

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Curso de Pós-Graduação em Engenharia Metalúrgica, Materiais e de Minas da
UFMG – Mestrado Profissional

Hugo Alexandre Garcia Gonçalves

INTERPOLAÇÃO NO ESPAÇO EVOLUTIVO: uma aplicação geotécnica em
estações totais robóticas

Belo Horizonte

2023

Hugo Alexandre Garcia Gonçalves

**INTERPOLAÇÃO NO ESPAÇO EVOLUTIVO: uma aplicação geotécnica em
estações totais robóticas**

Dissertação apresentada ao Curso de Pós-Graduação em Engenharia Metalúrgica, Materiais e de Minas – Mestrado Profissional da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do Grau de Mestre em Engenharia Metalúrgica, Materiais e de Minas

Orientador: Prof. Dr. Alizeibek Saleimen Nader

Belo Horizonte

2023

G635i

Gonçalves, Hugo Alexandre Garcia.

Interpolação no espaço evolutivo [recurso eletrônico] : uma aplicação geotécnica em estações totais robóticas / Hugo Alexandre Garcia Gonçalves. – 2023.

1 recurso online (73 f. : il., color.) : pdf.

Orientador: Alizeibek Saleimen Nader.

Dissertação (mestrado profissional) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 70-73.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia de minas – Teses. 2. Tecnologia mineral – Teses. 3.
Geotecnia – Automação – Teses. 4. Geomática – Teses. 5.
Geostatística – Teses. 6. Instrumentos de medição – Teses. 7.
Topografia – Medição – Teses. 8. Topografia – Instrumentos – Teses. 9.
Robótica – Teses. I. Nader, Alizeibek Saleimen. II. Universidade Federal de
Minas Gerais. Escola de Engenharia. III. Título.

CDU: 622(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
Curso de Pós-Graduação em Engenharia Metalúrgica,
Materiais e de Minas da UFMG - Mestrado Profissional

UFMG

***INTERPOLAÇÃO NO ESPAÇO EVOLUTIVO:
UMA APLICAÇÃO GEOTÉCNICA EM ESTAÇÕES TOTAIS
ROBÓTICAS***

HUGO ALEXANDRE GARCIA GONÇALVES

Dissertação de mestrado submetida à Comissão Examinadora designada pelo Colegiado do Curso de Pós-Graduação em Engenharia Metalúrgica, Materiais e de Minas da UFMG – Mestrado Profissional da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Mestre em Engenharia Metalúrgica, Materiais e de Minas.

Aprovada em 16 de junho de 2023.

Por:

Prof. Dr. Alizeibek Saleimen Nader (UFMG)
Orientador

Prof. Dr. Roberto Galery (UFMG)

Pedro Henrique Alves Campos

Prof. Dr. Pedro Henrique Alves Campos (UFMG)

Dedicatória

Dedico este trabalho a todos que buscam criar um ambiente de maior segurança em locais de incertezas geotécnicas. Que ele possa auxiliar nas tomadas de decisões de engenharia que podem salvar vidas

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao meu orientador, Beck, pela sua excelente orientação, apoio contínuo e insights valiosos ao longo deste projeto. Sua dedicação e expertise foram fundamentais para a minha jornada de pesquisa, e sou imensamente grato por sua orientação. Gostaria de estender meus agradecimentos também aos membros da banca por estarem aqui dedicando seu tempo para avaliar meu trabalho e fornecer suas contribuições. Agradeço também a minha noiva, Maíra, por seu apoio incondicional e por me motivar em todos os momentos que achei que não daria conta. Também expresso minha gratidão a minha família, que sempre me apoiou em todos os estudos que eu tive interesse em fazer. Por fim, agradeço à UFMG e ao DEMIN pela oportunidade que me foi dada de desenvolver esse projeto junto a pesquisadores excepcionais, professores da escola.

Em suma, esta pesquisa não teria sido possível sem o apoio e contribuição de tantas pessoas maravilhosas. Sou verdadeiramente grato a todos vocês por seu envolvimento e dedicação. Obrigado!

RESUMO

A análise de estabilidade de estruturas geotécnicas com base no monitoramento da deformação tem sido comumente utilizada para prever, evitar e mitigar danos. Um dos equipamentos mais utilizados para esse trabalho é a estação total robótica (ETR), capaz de medir de forma contínua todos os prismas posicionados em determinado local. Esse monitoramento contínuo é responsável pela geração de um imenso volume de dados, o que cria a necessidade de diferentes técnicas para visualização e interpretação dos resultados. Além disso, problemas no equipamento, ou em partes dele, prejudica a qualidade do monitoramento naqueles períodos. Devido a isso, surgiu a necessidade de interpolar os dados no tempo e no espaço. No presente trabalho foi desenvolvida uma técnica de interpolação espacial aplicada temporalmente aos dados obtidos de uma ETR. Para visualização do resultado, um vídeo contendo mapas de calor com os dados interpolados em cada intervalo temporal foi criado, facilitando a identificação de anomalias.

Palavras-chave: Interpolação; Estação Robótica Total; Interpolação no espaço evolutivo; Geoestatística; Geotecnia; Geomática

ABSTRACT

Stability analysis of geotechnical structures based on deformation monitoring has been commonly used to predict, prevent, and mitigate damages. One of the most used equipment for this task is the robotic total station (RTS), capable of continuously measuring all prisms positioned at a specific location. This continuous monitoring is responsible for generating a vast volume of data, which creates the need for different techniques for visualization and interpretation of the results. Additionally, equipment problems, or issues with its components, can compromise the quality of monitoring during those periods. As a result, there arose a need to interpolate the data in time and space. In this study, a spatial interpolation technique applied temporally to the data obtained from an RTS was developed. To visualize the results, a video containing heat maps with interpolated data at each time interval was created, facilitating the identification of anomalies.

Keywords: Interpolation; Total Robotic Station; Spatio-temporal Interpolation; Geostatistics; Geotechnics; Geomatics

Lista de Figuras

Figura 1: A ETR está representada pelo cubo e pirâmide roxos. O prisma azul está fincado em uma encosta. A linha que conecta os dois representa a distância inclinada medida pelo equipamento..	17
Figura 2: Vista lateral de uma ETR e um prisma. O ângulo entre o zênite do equipamento e o segmento de reta que liga a ETR ao prisma representa o ângulo vertical	18
Figura 3: Vista de topo do mesmo sistema anterior. O segmento de reta que sai da ETR aponta para a direção do seu limbo interno. O ângulo horizontal é representado pelo ângulo entre esse segmento de reta e o segmento de reta que liga o equipamento ao prisma. Ao se calcular o ângulo horizontal corrigido, a referência deixa de ser o limbo da ETR e passa a ser o norte..	18
Figura 4: Vista aérea de prismas de monitoramento e de referência posicionados numa barragem.	19
Figura 5: Medição do deslocamento usando radar interferométrico. Ele é calculado comparando mudanças de fase dos sinais refletidos entre duas medições de radar consecutivas.	22
Figura 6: Mapa de deslocamento obtido por um radar SAR. Cada pixel da imagem possui um valor de deslocamento atribuído a ele.	24
Figura 7: Mapa de amplitude do sinal obtido por um radar SAR. Cada pixel da imagem possui um valor relativo à perda de energia pela onda	25
Figura 8: Exemplo de deslocamento registrado por uma ETR.....	26
Figura 9: Comparação entre dois prismas. O prisma representado pela série em azul apresentou forte deslocamento no período, enquanto o vermelho permaneceu estável	27
Figura 10: Tipos de variáveis.	29
Figura 11: Correlação linear positiva.	35
Figura 12: Correlação linear negativa.	35
Figura 13: Sem correlação linear.....	36
Figura 14: Variograma e suas propriedades	37
Figura 15: Diagrama de Venn ilustrando a interseção de dois eventos A e B	39
Figura 16: Diagrama de Venn ilustrando a união de dois eventos A e B.....	40
Figura 17: Diagrama de Venn ilustrando dois eventos disjuntos, A e B.....	40
Figura 18: Diagrama de Venn ilustrando dois eventos complementares, A e B.	41
Figura 19: Função densidade de probabilidade da distribuição normal	45
Figura 20: Diferentes distribuições normais, demonstrando uma variação na média.....	46
Figura 21: Diferentes distribuições normais, demonstrando uma variação no desvio padrão .	46
Figura 22: Ilustração do viés na estimativa. O painel ABDC a ser explorado foi delimitado a partir das amostras ricas observadas ao longo de AB e CD. Como a verdadeira fronteira entre os minérios ricos e pobres segue uma linha semelhante à linha pontilhada em vez dos segmentos AC e BD, parte do minério pobre será explorado e parte do minério rico abandonado.....	48
Figura 23: Representação esquemática de alto nível, mostrando a comunicação entre os diferentes módulos do software.....	56

Figura 24: Dados originais com tempo de ciclo de 2 horas	58
Figura 25: Mesma janela temporal anterior com dados reamostrados a cada 5 minutos	59
Figura 26: Boxplot com função densidade de probabilidade de uma distribuição normal. Os outliers estão à esquerda e à direita do boxplot representado.....	60
Figura 27: Exemplo de variograma experimental gerado pela função.....	64
Figura 28: Log de execução com os parâmetros experimentais	64
Figura 29: Mapa de deslocamento gerado a partir da krigagem dos valores de deslocamento dos prismas.....	65
Figura 30: Evolução do deslocamento visualizado com a utilização da interpolação no espaço evolutivo	68

Sumário

1	INTRODUÇÃO	12
2	OBJETIVOS	13
3	REVISÃO BIBLIOGRÁFICA	13
3.1	Estações totais robóticas (ETR)	13
3.1.1	Funcionamento	15
3.1.2	Parâmetros gerados	16
3.1.3	Correções aplicadas	20
3.2	Uso de radares no monitoramento geotécnico	21
3.2.1	Radar de abertura sintética	21
3.2.2	Comparação entre os dados de ETR e SAR	25
3.3	Conceitos básicos de Estatística	28
3.3.1	Tipo de variáveis	28
3.3.2	Medidas de tendência central	30
3.3.3	Medidas de dispersão	31
3.3.4	Covariância	33
3.3.5	Correlação	34
3.3.6	Variograma	36
3.4	Conceitos de Probabilidade	38
3.4.1	Experimentos aleatórios	38
3.4.2	Axiomas da probabilidade	38
3.4.3	Interações entre eventos	39
3.4.4	Probabilidade condicional	41
3.4.5	Variáveis aleatórias	42
3.4.6	Distribuição discreta de probabilidade	43
3.4.7	Distribuição contínua de probabilidade	44
3.4.8	Distribuição Normal ou Gaussiana	44
3.4.9	Teorema central do limite	47
3.5	História da krigagem	47
3.5.1	Krigagem ordinária	49
3.5.2	Krigagem simples	50

3.6	Interpolação espaço-temporal	50
3.7	Modelagem da covariância e variogramas.....	52
3.7.1	Modelo separável.....	52
3.7.2	Modelo da soma do produto	53
3.7.3	Modelo métrico	53
3.7.4	Modelo métrico aditivo	54
3.7.5	Modelo métrico aditivo simplificado	54
3.8	Krigagem espaço-temporal	54
4	METODOLOGIA.....	55
4.1	Grupos de execução	55
4.1.1	Database.py.....	56
4.1.2	Pre_processing.py.....	57
4.1.3	Kriging.py.....	62
4.1.4	Render.py.....	66
4.1.5	Main.py.....	66
5	RESULTADOS	67
6	CONCLUSÃO.....	69
	REFERÊNCIAS	70

1 Introdução

Com a crescente demanda por recursos minerais no mundo e o consequente aumento dos empreendimentos minerários, há também um aumento no número de cavas a céu aberto e de barragens de rejeito. Devido a isso, tem crescido a preocupação com a segurança da produção e gerenciamento das minas, bem como da população vizinha aos locais de mineração (XIAO et al., 2020; DONATI, 2020). Para garantir a segurança dos operadores e da sociedade no entorno, têm se adotado diversas técnicas de monitoramento, a fim de analisar o comportamento geotécnico das estruturas e prevenir acidentes em tempo hábil de se evitar perdas humanas e de equipamentos (ZHOU, 2020).

Devido a restrições monetárias, de equipamento, tecnologia e condições de observação, muitas vezes os dados obtidos possuem uma qualidade questionável, o que afeta a efetiva análise de estabilidade. Para compensar a falta ou perda de dados de monitoramento e analisar com acurácia a tendência de deslocamento, é necessário interpolar os dados de monitoramento irregulares, descontínuos e limitados existentes, de forma a suportar o processamento posterior, a análise de estabilidade e pesquisas científicas correlatas (LIU et al., 2019)

Interpolação é um processo de estimativa de valores quando e/ou onde nenhum dado foi adquirido. A interpolação espacial é definida como um processo de análise de dados, no qual o investigador e os Sistemas de Informação Geográfica (GIS) tentam estimar valores contínuos em locais onde não houve medições. Existem três categorias principais que categorizam as técnicas comuns de interpolação: interpolação geoestatística, interpolação não geoestatística e métodos combinados (ELDRANDALY et al., 2017). Há também várias técnicas usadas para interpolar valores faltantes para dados de séries temporais.

Devido a essa necessidade de se estimar valores nos domínios do espaço e do tempo simultaneamente, várias metodologias têm sido propostas. Atualmente, a forma mais fácil de se gerar uma interpolação espaço temporal é através de operações geoestatísticas disponíveis através de pacotes em linguagens de programação como R ou Python (GRÄLER; PEBESMA; HEUVELINK, 2016). Existem também métodos

computacionais que trazem melhores resultados e são, conseqüentemente, mais complexos e com maior custo computacional, como algoritmos genéticos adaptativos (XIAO et al., 2020) e regressões não lineares usando processos gaussianos (WANG et al., 2014).

Um dos principais desafios enfrentados no dia a dia do monitoramento de estruturas com a utilização de estações totais robóticas (ETR) é o de visualização dos dados gerados. O equipamento realiza diversas medições em pontos pré-definidos na estrutura e gera diariamente uma grande quantidade de dados, que são analisados, em grande parte das vezes, através da observação das séries temporais geradas. Os problemas dessa metodologia são: identificar tendências regionais e separar os ruídos existentes dos comportamentos típicos de um deslocamento real.

Para sanar esses problemas observados, foi então desenvolvido um algoritmo que não só automatiza a geração do visual como também considere todo o histórico de dados registrados na estrutura (ou, pelo menos, uma parte representativa dele).

2 Objetivos

O presente trabalho tem como objetivo desenvolver uma nova metodologia de visualização dos dados de estações totais robóticas, permitindo visualizar de forma regional e ao longo do tempo a evolução de potenciais deslocamentos capturados no monitoramento feito por ETR.

3 Revisão bibliográfica

3.1 Estações totais robóticas (ETR)

ETRs são, basicamente, estações totais que realizam operação remota e automática, permitindo a realização de muito mais medições e cálculos do que sua versão manual. Devido a isso, esses equipamentos são extremamente úteis para o monitoramento de estruturas de mineração e de construção civil (GOPI et al, 2007).

As ETRs são projetadas para medir distâncias horizontais e inclinações, ângulos horizontais e verticais e elevações em levantamentos topográficos e trabalhos geodésicos,

levantamentos taqueométricos, bem como para soluções de levantamento (GOPI et al, 2007). Os resultados da medição podem ser gravados na memória interna da estação total e podem ser transferidos para um computador, bem como podem ser diretamente conectadas a um computador que armazena os dados diretamente em um banco de dados.

As propriedades básicas são alcance, velocidade e precisão incomparáveis das medições. As estações totais robóticas são desenvolvidas com vista à conveniência máxima do usuário. A ETR mede ângulos e distâncias para vários pontos sob levantamento e as coordenadas dos pontos observados (X, Y e Z ou norte, leste e elevação) em relação à posição da ETR são calculadas usando funções trigonométricas/geométricas (GOPI et al, 2007, LEICA, 2013).

Uma estação total é uma combinação de um teodolito eletrônico, um dispositivo eletrônico de medição de distância (EDM) e um microprocessador com unidade de memória (GOPI et al, 2007). A ETR adiciona a isso a capacidade de girar em seu próprio eixo e de buscar automaticamente por prismas em janelas de busca, pré-determinadas ao se cadastrar os pontos no software do equipamento. O teodolito eletrônico digital, introduzido pela primeira vez na década de 1960 pela Carl Zeiss Inc., ajudou a estabelecer as bases para a coleta e processamento de dados de campo modernos. Quando o teodolito eletrônico foi usado com uma unidade de EDM incorporada, nasceu o novo conceito de levantamento totalmente automatizado. O nome original para um instrumento desse tipo era taqueômetro eletrônico, mas a Hewlett-Packard introduziu o nome estação total há mais de 30 anos e o nome imediatamente se popularizou entre a profissão. Com este dispositivo, é possível determinar ângulos e distâncias do instrumento para os pontos a serem levantados. Com a ajuda da trigonometria, os ângulos e distâncias podem ser usados para calcular as posições reais (x, y e z ou norte, leste e elevação) dos pontos levantados em termos absolutos.

Uma ETR possui a capacidade de se comunicar e operar de forma autônoma, sem a necessidade de intervenção humana. Isso permite uma maior precisão e rapidez nas medições, além de reduzir a necessidade de equipes de campo maiores e aumentar a segurança nas áreas de trabalho. Além disso, essas estações totais também possuem a

capacidade de armazenar grandes quantidades de dados e transferi-los para computadores de campo para análise e processamento.

A ETR é utilizada em diversos setores, como construção, mineração e agricultura, pois é capaz de fornecer medidas precisas e confiáveis com rapidez e eficiência.

3.1.1 Funcionamento

Uma estação total padrão é basicamente um telescópio com miras cruzadas para mirar um alvo. O telescópio está conectado a escalas para medir o ângulo de rotação do telescópio (normalmente relativo ao norte como 0 grau) e o ângulo de inclinação do telescópio (relativo ao horizontal como 0 grau). Depois de girar o telescópio para apontar para um alvo, é possível ler o ângulo de rotação e o ângulo de inclinação a partir de um ponto de referência. Todas as estações totais têm um dispositivo EDM e varredura de ângulos eletrônicos. As escalas codificadas dos círculos horizontais e verticais são varridas eletronicamente e, em seguida, os ângulos e distâncias são exibidos digitalmente. A distância horizontal, a diferença de altura e as coordenadas são calculadas automaticamente e todas as medições e informações adicionais são registradas no banco de dados (GOPI et al, 2007). O equipamento eletrônico fornece uma leitura digital desses ângulos em vez de uma escala, o que é mais preciso e menos propenso a erros decorrentes da interpolação entre marcas na escala ou erro de gravação. A leitura também é contínua, e, portanto, os ângulos podem ser verificados a qualquer momento.

