



Proposição de um Estimador Pontual para Incerteza Vertical de Levantamentos Hidrográficos

Proposition of a punctual estimator for vertical uncertainty of hydrographic surveys

*Italo Oliveira Ferreira*¹
*Paulo César Emiliano*²
*Afonso de Paula dos Santos*¹
*Nilcilene das Graças Medeiros*¹
*Júlio César de Oliveira*¹

Recebido em agosto de 2018.
Aprovado em dezembro de 2018.

RESUMO

Ter o conhecimento da morfologia submersa dos ambientes aquáticos sempre foi um desafio, justamente pelas dificuldades inerentes em caracterizar e explorar as ocorrências não aparentes dessas superfícies. A descrição das características dos oceanos, rios, lagos e outros corpos d'água possibilitam a obtenção de informações batimétricas úteis a diversas áreas, como para a navegação marítima ou fluvial, obras civis, prospecção de recursos minerais, etc. A partir das informações de profundidade pode-se ter o estabelecimento de práticas voltadas para o planejamento e execução de inúmeras atividades de hidrografia. Os levantamentos hidrográficos, empregados atualmente, são realizados por sistemas acústicos como ecobatímetros monofeixe, multifeixe e sonares interferométricos. Contudo, independente da tecnologia utilizada, os dados coletados sempre conterão incertezas, podendo estas serem causadas por efeitos de natureza grosseira, sistemática ou aleatória. Se os dados estiverem eivados de incertezas não aceitáveis a uma dada tolerância definida por norma, tais informações podem não ser adequadas para determinados propósitos. Um dos estimadores geralmente utilizados, o RMSE (*Root Mean Square Error*) é altamente influenciado pela presença de *outliers* nas amostras, podendo não ser adequado para descrever a qualidade estatística do conjunto de observações. Desta forma, o objetivo dessa pesquisa é a proposição de um estimador pontual visando quantificar a incerteza vertical de levantamentos batimétricos, denominado Incerteza Robusta, que

¹Universidade Federal de Viçosa – Departamento de Engenharia Civil. Programa de Pós-graduação em Engenharia Civil, área de concentração em Informações Espaciais, Brasil. E-mail: italo.ferreira@ufv.br; afonso.santos@ufv.br; nilcilene.medeiros@ufv.br; oliveirajc@ufv.br

²Universidade Federal de Viçosa – Departamento de Estatística. Programa de Pós-graduação em Estatística Aplicada e Biometria, Brasil. E-mail: paulo.emiliano@ufv.br

diferentemente dos estimadores utilizados na maioria das vezes, é resistente a *outliers* e independe da distribuição de probabilidades da amostra. Para o estabelecimento do intervalo de confiança deste estimador foi usada a técnica *Bootstrap*. Foram realizados experimentos com dados simulados, bem como o uso de dados reais, referentes a duas áreas de estudo. A partir dos resultados obtidos, pôde-se verificar o desempenho do estimador proposto, que nitidamente mostrou-se resistente aos possíveis *outliers* presentes no conjunto de dados. Pôde-se perceber também, que a presença dos *outliers* nas bases de dados teve pouca influência nas estimativas pontuais de incerteza, mostrando sua eficiência e robustez.

PALAVRAS-CHAVE: Incerteza Vertical, Levantamentos hidrográficos, Controle de qualidade, Batimetria, outliers.

ABSTRACT

Knowledge of the submerged morphology of aquatic environments has always been a challenge, precisely due to the inherent difficulties in characterizing and exploring the non-apparent occurrences of these surfaces. The description of the characteristics of the oceans, rivers, lakes and other bodies of water makes it possible to obtain bathymetric information useful in several areas, such as maritime or fluvial navigation, civil works, prospecting of mineral resources, supply or generation of energy, etc. From the depth information it's possible to establish practices directed to the planning and execution of numerous hydrographic activities. The hydrographic surveys, currently used, are performed by acoustic systems such as single beam echo sounders, multibeam echo sounders and interferometric sonars. However, regardless of the technology used, the data collected will always contain uncertainties, that is caused by gross, systematic or random effect. If the data are subject to uncertainties not acceptable to a given standard tolerance, such information may not be accurate for certain purposes. One of the commonly used estimators, RMSE (Root Mean Square Error) is highly influenced by the presence of outliers in the samples, and may not be adequate to describe the statistical quality of the set of observations. Thus, the objective of this research is the proposition of a point estimator to quantify the vertical uncertainty of bathymetric surveys, called Robust Uncertainty, which, unlike the estimators used most of the time, is resistant to outliers and does not depend on the probability distribution of the sample. To establish the confidence interval of this estimator, the Bootstrap technique was used. Experiments were performed with simulated data, as well as the use of real data, referring to two study areas. From the results obtained, it was possible to verify the performance of the proposed estimator, which was clearly resistant to the possible outliers present in the data set. It was also observed that the presence of the outliers in the databases had little influence on the point estimates of uncertainty, showing their efficiency and robustness.

KEYWORDS: Vertical Uncertainty, Hydrographic surveys, Quality control, Bathymetry, outliers.

* * *

Introdução

Hidrografia consiste no ramo das ciências aplicadas que lida com a descrição das características dos oceanos, rios, lagos e outras formas de água na terra. Neste sentido, os levantamentos hidrográficos tratam da medição de profundidades, marés, correntes oceânicas, gravidade, magnetismo terrestre e a determinação das propriedades químicas e físicas da água (IHO, 2005).

O termo levantamento batimétrico, comumente encontrado na literatura, por vezes é confundido com o termo levantamento hidrográfico. Deve-se atentar, no entanto, que o primeiro se refere apenas a coleta de profundidades georreferenciadas que, por outro lado, é a tarefa mais importante de um levantamento hidrográfico (FERREIRA et al., 2016a).

O sensoriamento remoto acústico é, na atualidade, o principal meio de investigação de fundos submersos, visto que os métodos tradicionais de sensoriamento remoto são pouco eficientes no meio aquático (AYRES NETO, 2000; FERREIRA et al., 2017a). Sendo assim, têm-se visto uma preferência pela utilização de sistemas acústicos para medição de profundidade, como ecobatímetros monofeixe, multifeixe e sonares interferométricos. Estes sistemas são, geralmente, utilizados em conjunto com sistemas diferenciais de posicionamento GNSS (*Global Navigation Satellite System*) em tempo real e sensores inerciais.

Métodos alternativos para mapeamento batimétrico de águas rasas, como sistemas de sondagem laser aerotransportados (GUENTHER et al., 1996; IHO, 2005; ATHEARN et al., 2010; PASTOL, 2011; ELLMER et al., 2014) e o emprego de imagens orbitais (GAO, 2009; CHENG et al., 2015; MOURA et al., 2016; FERREIRA et al., 2016b), também são utilizados, principalmente, em levantamentos para fins ambientais, de reconhecimento ou em áreas que ofereçam riscos ao navegante. De acordo com Ferreira et al. (2017a), em áreas não navegáveis ou de difícil acesso, o uso de embarcações não tripuladas para execução de levantamentos batimétricos cresceu demasiadamente nos últimos anos.

Dentre os diversos produtos gerados por um levantamento hidrográfico, destacam-se as cartas e publicações náuticas. Os navegantes possuem uma confiança inquestionável nestas produções, de tal forma que, na ausência da representação cartográfica de um perigo, eles acreditam fielmente na sua inexistência (IHO, 2005; MIGUENS, 2000). Sob esta perspectiva, é notória a necessidade de coleta de dados batimétricos confiáveis, proporcionando, assim, o desenvolvimento de uma cartografia náutica de qualidade.

Independente da tecnologia empregada numa sondagem batimétrica, os dados coletados conterão incertezas, sejam elas causadas por efeitos de natureza grosseira, sistemática ou aleatória (FERREIRA et al., 2016a). Inúmeras fontes de erros, tais como: medição de offsets, estimativas da velocidade da propagação do som na água e correção do draft, são observadas na aquisição da profundidade e posição planimétrica, afetando, conseqüentemente, os produtos finais. Deve-se atentar que não é possível gerar modelos batimétricos confiáveis se os dados estiverem eivados de incertezas de magnitude maior que uma determinada tolerância definida por norma (IHO, 2005, 2008).

