerteza, para os proximoms meses o planejamento é também começar a trabalhar no interpolador. Os últimos meses serão dedicados à escrita da tese e publicação dos resultados.

# Referências bibliográficas |

Allard, D., D'Or, D., Biver, P., and Froidevaux, R. (2012). Non-parametric diagrams for plurigaussian simulations of lithologies. In *9th international geostatistical congress, Oslo, Norway*, volume 1115.

Astrakova, A., Oliver, D. S., and Lantuéjoul, C. (2015). Truncation map estimation based on bivariate probabilities and validation for the truncated plurigaussian model. *arXiv preprint arXiv:1508.01090*.

Boisvert, J. (2013). Automatic geological modeling with iterative refinement. *CCG Annual Report*.

Deutsch, J. L. and Deutsch, C. V. (2014). A multidimensional scaling approach to enforce reproduction of transition probabilities in truncated plurigaussian simulation. *Stochastic environmental research and risk assessment*, 28(3):707–716.

Froidevaux, R. (1993). Probability field simulation. In *Geostatistics tróia '92*, pages 73–83. Springer.

Proposta de tese para o exame de qualificação

2019-06-11

Referências bibliográficas

Referências bibliográficas

Referências bibliográficas

Por fim as referências bibliográficas e eu gostaria de agradecer novamente a presença de todos, ao LPM e à capes. Muito obrigado!

Referências bibliográficas

Allard, D., D'Or, D., Biver, P., and Froidevaux, R. (2012). Non-parametric diagrams for plurigaussian simulations of lithologies. In *9th international geostatistical congress, Oslo, Norway*, volume 1115.

Astrakova, A., Oliver, D. S., and Lantuéjoul, C. (2015). Truncation map estimation based on bivariate probabilities and validation for the truncated plurigaussian model. *arXiv preprint arXiv:1508.01090*.

Boisvert, J. (2013). Automatic geological modeling with iterative refinement. *CCG Annual Report*.

Deutsch, J. L. and Deutsch, C. V. (2014). A multidimensional scaling approach to enforce reproduction of transition probabilities in truncated plurigaussian simulation. *Stochastic environmental research and risk assessment*, 28(3):707–716.

Froidevaux, R. (1993). Probability field simulation. In *Geostatistics tróia '92*, pages 73–83. Springer.

## Referências bibliográficas II

Hosseini, A. and Deutsch, C. V. (2007). A distance function based algorithm to quantify uncertainty in areal limits. *CCG Annual Report*.

Lajaunie, C., Courrioux, G., and Manuel, L. (1997). Foliation fields and 3d cartography in geology: principles of a method based on potential interpolation. *Mathematical Geology*, 29(4):571–584.

Madani, N. and Emery, X. (2015). Simulation of geo-domains accounting for chronology and contact relationships: application to the río blanco copper deposit. *Stochastic environmental research and risk assessment*, 29(8):2173–2191.

Mallet, J.-L. (2004). Space-time mathematical framework for sedimentary geology. *Mathematical geology*, 36(1):1–32.

Manchuck, J. and Deutsch, C. (2018). MIs for boundary interpretations with uncertainty. *CCG Annual Report*.

Rolo, R. (2017). Modelagem geológica implícita com funções distância assinaladas. Master's thesis, Universidade Federal do Rio Grande do Sul.

## Proposta de tese para o exame de qualificação

### Referências bibliográficas

2019-06-11

### Referências bibliográficas

- Referências bibliográficas II
- Hosseini, A. and Deutsch, C. V. (2007). A distance function based algorithm to quantify uncertainty in areal limits. *CCG Annual Report*.
- Lajaunie, C., Courrioux, G., and Manuel, L. (1997). Foliation fields and 3d cartography in geology: principles of a method based on potential interpolation. *Mathematical Geology*, 29(4):571–584.
- Madani, N. and Emery, X. (2015). Simulation of geo-domains accounting for chronology and contact relationships: application to the río blanco copper deposit. *Stochastic environmental research and risk assessment*, 29(8):2173–2191.
- Mallet, J.-L. (2004). Space-time mathematical framework for sedimentary geology. *Mathematical geology*, 36(1):1–32.
- Manchuck, J. and Deutsch, C. (2018). MIs for boundary interpretations with uncertainty. *CCG Annual Report*.
- Rolo, R. (2017). Modelagem geológica implícita com funções distância assinaladas. Master's thesis, Universidade Federal do Rio Grande do Sul.

## Referências bibliográficas III

- Sadeghi, S. and Boisvert, J. (2012). Optimizing thresholds in truncated pluri-gaussian simulation. *CCG Annual Report*.
- Samson, M. and Deutsch, C. (2018). Estimation with machine learning. *CCG Annual Report*.
- Silva, D. and Deutsch, C. (2015). Speeding up the signed distance implicit geologic modeling method for multiple rock types. *CCG Annual Report*.
- Silva, D. and Deutsch, C. (2016). Hierarchical approach to truncated plurigaussian simulation. *CCG Annual Report*.
- Silva, D. M. A. (2015). *Enhanced geologic modeling with data-driven training images for improved resources and recoverable reserves*. PhD thesis, University of Alberta.

Proposta de tese para o exame de qualificação  
└ Referências bibliográficas

└ Referências bibliográficas

2019-06-11

- Sadeghi, S. and Boisvert, J. (2012). Optimizing thresholds in truncated plurigaussian simulations. *CCG Annual Report*.
- Samson, M. and Deutsch, C. (2018). Estimates with machine learning. *CCG Annual Report*.
- Silva, D. and Deutsch, C. (2015). Speeding up the signed distance implicit geologic modeling method for multiple rock types. *CCG Annual Report*.
- Silva, D. and Deutsch, C. (2016). Hierarchical approach to truncated plurigaussian simulations. *CCG Annual Report*.
- Silva, D. M. A. (2015). Enhanced geologic modeling with data-driven training images for improved resources and recoverable reserves. PhD thesis, University of Alberta.