Na maioria das estações totais usa-se diodos emissores de luz infravermelha modulada, que enviam um feixe do instrumento para o prisma (GOPI et al, 2007). O prisma reflete este feixe de volta para a estação total. A porção de comprimento de onda que sai do instrumento e retorna é avaliada e calculada. Em outras palavras, o instrumento EDM instalado dentro do telescópio da estação total transmite um feixe infravermelho, que é refletido de volta para a unidade com a ajuda de um prisma (após reflexão interna total), e o EDM usa medições de tempo para calcular a distância percorrida pelo feixe. Com poucas exceções, o instrumento EDM exige que o alvo seja altamente reflexivo, e normalmente é usado um prisma refletor como alvo.

3.1.2 Parâmetros gerados

Para o presente trabalho foram utilizadas estações totais robóticas Leica. Os dados gerados pelo equipamento podem ser divididos em dados brutos e dados calculados. Os dados brutos são todos aqueles que a ETR faz medição direta (LEICA, 2013). São eles a distância inclinada e os ângulos vertical e horizontal. Também são medidos diretamente temperatura, pressão e inclinação (nível de bolha) do equipamento. Todos os outros parâmetros de monitoramento são calculados a partir dessas medições. São eles as coordenadas (easting, northing e height), vetores deslocamento 2D e 3D, além de vetores deslocamento longitudinais e transversais a um plano pré-definido.

Como o equipamento possui precisão milimétrica (LEICA, 2013), variações naturais de temperatura e pressão são capazes de afetar consideravelmente o resultado. A fim de mitigar esse efeito, são também calculadas correções para os dados brutos, com base nos valores de pressão. No caso do ângulo horizontal corrigido, o valor também é corrigido para que o ângulo seja equivalente ao azimute (que tem como referência o norte) e não o limbo interno. Isso é fundamental nos casos em que a ETR precisa ser trocada. Os limbos internos podem ser ligeiramente diferentes e, caso o ângulo horizontal não seja com relação ao norte, isso ocasionaria um salto nos valores dos dados

3.1.2.1 Distância inclinada

A distância inclinada diz respeito à distância, em linha reta, entre a estação total robótica e o prisma medido. Ela é obtida através da utilização de um dispositivo eletrônico de medição de distâncias, que funciona com a utilização de um laser. A figura 2 abaixo ilustra esse parâmetro:

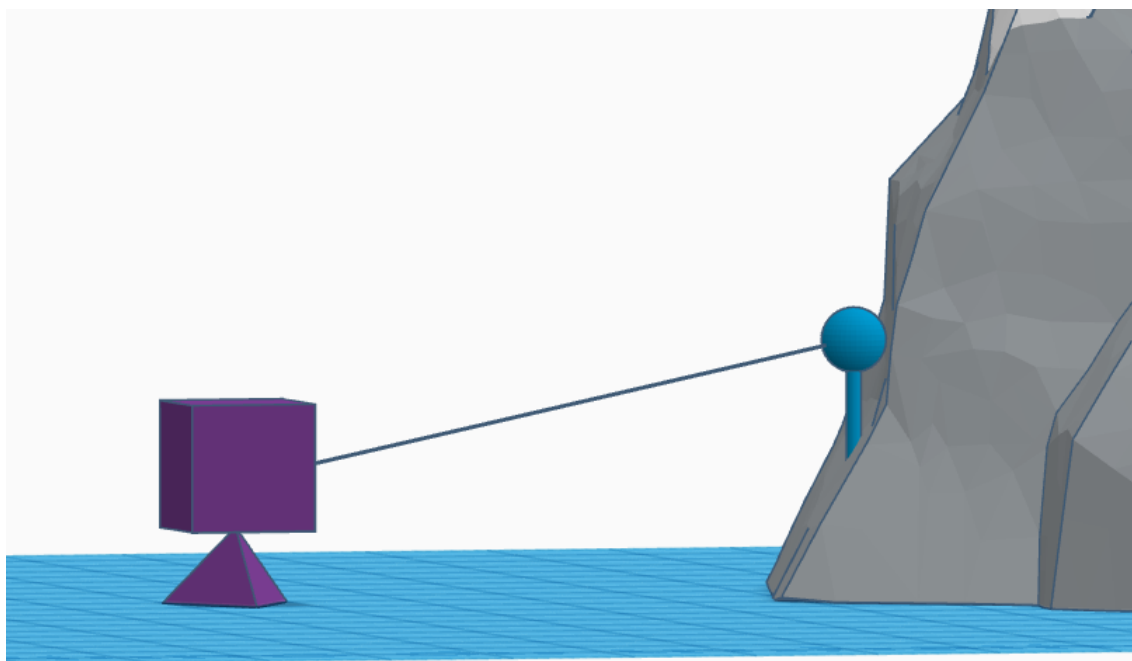


Figura 1: A ETR está representada pelo cubo e pirâmide roxos. O prisma azul está fincado em uma encosta. A linha que conecta os dois representa a distância inclinada medida pelo equipamento. Fonte: Elaborado pelo autor.

3.1.2.2 Ângulos

O equipamento também mede os ângulos zenitais e azimutais. O ângulo zenital diz respeito ao ângulo medido entre um vetor que aponta da ETR para o céu (zênite) e o prisma. Já o ângulo horizontal diz respeito ao ângulo feito entre uma referência interna do equipamento e o prisma, ou seja, considerando um vetor que sai da ETR e aponta para esse limbo interno, num plano definido por norte e leste. As figuras 3 e 4 abaixo ilustram essas grandezas:



Figura 2: Vista lateral de uma ETR e um prisma. O ângulo entre o zênite do equipamento e o segmento de reta que liga a ETR ao prisma representa o ângulo vertical. Fonte: Elaborado pelo autor.

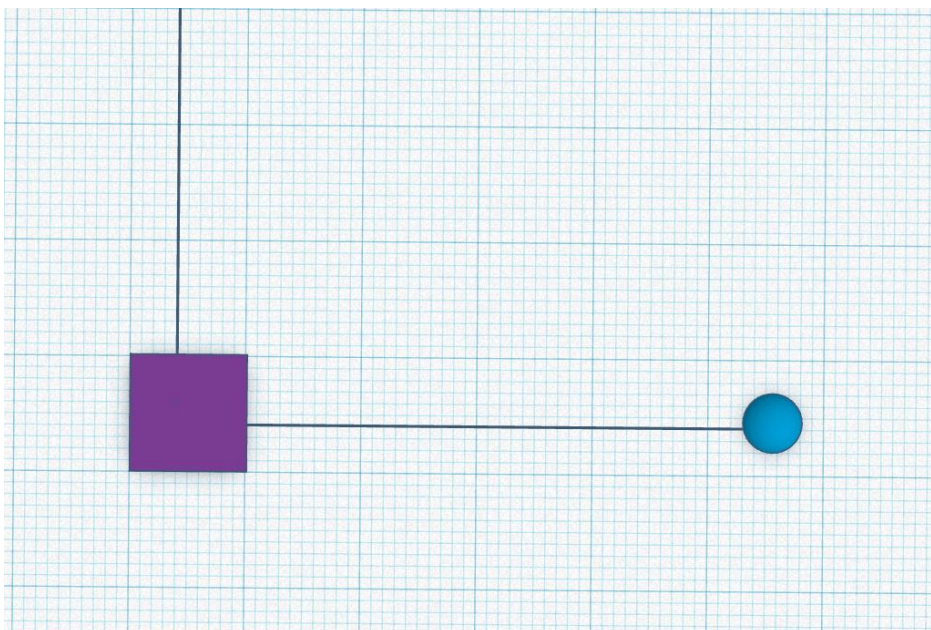


Figura 3: Vista de topo do mesmo sistema anterior. O segmento de reta que sai da ETR aponta para a direção do seu limbo interno. O ângulo horizontal é representado pelo ângulo entre esse segmento de reta e o segmento de reta que liga o equipamento ao prisma. Ao se calcular o ângulo horizontal corrigido, a referência deixa de ser o limbo da ETR e passa a ser o norte. Fonte: Elaborado pelo autor.

3.1.2.3 Easting, northing e height

A partir das medições obtidas, é possível calcular as coordenadas dos prismas medidos em um sistema de coordenadas qualquer, desde que sejam conhecidas as coordenadas da ETR em um sistema de coordenadas específico. É importante observar que, durante o cálculo das coordenadas dos prismas de monitoramento, é suposto que a ETR permaneça estável. No entanto, se houver suspeita de instabilidade na base do equipamento durante o monitoramento, pode-se utilizar prismas de referência (que também devem ser estáveis) para ajustar o posicionamento da estação. Essa metodologia tem como vantagem o fato de que, ao se utilizar múltiplos prismas de referência, é possível mitigar o efeito causado por deslocamentos individuais. Por exemplo, se somente um dos prismas de referência se deslocar e o restante permanecer estável, o impacto desse deslocamento será minimizado nos prismas de monitoramento. Além disso, é possível descartar a medição do prisma deslocado e orientar o sistema a partir das leituras dos outros prismas de referência. A figura 5 ilustra o posicionamento dos prismas numa barragem utilizando as coordenadas calculadas pelo equipamento.



Figura 4: Vista aérea de prismas de monitoramento e de referência posicionados numa barragem. Fonte: Elaborado pelo autor.

3.1.2.4 Vetores deslocamento

As coordenadas de todos os pontos de monitoramento são calculadas a cada ciclo de leitura do equipamento. Após pelo menos duas medições consecutivas é possível calcular seu vetor deslocamento. Por padrão, o equipamento calcula o vetor sempre com relação à primeira coordenada indicada para tal cálculo no sistema, conhecida como coordenada “null”. Com as variações em X, Y e Z (easting, northing e height) é possível calcular o vetor da seguinte forma:

$$\vec{3D} = \langle (X_f - X_i), (Y_f - Y_i), (Z_f - Z_i) \rangle$$

sendo o módulo do vetor calculado como:

$$\|\vec{3D}\| = \sqrt{(X_f - X_i)^2 + (Y_f - Y_i)^2 + (Z_f - Z_i)^2}$$

No presente trabalho, será interpolada espacialmente e temporalmente a componente 2D do vetor deslocamento, considerando somente as variações em easting e northing.

3.1.2.5 Tilts

O equipamento também realiza medições internas para garantia de estabilidade. É possível consultar a qualquer momento o nível de bolha interno, conhecido como tilt. Ele é medido no sentido transversal e longitudinal em relação ao norte.

É importante ressaltar que existem parâmetros internos que impedem que o equipamento continue a realizar medições caso os valores de tilt ultrapassem os limites previamente estabelecidos. Isso garante que as medições sempre ocorrerão dentro de um alinhamento aceitável do equipamento.

3.1.3 Correções aplicadas

Como o equipamento possui uma alta precisão, os efeitos de temperatura e pressão afetam fortemente o resultado. Como é previsível o efeito que essas condições causam nos dados, são feitas algumas correções. Assim, são criados os parâmetros de distância inclinada corrigida e ângulo vertical corrigidos. O ângulo horizontal também recebe uma

correção dos efeitos atmosféricos, além de uma transformação nos seus valores, que deixam de ter como referência o limbo interno do equipamento e passam a ter como referência o norte (LEICA, 2013).

3.2 Uso de radares no monitoramento geotécnico

Além das ETRs, um dos equipamentos também muito utilizados para o monitoramento geotécnico são os radares interferométricos. Como descrito por Nader (2012), os radares, através da sua capacidade de medir, quantificar e gerar indicadores chave de desempenho (KPIs), são capazes de gerar benefícios tangíveis na cadeia de valor mineral.

Quando utilizados no monitoramento da estabilidade de taludes, com a aplicação de metodologias de monitoramento e controle adequadas, as atividades da cadeia de valor mineral relacionadas à geotecnia são positivamente impactadas (NADER, 2012).

Nos últimos anos, tem se utilizado radares interferométricos não somente para o monitoramento de cavas em operação, mas também de barragens de rejeito. Para tais estruturas é comum encontrar ETRs e radares interferométricos na operação de monitoramento. Existem dois tipos gerais de radares: os tipo SAR (Synthetic aperture radar) e RAR (Real aperture radar). No presente trabalho, trataremos sobre os equipamentos tipo SAR

3.2.1 Radar de abertura sintética

3.2.1.1 Funcionamento

O equipamento funciona gerando uma forma de onda sinusoidal com uma frequência precisa que é transmitida através de uma antena como um feixe estreito de ondas eletromagnéticas. Essa onda é então direcionada para a superfície que está sendo monitorada, onde uma pequena fração da energia é refletida para o radar (também conhecido como retroespalhamento). A onda retroespalhada, ou eco, é interceptada pela antena e convertida em um sinal que é processado e analisado usando interferometria para determinar as diferenças nos valores de fase das ondas recebidas entre duas medições

consecutivas. Se uma mudança de fase for registrada, ocorreu um movimento do alvo entre as duas aquisições (MONSERRAT, 2014) A magnitude do deslocamento pode ser calculada com base no deslocamento de fase, conforme ilustrado na Figura 6.

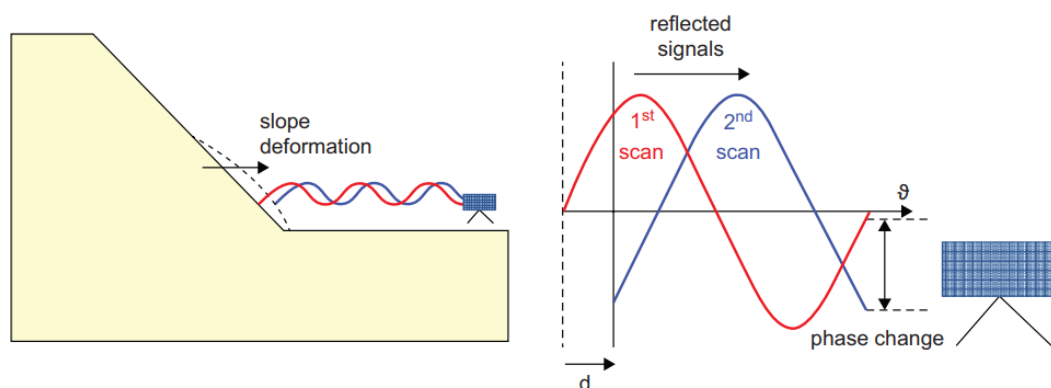


Figura 5: Medição do deslocamento usando radar interferométrico. Ele é calculado comparando mudanças de fase dos sinais refletidos entre duas medições de radar consecutivas. Fonte: Eberhardt & Sharon, 2020.

Os radares SAR terrestres mais comuns (radares GB-Ground Based) consistem em um sensor de radar conectado a um módulo de fonte de alimentação e um notebook que possui um software controlador instalado. Este computador é conectado via rede a outro notebook, longe do ambiente de mina para pós-processar os arquivos em outro software dedicado.

Esse software recebe os dados do radar no campo e atribui a coordenada espacial correta a cada pixel de radar que são então plotados para visualização otimizada em um modelo de terreno digital (DTM). Depois disso, o algoritmo de processamento calcula o deslocamento ocorrido entre varreduras consecutivas para cada pixel de radar, após a remoção da contribuição ruidosa causada pelas flutuações dos parâmetros atmosféricos. Como a posição do radar e o norte são conhecidos, o software converte as coordenadas do mapa de radar (R, θ) em coordenadas DTM (X, Y e Z). Em suma, para cada pixel de radar, o software lê um valor de alcance e azimuth, então o deslocamento desse pixel é atribuído à célula DTM correspondente nessa faixa e azimuth específicos.

3.2.1.2 Limitações e correções de medição

As capacidades de medição dos sistemas de radar são, em última análise, limitadas por distúrbios aleatórios sempre presentes, comumente chamados de ruído. A precisão de medição do radar depende da força do eco em relação ao ruído mínimo do radar. Isso é chamado de relação sinal-ruído (SNR – Signal to Noise Ratio) da medição e é determinado pela energia do eco retornado (EBERHARDT & SHARON, 2020). Em instrumentos modernos, a diferença de fase introduzida pelo ruído intrínseco é insignificante e geralmente abaixo da precisão do instrumento. As diferentes contribuições para a mudança de fase total ocorrida entre diferentes imagens de radar podem ser representadas pela seguinte equação:

$$\Delta\varphi = \Delta\varphi_{def} + \Delta\varphi_{atm} + \Delta\varphi_{noise}$$

onde $\Delta\varphi_{def}$ é o movimento real, $\Delta\varphi_{atm}$ é o efeito atmosférico e $\Delta\varphi_{noise}$ é um efeito de ruído aleatório.

É importante notar que, embora os radares possam trabalhar com qualquer condição meteorológica, noite e dia, a propagação das ondas eletromagnéticas pela atmosfera é afetada pela variação de temperatura, umidade e pressão.

Outro tipo de fator de perturbação é a presença de vegetação na superfície monitorada. Antes de iniciar qualquer discussão sobre o processo de interação entre vegetação e reflectância eletromagnética (REM), é importante entender o que pode ser considerado vegetação do ponto de vista da aplicação de técnicas de sensoriamento remoto: pode ser desde uma folha isolada até um conjunto de plantas.

Meneses et al (2019) descrevem que a caracterização espectral de um grupo de folhas, aqui chamada de dossel, é muito mais complexa do que a caracterização espectral de folhas individuais. Goel (1988) define tanto as bases conceituais da interação da REM com um grupo de folhas quanto os modelos matemáticos que tentam descrevê-lo. Para esse autor, a radiação solar em um dossel é espalhada e refletida de forma muito complexa, alterando as direções dos raios refletidos e a composição espectral da reflectância. Assim, a atmosfera ao redor do campo de monitoramento tem uma forte influência quando a evapotranspiração (ET) começa. Essas alterações são causadas tanto

por parâmetros inerentes ao próprio dossel quanto por fatores ambientais e geométricos relacionados à iluminação e ao posicionamento do sensor, que podem afetar a mudança de fase das aquisições que são feitas continuamente.

A presença de vegetação reduz o SNR e pode aumentar o nível de ruído nos dados a serem processados. Como consequência, pode gerar uma mudança na fase do sinal retroespalhado e uma descorrelação dos dados do radar que tornará mais desafiador o cálculo de informações precisas de deslocamento ocorridas no cenário monitorado.