Nesse sentido, avaliar a qualidade dos dados coletados tornou-se uma tarefa indispensável, isto é, a descrição georreferenciada deve sempre ser fornecida conjuntamente com um nível de confiabilidade estatisticamente comprovado. Sem uma estimativa da incerteza do produto final, até mesmo a destinação das informações fica comprometida.

No ramo da inferência estatística, é comum a estimativa de um parâmetro a partir de amostras coletadas com um certo grau de incerteza. Parâmetro pode ser definido como uma medida usada para descrever características de uma população, como por exemplo, a média, a variância e a amplitude (MOOD et al., 1974). Para estimar estes parâmetros são empregadas expressões matemáticas, ou seja, estimadores. Na obtenção destas expressões são utilizados diferentes métodos, como: método dos momentos, método da máxima verossimilhança e, mais comum nas ciências

geodésicas, o método dos mínimos quadrados (MOOD et al., 1974; MIKHAIL e ACKERMAN, 1976; COOPER, 1987; LEHMANN e CASELLA, 1998; MORETTIN e BUSSAB, 2004).

Estimativas podem ser pontuais ou intervalares. Como a estimação pontual não é capaz de fornecer o nível de confiabilidade do parâmetro, idealmente, deve-se fornecer a estimativa pontual conjuntamente com o seu intervalo de confiança.

O estimador RMSE (Root Mean Square Error) é utilizado em vários ramos da ciência como a Cartografia e Geodésia (GREENWALT e SCHULTZ, 1962; MIKHAIL E ACKERMANN, 1976; FGDC, 1998); Meteorologia (HALLAK e PEREIRA FILHO, 2011); Sensoriamento Remoto (KAFRAWY et al., 2017); Estatística e Geoestatística (VIEIRA, 2000; FERREIRA et al., 2013; CARMO et al., 2015; SANTOS, 2015), Modelagem computacional de Relevo (FERREIRA et al., 2017b), dentre inúmeras outras, como medida pontual de incerteza. Nas ciências hidrográficas, também se verifica o uso frequente do estimador RMSE como indicador da qualidade vertical das profundidades coletadas (SUSAN e WELLS, 2000; EEG, 2010; SEKELLICK e BANKS, 2010).

Segundo Santos (2015), o RMSE é um estimador robusto, pois pode ser utilizado para descrever um conjunto de dados independente da sua distribuição de probabilidade. No entanto, esse estimador é altamente influenciado pela presença de outliers na amostra (LI et al., 2005). De acordo com Armstrong e Collopy (1992), o RMSE é extremamente sensível a outliers. Os mesmos autores afirmam que este estimador é inadequado para descrever a qualidade estatística de conjuntos de observações, especialmente, séries temporais.

Nos últimos tempos diversos trabalhos na área de levantamentos hidrográficos foram desenvolvidos, especialmente, aqueles com foco na filtragem e detecção de spikes (*outliers*). Ware et al. (1991), apresentaram um processo de detecção de outliers baseado na análise das propriedades estatísticas da amostra. Seguindo linhas similares, Eeg (1995) propôs um

teste estatístico para validar o tamanho dos agrupamentos que são utilizados para detectar spikes. Debese e Bisquay (1999), Motao et al. (1999), Debese (2007) e Debese et al. (2012) aplicaram, basicamente, estimadores M. Calder e Mayer (2003) utilizaram o filtro de Kalman para processar dados batimétricos automaticamente. De forma análoga, Bottelier et al. (2005) empregaram técnicas de krigagem e Bjørke e Nilsen (2009) apresentaram uma técnica para detecção de spikes fundamentada na construção de superfícies de tendência. Já Lu et al. (2010) desenvolveram um algoritmo baseado no estimador robusto LTS (*Least Trimmed Squares*). Contudo, são raros os trabalhos que propõem metodologias para o controle de qualidade dos dados coletados. Portanto, é nítida a necessidade do desenvolvimento de metodologias para uma avaliação estatística adequada e consistente dos levantamentos hidrográficos, especialmente da componente vertical, isto é, da profundidade.

Diante do exposto, o objetivo desse artigo é propor um estimador pontual para quantificação da incerteza vertical de levantamentos batimétricos, denominado Incerteza Robusta (Φ_{Robs}). Este estimador é resistente a outliers e independe da distribuição de probabilidades da amostra. Para construção dos intervalos de confiança é utilizada a técnica Bootstrap.

1 Proposição do estimador

A avaliação da qualidade das profundidades coletadas num levantamento hidrográfico é realizada, idealmente, através da estimativa intervalar da acurácia vertical, assim como é feito, por exemplo, nas ciências geodésicas e cartográficas (MIKHAIL e ACKERMANN, 1976; IHO, 2005; HÖHLE e HÖHLE, 2009). Não raro, encontra-se autores, como Mezine et al. (2016), utilizando o termo precisão para descrever quantitativamente a qualidade vertical e, por vezes, posicional, das sondagens batimétricas.

De acordo com Mikhail e Ackermann (1976), a acurácia pode ser definida como sendo o grau de proximidade de uma estimativa com seu valor de referência, incorporando, assim, os efeitos aleatórios e sistemáticos. Já a precisão expressa o grau de consistência da grandeza medida com sua média, estando unicamente vinculada aos efeitos aleatórios. Ambas, precisão e acurácia, são estimadas como base em informações redundantes.

Num levantamento batimétrico, é pouco provável que a mesma feição submersa seja reamostrada, uma vez que a coleta de dados é realizada com a embarcação em movimento e pontos de controle, como marcos de concreto, não podem ser materializados. Diante desta dificuldade, é comum a sondagem de linhas de verificação planejadas de forma aproximadamente perpendicular às linhas regulares de sondagem. Empregando artifícios teóricos e práticos, determinam-se “pontos homólogos” e, conseqüentemente, as discrepâncias (d_i), entre as profundidades calculadas pelas respectivas linhas regulares e de verificação. Conforme exposto por Ferreira (2018), esta técnica é exaustivamente utilizada nos levantamentos com ecobatímetros monofeixe, entretanto, quando as sondagens são realizadas com sistemas de varrimento, maiores adaptações são requeridas.

Nesta filosofia, é teoricamente inconsistente empregar termos como: erro, exatidão, precisão, desvio padrão, repetibilidade e acurácia, em referência a qualidade vertical da batimetria (MONICO, 2009; FERREIRA et al, 2016a). Estes parâmetros são improváveis de serem quantificados nestes tipos de levantamentos, visto que informações redundantes estão, quase sempre, indisponíveis. Neste sentido, para avaliação da qualidade de uma sondagem batimétrica, é preferível o uso do termo incerteza vertical (VICENTE, 2011; FERREIRA, 2018).

Outrem é frequente na literatura o emprego do estimador RMSE como indicador da qualidade vertical das profundidades coletadas numa sondagem batimétrica, ou seja, a incerteza vertical (SUSAN e WELLS, 2000; EEG, 2010; SEKELLICK e BANKS, 2010). Este estimador, consiste na raiz

quadrada do MSE (Mean Square Error), medida de acurácia proposta por Gauss (MONICO, 2009). Nota-se, então, indícios de inconsistências teóricas.

Segundo Hallak e Pereira Filho (2011), a única diferença entre o MSE e o RMSE, reside no fato deste último apresentar valores do “erro” nas mesmas dimensões da variável analisada. Outra inconsistência observada na literatura, está presente na expressão matemática do RMSE, ora apresentada como o exposto na Equação 1, ora como exposto na Equação 2.

$$\text{RMSE}_1 = \sqrt{\frac{\sum_{i=1}^n (d_i)^2}{n}} \quad (1)$$

$$\text{RMSE}_2 = \sqrt{\frac{\sum_{i=1}^n (d_i)^2}{n-1}} \quad (2)$$

Em que d_i é a diferença entre o valor esperado e o valor observado e n corresponde ao tamanho amostral de d_i .

De acordo com Mood et al. (1974), a principal diferença entre as expressões (1) e (2) está em seu valor esperado. O estimador dado em (2) é não viesado, isto é, $E[\text{RMSE}_2] = \sigma^2$, enquanto que o primeiro estimador tem um valor esperado $E[\text{RMSE}_1] = \frac{(n-1)}{n} \sigma^2$, e, embora um estimador não viesado seja desejado, muitas vezes o estimador dado em (1) é utilizado, pois se trata de um estimador de máxima verossimilhança de σ^2 , isto é, da variância populacional, método este muito empregado e baseado em bons princípios estatísticos

Segundo Monico et al. (2009), o RMSE reflete a acurácia teórica da amostra, que, para todos os casos, engloba efeitos aleatórios ($\sigma_{\text{aleatório}}$) e sistemáticos ($\sigma_{\text{sistemático}}$). Sendo assim, uma avaliação estatística mais fidedigna e conveniente seria em termos de dois parâmetros independentes, possibilitando que haja discriminação entre estes efeitos, tal como apresentado na Equação 3.

$$\Phi = \sqrt{\sigma_{\text{aleatório}}^2 + \sigma_{\text{sistemático}}^2} \quad (3)$$

Esta equação é frequentemente utilizada no meio hidrográfico para estimação da incerteza vertical (Φ) (SOUZA e KRUEGER, 2009; FERREIRA et al., 2015; MEZINE et al., 2016). Nestes casos, geralmente os efeitos aleatórios, parcela variável com a profundidade observada, são estimados através do desvio padrão populacional (σ) das discrepâncias entre as profundidades, enquanto que os efeitos sistemáticos, parcela que não depende da profundidade, são quantificados através da média populacional (μ) das discrepâncias. Höhle e Höhle (2009), também sugerem o uso da média e do desvio padrão para quantificar, respectivamente, efeitos sistemáticos e aleatórios, porém, em amostras normais.

O emprego da Equação 3 tem raízes nas especificações previstas na Publicação Especial S-44, 5ª edição (IHO, 2008), em que a tolerância prevista para a incerteza vertical, é calculada através de uma expressão matemática que faz distinção entre efeitos aleatórios e sistemáticos, isto é, uma componente da incerteza vertical que varia de acordo com a profundidade e uma componente que independe da profundidade observada (IHO, 2008).

Entretanto, na prática, as Equações 1 e 3 fornecem os mesmos resultados, como poderá ser confirmado na seção 3.

Num contexto estatístico, é sempre preferível que uma estimativa seja apresentada conjuntamente com seu grau de confiabilidade. Isso pode ser realizado através da construção de intervalos de confiança (IC), quase sempre, baseados na distribuição amostral do estimador pontual (MORETTIN e BUSSAB, 2004). Nesta etapa, outro equívoco teórico associado ao uso das expressões acima é corriqueiramente encontrado na literatura, conforme exposto por Ferreira (2018).

Todavia, a maior problemática do uso das expressões supracitadas, reside no fato de ambas serem altamente influenciadas pela presença de

outliers (ARMSTRONG e COLLOPY, 1992; LI et al., 2005). Deste modo, o emprego destas equações fica limitado a amostras sabidamente não contaminadas por outliers. Assim sendo, visando confirmar a presença, ou não, de valores anômalos, uma pesquisa por outliers deve ser conduzida.

Para a detecção e eliminação de valores anômalos, recorre-se, geralmente, a técnicas de detecção de outliers, tais como: boxplot e z-score (SEO, 2006). Estas técnicas, apesar de na sua grande maioria serem baseadas em estatísticas robustas, não levam em consideração a posição geográfica dos dados, o que, a priori, as torna ineficientes no controle de qualidade de dados batimétricos (TUKEY, 1977; CHAMBERS et al., 1983; HOAGLIN et al., 1983; IGLEWICZ e HOAGLIN, 1993). Soma-se a isto, o fato destas técnicas assumirem que as observações são variáveis aleatórias independentes e identicamente distribuídas (MORETTIN e BUSSAB, 2004; SEO, 2006), pressupostos indispensáveis para um tratamento estatístico clássico e coerente.

Outra problemática da maioria dessas metodologias reside no fato dos valores de corte para detecção de outliers serem derivados da distribuição normal, o que reduz a eficiência da metodologia quando a amostra não é simétrica (HUBERT e VANDERVIEREN, 2008). A busca e eliminação manual de outliers, também pode ser realizada, porém, a alta subjetividade deste procedimento, o torna inadequado, especialmente, para grandes volumes de dados. Em todos os casos, a detecção e eliminação de dados anômalos deve ser realizada com cautela, uma vez que na mesma proporção que a presença de outliers compromete as análises, a exclusão indiscriminada de discrepâncias pode omitir informações importantes.

Destaca-se ainda que, num levantamento batimétrico a superfície de fundo não é visível, deste modo, a busca por profundidades anômalas, que possam estar eivando com outliers a amostra de discrepâncias, torna-se uma tarefa ainda mais difícil.

Em vista disso, este artigo buscou o desenvolvimento de um estimador pontual robusto, ou seja, resistente a outliers, chamado Incerteza Robusta

(Φ_{Robs}). O desenvolvimento deste estimador foi baseado nas seguintes premissas principais:

- A) Resistente a outliers;
- B) Independente da distribuição de probabilidade da amostra analisada;
- C) De aplicação e implementação computacional simples, e
- D) Descrito por meio de dois parâmetros: efeito sistemático e efeito aleatório.

A eficiência do Φ_{Robs} foi estabelecida a partir da aplicação de estatísticas sabidamente robustas para estimativa dos efeitos aleatórios e sistemáticos da amostra de discrepâncias (MOOD, 1913). Neste sentido, adotou-se a mediana (Q2), como estimador para o possível efeito sistemático presente nos dados, e o NMAD (Desvio Absoluto da Mediana Normalizado, do Inglês, *Normalized Median Absolute Deviation*), para estimar o efeito aleatório (MOOD, 1913; HÖHLE e HÖHLE, 2009).

A mediana é uma medida de tendência central mais adequada quando comparada à média amostral, pois indica exatamente o valor central de um conjunto de dados quando organizados em ordem crescente ou decrescente, além de ser pouco afetada por outliers (MOOD, 1913). Segundo Höhle e Höhle (2009), a mediana também é uma medida de localização mais apropriada para dados assimétricos.

O NMAD corresponde a $1,4826 \cdot \text{mediana}\{|x_i - Q2|\}$, em que x_i corresponde a discrepância i . É considerado uma estimativa para a dispersão dos dados mais resistente a outliers que o tradicional desvio padrão. Nos casos em que a distribuição normal for verificada, a Q2 e o NMAD, serão equivalentes, respectivamente, à média e ao desvio padrão (HOAGLIN et al., 1983; HÖHLE e HÖHLE, 2009). Nesta linha, o estimador proposto na Equação 4, torna-se mais robusto que aqueles apresentados até o momento.

$$\Phi_{\text{Robs}} = \sqrt{(Q2)^2 + (\text{NMAD})^2}. \quad (4)$$

Para definição do grau de confiabilidade do Φ_{Robs} sugere-se o emprego da técnica Bootstrap.

O Bootstrap consiste numa técnica de reamostragem que permite aproximar a distribuição de uma função das observações a partir da distribuição empírica dos dados baseado em uma amostra de tamanho finito (EFRON, 1979; EFRON e TIBSHIRANI, 1993). O método foi proposto por Efron (1979), inicialmente, como uma ferramenta para estimar o erro padrão da estimativa de um parâmetro. Atualmente esta metodologia possui uma série de aplicações, destacando-se, a obtenção de intervalos de confiança (HESTERBERG et al., 2003; FRANCO e REISEN, 2007).

2 Experimentos e resultados

Com o objetivo de avaliar a robustez e eficiência do estimador proposto, em um primeiro momento, recorreu-se a simulação computacional. Posteriormente, utilizou-se dados batimétricos reais, provenientes de sondagens batimétricas monofeixe e multifeixe. As simulações e demais estimativas, foram realizadas num algoritmo implementado no software livre R (R Core Team, 2017).

2.1. Dados simulados

A variável aleatória discrepância (d_i) foi simulada a partir de uma distribuição conjunta de probabilidade conhecida, a saber, um processo gaussiano, com média 0,250 metros, desvio padrão de 0,050 metros e tamanho amostral igual a 900. A unidade metro foi adotada unicamente com fins ilustrativos, uma vez que os dados advêm de uma simulação computacional. Destaca-se que, o conjunto de dados simulado está livre de valores anômalos. A Tabela 1 resume as informações descritivas do conjunto

de dados gerado.