3.2.1.3 *Dados gerados*

Após duas aquisições consecutivas do SAR são gerados os dados de deslocamento, obtidos através do cálculo da diferença de fase em cada um dos pixels observados. Esses valores são plotados em um mapa de deslocamento, em que cada célula do DTM recebe o valor do deslocamento acumulado entre um período inicial e final definido pelo usuário do sistema. É importante lembrar que o deslocamento medido por um radar SAR é somente uma componente do vetor deslocamento real, na direção da linha de visada do radar (LOS - line of sight) até a estrutura. A figura 7 abaixo ilustra um mapa de deslocamento.

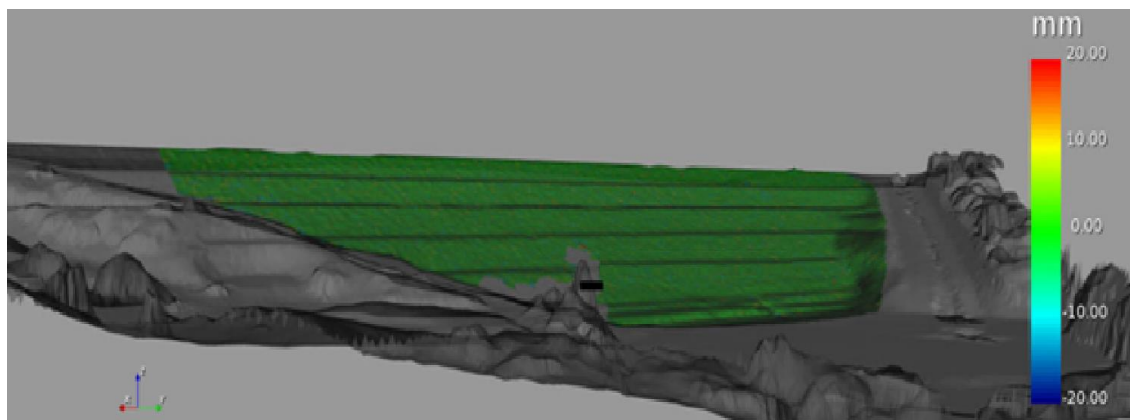


Figura 6: Mapa de deslocamento obtido por um radar SAR. Cada pixel da imagem possui um valor de deslocamento atribuído a ele. Fonte: Extraído do software Ibis Guardian pelo autor.

Simultaneamente é calculado o mapa de amplitude do sinal, que mostra também para cada célula do DTM a perda de energia da onda retroespalhada. Entretanto,

diferentemente do mapa de deslocamento, a amplitude é visualizada somente para um período único (usualmente o período final definido ao se visualizar o mapa de deslocamento). A figura 8 ilustra a mesma região visualizada pelo mapa de amplitude.

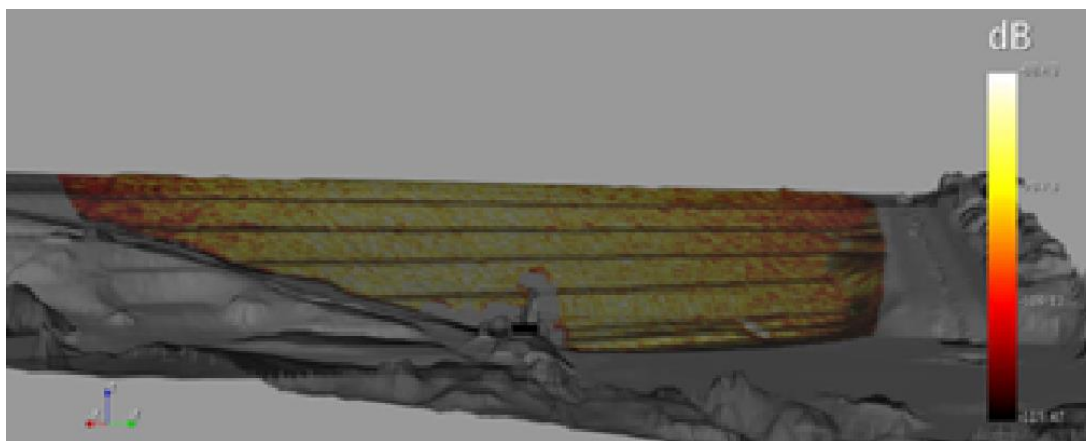


Figura 7: Mapa de amplitude do sinal obtido por um radar SAR. Cada pixel da imagem possui um valor relativo à perda de energia pela onda. Fonte: Extraído do software Ibis Guardian pelo autor.

Obtidos os dados de deslocamento, também são gerados dados de velocidade e aceleração, que são mostrados em forma de mapa 3D assim como os outros dados. Por fim, é gerado também uma informação de sensibilidade. Esse dado é apresentado como um valor entre 0 e 1 e representa a posição da LOS em relação à direção preferencial de movimento, sendo 0 a pior posição e 1 a melhor. Por exemplo, um radar que olha horizontalmente para a berma de uma barragem apresentaria sensibilidade igual a 0, pois a direção preferencial de movimento (verticalmente) é perpendicular à LOS. Por outro lado, se esse mesmo radar observa um talude com ângulo de 30° , a sensibilidade seria igual a 0,5.

3.2.2 Comparação entre os dados de ETR e SAR

Tanto ETRs quanto radares SAR têm sido amplamente utilizados no monitoramento geotécnico. As ETRs oferecem uma grande vantagem que é a capacidade de medir o deslocamento inteiro (o vetor 3D do deslocamento) com uma precisão milimétrica (LEICA, 2013), porém essa medição é feita de forma pontual em cada um dos prismas e os dados são visualizados em forma de gráfico de linhas. Elas também são

equipamentos mais baratos do que os radares. A figura 9 abaixo exemplifica a visualização de um deslocamento que foi registrado por uma ETR. Cada medição é plotada como um ponto numa dispersão em que o eixo X é a data e o eixo Y o valor da distância inclinada corrigida. A figura 10 mostra a comparação da medição de um prisma que se movimentou com outro que permaneceu estável no mesmo período. Importante ressaltar que a escala do eixo Y na figura 9 está absoluta e na figura 10 está relativa, mostrando somente a variação do parâmetro dentro do intervalo definido.

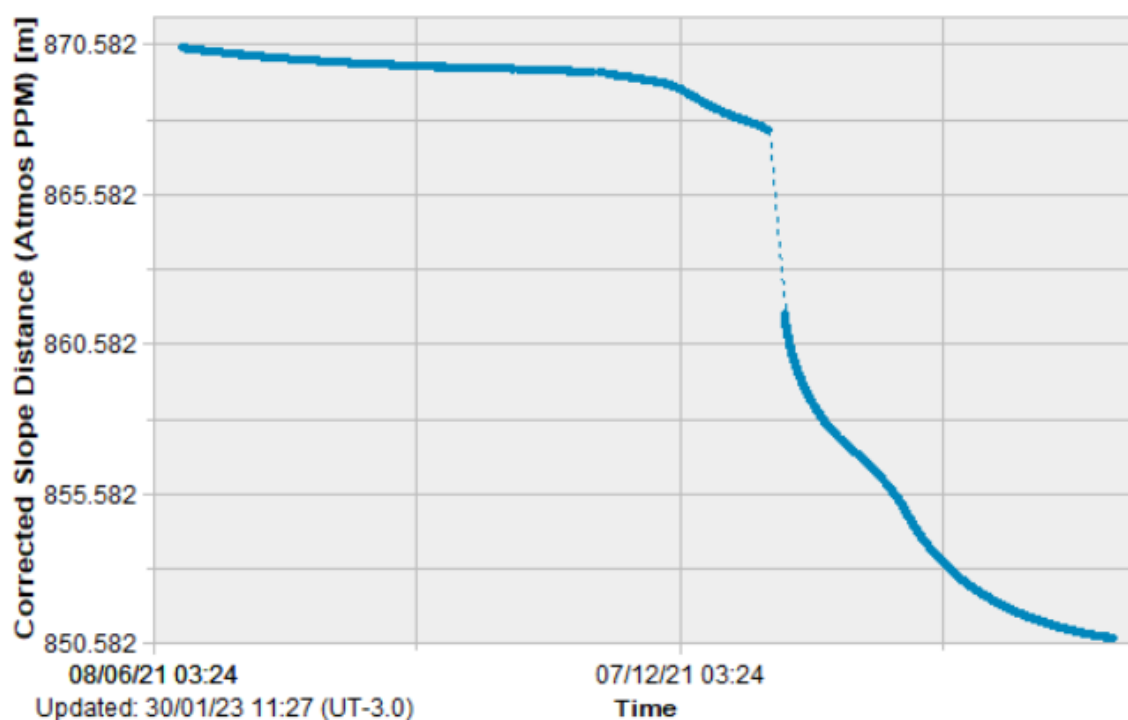


Figura 8: Exemplo de deslocamento registrado por uma ETR. Fonte: Extraído do software Geomos Now, pelo autor.

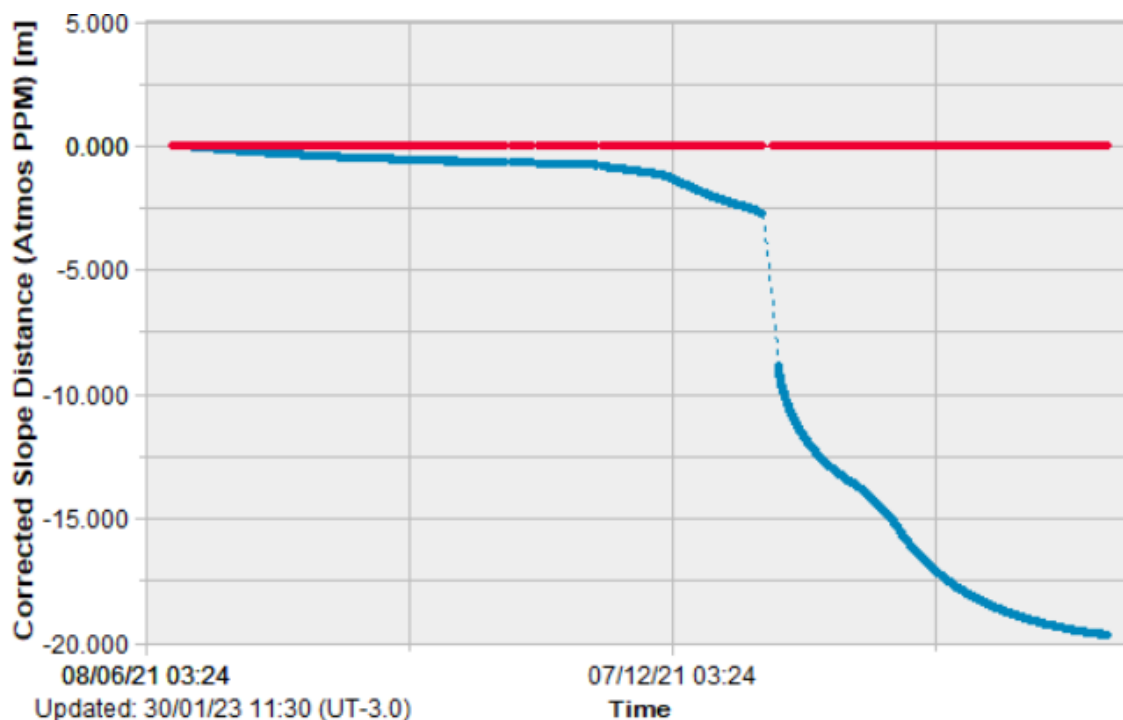


Figura 9: Comparação entre dois prismas. O prisma representado pela série em azul apresentou forte deslocamento no período, enquanto o vermelho permaneceu estável. Fonte: Extraído do software Geomos Now, pelo autor.

Por outro lado, os radares oferecem uma cobertura por toda a região monitorada, sendo a distância entre medições determinada pela resolução atingida pelo equipamento. Entretanto, as medições de deslocamento são limitadas pela LOS e podem subdimensionar eventos geotécnicos, caso o posicionamento não seja o ideal.

Ambos os equipamentos apresentam suas aplicações, vantagens e desvantagens e, por isso, muitas vezes são utilizados em conjunto no monitoramento de estruturas. Nesses casos, é possível propor uma análise conjunta dos dados, em que o resultado proveniente da ETR pode ser utilizado para validar o resultado do radar e vice-versa. No entanto, tal análise não é feita de forma trivial: a medição de um mesmo fenômeno geotécnico é feita de forma diferente por cada um deles (a ETR mede ângulos e distâncias de um prisma fincado no solo, o radar mede diferença de fase de uma onda retroespalhada que atingiu a estrutura em momentos diferentes), gerando dados diferentes (a ETR gera o dado tridimensionalmente, o radar somente a componente na direção da LOS), que são

mostrados de forma diferente (gráficos de linha na ETR versus imagens com valores por pixel no radar) e com ruídos específicos de cada equipamento.

A fim de se comparar os resultados, é preciso contornar alguns desses problemas. A interpolação dos dados da ETR aparece nesse contexto como um complemento a essa análise, oferecendo uma visualização similar à visualização apresentada pelo radar. Outra técnica que também pode ser aplicada para auxiliar nessa análise é a projeção do vetor 3D do deslocamento calculado pela ETR na LOS do radar. Conhecendo-se as coordenadas do radar e as coordenadas iniciais do prisma antes do movimento, isso pode ser feito utilizando-se a seguinte operação vetorial:

$$\text{proj}_V U = \frac{U \cdot V}{|V|^2} V$$

onde U é o vetor 3D calculado pela ETR, V é o vetor que aponta do radar para o prisma medido e $\text{proj}_V U$ é a projeção do vetor calculado pela ETR na LOS do radar (HEFFERON, 2018).

No presente trabalho será tratada somente a interpolação dos dados da ETR nos domínios temporal e espacial, a fim de gerar uma visualização semelhante à dos radares.

3.3 Conceitos básicos de Estatística

O conhecimento estatístico é construído de forma dependente: não é possível entender conceitos avançados sem antes estudar os conceitos básicos que estão por trás de determinada análise. Assim, até ser possível conceituar a krigagem, a análise espaço-temporal e todas as suposições por trás dos métodos utilizados, é necessário criar uma base de conhecimento estatístico forte. Nos próximos tópicos serão apresentadas as principais definições estatísticas necessárias ao estudo.

3.3.1 Tipo de variáveis

São chamadas de variáveis quaisquer valores que representem uma característica de interesse numa pesquisa. Elas são medidas em cada elemento da amostra ou população, apresentando valores que podem ser numéricos ou não, sendo classificadas da seguinte forma:

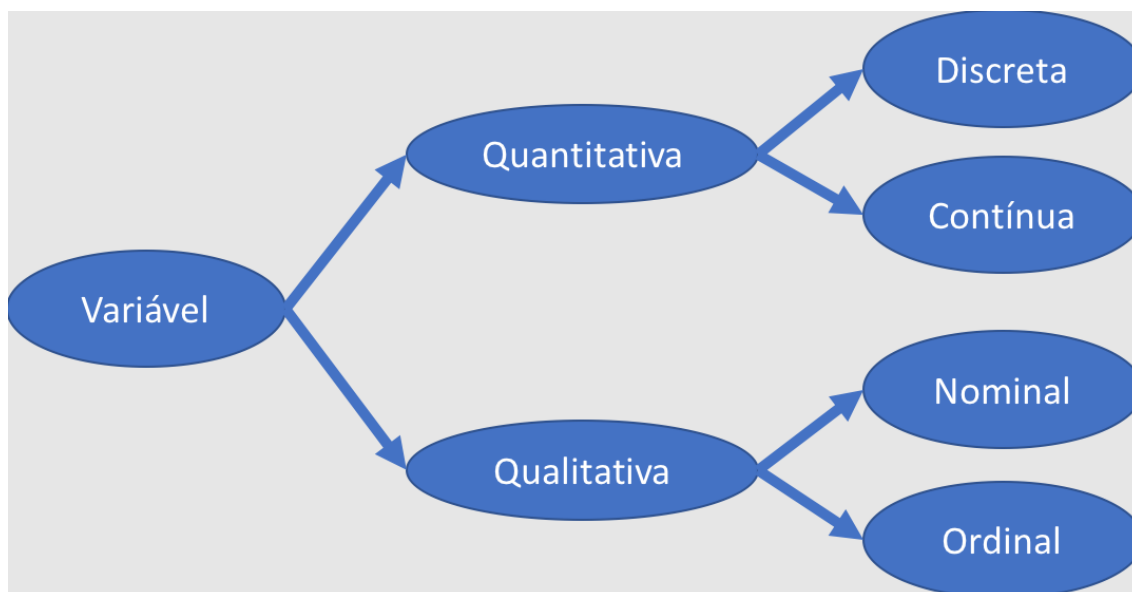


Figura 10: Tipos de variáveis. Fonte: BUSSAB, 2010.

3.3.1.1 *Quantitativa Discreta*

Variáveis quantitativas são todas aquelas representadas por meio de números, que podem ser resultantes de uma medida ou contagem. No caso das variáveis quantitativas discretas, os valores são determinados por um conjunto finito de números, resultado de uma contagem (como número de moradores de uma cidade ou número de filhos) (BUSSAB, 2010).

3.3.1.2 *Quantitativa Contínua.*

As variáveis quantitativas contínuas são pertencentes a um intervalo de números reais, podendo assumir infinitos valores. Elas representam uma mensuração, como distância, altura ou peso (BUSSAB, 2010).

3.3.1.3 *Qualitativa Nominal*

Variáveis qualitativas são aquelas que representam um atributo, uma qualidade, de um elemento pesquisado. Não possuem valores numéricos e são utilizadas para categorizar os dados. As variáveis qualitativas nominais são aquelas em que não existe

nenhuma ordem intrínseca à representação, como, por exemplo, cor dos olhos, sexo, cor do cabelo (BUSSAB, 2010).

3.3.1.4 Qualitativa Ordinal

Quando uma variável qualitativa apresenta uma ordem nos resultados, ela é chamada de ordinal. São exemplos de variáveis qualitativas ordinais o nível de escolaridade ou ano da observação (BUSSAB, 2010)

3.3.2 Medidas de tendência central

Medidas de tendência central, como o nome revela, descrevem e definem o centro dos dados. As principais medidas de tendência central são a média, moda e mediana. São simples de calcular, porém a dificuldade da análise reside em qual das medidas escolher, pois não existem regras definidas para isso. Apesar disso, algumas normas de boas práticas foram definidas. Uma medida de tendência central é satisfatória caso possua as seguintes características (AGARWAL, 2006):

1. Ela deve ser baseada em todas as observações;
2. Valores extremos não devem estar afetando muito a medida;
3. Deve ser rigidamente definida, o que significa que precisa ter um valor definido. O investigador não deve precisar fazer escolhas;
4. Deve ser capaz de um maior tratamento algébrico, o que significa que essas medidas devem ser capazes de serem utilizadas na formulação de outros cálculos;
5. Devem ser estáveis quanto à amostragem. Isso significa que, se um número de amostras de mesmo tamanho são retiradas de uma população, a melhor medida de tendencia central deve ser aquela que tem a menor variação entre as amostras.