Tabela 1 – Análise exploratória dos dados simulados.

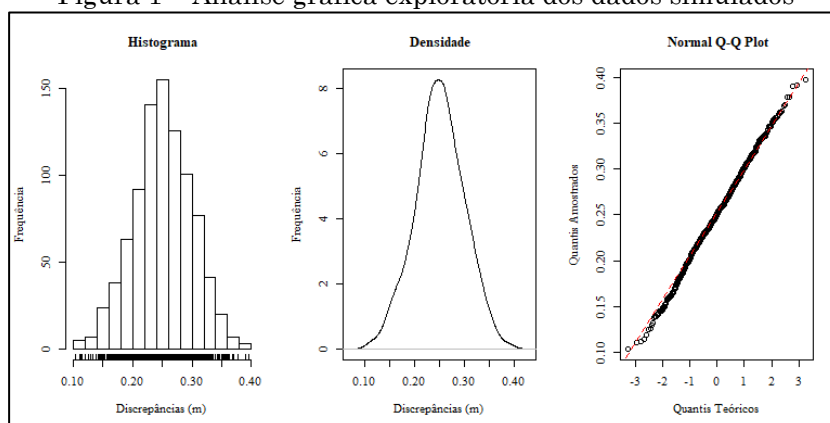
Número de Discrepâncias	900
Média (m)	0,250
Mínimo (m)	0,107
Máximo (m)	0,391
Variância (m²)	0,0025
Desvio-padrão (m)	0,050
Coefficiente de Curtose	2,910
Coefficiente de Assimetria	0,004

Fonte: Elaborada pelos autores.

Conforme esperado, os coeficientes de assimetria e curtose que quantificam, respectivamente, o desvio da distribuição das discrepâncias em relação a uma distribuição simétrica e o grau de achatamento da distribuição, indicam uma distribuição basicamente simétrica e leptocúrtica.

O histograma das discrepâncias, a curva de densidade e o gráfico Q-Q Plot, são apresentados na Figura 1, que são ferramentas muito úteis na análise exploratória, pois fornecem uma ideia sobre distribuição de probabilidade e variabilidade dos dados.

Figura 1 – Análise gráfica exploratória dos dados simulados



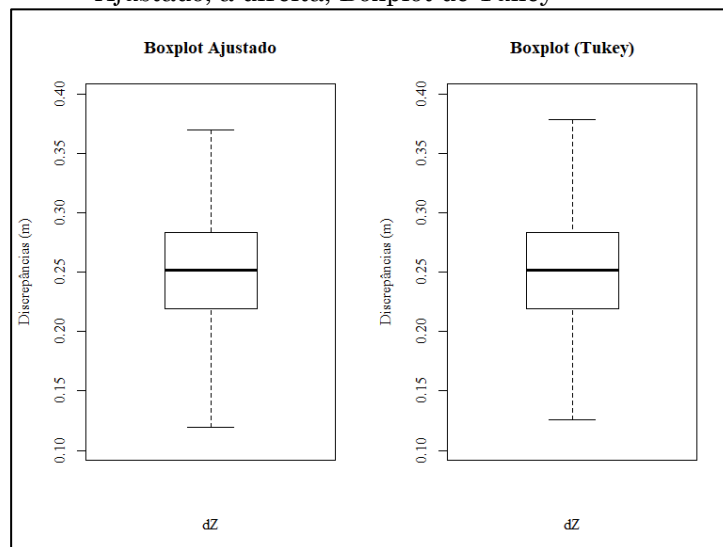
Fonte: Elaborada pelos autores.

A partir de uma análise visual, verifica-se que os dados são normais. Evidencia-se que, na prática, tal constatação deve ser precedida por testes

de normalidade, que por sua vez, supõem independência espacial (Santos, 2015).

Para confirmação da inexistência de outliers foram aplicados dois testes de detecção de outliers, a saber: Boxplot de Tukey e o Boxplot Ajustado (TUKEY, 1977; CHAMBERS et al., 1983; HOAGLIN et al., 1983; VANDERVIEREN e HUBERT, 2004; HUBERT e VANDERVIEREN, 2008). A Figura 2 ilustra os gráficos boxplot, em que se pode constatar que o conjunto de dados não está eivado de outliers, como esperado.

Figura 2 – Gráficos Boxplot do conjunto de dados simulados. À esquerda, Boxplot Ajustado, à direita, Boxplot de Tukey



Fonte: Elaborada pelos autores.

Por fim, a incerteza vertical foi estimada através do *RMSE* (Equação 1), Φ (Equação 3) e Φ_{RobS} (Equação 4). Os intervalos de confiança, ao nível de 95% ($IC_{95\%}$), foram determinados com base na técnica Bootstrap. A Tabela 2 apresenta os resultados.

Tabela 1 – Estimativa intervalar da incerteza vertical dos dados simulados.

Estimador	Incerteza Vertical	$IC_{95\%}$
-----------	--------------------	-------------

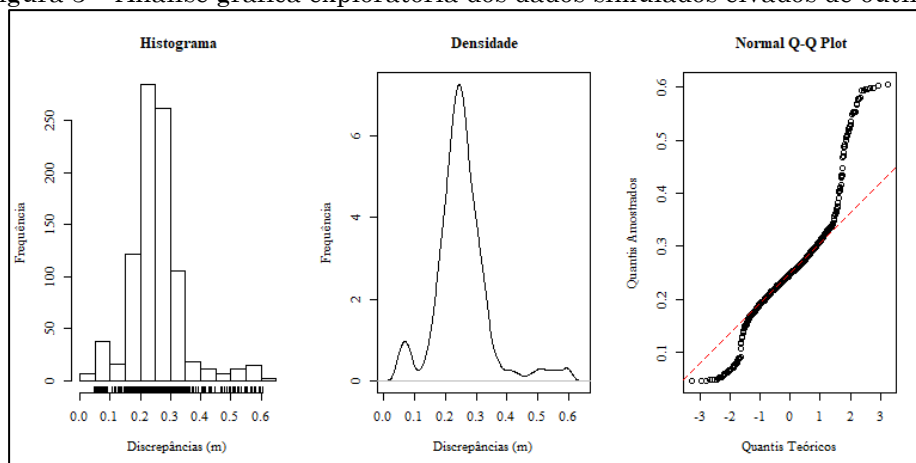
$RMSE (m)$	0,254	[0,250; 0,259]
$\Phi (m)$	0,254	[0,250; 0,259]
$\Phi_{Robs} (m)$	0,253	[0,247; 0,258]

Fonte: Elaborada pelos autores.

Conforme discutido na seção 2, os estimadores RMSE e Φ são teoricamente equivalentes, ou seja, na grande maioria das vezes, fornecem resultados idênticos (Tabela 2). Analisando a Tabela 2, pode-se concluir que nas condições simuladas, isto é, dados normais e sem outliers, o estimador Φ_{Robs} também fornece resultados similares. Em ambos os casos, os intervalos de confiança mostraram-se ótimos.

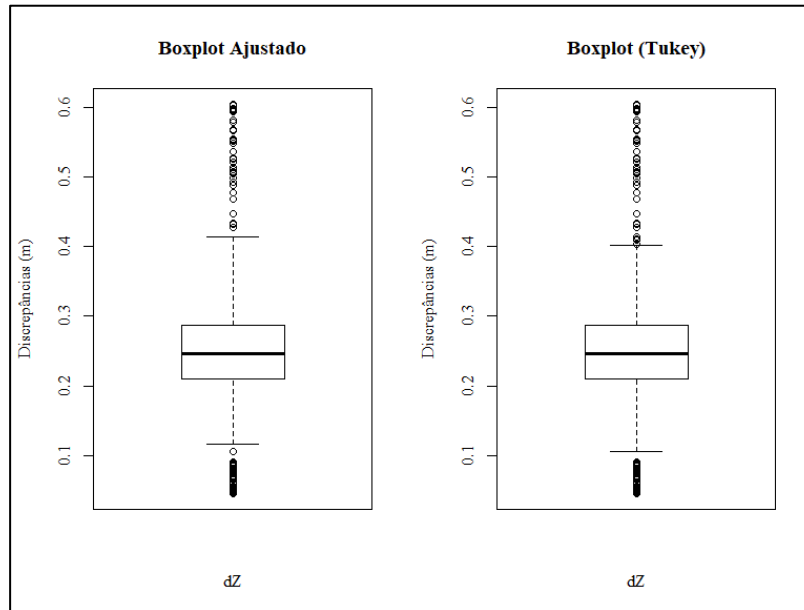
Num segundo momento, mantendo o mesmo tamanho amostral, foram introduzidos aleatoriamente 90 outliers ao conjunto de dados, ou seja, a amostra foi contaminada com 10% de dados anômalos. Como suposto, o novo conjunto de dados apresentou uma distribuição não-normal. Assumindo independência espacial, apenas para fins explicativos, aplicou-se os testes de normalidade Kolmogorov-Smirnov (DOOB, 1949) e Shapiro-Wilk (SHAPIRO e WILK, 1965), obtendo valores-p muito próximos de zero ($\sim 10^{-16}$). As Figuras 3 e 4 resumem o discutido.