3.3.2.1 Mediana

A mediana, também denotada \tilde{X} , é o valor que divide um conjunto de dados ao meio, deixando os 50% menores valores de um lado e os 50% maiores valores do outro

lado. Para seu cálculo, é necessário organizar o conjunto de dados em ordem crescente. Caso n seja par, a mediana é dada pelo valor médio dos dois elementos centrais. Caso seja ímpar, ela é o próprio valor que ocupa a posição central (AGARWAL, 2006).

3.3.2.2 *Moda*

Dado um conjunto de valores, a Moda desses valores será aquele que se repetir o maior número de vezes. uma amostra que não tem moda é chamada de amodal, também podem ocorrer duas modas (chamada bimodal) ou mais de duas modas (chamada multimodal). A moda é a única medida de tendência central que pode ser usada para descrever dados qualitativos (AGARWAL, 2006).

3.3.2.3 *Média aritmética*

A média aritmética, comumente referida como simplesmente média, é, de um modo geral, a mais importante de todas as medidas de tendência central, devido à sua facilidade de compreensão. Usualmente, a média amostral é denotada por \bar{x} e a média populacional é denotada por μ . Sendo $(x_1, x_2, x_3, \dots, x_n)$ uma amostra de n observações de uma variável x , para calcular a média é necessário somar todos os valores obtidos de x e dividir por n , segundo a seguinte fórmula matemática:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

A média possui quase todas as características de uma boa medida de tendência central citadas anteriormente, com exceção do número 2, pois ela é fortemente afetada por valores extremos (AGARWAL, 2006).

3.3.3 **Medidas de dispersão**

Medidas de tendência central sozinhas não são capazes de produzir uma imagem clara dos dados, pois não definem como é a variabilidade das observações. É possível que dois conjuntos de dados bem diferentes possuam valores iguais em suas medidas centrais. Devido a isso, se faz necessário estudar também a dispersão (ou variabilidade) dos dados em questão. As medidas mais utilizadas são a amplitude, a variância, o desvio padrão e o

coeficiente de variação. Segundo Agarwal, as propriedades de uma boa medida de dispersão são:

1. Deve ser baseada em todos os valores de uma série;
2. Não deve ser susceptível a flutuações por causa da amostragem;
3. Deve ser capaz de maior tratamento algébrico;
4. Deve ser rigidamente definida: todos os investigadores devem chegar ao mesmo valor para um mesmo conjunto de dados;
5. É preferível que a unidade de medida da dispersão seja igual à unidade de medida das observações;
6. Deve ser pouco afetada por valores extremos.

3.3.3.1 *Amplitude*

É usualmente denotada como simplesmente A. Seu cálculo é a diferença entre o maior e o menor valor do conjunto de dados. A maior vantagem da amplitude é a facilidade do cálculo, porém tem a desvantagem de levar em conta apenas dois valores de todo o conjunto de dados, o que também a torna extremamente sensível a valores extremos (AGARWAL, 2006).

3.3.3.2 *Variância e Desvio padrão*

A variância mede a variação do conjunto de dados em torno da média. Pode ser calculada para a população, sendo denotada σ^2 , ou para a amostra, denotada por s^2 . Ela é calculada com base na média dos quadrados dos desvios da média, seguindo as fórmulas matemáticas:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

A variância se mostra ótima para representar a dispersão dos dados, porém ela possui uma grande desvantagem: sua unidade de medida é igual ao quadrado da unidade de medida das observações. Isso atrapalha o entendimento e a comparação entre o

resultado obtido e os dados reais. Para se contornar esse problema, faz a raiz quadrada, chegando-se no desvio padrão (AGARWAL, 2006), que é uma das medidas de dispersão mais utilizadas.

O desvio padrão populacional é denotado por σ e o amostral é denotado por s :

$$s = \sqrt{s^2}$$

$$\sigma = \sqrt{\sigma^2}$$

3.3.3.3 Coeficiente de variação

Embora o desvio padrão seja uma boa medida de dispersão, amplamente utilizada, ele mede a variabilidade dos dados em termos absolutos. Isso impede a comparação de dados com médias diferentes ou unidades de medida diferentes. Para suprir essa necessidade, utiliza-se o coeficiente de variação, comumente chamado de CV. Ele calcula a variabilidade em termos relativos, seguindo a fórmula:

$$CV = \frac{s}{\bar{x}} * 100$$

Ele mede a dispersão em termos relativos, sendo muito útil ao se comparar a variabilidade de diferentes objetos de estudo (AGARWAL, 2006).

3.3.4 Covariância

A covariância, também conhecida como variância conjunta, é uma medida do grau de interdependência numérica entre duas variáveis aleatórias. Devido a isso, variáveis independentes tem covariância zero (MILONE, 2009). Ela mede a relação linear entre duas variáveis. É possível utilizar a covariância para compreender a direção da relação entre as variáveis. Valores de covariância positivos indicam que valores acima da média de uma variável estão associados a valores acima da média da outra variável, bem como valores abaixo da média de uma variável estão associados a valores abaixo da média da outra variável. Valores de covariância negativos indicam que valores acima da média de uma variável estão associados com valores abaixo da média da outra variável.

A covariância é uma boa medida para determinar a relação linear entre duas variáveis, porém seus resultados não são padronizados, podendo variar de menos infinito a mais infinito. Isso torna a relação linear ideal dependente dos dados e, devido a essa falta de padronização, é difícil determinar a força da relação entre as variáveis. As fórmulas matemáticas para a covariância amostral s_{xy} e populacional σ_{xy} são:

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) * (y_i - \bar{y})}{n - 1}$$

$$\sigma_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) * (y_i - \bar{y})}{n}$$

3.3.5 Correlação

A covariância, como mostrado anteriormente, busca mostrar o comportamento de interdependência linear entre duas variáveis, mas é uma medida dimensional que é afeta pelas unidades de medida das séries X e Y. A fim de corrigir esse problema foi criada a medida de correlação, que nada mais é do que uma normalização da covariância e representada por um número sem unidade de medida que varia entre -1 e 1 (sendo capaz de fornecer a força da relação).

Também conhecido como coeficiente de correlação linear de Pearson, suas fórmulas matemáticas para correlação populacional e amostral são dadas por:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x * \sigma_y}$$

$$r_{xy} = \frac{s_{xy}}{s_x * s_y}$$

Assim, é possível perceber que a correlação é a covariância de X e Y dividida pelo produto de seus desvios padrão.

Caso o coeficiente assumo o valor -1, significa que as variáveis possuem correlação negativa perfeita, ou seja, quando X aumenta, Y diminui na mesma magnitude. Um coeficiente de correlação igual a zero significa que não há correlação linear entre as variáveis. Por fim, um coeficiente igual a 1 significa que as variáveis possuem correlação positiva perfeita, ou seja, quando X aumenta, Y também aumenta na mesma magnitude.

As figuras 3, 4 e 5 mostram o comportamento esperado de duas variáveis nas diferentes situações descritas

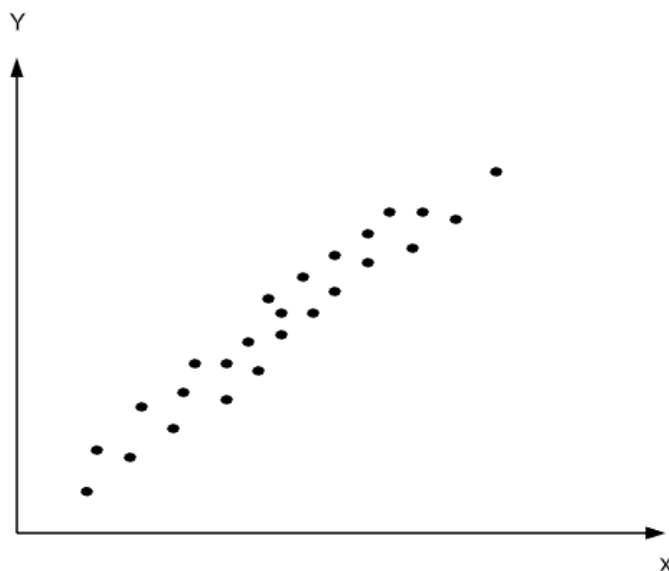


Figura 11: Correlação linear positiva. Fonte: Elaborado pelo autor

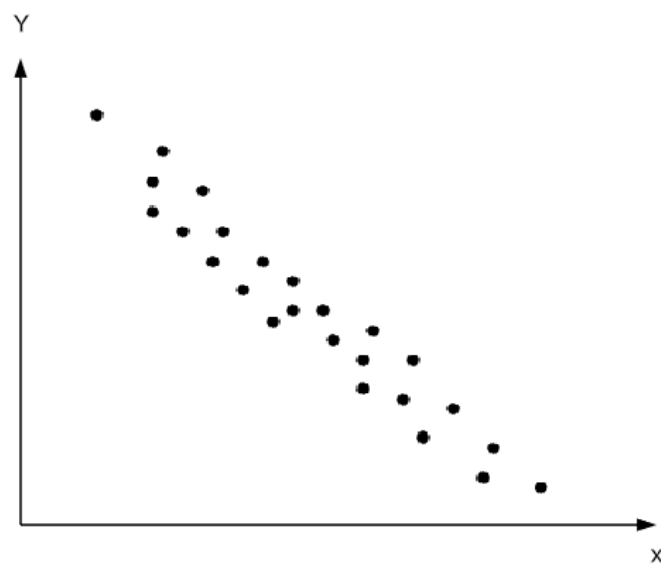


Figura 12: Correlação linear negativa. Fonte: Elaborado pelo autor

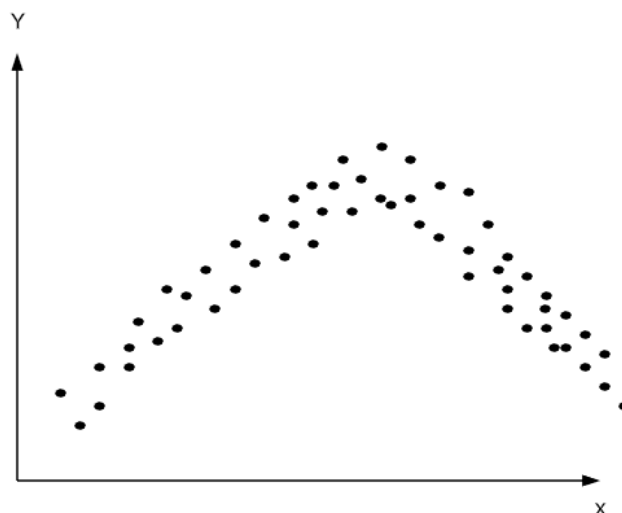


Figura 13: Sem correlação linear. Fonte: Elaborado pelo autor

3.3.6 Variograma

A análise geoestatística busca determinar um modelo de correlação espacial dos dados e modelar como a variável (ou variáveis) de interesse se comporta num campo espacial pré-definido. Devido a isso, utiliza-se a função variograma, que é responsável por medir a variação do valor de uma variável em relação às restantes de uma mesma amostragem, porém em locais diferentes no espaço (YAMAMOTO, 2013). Para seu cálculo, primeiro escolhe-se uma direção preferencial. Após isso, são re-amostrados pares de pontos separados por uma distância h . Sendo 2γ a função variograma, sua formulação é:

$$\begin{aligned} 2\gamma(x, x + h) &= \text{var}(Z(x) - Z(x + h)) = \\ &= E\left(\left((Z(x) - E(x)) - (Z(x + h) - E(x + h))\right)^2\right) \end{aligned}$$

Considerando a estacionariedade da média (ou seja, ela é constante por todo o campo espacial considerado), a formulação da função variograma pode ser reduzida a:

$$2\gamma(x, x + h) = \frac{1}{N(h)} * \sum_{N(h)} [Z(x) - Z(x + h)]^2$$

Essa equação possui extrema similaridade com a equação de variância, porém, ao substituir o valor médio pelo valor separado a uma distância h em uma determinada direção, ela passa a considerar a variabilidade espacial da variável analisada.

O variograma experimental é calculado a partir dos valores amostrais de uma variável de interesse, sendo comumente visualizado a partir de um gráfico de dispersão, onde é possível perceber as diversas propriedades de interesse da função. A figura 3 mostra as principais estruturas de um variograma: O patamar (ou sill), o efeito pepita (ou nugget effect) e a amplitude (ou range).

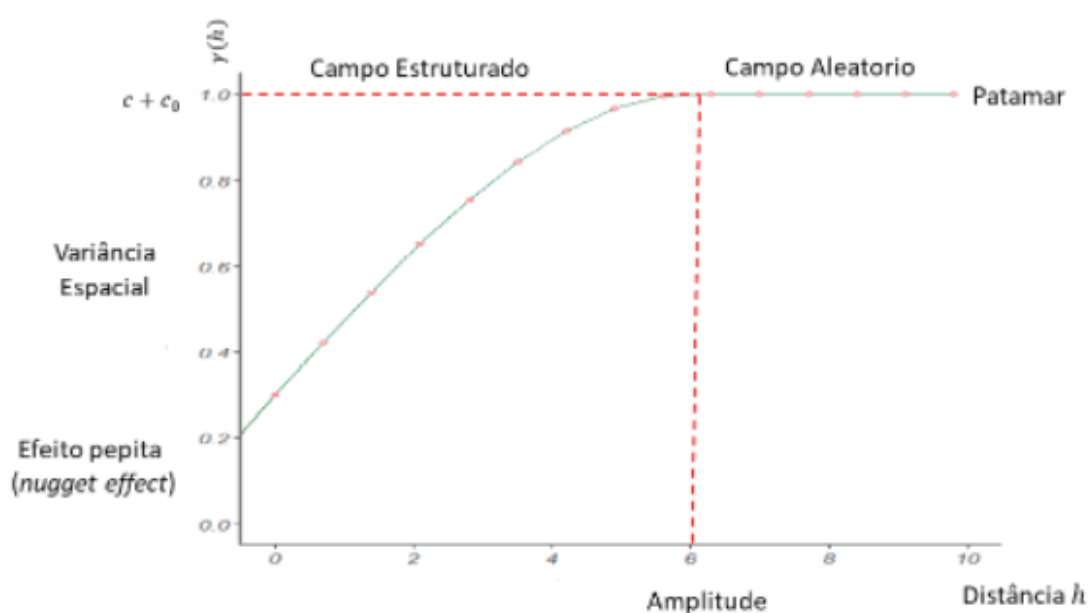


Figura 14: Variograma e suas propriedades. Fonte: YAMAMOTO, 2013

O patamar é o valor a partir do qual não existe mais correlação espacial na variável analisada. Ele pode ser reconhecido pelo comportamento assintótico da curva da função variograma: Por mais que a distância h aumente, não é vista uma variação no valor da função. Essa distância “máxima” h é denominada amplitude.

Ao se calcular o variograma experimental é necessário escolher uma direção onde o h irá crescer. Quando o variograma apresenta configurações similares para as várias direções possíveis, o fenômeno é dito ser isotrópico. Caso contrário, ele possui algum tipo de anisotropia, que é a direção preferencial de ocorrência do fenômeno. Se os variogramas apresentam a mesma forma, mas com diferentes amplitudes, denomina-se anisotropia

geométrica. De outra forma, o zoneamento da variável ou mistura de populações distintas geram uma anisotropia mais complexa, conhecida como anisotropia zonal (STURARO, 2015).

Em suma, a função variograma caracteriza a variabilidade espacial de uma variável, sendo capaz de encontrar direções preferenciais de ocorrência do fenômeno.

3.4 Conceitos de Probabilidade

Assim como os conhecimentos sobre estatística, entender sobre os conceitos de probabilidade é fundamental para o embasamento teórico de quaisquer metodologias que lidem com inferência estatística. Nos próximos tópicos serão apresentados os principais conceitos básicos de probabilidade.

3.4.1 Experimentos aleatórios

São classificados como experimentos aleatórios todos aqueles cujos resultados podem não ser iguais quando repetidos sob condições essencialmente idênticas. Não é possível conhecer um particular valor do experimento antes dele acontecer, porém é possível descrever todas as possibilidades, todos os possíveis resultados, o que se chama espaço amostral. Usualmente, após repetir o experimento muitas vezes, surge uma regularidade nas ocorrências (AGARWAL, 2006).

3.4.1.1 Espaço amostral

O conjunto de todos os possíveis resultados de um experimento aleatório é chamado espaço amostral. Cada experimento aleatório sempre estará associado a um conjunto de resultados possíveis, que é seu espaço amostral. Quaisquer subconjuntos do espaço amostral são chamados de eventos

3.4.2 Axiomas da probabilidade

Dado um espaço amostral Ω e um evento A , a probabilidade do evento A ocorrer é denotada por $P(A)$. A função $P(A)$ só será uma probabilidade se ela satisfizer três condições básicas:

$$0 \leq P(A) \leq 1;$$

$$P(\Omega) = 1 \text{ e};$$

$P(A1 \cup A2 \cup A3 \cup \dots) = P(A1) + P(A2) + P(A3) + \dots$, se os eventos A_n forem disjuntos, ou seja, mutuamente exclusivos (AGARWAL, 2006).

A probabilidade de ocorrência de um evento A pode ser calculada como a relação entre o número de pontos favoráveis (f) à realização do evento A e o número total de pontos (n), ou seja:

$$P(A) = \frac{f}{n}$$

Caso o fenômeno sob análise puder ser repetido nas mesmas condições e de forma independente, a probabilidade de A pode então ser calculada da seguinte forma:

$$P(A) = \frac{\text{Número de ocorrências de A}}{\text{Total de repetições do evento}}$$

3.4.3 Interações entre eventos

3.4.3.1 Interseção de eventos

A interseção de dois eventos A e B corresponde à ocorrência simultânea dos eventos A e B, contendo todos os pontos do espaço amostral comuns a A e B. É denotada por $A \cap B$ (AGARWAL, 2006). A interseção é ilustrada pelo diagrama de Venn na figura 7.

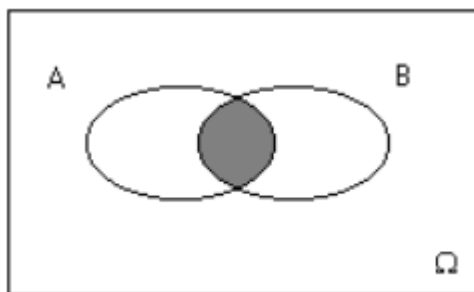


Figura 15: Diagrama de Venn ilustrando a interseção de dois eventos A e B. Fonte: Elaborado pelo autor

3.4.3.2 União de eventos

A união dos eventos A e B , denotada por $A \cup B$, retrata a ocorrência de A ou B ou de ambos, ou seja, é a probabilidade de ocorrência de pelo menos um dos eventos (AGARWAL, 2006). A figura 8 ilustra a união dos dois eventos A e B .