Figura 3 – Análise gráfica exploratória dos dados simulados eivados de outliers



Fonte: Elaborada pelos autores.

Figura 4 – Gráficos Boxplot do conjunto de dados simulados eivados de outliers. À esquerda, Boxplot Ajustado, à direita, Boxplot de Tukey



Fonte: Elaborada pelos autores.

Os resultados das estimativas intervalares, ao nível de confiança de 95%, são apresentados na Tabela 3.

Tabela 2 – Estimativa intervalar da incerteza vertical dos dados simulados eivados de outliers

Estimador	Incerteza Vertical	IC_{95%}
<i>RMSE (m)</i>	0,268	[0,260; 0,278]
$\Phi (m)$	0,268	[0,261; 0,277]
$\Phi_{Robs} (m)$	0,253	[0,246; 0,259]

Fonte: Elaborada pelos autores.

Analisando a Tabela 3, nota-se que os intervalos de confiança novamente se mostraram ótimos, comprovando que a técnica Bootstrap, para estes fins, é uma ferramenta muito útil e eficiente.

Observando as Tabelas 2 e 3, é nítida a problemática causada pelos outliers aos estimadores $RMSE$ e Φ . Em ambos os casos, estes estimadores apresentaram resultados idênticos, porém divergentes para os conjuntos de dados com e sem dados anômalos.

Em contrapartida, o estimador Φ_{Robs} mostrou-se bastante eficiente, visto que a presença de outliers na base de dados não interferiu na estimativa pontual da incerteza. Desse modo, confirma-se o alto grau de

robustez da estatística proposta, no que concerne o tratamento de dados, sabidamente ou duvidosamente, eivado de outliers.

Comparando as incertezas obtidas através dos dois diferentes conjuntos de dados, observa-se que a computação da incerteza vertical através do RMSE (Equação 1) ou Φ (Equação 3) pode subestimar a qualidade do levantamento hidrográfico. Uma análise prévia e eficiente da presença de outliers pode, a priori, solucionar este problema. Entretanto, conforme discutido na seção 2, é mais confiável empregar o estimador Φ_{Robs} , visto que o mesmo independente da presença de dados anômalos. Deste modo, para estes fins específicos, as aplicações de técnicas de detecção de outliers, ou mesmo a busca manual, podem ser descartadas ou realizadas com menos afinco. Um resultado bastante relevante, especialmente, no tratamento de dados batimétricos.

2.2. Dados reais

Os conjuntos de dados reais que serviram de base para a avaliação do estimador proposto são provenientes de dois levantamentos batimétricos.

O primeiro conjunto de dados foi coletado em setembro de 2013 através de um levantamento batimétrico monofeixe realizado em uma das lagoas do município de Capitólio – centro oeste do Estado de Minas Gerais. Foi utilizado um sistema de batimetria monofeixe composto por um ecobatímetro de dupla frequência (33/210kHz) e um receptor GNSS operando no modo RTK (*Real Time Kinematic*).

Na fase de planejamento da batimetria estipulou-se um espaçamento de 20 metros para as linhas regulares de sondagem dispostas de modo perpendicular às linhas isobatimétricas da área e 100 metros para as linhas de verificação. Esse levantamento não adotou, a priori, espaçamentos baseados em normativas.

O segundo conjunto de dados resultou de uma sondagem batimétrica multifeixe realizada em 18 de maio de 2017, na região do Porto do Rio de

Janeiro, entre a Ilha das Enxadas e a Ilha das Cobras. Os dados foram coletados por um sistema de sondagem por faixa, composto pelo ecobatímetro multifeixe modelo Sonic 2022 da marca R2 Sonic, integrado com o sistema inercial, modelo I2NS (*Integrated Inertial Navigation System*) da marca Applanix. Destaca-se que o planejamento, execução, bem como o processamento dos dados seguiram recomendações da NORMAM-25 (DHN, 2014) e S-44 (IHO, 2008) para a categoria A e ordem Especial, respectivamente.

Destaca-se que para obtenção do arquivo de discrepâncias dos levantamentos multifeixe foram empregadas abordagens bastante conhecidas entre a comunidade hidrográfica e, por isso, de larga aplicação (SUSAN e WELLS, 2000; SOUZA e KRUEGER, 2009; EEG, 2010).

A tabela a seguir resume as informações descritivas dos conjuntos de dados.

Tabela 3 – Análise exploratória dos dados reais.

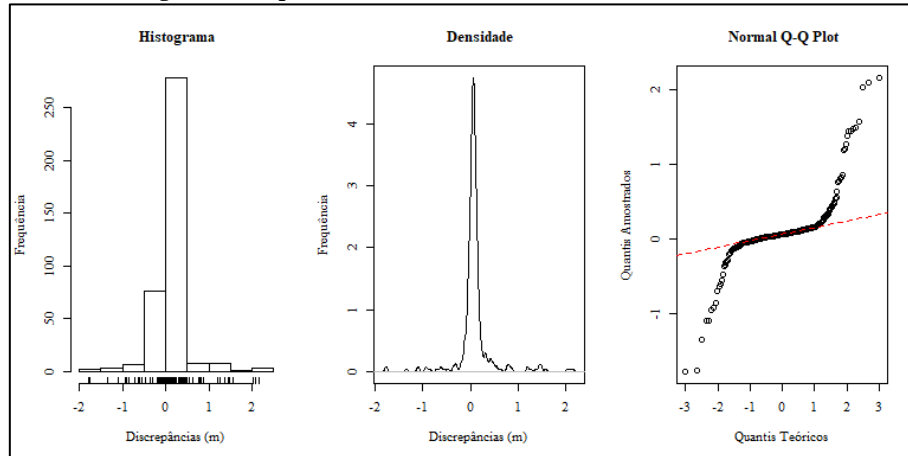
Estatísticas	Dados Monofeixe	Dados Multifeixe
Nº de Discrepâncias	386	259.435
Média (m)	0,091	-0,173
Mínimo (m)	-1,773	-3,725
Máximo (m)	2,150	0,226
Variância (m ²)	0,1363	0,5554
Coef. de Curtose	15,790	15,350
Coef. de Assimetria	1,150	-3,610

Fonte: Elaborada pelos autores.