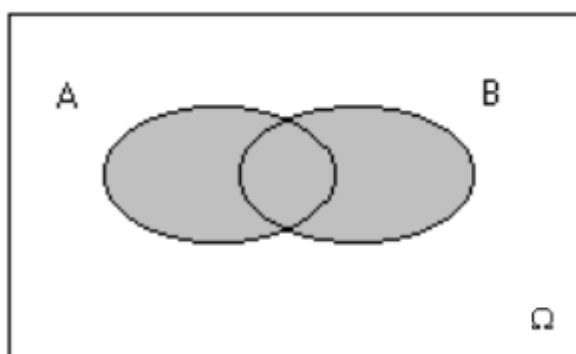


Figura 16: Diagrama de Venn ilustrando a união de dois eventos A e B . Fonte: Elaborado pelo autor

3.4.3.3 Eventos disjuntos ou mutuamente exclusivos

Quando não há possibilidade de dois eventos ocorrerem de forma mútua, eles são chamados de eventos disjuntos ou mutuamente exclusivos. Eles não têm elementos em comum, o que significa dizer que $A \cap B = \emptyset$ (AGARWAL, 2006). A figura 9 ilustra essa situação.

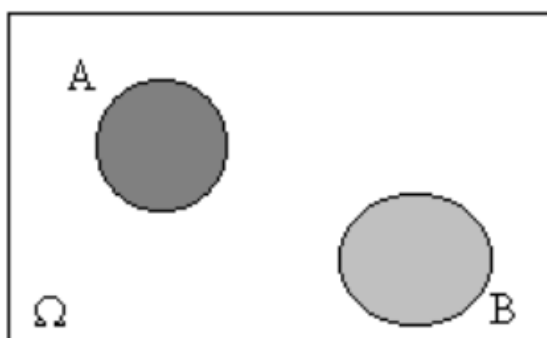


Figura 17: Diagrama de Venn ilustrando dois eventos disjuntos, A e B . Fonte: Elaborado pelo autor

3.4.3.4 Eventos complementares

Dois eventos são ditos complementares quando sua união corresponde ao espaço amostral e sua interseção é vazia, ou seja, $A \cup B = \Omega$ e $A \cap B = \emptyset$ (AGARWAL, 2006). É possível postular também que $P(A) = 1 - P(B)$ e que $P(A^c) = P(B) = 1 - P(A)$ (AGARWAL, 2006). A figura 10 ilustra essa situação.

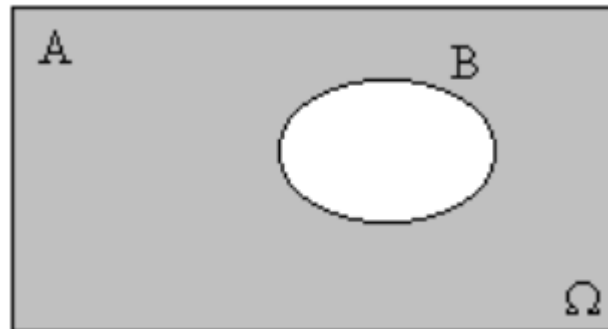


Figura 18: Diagrama de Venn ilustrando dois eventos complementares, A e B. Fonte: Elaborado pelo autor

3.4.4 Probabilidade condicional

Em diversas situações, saber da ocorrência de um evento B modifica a probabilidade de ocorrência do evento A. A esses casos dá-se o nome de probabilidade condicional (AGARWAL, 2006).

A probabilidade de ocorrência de um evento A dado B, é a probabilidade de ocorrência de A, sendo que B já ocorreu. Ela pode ser calculada através da divisão da probabilidade de ocorrência de A e B (interseção de A e B) pela probabilidade de ocorrência de B, seguindo a formulação:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{sendo } P(B) > 0$$

3.4.4.1 Independência de eventos

Dois eventos são ditos independentes caso a ocorrência de um deles não afete a probabilidade de ocorrência do outro, ou seja, $P(A|B) = P(A)$ e $P(B|A) = P(B)$. Assim, se dois eventos são independentes, pode-se formular:

$$P(A \cap B) = P(A) * P(B)$$

3.4.4.2 Teorema de Bayes

O teorema de Bayes é uma fórmula matemática utilizada para calcular a probabilidade de um evento, dado que outro ocorreu. Suponha que um evento B só pode ocorrer devido a uma das causas complementares E_1, E_2, \dots, E_n eventos de um mesmo espaço amostral Ω (AGARWAL, 2006). Dado que um evento B tenha ocorrido, a probabilidade que tenha se manifestado devido a uma das causas E_1, E_2, \dots , ou E_n pode ser calculada por:

$$P(E_i|B) = \frac{P(E_i \cap B)}{P(B)} = \frac{P(E_i) * P(B|E_i)}{\sum_i^n [P(E_i) * P(B|E_i)]}$$

De forma mais concisa, dado que um evento B ocorreu, para se calcular a $P(A|B)$, utiliza-se:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

3.4.5 Variáveis aleatórias

Sempre que o resultado de um experimento probabilístico é uma contagem ou medida, ele é chamado de variável aleatória, que podem ser discretas ou contínuas. Variáveis aleatórias sempre devem ser denotadas com letra maiúscula e os valores assumidos pelas variáveis serão denotados pelas letras minúsculas correspondentes.

Uma variável aleatória discreta assume valores que podem ser contados, ou seja, possui um número finito de resultados possíveis que podem ser enumerados. Seus valores são associados aos números naturais (0, 1, 2, 3...) (AGARWAL, 2006).

Já uma variável aleatória contínua assume valores que podem ser medidos, sendo sua principal característica poder assumir infinitos valores. Por exemplo, a distância do ponto A ao ponto B pode ser infinitamente subdividida, tornando a probabilidade de um valor pontual igual a 0.

3.4.6 Distribuição discreta de probabilidade

Uma distribuição de probabilidades pode ser entendida como o conjunto de todos os valores que uma variável aleatória pode assumir, com suas respectivas probabilidades. A probabilidade de que uma variável aleatória X assumira um valor x é descrito por meio de um modelo matemático chamado distribuição de probabilidades de x , representado por $P(X=x)$ ou, simplesmente, $P(x)$ (AGARWAL, 2006).

No caso das variáveis discretas, uma distribuição de probabilidades deve satisfazer às seguintes condições:

1. A probabilidade de cada valor da variável discreta deve estar no intervalo de 0 a 1;
2. A soma de todas as probabilidades deve ser 1.

Ao se tratar variáveis aleatórias, têm-se como objetivo modelar matematicamente a realidade considerando os componentes aleatórios intrínsecos ao fenômeno analisado. Dessa forma, foram criados vários modelos que se adaptam bem a uma série variáveis aleatórias, sendo capazes de resolver problemas práticos e que têm regularidade em seu comportamento básico. Dentro das variáveis aleatórias discretas, o modelo mais famoso e mais utilizado é o binomial.

3.4.6.1 Modelo Binomial

É um dos mais populares modelos discretos de distribuição de probabilidades. A origem dessa distribuição se dá nas “tentativas de Bernoulli”. Uma tentativa de Bernoulli é um experimento com somente dois resultados possíveis, que são o sucesso e a falha. Por exemplo, ao se jogar uma moeda só se pode obter cara ou coroa; Um item produzido pode ser defeituoso ou não; uma pessoa pode estar doente ou saudável.

Em um experimento aleatório consistindo em n tentativas independentes, a probabilidade de sucesso em cada uma das tentativas n é sempre igual a p e a de fracasso igual a q , sendo $p + q = 1$. Para cada tentativa, as probabilidades de sucesso e fracasso são sempre as mesmas (AGARWAL, 2006).

Sendo X o número de sucesso em n tentativas, então x pode assumir valores $x=0,1,2, \dots, n$. Dadas essas condições, X tem distribuição binomial com parâmetro n e p , sendo denotada $X \sim \text{Bi}(n;p)$. Para calcular a probabilidade de um evento x , faz-se:

$$P(X = x) = \binom{n}{x} * p^x * (1 - p)^{n-x}$$

onde x assume valores $0, 1, 2, \dots, n$. O coeficiente binomial pode ser calculado por

$$\binom{n}{x} = \frac{n!}{x! (n - x)!}$$

3.4.7 Distribuição contínua de probabilidade

Da mesma forma que as distribuição discretas de probabilidade, as contínuas visam descrever e modelar variáveis aleatórias, porém agora o objeto de estudo são as variáveis aleatórias contínuas. Devido à sua natureza contínua, a probabilidade relacionada a valores pontuais sempre será 0. Dessa forma, ao analisar esse tipo de variável, sempre será necessário trabalho com intervalos (AGARWAL, 2006).

Dentre os diversos modelos utilizados para descrever variáveis contínuas, o mais conhecido e amplamente utilizado é o normal, também conhecido como gaussiano.

3.4.8 Distribuição Normal ou Gaussiana

A distribuição normal conhecida também como distribuição gaussiana é sem dúvida a mais importante distribuição contínua. Ela permite modelar uma infinidade de fenômenos naturais e, além disso, possibilita realizar aproximações para calcular probabilidades de muitas variáveis aleatórias que têm outras distribuições, através do teorema central do limite.

O requisito mínimo para adotarmos o modelo normal para um conjunto de dados contínuos é que seu histograma seja aproximadamente simétrico em torno do ponto central, que também deve ser o ponto de máximo (AGARWAL, 2006).

A variável aleatória X , que toma todos os valores reais $-\infty < x < +\infty$, possui uma distribuição normal (ou gaussiana) se sua função densidade de probabilidade for da forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty$$

A notação $X \sim N(\mu, \sigma^2)$ representa que a variável aleatória X possui uma distribuição normal com parâmetros μ (média) e σ^2 (variância). Essa função pode ser plotada de forma gráfica, como mostra a figura 11.

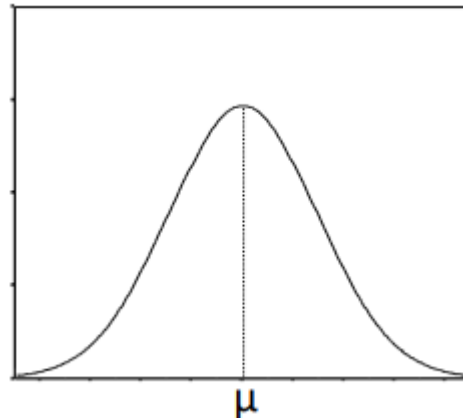


Figura 19: Função densidade de probabilidade da distribuição normal. Fonte: Elaborado pelo autor.

As principais propriedades dessa distribuição são:

1. A área sob a curva corresponde a 1;
2. O ponto de máximo da curva ocorre quando x é igual à média, o que significa uma maior probabilidade de ocorrência quando os valores avaliados estão mais próximos da média;
3. A distribuição é simétrica em torno da média. Dessa forma, as medidas de tendência central apresentam o mesmo valor;
4. A distribuição normal é definida pelo seu desvio padrão e média. Para diferentes combinações de médias e desvios padrão existem diferentes distribuições normais. As figuras 12 e 13 mostram variações na média e no desvio padrão, respectivamente.

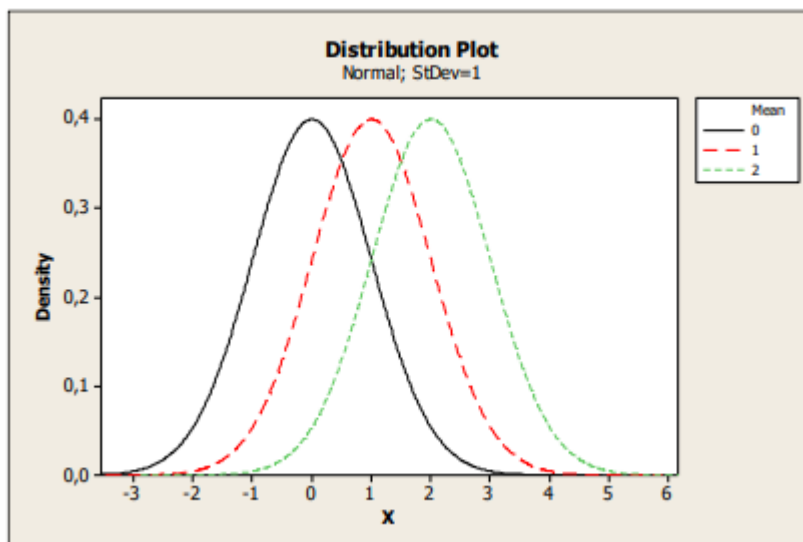


Figura 20: Diferentes distribuições normais, demonstrando uma variação na média.

Fonte: Elaborado pelo autor

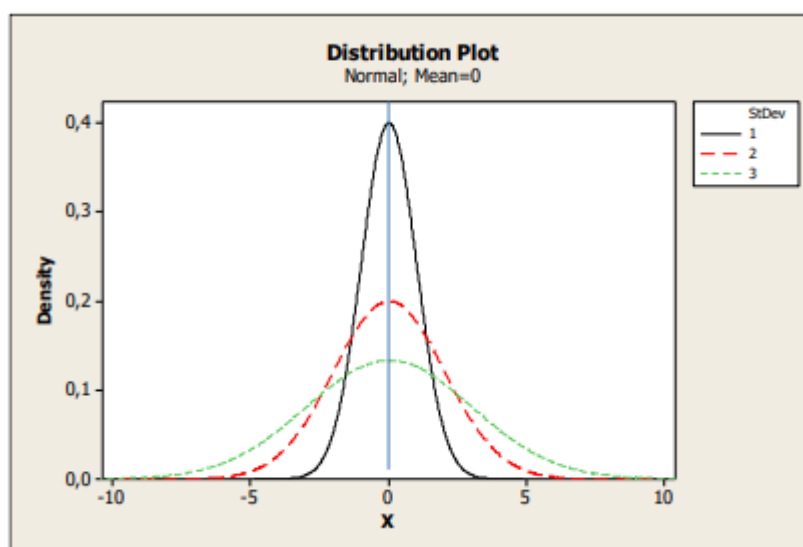


Figura 21: Diferentes distribuições normais, demonstrando uma variação no desvio padrão. Fonte: Elaborado pelo autor

Para se obter a probabilidade da variável X assumir um valor x entre os valores a e b , pode-se integrar a função densidade de probabilidade apresentada, dentro do intervalo a e b . Para se facilitar o cálculo das probabilidades de variáveis adequadamente descritas pela distribuição normal, é utilizada a distribuição normal padrão, que considera uma média igual a 0 e uma variância igual a 1. Quando se tem uma variável X com média e

desvio padrão diferentes da normal padrão, torna-se necessário reduzi-la efetuando o seguinte cálculo:

$$z = \frac{x - \mu}{\sigma}$$

onde z é o valor da variável normal padronizada, x é o valor da variável normal, μ é a média da variável aleatória e σ o desvio padrão da variável aleatória.

O grande poder e fama da distribuição gaussiana vem do fato dela ser capaz de modelar também diversas outras variáveis aleatórias com distribuições diferentes da normal, comportamento que é descrito pelo teorema central do limite.

3.4.9 Teorema central do limite

O teorema central do limite afirma que, para amostras aleatórias simples retiradas de uma população com média μ e variância σ^2 finita, a distribuição amostral da média se aproxima de uma distribuição normal com média μ e variância $\sigma^2/2$, à medida que o número de amostras n se aproxima de infinito.

Uma razão para a distribuição normal ser considerada tão importante é porque qualquer que seja a distribuição da variável de interesse, para grandes amostras a distribuição das médias amostrais serão aproximadamente normalmente distribuídas, e tenderão a uma distribuição normal à medida que o tamanho de amostra crescer. Assim, é possível existir uma variável original com uma distribuição muito diferente da Normal (pode até mesmo ser discreta), mas se tomarmos várias amostras grandes desta distribuição e então fizermos um histograma das médias amostrais, a forma se parecerá como uma curva normal. Isso torna possível modelar e fazer inferências de quaisquer distribuições utilizando a normal, que é um modelo extensamente estudado (AGARWAL, 2006).

3.5 História da krigagem

Uma das incumbências dos engenheiros de minas é selecionar os blocos a serem extraídos e até mesmo delimitá-los, se o método de extração usado lhe permitir. Em resumo, um bloco só deve ser explorado se o custo de sua extração e processamento não

exceder o valor do metal que pode ser extraído dele. Para determinados parâmetros técnico-econômicos, isso significa que o teor do bloco tem que ser maior que algum teor de corte. Na prática, o verdadeiro teor de um bloco não é conhecido antes de sua extração, de modo que a escolha dele é feita com base em estimativas. No início da década de 1950, a estimativa era simplesmente o teor médio dos dados pertencentes ao painel ou situados em sua fronteira. Krige (KRIGE, 1952), estudando dados de extração de vários corpos de minério, observou que para teores de corte elevados os blocos selecionados dessa forma eram, em média, menos ricos do que o esperado.

Como mostra a figura 3, isso não é nada surpreendente. Duas galerias paralelas em depósito sub-horizontal apresentam segmentos AB e CD com teores acima do corte, diferente das partes vizinhas das galerias. Portanto, a decisão é de explorar o trapezoide ABDC, e seu teor deverá ser igual à média ponderada dos segmentos AB e CD. Entretanto, os segmentos AC e BD não representam a verdadeira fronteira entre minérios ricos e pobres. Os verdadeiros limites (desconhecidos) se parecem mais com as linhas pontilhadas. Portanto, o minério pobre é explorado (e o minério rico abandonado), de modo que o teor do minério explorado é menor do que o esperado.

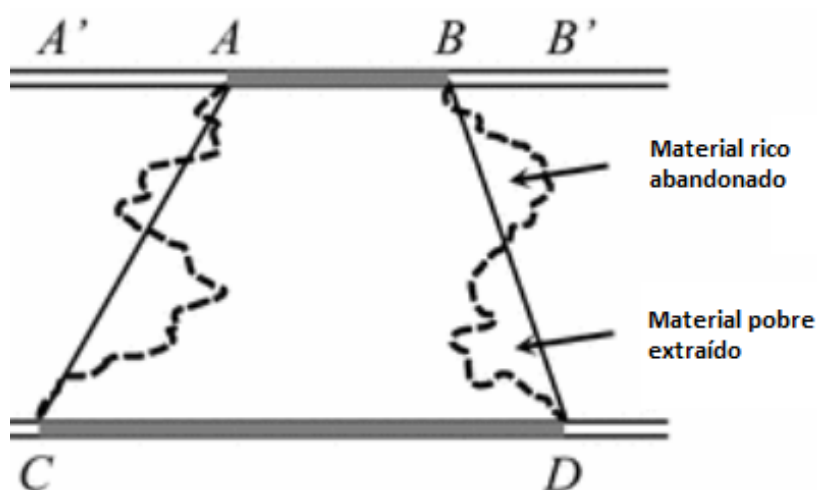


Figura 22: Ilustração do viés na estimativa. O painel ABDC a ser explorado foi delimitado a partir das amostras ricas observadas ao longo de AB e CD. Como a verdadeira fronteira entre os minérios ricos e pobres segue uma linha semelhante à linha pontilhada em vez dos segmentos AC e BD, parte do minério pobre será explorado e parte do minério rico abandonado. Fonte: CHILÈS, 2018.

Matematicamente, isso expressa um viés condicional: Sendo Z_v o teor do bloco e \bar{Z} o teor médio do bloco, a expectativa condicional $E[Z_v|\bar{Z}]$ não é igual a Z_v .

Para evitar esse viés, Krige confere um peso λ ao teor médio dos dados situados no bloco e o peso complementar $1 - \lambda$ ao grau médio do corpo de minério, sendo λ determinado por regressão linear.

Também enfrentando problemas de estimativa, Matheron estudou o trabalho de Krige e generalizou sua abordagem atribuindo um peso adequado a cada amostra, sendo esses pesos determinados de modo a minimizar a variância da estimativa sob a condição de que os pesos totalizem 1 (esta condição expressa que o estimador é uma média ponderada dos dados).

Matheron chamou este método de "kriging" em homenagem a Danie Krige. Para ser preciso, de acordo com Cressie (1990), o termo francês "krigeage" foi cunhado por Pierre Carlier e usado pela primeira vez no "Commissariat à l'énergie atomique" no final da década de 1950, e Matheron o traduziu por "kriging" (MATHERON, 1963).

3.5.1 Krigagem ordinária

A teoria da krigagem como é geralmente apresentada aparece a primeira vez em Matheron, 1962. Ela ocorre num modelo de função aleatória estacionária de ordem 2 (SRF). A variável regionalizada de interesse (aqui um teor) é considerada como uma realização de um SRF $Z(x)$, onde x denota um ponto em um espaço bi ou tridimensional. N pontos de dados estão disponíveis, nos locais x_α , $\alpha = 1, 2, \dots, N$, com valores $Z_\alpha = Z(x_\alpha)$. O alvo, Z_0 , é o valor $Z(x_0)$ de Z em um ponto x_0 não observado, ou de forma geral o valor médio $Z(v)$ de Z em uma determinada célula ou bloco v . O estimador de krigagem de Z_0 é, por definição, da forma

$$Z^* = \sum_{\alpha=1}^N \lambda_\alpha Z_\alpha$$

com pesos λ_α somando 1. Os pesos são escolhidos de modo a minimizar a variância do erro de estimativa $Z^* - Z_0$, sujeito à condição da soma. Isso leva a um sistema linear de equações $N + 1$ com $N + 1$ variáveis (os N pesos λ_α e um parâmetro de Lagrange μ):

$$\begin{cases} \sum_{\beta} \lambda_{\beta} \sigma_{\alpha\beta} + \mu = \sigma_{\alpha 0} \\ \sum_{\beta} \lambda_{\beta} = 1 \end{cases} \quad \alpha = 1, \dots, N$$

onde $\sigma_{\alpha\beta}$ denota a covariância das observações Z_{α} e Z_{β} e $\sigma_{\alpha 0}$ a covariância de Z_{α} e do alvo Z_0 . Assim forma-se o sistema da krigagem ordinária.

3.5.2 Krigagem simples

É possível perceber que o sistema de krigagem apresentado não requer conhecimento da média. Caso a média m seja conhecida, é possível usar um estimador da forma:

$$Z^* = \sum_{\alpha} \lambda_{\alpha} Z_{\alpha} + \left(1 - \sum_{\alpha} \lambda_{\alpha}\right) m$$

sem restrição sobre os pesos, e a minimização da variância da estimativa leva ao sistema de krigagem simples

$$\sum_{\beta} \lambda_{\beta} \sigma_{\alpha\beta} = \sigma_{\alpha 0} \quad \alpha = 1, \dots, N$$

A krigagem simples possui aplicações limitadas. É, no entanto, importante, porque tem propriedades que não são compartilhadas pela krigagem comum e nem pela universal (CHILÈS, 2009). De um ponto de vista computacional, sendo a matriz de krigagem definida positiva, o sistema pode ser resolvido pelo método de Cholesky.

3.6 Interpolação espaço-temporal

Atualmente, existem três classes principais de interpolação dos dados de monitoramento (XIAO et al., 2020). A primeira classe considera somente a relevância no domínio espacial, como o método inverso do quadrado da distância (LU, 2008) e o método de interpolação da krigagem ordinária (CRESSIE, 1988). Esses métodos utilizam a correlação espacial entre as medições para inferir os valores entre eles. Apesar desses métodos considerarem a relação entre os pontos vizinhos, eles não são capazes de utilizar

a variação no domínio temporal. A segunda classe considera apenas o domínio do tempo, como o método de interpolação Lagrange (HE, 2020), o método de interpolação Hermite (DE BOOR, 1987) e o método de interpolação spline (FREEDEN, 1987). Esses métodos fazem uso da correlação entre dados dos pontos danificados ou perdidos, coletados em tempos diferentes. Por considerarem somente um dos atributos dos dados, essas duas classes de interpoladores podem afetar a qualidade da estimativa. Já a terceira, combina o espaço e o tempo, gerando uma interpolação espaço-temporal.

Graças aos avanços tecnológicos das últimas décadas, a coleta e processamento de dados espaço-temporais tem aumentado rapidamente. Diversos parâmetros que variam no espaço e no tempo (como variáveis meteorológicas, de qualidade do ar, produção agrícola ou, até mesmo, deslocamento em estruturas monitoradas) se beneficiam desses avanços. As análises de correlações espaciais e temporais são muito úteis para se obter informações sobre as causas da variabilidade, sendo também fundamentais para a predição de valores em observações vizinhas (especialmente e/ou temporalmente). Interpolações espaço-temporais são potencialmente capazes de gerar melhores previsões do que as interpolações espaciais tradicionais, pois levam em conta uma série temporal de observações espaciais (GRÄLER; PEBESMA; HEUVELINK, 2016).

Apesar das vantagens em precisão, adicionar o domínio temporal à análise representa um aumento considerável em sua complexidade, pois implica a necessidade de se modelar a variabilidade no espaço e no tempo simultaneamente, o que é mais complicado do que modelar somente espacialmente ou somente temporalmente.

Dados de deslocamento de estruturas geotécnicas, obtidos através de estações totais robóticas, são amostrados pontualmente em diversos locais pela estrutura, em intervalos aproximadamente regulares de tempo. Assim, entender o comportamento das regiões vizinhas, onde não há coleta direta de dados, beneficia a análise geotécnica. De acordo com resultados empíricos e de pesquisa, as variações observadas nos dados de monitoramento de deformações estão relacionadas tanto aos atributos do tempo quanto aos atributos espaciais, ou seja, possuem forte correlação espaço-temporal (BOGAERT, 1996; RAHIMI, 2019). Liu et al. adotaram um algoritmo de interpolação espaço-temporal não uniforme para analisar o deslocamento em deslizamentos de terra (LIU et al., 2019).

Lenda et al. utilizaram quatro algoritmos de interpolação diferentes para construir modelos superficiais e analisar o movimento progressivo dos deslizamentos (LENDA et al., 2016). Wang et al. usaram a krigagem espaço-temporal com um processo gaussiano para interpolar os dados de monitoramento para análise de deformação em encostas (WANG et al., 2014). Xiao et al. desenvolveram um algoritmo melhorado de krigagem espaço-temporal baseado em um algoritmo genético adaptativo, obtendo uma raiz do erro quadrático médio (RMSE) melhor que os outros métodos presentes na literatura até o momento (XIAO et al., 2020).

Em 2016, Benedikt Gräler, Edzer Pebesma e Gerard Heuvelink elaboraram funções espaço-temporais para modelagem geoestatística no pacote “gstat”, da linguagem R. Como o manuseio de dados espaço-temporais em R é dado pelo pacote “spacetime”, suas classes foram utilizadas junto a funções originais do “gstat” para estimar a covariância espaço-temporal/modelos de variograma e para gerar a interpolação espaço-temporal (GRÄLER; PEBESMA; HEUVELINK, 2016).

3.7 Modelagem da covariância e variogramas

Assumindo-se um campo espaço-temporal aleatório Gaussiano Z , estacionário e espacialmente isotrópico, é possível caracterizá-lo através de uma média μ e uma função covariância C_{st} , onde a covariância espaço-temporal depende das distâncias no espaço $h \in \mathbb{R} \geq 0$ e tempo $u \in \mathbb{R} \geq 0$ (GRÄLER; PEBESMA; HEUVELINK, 2016).

Para estimar a covariância e seu variograma foram implementados diversos modelos no gstat, que são apresentados a seguir.

3.7.1 Modelo separável

O modelo de covariância separável assume que a função de covariância espaço-temporal pode ser representada como produto de um termo espacial e temporal:

$$C_{sep}(h, u) = C_s(h) C_t(u)$$

Seu variograma é dado por:

$$\gamma_{sep}(h, u) = sill * (\bar{\gamma}_s(h) + \bar{\gamma}_t(u) - \bar{\gamma}_s(h)\bar{\gamma}_t(u)),$$

onde $\bar{\gamma}_s$ e $\bar{\gamma}_t$ são variogramas espaciais e temporais padronizados com efeitos de pepita separados e sill de 1.

3.7.2 Modelo da soma do produto

O modelo de covariância da soma do produto possui uma notação muito similar à do modelo anterior, sendo a seguinte:

$$C_{ps}(h, u) = kC_s(h)C_t(u) + C_s(h) + C_t(u),$$

considerando $k > 0$.

O variograma correspondente é dado por:

$$\gamma_{sep}(h, u) = (k * sill_t + 1)\gamma_s(h) + (k * sill_s + 1)\gamma_t(u) - \gamma_s(h)\gamma_t(u),$$

Onde γ_s e γ_t são variogramas espaciais e temporais. O parâmetro k precisa ser positivo. A seguinte identidade define o sill geral do modelo em termos dos sill temporal e espacial.

$$sill_{st} = k * sill_s * sill_t + sill_s + sill_t$$

O k também pode ser estimado a partir da equação anterior e dos três valores de sill, caso conhecidos.

3.7.3 Modelo métrico

Assumindo funções de covariância idênticas nos domínios do tempo e espaço, exceto para anisotropia espaço-temporal, é possível utilizar o modelo métrico de covariância, onde, após regularizar o espaço e o tempo por uma correção de anisotropia k , as distancias espacial, temporal e espaço-temporal são tratadas igualmente, resultando em um modelo único de covariância, denotado por C_{joint} :

$$C_m(h, u) = C_{joint}(\sqrt{h^2 + (k * u)^2})$$

Sendo o variograma definido por:

$$\gamma_m(h, u) = \gamma_{joint}(\sqrt{h^2 + (k * u)^2})$$

3.7.4 Modelo métrico aditivo

Uma combinação dos modelos temporal, espacial e métrico, incluindo um parâmetro de anisotropia k é o modelo de covariância métrico aditivo:

$$C_{sm}(h, u) = C_s(h) + C_t(u) + C_{joint}(\sqrt{h^2 + (k * u)^2}),$$

Para ele, o variograma é definido por:

$$\gamma_{sm}(h, u) = \gamma_s(h) + \gamma_t(u) + \gamma_{joint}(\sqrt{h^2 + (k * u)^2})$$

onde γ_s , γ_t e γ_{joint} são variogramas espaciais, temporais e conjuntos com efeitos de pepita separados.

3.7.5 Modelo métrico aditivo simplificado

Uma simplificação do modelo anterior é restringir os variogramas espacial, temporal e conjuntos em modelos sem efeito de pepita. Após isso, um único efeito de pepita espaço-temporal é introduzido e o variograma obtém a seguinte representação:

$$\gamma_{ssm}(h, u) = nugget + \gamma_s(h) + \gamma_t(u) + \gamma_{joint}(\sqrt{h^2 + (k * u)^2})$$

3.8 Krigagem espaço-temporal

A krigagem espaço-temporal representa um conjunto complexo de cálculos, que pode se tornar computacionalmente cara quando executada em um conjunto de dados completo. Devido a isso, krigagem local é uma alternativa atrativa, capaz também de relaxar um pouco a suposição de estacionaridade, pois se torna possível respeitar variações da média pelo domínio (GRÄLER; PEBESMA; HEUVELINK, 2016). Sem selecionar uma vizinhança, o algoritmo da krigagem implementado em R usa todos os dados. Para selecionar os vizinhos a serem utilizados na análise, utiliza-se a função “k-nearest neighbours”, do pacote FNN (BEYGELZIMER et al., 2013). Dessa forma, a interpolação ocorre de forma iterativa para cada subconjunto local de dados.

4 Metodologia

Para o desenvolvimento desse projeto, será utilizada uma krigagem ordinária aplicada sequencialmente em todos os espaços no tempo, gerando uma sequência de frames espaciais.

Todo o processamento dos dados para a interpolação foi feito utilizando-se python. Foram utilizados os módulos: pyodbc (para conexão com o banco de dados SQL), pandas (para manipulação dos dados obtidos), numpy (para manipulação de matrizes), math (para utilização de funções e constantes matemáticas), sklearn (para pré-processamento dos dados), pykrige (para a krigagem ordinária), tqdm (para criação de barras de progresso para acompanhamento da execução do software), matplotlib (para gerar as imagens resultantes), cv2 e os (ambos para iterar por todas as imagens geradas e criar um vídeo). Para facilitar o gerenciamento, manutenção e versionamento das funções, o projeto foi dividido em cinco grandes grupos.

4.1 Grupos de execução

Cada etapa do projeto foi implementada como um módulo Python específico, com funcionalidades bem definidas. Os módulos incluem: "database.py", "pre_processing.py", "kriging.py", "render.py" e "main.py". De maneira geral, o "database.py" é responsável por fornecer os dados necessários para o restante do programa. O "pre_processing.py" realiza as operações de pré-processamento necessárias nos dados, a fim de prepará-los corretamente para a aplicação da krigagem. O "kriging.py" implementa a classe "Kriger", que recebe os dados, realiza os processamentos, calcula a krigagem e salva as imagens geradas. O "render.py" é responsável por criar o vídeo, utilizando as imagens geradas pela classe "Kriger". Já o "main.py" desempenha o papel de coordenar todas as operações de alto nível, chamando os outros módulos na ordem correta. A figura abaixo ilustra o fluxo dos dados, desde a entrada do usuário até a geração do vídeo. A seguir, cada módulo será detalhado, descrevendo suas funções específicas.

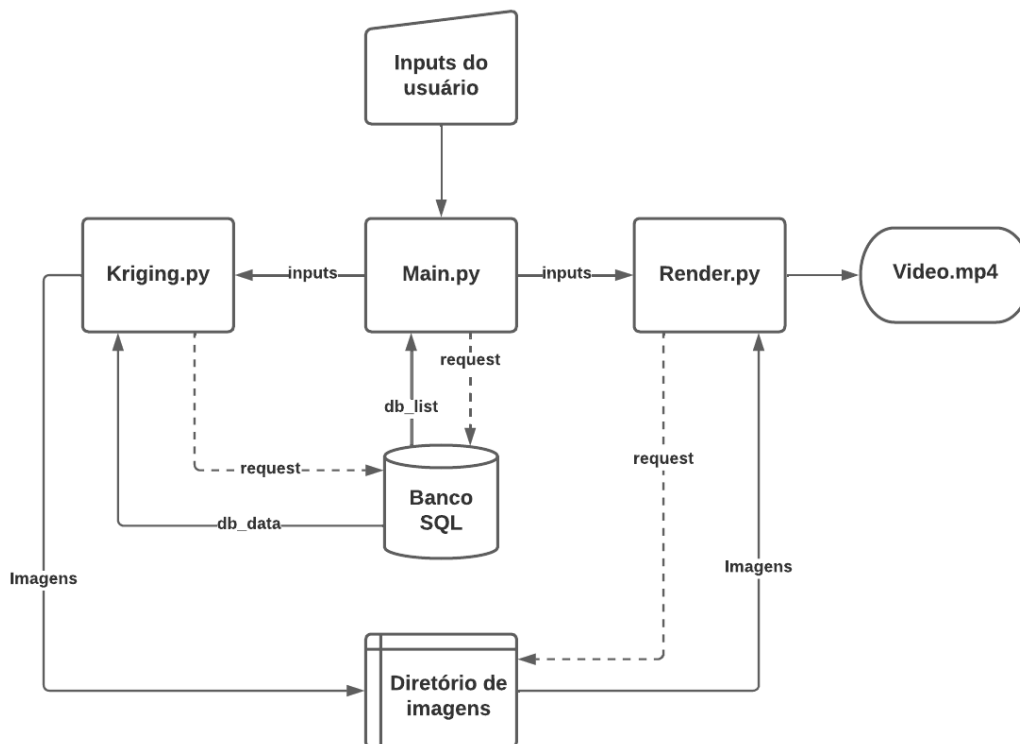


Figura 23: Representação esquemática de alto nível, mostrando a comunicação entre os diferentes módulos do software. Fonte: Elaborado pelo autor.

4.1.1 Database.py

O primeiro módulo, chamado aqui de “database.py”, é responsável por implementar um conector que faz a parte inicial de extração dos dados para análise pelos outros scripts. Para o presente trabalho foram utilizados dados disponibilizados através de um SQL Server e, portanto, o conector implementado é composto por três funções diferentes que realizam a comunicação com o banco e recuperação dos dados necessários.