Em termos de tamanho amostral, o arquivo de discrepâncias multifeixe é muito maior, cerca de 700 vezes. Possui também, uma maior variabilidade e uma conseqüente menor simetria. Entretanto, o achatamento da distribuição das amostras é basicamente o mesmo (Tabela 4). Por outro lado, a partir da construção dos gráficos: histograma, curva de densidade e Q-Q Plot, nota-se que ambos os conjuntos de dados são distorcidos, possuem outliers e não seguem distribuição normal (Figura 5 e

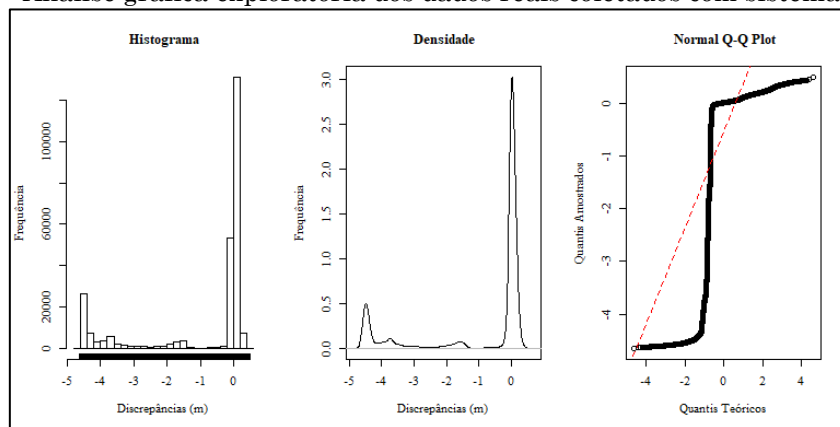
Figura 6). Esta problemática é ainda mais visível para o conjunto de discrepâncias provenientes da sondagem multifeixe. Embora, vale destacar que o método empregado para obtenção dos arquivos de discrepâncias tem grande influência sobre estas conclusões (Ferreira, 2018).

Figura 5 – Análise gráfica exploratória dos dados reais coletados com sistema monofeixe



Fonte: Elaborada pelos autores.

Figura 6 – Análise gráfica exploratória dos dados reais coletados com sistema multifeixe

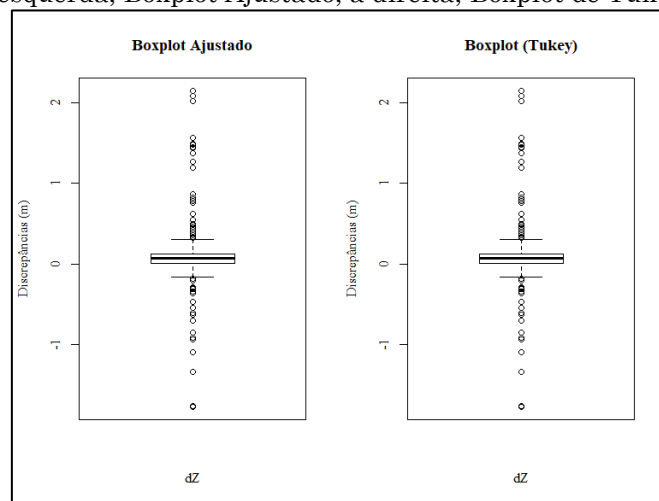


Fonte: Elaborada pelos autores.

Nas Figuras 7 e 8 apresentam-se os gráficos boxplot para os dados monofeixe e multifeixe. Conforme explanado por Ferreira (2018) e discutido na seção 2, estes mecanismos de detecção de outliers apresentam falhas quando os dados são georreferenciados. Tais falhas, isto é, suposições estatísticas dificilmente atendidas e/ou verificadas, são ainda mais agravantes para dados espacialmente densos, tais como os dados adquiridos

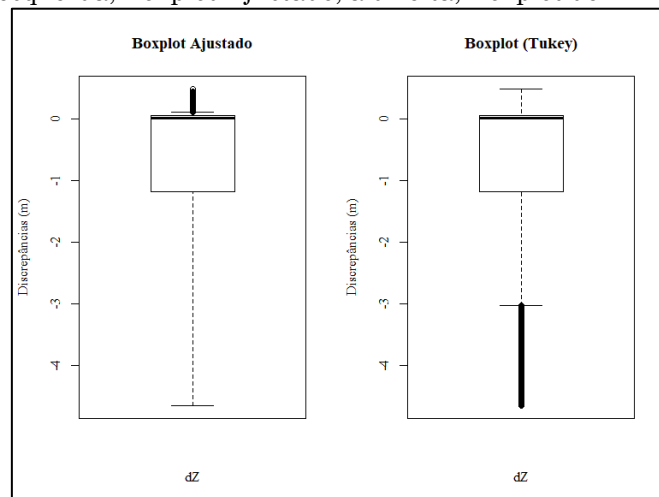
num levantamento batimétrico multifeixe, como pode ser visto na Figura 8, apresentada aqui apenas com fins ilustrativos. Uma metodologia chamada AEDO (Algoritmo Espacial Para detecção de Outliers), foi apresentada por Ferreira (2018), para tratar este tipo de problema.

Figura 7 – Gráficos Boxplot do conjunto de dados reais coletados com sistema monofeixe. À esquerda, Boxplot Ajustado, à direita, Boxplot de Tukey



Fonte: Elaborada pelos autores.

Figura 8 – Gráficos Boxplot do conjunto de dados reais coletados com sistema multifeixe. À esquerda, Boxplot Ajustado, à direita, Boxplot de Tukey



Fonte: Elaborada pelos autores.

O AEDO foi utilizado como metodologia para busca e eliminação de outliers possivelmente presentes na amostra de discrepâncias da sondagem multifeixe. Dado a baixa densidade espacial da amostra de discrepâncias

monofeixe, optou-se por utilizar a técnica boxplot, conforme sugerido por Ferreira (2018).

Ressalta-se, mais uma vez, que esta é uma fase importante e, por esse motivo, deve ser realizada com muita atenção. Diante disso, os dados processados, correspondentes às discrepâncias localizadas como outliers, foram reavaliados, buscando-se falhas no processamento durante as fases de eliminação de spikes e tops, redução de maré e aplicação da velocidade de propagação do som às sondagens. Por fim, todos os possíveis outliers foram excluídos, num total de 55 (~15%) para os dados monofeixe e 52.108 (~20%) no conjunto de discrepância do levantamento multifeixe.

Finalmente, a incerteza vertical foi estimada, para ambos os conjuntos de dados e cenários. A Tabela 5 resume os resultados.

Tabela 4 – Estimativa intervalar da incerteza vertical dos dados reais.

Amostra	Estimador	Dados com outliers	Dados sem outliers	
		Incerteza Vertical	Incerteza Vertical	IC _{95%}
Dados Monofeixe	RMSE (m)	0,380	0,100	[0,090; 0,107]
	Φ (m)	0,380	0,100	[0,090; 0,108]
	Φ_{Robs} (m)	0,109	0,091	[0,080; 0,102]
Dados Multifeixe	RMSE (m)	1,937	0,765	[0,743; 0,795]
	Φ (m)	1,936	0,766	[0,740; 0,796]
	Φ_{Robs} (m)	0,100	0,061	[0,055; 0,076]

Fonte: Elaborada pelos autores.

Analisando a Tabela 5, percebe-se que os estimadores RMSE e Φ apresentaram equivalência em todos os casos, observando-se diferenças milimétricas para os dados multifeixe. No entanto, estes estimadores são altamente influenciados por dados anômalos. Para os dados analisados, a computação da incerteza vertical utilizando estes estimadores e sem uma análise prévia da presença de outliers, comprometeria excessivamente a avaliação da qualidade das profundidades coletadas, podendo alterar a

destinação das informações, principalmente aquelas obtidas através do levantamento multifeixe. Em contrapartida, o estimador Φ_{Robs} apresentou resultados bastante promissores e significativos. Para os dados monofeixe, a diferença entre as incertezas calculadas com base nos conjuntos de dados com outliers e sem outliers, foi menor que 2 centímetros, um valor insignificante, para metodologia de coleta de profundidades empregada. Para o conjunto de dados multifeixe, essa diferença foi em torno de 4 centímetros, o que também é um resultado satisfatório conhecida a natureza de coleta de dados batimétricos.

Por fim, os estimadores apresentaram intervalos de confiança bastante estreitos, mostrando que os dados analisados (sem outliers) possuem uma boa confiabilidade, isto é, tais amostras representam com fidelidade a população de origem e dessa forma, pode-se confiar no julgamento acerca da qualidade vertical do levantamento hidrográfico analisado.

Deve-se destacar que neste estudo, as estimativas a partir dos dados brutos foram comparadas com as incertezas calculadas através dos dados com possíveis outliers eliminados com auxílio de técnicas que não são 100% exatas. Assim, estes resultados são altamente dependentes dos métodos Boxplot e AEDO. Acredita-se que o Boxplot tenha influenciado as estimativas de incerteza vertical do levantamento batimétrico monofeixe, visto que o mesmo é uma técnica destinada ao tratamento de dados univariados.