A primeira função, chamada de “db_list()”, não recebe nenhum argumento e conecta no banco de dados chamado “master”, existente em servidores SQL, e recupera o nome de todos os bancos de dados de ETR disponíveis no servidor especificado (criado no localhost, para o presente trabalho, mas poderia ser qualquer servidor remoto). A segunda, chamada “database_conn()”, recebe o nome de um banco de dados como

argumento e, opcionalmente, o nome de uma das variáveis observadas pela ETR. Caso uma observação seja definida, a consulta retornará uma tabela somente com as colunas com o momento da leitura, o prisma e o valor da variável naquele momento. Caso não se defina nenhum nome de variável, todas as disponíveis serão retornadas. Em ambos, o objeto retornado será um dataframe do pandas, que é amplamente utilizado para análise de dados em python. A terceira função, nomeada como “points_null()”, recebe somente o nome de um banco de dados como argumento. Ela é responsável por calcular para todos os prismas a coordenada “null”, que é a coordenada utilizada como referência para cada ponto. Usualmente, ela é uma das primeiras coordenadas calculadas para cada prisma. Ela possui um papel extremamente importante no presente trabalho, pois será utilizada como ponto base de cada prisma.

O módulo database.py utiliza somente as bibliotecas pyodbc e pandas, não consumindo nenhum modulo customizado desenvolvido unicamente para esse trabalho. Ele é consumido por dois outros módulos que serão detalhados mais adiante: o kriging.py e o main.py

4.1.2 Pre_processing.py

O segundo módulo, chamado de pre_processing.py, é onde está a implementação de diversas funções que fazem transformações necessárias aos dados antes da krigagem ser de fato aplicada. Existem três funções que foram efetivamente utilizadas no resto do projeto, porém outras funções também foram definidas para uso posterior.

4.1.2.1 Funções utilizadas no projeto

A primeira, chamada de “resample_df()”, recebe um dataframe e um tempo de amostragem como argumentos e faz a amostragem dos dados recebidos. Ela é necessária pois cada projeto pode possuir seu próprio tempo de ciclo definido para a ETR. Mesmo que ele seja definido como “contínuo”, ainda haverá um intervalo de tempo entre medições consecutivas de um mesmo prisma, que é o tempo do equipamento ler todos os prismas cadastrados e retornar ao primeiro. Esse tempo médio entre leituras costuma ter baixa variância, porém, a fim de eliminá-la e padronizar o tempo entre leituras, é utilizada

essa técnica de reamostragem dos dados. Essa técnica também permite agrupar todos os prismas que foram lidos em um mesmo ciclo de leitura dentro de um índice de data/hora único, aqui chamados de imagens temporais do monitoramento da ETR, que serão utilizadas posteriormente pelo módulo de krigagem.

A reamostragem consiste em criar um eixo temporal onde cada entrada varia por um tempo fixo (por exemplo, de 20 em 20 minutos ou de 10 em 10 minutos). Caso exista mais de uma medição dentro do intervalo definido, a média dos valores é atribuída ao tempo inicial do intervalo. Caso não exista nenhuma medição no período, os dois valores mais próximos são interpolados linearmente para estimar o dado faltante. Feita a reamostragem, obtém-se um dataframe pandas com o valor da variável de interesse, separado por prisma e com um índice de data/hora homogêneo.

As figuras abaixo mostram uma dispersão plotada com o dado original e com o dado após reamostragem.

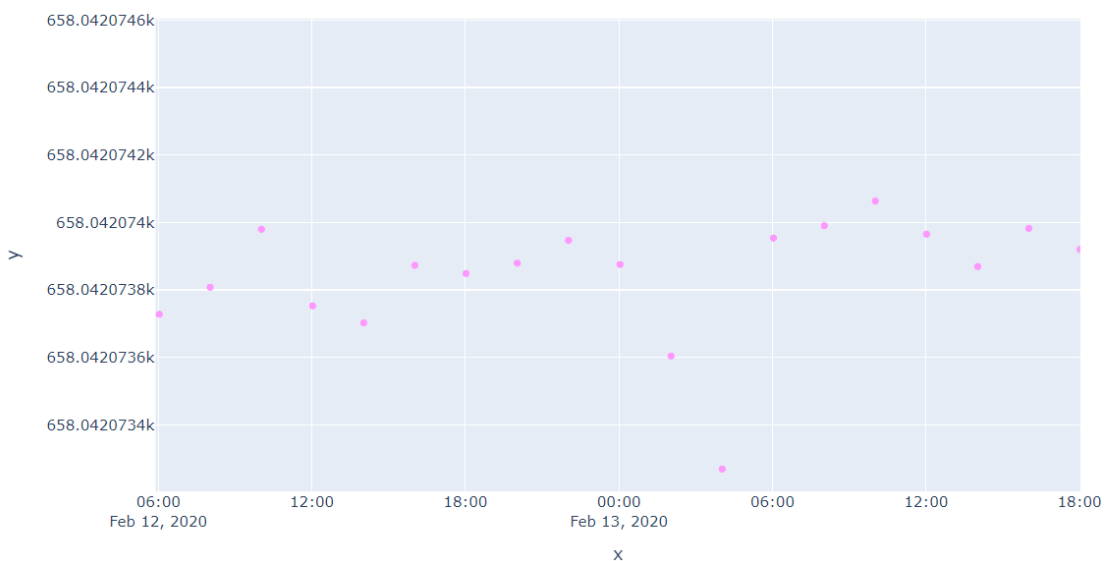


Figura 24: Dados originais com tempo de ciclo de 2 horas. Fonte: Elaborado pelo autor.

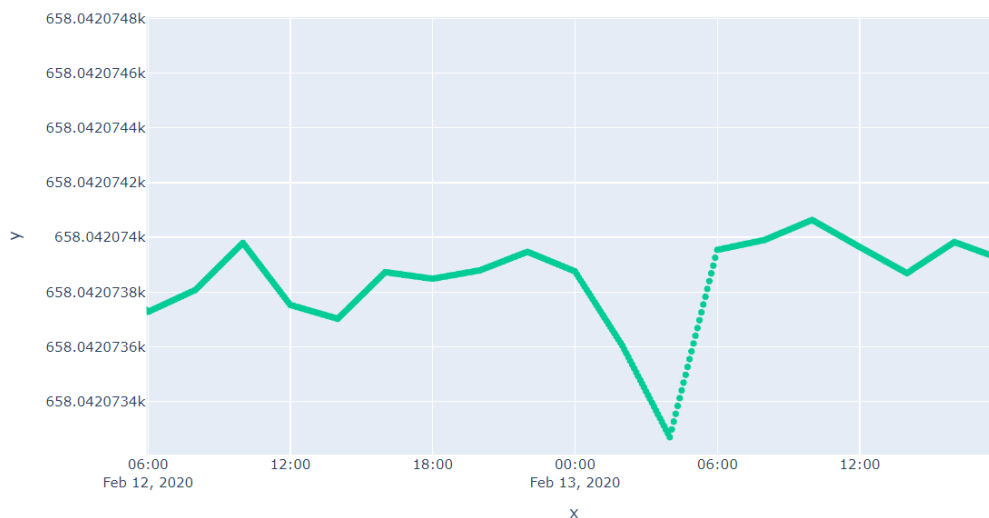


Figura 25: Mesma janela temporal anterior com dados reamostrados a cada 5 minutos.
Fonte: Elaborado pelo autor.

A segunda função, chamada de “`first_value()`”, é usada para obter o primeiro valor registrado para a observação requerida pelo usuário para cada ponto no conjunto de dados. Na prática, isso é usado para se obter uma medida de alteração ao longo do tempo, pois é criada a referência na primeira leitura de cada prisma. Posteriormente, dentro do módulo `kriging.py`, o resultado dessa função é mesclado com o conjunto de dados contendo as coordenadas “null” dos pontos. Este processo garante que cada ponto no conjunto de dados tenha não apenas suas coordenadas geográficas, mas também o primeiro valor registrado associado a ele. No contexto do projeto, a utilização da função `first_value` é crucial para preparar os dados para a krigagem, pois estabelece um ponto de referência que é utilizado no cálculo da variação do atributo ao longo do tempo.

A terceira função, “`remove_outliers()`” é uma parte crítica do pré-processamento de dados. Ela é usada para limpar os dados, removendo valores aberrantes que podem distorcer os resultados da análise. Os outliers são observações que se desviam tanto das demais observações que não podem ser considerados válidos, principalmente em um contexto de monitoramento onde se busca por movimentações lentas da estrutura.

Esta função implementa o método de detecção de outliers baseado no intervalo interquartil (IQR), que é uma medida de dispersão estatística, sendo igual à diferença entre os percentis 75 (Q3) e 25 (Q1). O IQR é uma medida de dispersão semelhante ao

desvio padrão, mas é mais robusta em relação aos outliers. A função primeiro calcula o Q1 (25º percentil) e o Q3 (75º percentil) do conjunto de dados. O IQR é então calculado como a diferença entre Q3 e Q1. Qualquer valor que esteja abaixo de $(Q1 - 1.5 * IQR)$ ou acima de $(Q3 + 1.5 * IQR)$ é considerado um outlier.

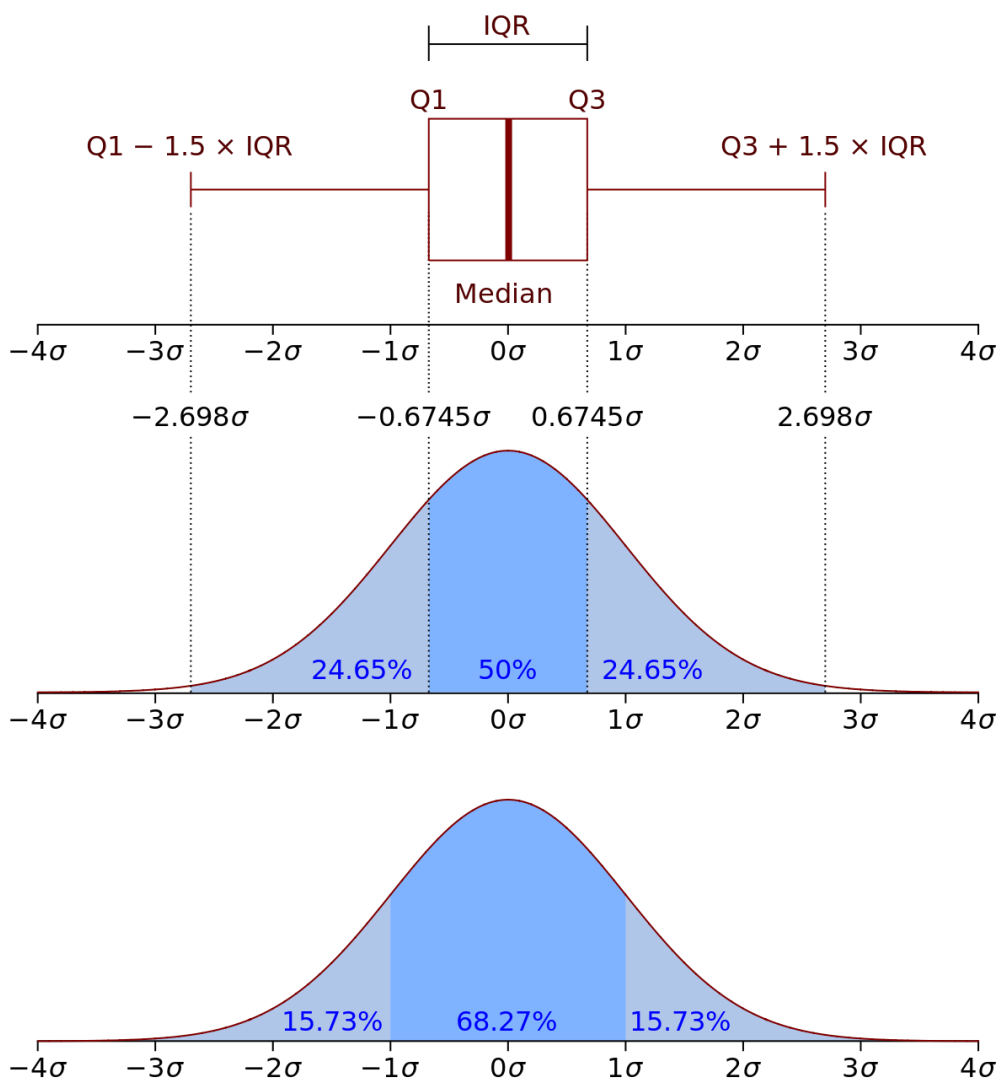


Figura 26: Boxplot com função densidade de probabilidade de uma distribuição normal. Os outliers estão à esquerda e à direita do boxplot representado. Fonte: Wikipedia, 2023.

Os índices dos valores classificados como outliers são identificados utilizando a função “np.where” e, em seguida, esses índices são usados para remover os outliers do DataFrame original com a função “drop”. Portanto, esta função remove os outliers do DataFrame com base no método IQR, que é uma técnica comum e robusta para lidar com outliers em análise de dados, sendo usada no projeto para garantir que os dados usados para a krigagem sejam representativos e não sejam distorcidos por valores anômalos.

4.1.2.2 Funções implementadas para uso posterior

Uma das funções implementadas para ser utilizada em uma continuação do projeto é a `azimute.py`. Ela recebe coordenadas X e Y num momento inicial e final e calcula o azimute do vetor formado por esse deslocamento. O objetivo futuro é interpolar também a direção dos vetores deslocamento encontrados para que, ao gerar o vídeo final com a krigagem, seja possível também enxergar o campo vetorial do movimento.

Por fim, são implementadas algumas funções de um pipeline que têm como objetivo a criação de uma modelagem de séries temporais usando o método SARIMA (Seasonal AutoRegressive Integrated Moving Average). O objetivo dessa modelagem é tentar prever como deve ser o movimento esperado da estrutura monitorada para um tempo futuro determinado. Com essa previsão sendo feita continuamente, é possível definir alarmes que avisem se algum prisma ou até mesmo a estrutura como um todo está se comportando diferentemente do esperado. As funções desse pipeline são detalhadas abaixo.

`Normalize_MinMax(df)`: Esta função normaliza um DataFrame pandas usando o método Min-Max. Esse processo redimensiona os valores em um intervalo de 0 a 1, o que pode ser útil para alguns algoritmos que esperam que os dados de entrada estejam nesse intervalo. O escalonador `MinMaxScaler` do Scikit-learn é usado para realizar a normalização.

`Standardize(df)`: Esta função padroniza um DataFrame pandas usando o `StandardScaler` do Scikit-learn, que transforma os dados para ter média zero e desvio padrão um. Isso pode ser útil para algoritmos que esperam dados com uma distribuição normal.

`Create_features(df)`: Esta função cria várias colunas em um DataFrame pandas com base no índice de data e hora. Esses recursos podem ser úteis para modelar efeitos sazonais e tendências temporais em uma série temporal.

`To_pickle(object,file_name)`: Esta função salva um objeto Python em um arquivo usando o módulo pickle. Isso permite que os dados ou modelos treinados sejam salvos e carregados posteriormente.

`Train_models(arima_df, obs, points, database, save_folder='arima_models', end_date='2022-01-01')`: Esta é a função principal do pipeline, que treina um modelo SARIMA para cada ponto em um conjunto de dados de séries temporais. A função usa a biblioteca pmdarima para ajustar os modelos SARIMA. Para cada ponto, o conjunto de dados é dividido em conjuntos de treinamento e teste, e o modelo é treinado no conjunto de treinamento. Os modelos treinados são salvos em um arquivo usando a função `to_pickle`.

4.1.3 Kriging.py

O módulo `Kriging.py` implementa uma classe central chamada "Kriger", responsável por realizar a krigagem nos dados já pré-processados, criando uma interpolação para cada ponto no tempo. A inicialização da classe "Kriger" exige parâmetros específicos, incluindo um banco de dados de ETR, a característica a ser interpolada na krigagem (com "Vector2D" sendo a opção padrão), o intervalo de tempo para a reamostragem dos dados e o índice de tempo inicial. Todas essas variáveis são inicializadas pelo construtor "`__init__`", utilizado por classes python ao inicializar a instância de uma classe.

A classe "Kriger" executa uma consulta a um banco de dados SQL correspondente à estrutura monitorada por meio do método "`connect`". Esse método utiliza funções do módulo "`database.py`", já detalhado anteriormente. O resultado desta consulta é uma tabela contendo todos os parâmetros da ETR disponíveis no banco de dados ou de uma observação específica, especificada pelo segundo parâmetro da classe "Kriger". Esta tabela é processada como um DataFrame pandas.

O método “add_coordinates_and_resample” é empregado para adicionar as coordenadas null e os valores iniciais aos prismas identificados no banco de dados e realizar a reamostragem de acordo com o intervalo de tempo definido. Ela utiliza as funções também já detalhadas anteriormente do módulo “pre_processing.py”.

O método “create_spatial_image” é utilizado para filtrar os dados por um índice de data/hora específico, criando uma "imagem temporal" contendo dados de deslocamento para todos os prismas no instante especificado. A krigagem será posteriormente executada individualmente para cada um dos instantes existentes no banco de dados.

Para preparar os dados para o processo de krigagem, o método “define_xyz_arrays” é usado para gerar matrizes com os dados de easting, northing e deslocamento. Ele simplesmente transforma as “colunas” do DataFrame em matrizes numpy, que é o formato solicitado pelo pykrige. Nesse momento também é possível definir os argumentos “x_filter” e “y_filter”, que podem ser utilizados para filtrar espacialmente os dados, caso uma determinada região necessite ser ignorada da análise.

O processo de krigagem em si é finalmente realizado pelo método “ordinary_krige”, que utiliza a função OrdinaryKriging do módulo pykrige. Este método utilizará os dados da imagem temporal criada anteriormente para realizar a interpolação com os dados contidos nela. Ele aceita diferentes modelos de variograma e permite a personalização dos parâmetros slope, nugget e sill para um ajuste ideal do modelo. É também possível definir a opção de plotar um variograma experimental, onde pode-se analisar o comportamento dos dados e o “fit” do modelo de variograma selecionado. A figura abaixo ilustra o resultado de um variograma experimental.

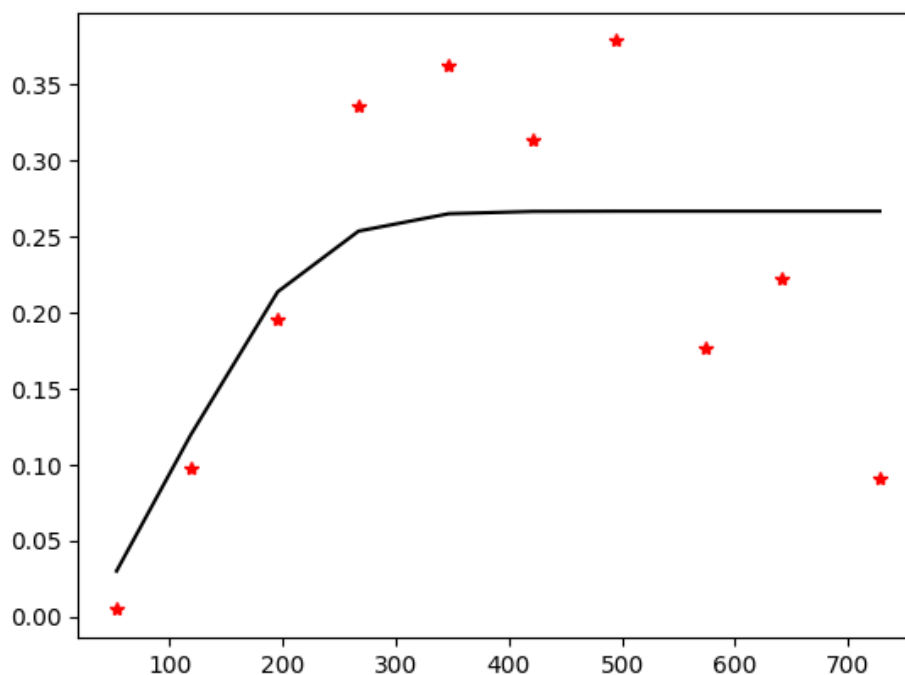


Figura 27: Exemplo de variograma experimental gerado pela função. Fonte: Elaborado pelo autor.

Os parâmetros experimentais encontrados também são mostrados como resultado. O log de execução é apresentado como mostra a figura abaixo.

```
Initializing variogram model...  
Coordinates type: 'euclidean'  
  
Using 'gaussian' Variogram Model  
Partial Sill: 0.2666036654794328  
Full Sill: 0.2666036654794328  
Range: 269.14884195098495  
Nugget: 9.666531794833644e-30
```

Figura 28: Log de execução com os parâmetros experimentais. Fonte: Elaborado pelo autor

Por último, a classe "Kriger" disponibiliza métodos para a visualização e o armazenamento dos resultados da interpolação. Estes resultados são fornecidos na forma de uma matriz, que contém os valores interpolados para cada ponto do grid estabelecido. Esta matriz é então visualizada como uma imagem espacial correspondente à imagem temporal usada para a interpolação.

O método "plot_result_image" é responsável pela exibição desta imagem interpolada ao usuário na forma de um mapa de calor. Este mapa apresenta os valores interpolados de deslocamento em todos os pontos definidos, oferecendo assim uma representação visual da interpolação realizada.

Já o método save_result_image é destinado ao armazenamento desta imagem. Através dele, a imagem é salva em um arquivo no diretório especificado, permitindo sua consulta posterior e utilização para geração do vídeo final.

A figura a seguir ilustra um exemplo de mapa gerado a partir de uma imagem temporal. Com a utilização da classe "Kriger", é possível gerar imagens semelhantes para todas as imagens temporais contidas em um banco de dados, proporcionando uma análise abrangente dos dados.

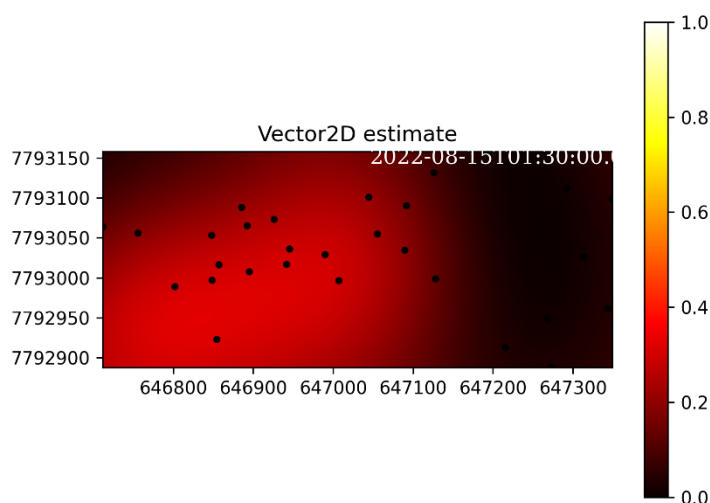


Figura 29: Mapa de deslocamento gerado a partir da krigagem dos valores de deslocamento dos prismas. Fonte: Elaborado pelo autor.

Ao operar em conjunto, esses métodos permitem a execução de um processo completo de krigagem, desde a consulta inicial aos dados até a geração e armazenamento da imagem resultante.

4.1.4 Render.py

O módulo `render.py` utiliza as bibliotecas “`cv2`” e “`os`” para gerar um vídeo com base em imagens no formato PNG. A função “`render_video`” recebe como argumentos o nome do arquivo de vídeo final, a pasta onde as imagens estão localizadas e a taxa de quadros por segundo (fps). Ela itera em todas as imagens e salva o vídeo final no diretório onde o código está sendo executado.

No projeto do presente trabalho, é gerada uma lista com todos os arquivos dentro da pasta definida para salvar as imagens geradas pelo módulo de krigagem, apresentado anteriormente. A função percorre as imagens presentes nesse diretório e gera um vídeo utilizando a taxa de quadros por segundo especificada. Para descobrir a escala de tempo do vídeo em relação à escala normal, basta levar em consideração também o tempo de reamostragem. Por exemplo, caso a reamostragem tenha sido definida em 30 minutos e o fps do vídeo em 10 (ou seja, 10 imagens por segundo), cada segundo do vídeo final mostrará o deslocamento ocorrido em 300 minutos (ou 5 horas) na estrutura.

4.1.5 Main.py

O módulo `main.py` atua como ponto de partida do programa, coordenando todo o processo de interpolação temporal e krigagem ordinária nos dados da estação total robótica e a subsequente geração de um vídeo visualizando os resultados.

O processo começa com a invocação da função `db_list()` do módulo `database.py`, gerando uma lista de todos os bancos de dados disponíveis. Uma vez que o usuário seleciona um banco de dados, o próximo passo é definir o atributo que será interpolado. Para todos os testes conduzidos neste estudo, o “`Vector2D`” foi empregado. Com a observação definida, o usuário tem a opção de excluir da análise prismas indesejados. Neste caso particular, como o objetivo era medir o deslocamento do maciço, prismas usados para medir o nível d’água dentro da cava foram excluídos.

Uma instância da classe `Krigger`, definida no módulo `kriging.py`, é então criada. Esta instância é inicializada com o banco de dados selecionado, o atributo a ser interpolado e o intervalo de reamostragem a ser aplicado na fase de pré-processamento. A conexão com o banco de dados e a reamostragem de coordenadas são executadas pelos métodos `connect` e `add_coordinates_and_resample`, respectivamente, preparando os dados para a análise subsequente.

Em seguida, um limite mínimo de prismas necessários para uma imagem temporal é estabelecido para a execução da krigagem. Empiricamente, decidiu-se que pelo menos 40% dos prismas válidos precisam estar presentes para que uma imagem temporal seja computada.

Com esse parâmetro estabelecido, o script itera por todos os instantes do banco de dados. Em cada iteração, a imagem temporal a ser interpolada é definida, e as matrizes `X`, `Y` e observação são criadas. Se houver dados suficientes na iteração atual (ou seja, o número de prismas válidos é maior que o limite definido), a krigagem é executada com os parâmetros especificados pelo usuário. O resultado é então salvo na forma de um mapa de calor, com uma escala ajustável pelo usuário.

Após a iteração completa do conjunto de dados, a função `render_video` é chamada para gerar um vídeo visualizando a evolução do deslocamento na estrutura ao longo do tempo.

5 Resultados

Foi desenvolvido um algoritmo em python capaz de executar a krigagem nos dados de uma ETR qualquer, bastando passar para a classe “`Krigger`” o banco de dados a ser estudado, o parâmetro avaliado e o tempo desejado para reamostragem. Feito isso, os dados são carregados e processados com os parâmetros da krigagem ordinária, também definidos pelo usuário.

Para este trabalho, foi processado um projeto de ETR em uma cava a céu aberto, que não pode ser identificada por motivos de privacidade dos dados. Nesse local são

frequentemente reportadas movimentações estruturais reais que, num histórico acumulado, chegam à magnitude de metros, deslocados sem a ocorrência de falha do maciço.

Os dados de aproximadamente seis meses foram processados com a utilização da krigagem, gerando uma matriz com os valores em cada pixel (escala dos pixels também definida no momento da execução do algoritmo da krigagem), a cada trinta minutos. Cada uma dessas matrizes são então renderizadas como uma imagem, utilizando-se uma escala de cores. A figura abaixo mostra as imagens obtidas para a evolução desse deslocamento, que atinge a magnitude de 1 metro na componente 2D do vetor deslocamento

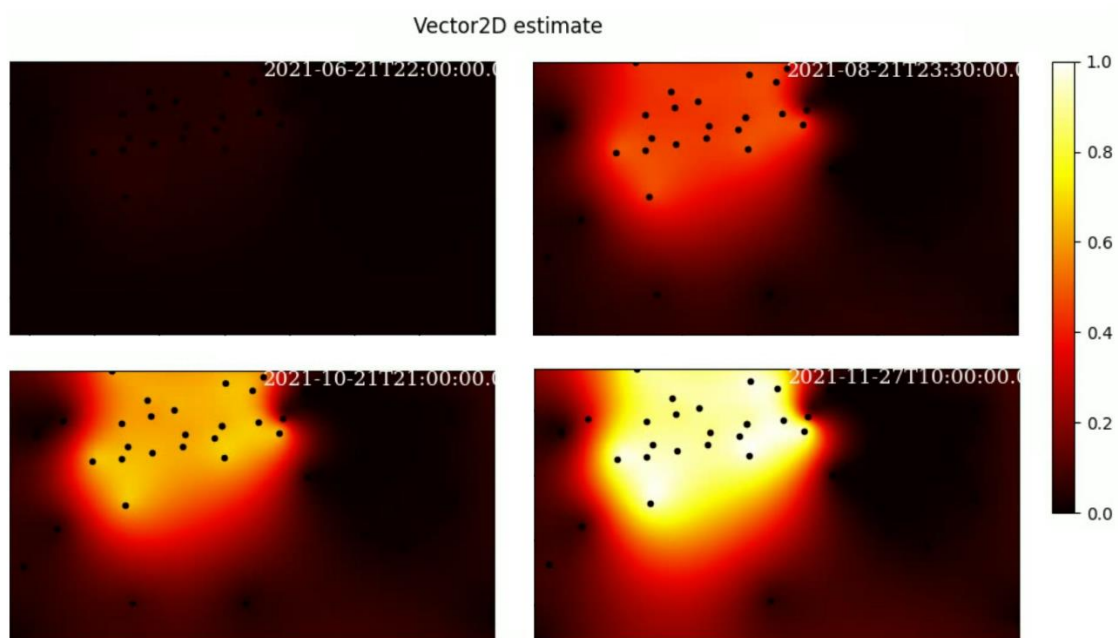


Figura 30: Evolução do deslocamento visualizado com a utilização da interpolação no espaço evolutivo. Fonte: Elaborado pelo autor

Todas as imagens geradas foram então renderizadas em um vídeo, que mostra a evolução do deslocamento dentro do período selecionado. A duração desse vídeo também pode ser especificada, variando-se a frequência de frames por segundo.

6 Conclusão

O presente trabalho realiza uma nova forma de visualização dos dados de estação total robótica. Series temporais, que são usualmente avaliadas através de gráficos de linha individuais por prisma, podem, com isso, serem estudadas levando em consideração a correlação espacial e temporal entre todos os elementos do sistema de monitoramento, facilitando a visualização de padrões que podem indicar a presença de deformações.

Tais imagens se assemelham muito às obtidas por radares interferométricos e laser scan, que apresentam seus dados através de mapas de deslocamento. O uso de tecnologias avançadas como radares SAR e ETRs tem permitido a medição de deslocamentos em estruturas geotécnicas com grande precisão. No entanto, apesar de ambos os métodos terem sido amplamente utilizados, eles medem diferentes tipos de deslocamentos. Enquanto as ETRs medem um vetor 3D de deslocamento, os radares SAR medem a projeção do deslocamento real na linha de visada do sensor. É importante comparar os dados obtidos pelos dois métodos para uma melhor compreensão do comportamento das estruturas geotécnicas. Devido a isso, uma continuação do presente trabalho seria comparar, em conjunto, os resultados obtidos através do método aqui descrito com os dados obtidos por radares e/ou laser scan, em um mesmo período de análise. Ao realizar as devidas correções aos dados (como a rotação do vetor 3D obtido pela ETR para a linha de visada do radar e a correção dos ruídos de cada equipamento), seria possível realizar uma comparação entre os diferentes conjuntos de dados, aumentando a acurácia das informações sobre a ocorrência de deslocamentos em uma determinada região. Isso proporcionaria uma análise mais completa e abrangente das deformações monitoradas, contribuindo para uma melhor compreensão dos fenômenos observados.

REFERÊNCIAS

- AGARWAL, Basant Lal. Basic statistics. New Age International, 2006.
- BEYGELZIMER, Alina et al. FNN: fast nearest neighbor search algorithms and applications. R package version, v. 1, n. 1, 2013.
- BOGAERT, Patrick. Comparison of kriging techniques in a space-time context. *Mathematical Geology*, v. 28, n. 1, p. 73-86, 1996.
- BUSSAB, Wilton de O.; MORETTIN, Pedro A. Estatística básica. In: Estatística básica. 2010. p. xvi, 540-xvi, 540.
- CHILÈS JP., Desassis N. (2018) Fifty Years of Kriging. In: Daya Sagar B., Cheng Q., Agterberg F. (eds) *Handbook of Mathematical Geosciences*. Springer, Cham. https://doi.org/10.1007/978-3-319-78999-6_29
- CHILÈS, Jean-Paul; DELFINER, Pierre. *Geostatistics: modeling spatial uncertainty*. John Wiley & Sons, 2009.
- CRESSIE, Noel. Spatial prediction and ordinary kriging. *Mathematical geology*, v. 20, n. 4, p. 405-421, 1988.
- CRESSIE, Noel. The origins of kriging. *Mathematical geology*, v. 22, n. 3, p. 239-252, 1990.
- DE BOOR, Carl; HÖLLIG, Klaus; SABIN, Malcolm. High accuracy geometric Hermite interpolation. *Computer Aided Geometric Design*, v. 4, n. 4, p. 269-278, 1987.
- DONATI, Davide et al. Spatio-temporal characterization of slope damage: insights from the Ten Mile Slide, British Columbia, Canada. *Landslides*, v. 17, n. 5, p. 1037-1049, 2020.
- ELDRANDALY, Khalid; ABDELMOUTY, Ahmed. Spatio-temporal interpolation: Current practices and future prospects. *International Journal of Digital Content Technology and its Applications*, v. 11, n. 06, p. 2017, 2017.

FREEDEN, Willi. A spline interpolation method for solving boundary value problems of potential theory from discretely given data. *Numerical Methods for Partial Differential Equations*, v. 3, n. 4, p. 375-398, 1987.

FUNÇÃO DE DENSIDADE DE PROBABILIDADE. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2023. Disponível em: <https://pt.wikipedia.org/w/index.php?title=Fun%C3%A7%C3%A3o_de_densidade_de_probabilidade&oldid=65831533>. Acesso em: 7 mai. 2023.

Goel, N. S. (1988). Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. *Remote sensing reviews*, 4(1), 1-212.

GOPI, Satheesh et al. *Advanced surveying: total station, GIS and remote sensing*. Pearson Education India, 2007.

GRÄLER, Benedikt; PEBESMA, Edzer J.; HEUVELINK, Gerard BM. Spatio-temporal interpolation using gstat. *R J.*, v. 8, n. 1, p. 204, 2016.

HE, Ren; ZHANG, Liwei. Interval recognition algorithm of the pavement surface condition based on Lagrange interpolation method. *Mathematical Problems in Engineering*, v. 2020, 2020.

HEFFERON, Jim. *Linear algebra third edition*. 2018.

KRIGE, D. G. A statistical analysis of some of the borehole values in the Orange Free State goldfield. *Journal of the Southern African Institute of Mining and Metallurgy*, v. 53, n. 3, p. 47-64, 1952.

LEICA GEOSYSTEMS. *Leica MS50/TS50/TM50 User Manual*. Heerbrugg: LEICA GEOSYSTEMS AG, 2013. 84 p.

LENDA, Grzegorz et al. The use of surface interpolation methods for landslides monitoring. *KSCE Journal of Civil Engineering*, v. 20, n. 1, p. 188-196, 2016.

LIU, Yong et al. A non-uniform spatiotemporal kriging interpolation algorithm for landslide displacement data. *Bulletin of Engineering Geology and the Environment*, v. 78, n. 6, p. 4153-4166, 2019.

LU, George Y.; WONG, David W. An adaptive inverse-distance weighting spatial interpolation technique. *Computers & geosciences*, v. 34, n. 9, p. 1044-1055, 2008.

MATHERON, Georges. Principles of geostatistics. *Economic geology*, v. 58, n. 8, p. 1246-1266, 1963.

MATHERON, Georges. *Traité de géostatistique appliquée*. Editions Technip, 1962.

Meneses, P. R., de Almeida, T., & de Mello Baptista, G. M. (2019). Reflectância dos materiais terrestres. *Oficina de textos*, pg 190.

MILONE, Giuseppe. *Estatística: geral e aplicada*. Pioneira Thomson Learning, 2004.

MONSERRAT, O.; CROSETTO, M.; LUZI, G. A review of ground-based SAR interferometry for deformation measurement. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 93, p. 40-48, 2014.

NADER, Alizeibek Saleimen. Monitoramento de taludes via radar SSR como indicador chave de desempenho geotécnico integrado às atividades primárias da cadeia de valor mineral. 2012. Tese de Doutorado. Universidade de São Paulo.

RAHIMI, Mehrzad et al. Bayesian calibration of multi-response systems via multivariate Kriging: Methodology and geological and geotechnical case studies. *Engineering Geology*, v. 260, p. 105248, 2019.

SHARON, Robert; EBERHARDT, Erik (Ed.). *Guidelines for Slope Performance Monitoring*. CSIRO PUBLISHING, 2020.

STURARO, JOSÉ RICARDO. *Apostila de geoestatística básica*. Rio Claro, UNESP, IGCE, 34p, 2015.

WANG, Jian-Min et al. Slope deformation analyses with space-time Kriging interpolation method. *Journal of China Coal Society*, v. 39, n. 5, p. 874-879, 2014.

XIAO, Haiping et al. An improved spatio-temporal kriging interpolation algorithm and its application in slope. *IEEE Access*, v. 8, p. 90718-90729, 2020.

YAMAMOTO, Jorge Kazuo; LANDIM, Paulo M. Barbosa. Geostatística: conceitos e aplicações. Oficina de textos, 2013.

ZHOU, Jianguo et al. Automatic subway tunnel displacement monitoring using robotic total station. Measurement, v. 151, p. 107251, 2020.