Por outro lado, o AEDO teve sua eficiência comprovada através de uma série de experimentos descritos em Ferreira (2018). Assim, atribui-se as diferenças obtidas nas computações dos Φ_{Robs} , cerca de 4 centímetros, ao método utilizado para obtenção das discrepâncias.

Nitidamente o estimador proposto é, de fato, resistente aos possíveis outliers presentes no conjunto de dados. Pode-se perceber que a presença dos outliers nas bases de dados interferiu minimamente nas estimativas pontuais de incerteza para este estimador, mostrando a eficiência e robustez

do Φ_{Robs} . Tais resultados podem ser de grande relevância no que concerne ao tratamento de dados de batimetria. A principal característica do estimador proposto é a resistência a mudanças de uma pequena porção nas observações.

3 Conclusões

A partir dos resultados obtidos empregando o estimador proposto, acredita-se que os objetivos deste trabalho foram atingidos, visto que o mesmo apresentou melhor desempenho na estimação da incerteza vertical dos dados utilizados, em comparação aos estimadores convencionais.

Os estimadores RMSE e Φ apresentaram equivalência nos resultados do conjunto de dados monofeixe, e diferenças mínimas para o caso dos dados multifeixe. Porém, apresentaram-se altamente influenciados por dados anômalos, o que pode comprometer seriamente a qualidade das profundidades levantadas.

Nesse sentido, para o estimador proposto a diferença entre as incertezas calculadas com base nos conjuntos de dados monofeixe com ou sem a presença de outliers, foi menor que 2 centímetros. Já para o conjunto de dados multifeixe, essa diferença foi em torno de 4 centímetros. Ressalta-se que esses valores podem ser considerados insignificantes, mesmo para o caso dos dados multifeixes, que foram um pouco maiores, considerando um procedimento de coleta de profundidades, conhecida sua natureza complexa de obtenção, indicando resultados bastante promissores e significativos.

Comparando as incertezas obtidas através dos dois diferentes conjuntos de dados, observa-se que, tais resultados subestimariam a qualidade do levantamento. Tal fato pode colocar em dúvida a classificação da sondagem batimétrica de acordo com a norma utilizada e/ou requerida para o projeto em questão, levando, inclusive, a uma rejeição da batimetria pelo executor ou contratante.

Por fim, vale destacar que, o estimador proposto apresentou-se eficiente e robusto, no sentido em que se mostrou resistente aos possíveis outliers presentes no conjunto de dados, e mesmo com a presença de outliers na amostra, os resultados mostraram uma influência mínima desses dados anômalos nas estimativas pontuais de incerteza.

Referências

- ARMSTRONG, J. S.; COLLOPY, F. Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons. **International Journal of Forecasting**, v. 8, n. 1, p. 69–80, 1992.
- ATHEARN, N.; TAKEKAWA, J. Y.; JAFFE, B.; HATTENBACH, B. J.; A. C. FOXGROVER. Mapping Elevations of Tidal Wetland Restoration Sites in San Francisco Bay: Comparing Accuracy of Aerial Lidar with a Singlebeam Echosounder. **Journal of Coastal Research** – v. 26. p.312-319, 2010.
- AYRES NETO, A. Uso da sísmica de reflexão de alta resolução e da sonografia na exploração mineral submarina. **Brazilian Journal of Geophysics**, v. 18, n. 3, p. 241-256, 2000.
- BJØRKE, J. T.; NILSEN, S. Fast trend extraction and identification of spikes in bathymetric data. **Computers & Geosciences**, v. 35, n. 6, p. 1061-1071, 2009.
- BOTTELIER, P.; BRIESE, C.; HENNIS, N.; LINDENBERGH, R.; PFEIFER, N. Distinguishing features from outliers in automatic Kriging-based filtering of MBES data: a comparative study. **Geostatistics for Environmental Applications**, Springer, p. 403-414, 2005.
- CARMO, E. J.; RODRIGUES, D. D.; SANTOS, G. R. Avaliação dos interpoladores krigagem e topo to raster para geração de modelos digitais de elevação a partir de um “as built”. **Boletim de Ciências Geodésicas**, v. 21, n. 4, p. 674-690, 2015.
- CALDER, B. R.; MAYER, L. A. Automatic processing of high-rate, high-density multibeam echosounder data. **Geochemistry, Geophysics, Geosystems**, v. 4, n. 6, 2003.

- CHAMBERS, J. M.; CLEVELAND, W. S.; KLEINER, B.; TUKEY, P. A. **Graphical Methods for Data Analysis**. Pacific Grove, CA: Wadsworth & Brooks/Cole, 1983.
- CHENG, L.; MA, L.; CAI, W.; TONG, L.; LI, M.; DU, P. Integration of Hyperspectral Imagery and Sparse Sonar Data for Shallow Water Bathymetry Mapping. **Geoscience and Remote Sensing. IEEE Transactions on**, v. 53, n. 6, p. 3235-3249, 2015.
- COOPER, M. A. R. **Control surveys in civil engineering**. Nichols Pub Co, 381p., 1987.
- DHN – Diretoria de Hidrografia e Navegação. **NORMAM 25 – Normas da Autoridade Marítima para Levantamentos Hidrográficos**. Marinha do Brasil, 2014.
- DEBESE, N.; BISQUAY, H. Automatic detection of punctual errors in multibeam data using a robust estimator. **The International Hydrographic Review**, v. 76 n. 1, p. 49-63, 1999.
- DEBESE, N. Multibeam Echosounder Data Cleaning Through an Adaptive Surface-based Approach. **In: US Hydro 07 Norfolk**, 18p., 2007.
- DEBESE, N.; MOITIÉ, R.; SEUBE, N. Multibeam echosounder data cleaning through a hierarchic adaptive and robust local surfacing. **Computers & Geosciences**, v. 46, p. 330-339, 2012.
- DOOB, J. L. Heuristic Approach to the Kolmogorov-Smirnov Theorems. **Ann. Math. Statist.**, v. 20, n. 3, p. 393–403, 1949.
- EEG, J. On the identification of spikes in soundings. **The International Hydrographic Review**, v. 72, n. 1, p. 33-41, 1995.
- EEG, J. Multibeam Crosscheck Analysis: A Case Study. **The International Hydrographic Review**, n. 4, p. 25-33, 2010.
- EFRON, B. Bootstrap Methods: Another Look at the Jackknife. **The Annals of Statistics**, v. 7, n. 1, p. 1–26, 1979.
- EFRON, B.; TIBSHIRANI, R. J. **An introduction to the bootstrap**. Boca Raton: CRC, USA, 436p, 1993.
- ELLMER, W.; ANDERSEN, R. C.; FLATMAN, A.; MONONEN, J.; OLSSON, U.; ÖIÅS, H. Feasibility of Laser Bathymetry for Hydrographic Surveys on the

- Baltic Sea. **The International Hydrographic Review**, n. 12, p. 33-50, 2014.
- FERREIRA, Í. O. **Controle de qualidade em levantamentos hidrográficos**. 2018. 233 f. Tese (Doutorado em Engenharia Civil) – Universidade Federal de Viçosa, 2018.
- FERREIRA, Í. O.; NETO, A. A.; MONTEIRO, C. S. O uso de embarcações não tripuladas em levantamentos batimétricos. **Revista Brasileira de Cartografia**, v. 68, n. 10, p. 1885-1903, 2017a.
- FERREIRA, Í. O.; RODRIGUES, D. D.; NETO, A. A.; MONTEIRO, C. S. Modelo de incerteza para sondadores de feixe simples. **Revista Brasileira de Cartografia**, v. 68, n. 5, p. 863-881, 2016a.
- FERREIRA, Í. O.; RODRIGUES, D. D.; SANTOS, G. R.; **Coleta, processamento e análise de dados batimétricos**. 1^a ed. Saarbrücken: Novas Edições Acadêmicas, v. 1, 100p., 2015.
- FERREIRA, Í. O.; RODRIGUES, D. D.; SANTOS, G. R.; ROSA, L. M. F. In bathymetric surfaces: IDW or Kriging? **Boletim de Ciências Geodésicas**, v. 23, n. 3, p. 493-508, 2017b.
- FERREIRA, Í. O.; SANTOS, G. R.; RODRIGUES, D. D. Estudo sobre a utilização adequada da krigagem na representação computacional de superfícies batimétricas. **Revista Brasileira de Cartografia**, v. 65, n. 5, p. 831-842, 2013.
- FERREIRA, I. O.; ZANETTI, J.; GRIPP, J. S.; MEDEIROS, N. G. Viabilidade do uso de imagens do sistema Rapideye na determinação da batimetria de águas rasas. **Revista Brasileira de Cartografia**, v. 68, n. 7, p. 1331-1340, 2016b
- FGDC – Federal Geographic Data Committee. **National Standard for Spatial Data Accuracy, Part 3: National Standard for Spatial Data Accuracy**. Federal Geographic Data Committee: Reston, USA, 25p., 1998.
- FRANCO, G. C.; REISEN, V. A. Bootstrap approaches and confidence intervals for stationary and non-stationary long range dependence processes. **Physica A: Statistical Mechanics and its Applications**, v. 375, p. 546-562, 2007.
- GAO, J. Bathymetric mapping by means of remote sensing: methods, accuracy and limitations. **Physical Geography**, v. 33, n. 1, p. 103-116, 2009.

- GREENWALT, C. R.; SHULTZ, M. E. **Principles of error theory and cartographic applications**. 1. ed. St. Louis: Aeronautical Chart and Information Center, 1962.
- GUENTHER, G. C.; THOMAS, R. W. L. ; LAROCQUE, P. E. Design considerations for achieving high accuracy with the Shoals bathymetric Lidar system. **In: CIS Selected Papers: Laser Remote Sensing of Natural Waters-From Theory to Practice**. International Society for Optics and Photonics, p. 54-71, 1996.
- HALLAK, R.; PEREIRA, A. J. Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de são paulo com o modelo arps: sensibilidade a variações com os esquemas de advecção e assimilação de dados. **Revista Brasileira de Meteorologia**, v. 26, n.5, p. 591-608, 2011.
- HESTERBERG, T.; MONAGHAN S.; MOORE, D. S.; CLIPSON, A.; EPSTEIN R. **Bootstrap methods and permutation tests: companion chapter 18 to the practice of business statistics**. New York: W. H. Freeman and Company, USA, 85 p., 2003.
- HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J. W. **Understanding robust and exploratory data analysis**. New York: Wiley, 433p., 1983.
- HÖHLE, J.; HÖHLE, M. Accuracy assessment of digital elevation models by means of robust statistical methods. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 64, n. 4, p. 398-406, 2009.
- HUBERT, M.; VANDERVIEREN, E. An adjusted boxplot for skewed distributions. **Journal of Computational statistics & data analysis**, v. 52, n. 12, p. 5186-5201, 2008.
- IGLEWICZ, B.; HOAGLIN, D. **How to detect and handle outliers**. Milwaukee, Wis.: ASQC Quality Press, 87p., 1993.
- IHO–International Hydrographic Organization. C-13: **IHO Manual on Hydrography**. Mônaco: International Hydrographic Bureau, 2005, 540p.
- IHO–International Hydrographic Organization. S-44: **IHO Standards for Hydrographic Surveys**. Special Publication n. 44–5th. Mônaco: International Hydrographic Bureau, 36p., 2008.

- KAFRAWY, S. B.; DONIA, N. S.; MOHAMED, A. M. Monitoring the Environmental Changes of Mariout Lake during the Last Four Decades Using Remote Sensing and GIS Techniques. **MOJ Ecology & Environmental Sciences**. n. July, p. 2–6, 2017.
- LEHMANN, E. L.; CASELLA, G. **Theory of point estimation**. 2. ed. New York: Springer-Verlag, 1998.
- LI, Z.; ZHU, Q.; GOLD, C. M. **Digital terrain modelling. Principles and methodology**. New York: CRC Press, 319 p., 2005.
- LU, D.; LI, H.; WEI, Y.; ZHOU, T. Automatic outlier detection in multibeam bathymetric data using robust LTS estimation. In: 3rd International Congress on Image and Signal Processing (CISP), **IEEE**, v. 9, p. 4032-4036, 2010.
- MEZINE, W. A.; SOARES, M. A.; FAGGION, P. L.; VEIGA, L. A. K.; NADAL, C. A. Evaluation methods of reducing bathymetric surveys to the water level instantaneous water for monitoring siltation: case study of the hydroelectric power plant reservoir Mauá. **Boletim de Ciências Geodésicas**. v.22, n.4, p. 736–750, 2016.
- MIGUENS, A. P. **Navegação: a Ciência e a Arte**. v. 3-Navegação Eletrônica e em Condições Especiais. Rio de Janeiro: DHN, Brasil, p. 1221-1822, 2000.
- MIKHAIL, E.; ACKERMAN, F. **Observations and Least Squares**. University Press of America, 497p., 1976.
- MONICO, J. F. G.; DAL POZ, A. P.; GALO, M.; DOS SANTOS, M. C.; DE OLIVEIRA, L. C. Acurácia e precisão: Revendo os conceitos de forma acurada. **Boletim de Ciências Geodesicas**, v. 15, n. 3, 2009.
- MOOD, A. M. **Introduction to the theory of statistics**. McGraw-Hill series in probability and statistics, 564p., 1913.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. McGraw-Hill International, 577p., 1974.
- MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. 1. ed. São Paulo: Saraiva, 2004.
- MOTAO, H.; GUOJUN, Z.; RUI, W.; YONGZHONG, O.; ZHENG, G. Robust method for the detection of abnormal data in hydrography. **The International Hydrographic Review**, v. 76, n. 2, p. 93-102, 1999.

- MOURA, A.; GUERREIRO, R.; MONTEIRO, C. As potencialidades da derivação de batimetria a partir de imagens de satélite multiespectrais na produção de cartografia náutica. **4as Jornadas de Engenharia Hidrográfica**. Instituto Hidrográfico Português, Lisboa, Portugal, 2016.
- PASTOL, Y. Use of Airborne LIDAR Bathymetry for Coastal Hydrographic Surveying: The French Experience. **Journal of Coastal Research**. n. 62, p. 6-18, 2011.
- SANTOS, A. P. **Controle de qualidade cartográfica: metodologias para avaliação da acurácia posicional em dados espaciais**. Tese (Doutorado). Programa de PósGraduação em Engenharia Civil, Departamento de Engenharia Civil, Universidade Federal de Viçosa, Viçosa, Minas Gerais, 172p., 2015.
- SEKELLICK, A. J.; BANKS, W. S. L. Water Volume and Sediment Accumulation in Lake Linganore, Frederick County , Maryland , 2009. **U. S. Geological Survey Scientific Investigations Report 2010-5174**. p. 1–14, 2010.
- SEO, S. **A review and comparison of methods for detecting outliers in univariate data sets**. Master Of Science, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, USA, 59p., 2006.
- SHAPIRO, A. S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika Trust**, v. 52, n. 3, p. 591–611, 2009.
- SOUZA, A. V.; KRUEGER, C. P. **Avaliação da qualidade das profundidades coletadas por meio de ecobatímetro multifeixe**. Anais Hidrográficos, Rio de Janeiro, n. 66, p. 90-97, 2009.
- SUSAN, S.; WELLS, D. Analysis of Multibeam Crosschecks Using Automated Methods. **US Hydro 2000 Conference paper**, Biloxi, Mississippi. 2000.
- TUKEY, J.W. **Exploratory Data Analysis**. Princeton, Ed. Pearson (1977).
- VANDERVIEREN, E.; HUBERT, M. An adjusted boxplot for skewed distributions. **Proceedings in Computational Statistics**, p. 1933-1940, 2004.
- VICENTE, J. P. D. **Modelação de dados batimétricos com estimação de incerteza**. 2011. 158 f. Dissertação (Mestrado em Sistemas de informação geográfica tecnologia e aplicações)-Universidade de Lisboa, Lisboa, 2011.

- VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F. et al (Ed). **Tópicos em ciência do solo**. Viçosa: SBGS, 2000. v.1, cap.1, p.1-54.
- WARE, C.; KNIGHT, W.; WELLS, D. Memory intensive statistical algorithms for multibeam bathymetric data. **Computers & Geosciences**, v. 17, n. 7, p. 985-993, 1991.