



**MARINHA DO BRASIL  
INSTITUTO DE ESTUDOS DO MAR ALMIRANTE PAULO MOREIRA  
UNIVERSIDADE FEDERAL FLUMINENSE  
PROGRAMA ASSOCIADO DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA MARINHA**

**THIAGO DA SILVA MATOS**

**ELABORAÇÃO DE UM ALGORITMO DE ANÁLISE SEMI-AUTOMÁTICA DE  
COPÉPODA POR IMAGENS *IN SITU* NO AMBIENTE MARINHO.**

**ARRAIAL DO CABO/RJ**

**2018**



**MARINHA DO BRASIL  
INSTITUTO DE ESTUDOS DO MAR ALMIRANTE PAULO MOREIRA  
UNIVERSIDADE FEDERAL FLUMINENSE  
PROGRAMA ASSOCIADO DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA MARINHA**

**THIAGO DA SILVA MATOS**

**ELABORAÇÃO DE UM ALGORITMO DE ANÁLISE SEMI-AUTOMÁTICA DE  
COPÉPODE POR IMAGENS *IN SITU* NO AMBIENTE MARINHO.**

Dissertação apresentada ao Instituto de Estudos do Mar Almirante Paulo Moreira e à Universidade Federal Fluminense, como requisito parcial para a obtenção do grau de Mestre em Biotecnologia Marinha.

**Orientador: Prof. Dr. Lohengrin Dias de Almeida Fernandes**

**ARRAIAL DO CABO/RJ**

**2018**

**THIAGO DA SILVA MATOS**

**ELABORAÇÃO DE UM ALGORITMO DE ANÁLISE SEMI-AUTOMÁTICA DE  
COPÉPODE POR IMAGENS *IN SITU* NO AMBIENTE MARINHO**

Dissertação apresentada ao Instituto de Estudos do Mar Almirante Paulo Moreira e a Universidade Federal Fluminense, como requisito parcial para a obtenção do título de Mestre em Biotecnologia Marinha.

**COMISSÃO JULGADORA:**

---

Prof. Dr. Lohengrin Dias de Almeida Fernandes (Orientador)  
IEAPM – Instituto de Estudos do Mar Almirante Paulo Moreira

---

Prof. Dr. Eduardo Vianna de Almeida (Membro Externo)  
FAMATH – Faculdades Integradas Maria Thereza

---

Prof. Dr. Leandro Calado (Membro Interno)  
IEAPM – Instituto de Estudos do Mar Almirante Paulo Moreira

---

Prof. Dr. Alexandre Dias Kassuga (Membro Suplente)  
IEAPM – Instituto de Estudos do Mar Almirante Paulo Moreira

Arraial do Cabo, 09 de Março de 2018.

Dedico esta dissertação à minha querida mãe Elizabete, meu pai Hélio e minha esposa Anielly, por todo incentivo e apoio em todas as minhas ações profissionais, em especial nesse período de curso, o incentivo de vocês impulsionou o rompimento das dificuldades, e favoreceu esta conquista, obrigado.  
Amo Vocês!

## **AGRADECIMENTOS**

A Deus toda primazia destes agradecimentos, pois sem Ele não seria possível alcançar um passo adiante.

Aos meus pais por todo amor e carinho dedicado em todo tempo, em especial à minha mãe, que sempre me incentivou sem ao menos saber e entender o que eu estudava, mas confiava ser um passo importante, pois via o quanto eu dedicava tempo à leitura e estudo em livros.

À minha esposa, por sempre incentivar, e entender os momentos de dedicação exclusiva ao trabalho, demonstrando paciência e sabedoria nos meus momentos de fraqueza e cansaço. Palavras amigas, atenção, e carinho foram fundamentais para a concretização deste trabalho.

Aos meus irmãos, que direta ou indiretamente sempre participaram das minhas conquistas, sempre incentivando e ajudando.

Aos meus amigos Lorrene, Augusto, Gabriela Nalini que me presentearam com a inscrição do processo seletivo do curso de mestrado.

À tenente Natalia Saraiva, que sempre incentivou os meus passos nos estudos e pesquisa, por inúmeras vezes dedicou tempo para me ajudar e até ensinar.

Ao Dr. Leandro Calado, que Juntamente com o Dr. Ricardo Coutinho me aceitou como bolsista no projeto Peld Ressurgência, me dando oportunidade dentro do Instituto de Pesquisa do Mar Almirante Paulo Moreira (IEAPM).

Ao Capitão de Fragata Fernando Marin, que atenciosamente corrigiu minha carta de intenção, e sempre oferecendo qualquer ajuda necessária.

Ao Capitão de Corveta Gilberto e ao Flávio da Costa, por sempre demonstrarem companheirismo e amizade nos momentos de trabalho.

Ao meu amigo biólogo Paulo Cordeiro, por amizade e companheirismo em diversos estudos e trabalhos biológicos.

Ao Departamento de Biotecnologia Marinha, por proporcionar suporte necessário a realização deste trabalho. Em especial a divisão de Bioincrustação e Bioinvasão, aos amigos, Luciana Altvater e Alexandre Kassuga, ao Rafael Menezes,

aos meus amigos Lais Naval e Fábio Contrera que incansavelmente sempre ajudou tirando dúvidas e até mesmo na realização de etapas fundamentais deste trabalho.

Ao Rafael Freire, Marcio, Debora Costa, Pedro, por companheirismo e amizade. A minha amiga Aline que entrou junto comigo no mestrado e em todo tempo de curso foi companheira nesta caminhada árdua. Ao grande amigo Marcio Tenório, que mesmo distante manteve-se amigo e companheiro em todo o tempo.

À minha querida amiga Wanda Ribas, por dedicar seu precioso tempo em me ensinar os conhecimentos taxonômicos, e sempre dispostos a tirar qualquer dúvida na identificação de copépoda.

À todos os amigos da igreja, por companheirismo e auxílio, palavras de incentivos e orações.

A todos os gestores e professores do PPGBM por grande oportunidade recebida no ingresso do curso.

Ao IEAPM, por ter aberto as portas de estágio, trabalho e desenvolvimento das pesquisas de mestrado. Este trabalho foi apoiado pelo conselho nacional de desenvolvimento científico e tecnológico (CNPq)

Aos membros da banca, Dr. Leandro Calado do IEAPM, Dr. Eduardo Vianna da FAMATH por gentilmente aceitarem o convite.

Ao meu Orientador Dr. Lohengrin Dias de Almeida Fernandes, agradeço a Deus por sua vida, por ter aceitado me orientar e incansavelmente dedicou seu tempo, atenção, e conhecimento na orientação deste trabalho, sem medir esforços ajudou a romper as inúmeras dificuldades. Em todo tempo se ocupou em ensinar, orientar, e algumas vezes quando preciso, foi um grande conselheiro, demonstrando sempre grande amizade para comigo e com todos no laboratório. É um profissional impecável e um ser humano de caráter, sempre cobrou resultados, porém sempre deu mais suporte do que cobrou, se preocupando no aprendizado e apropriação dos conhecimentos necessários para minha formação profissional.

# Índice

RESUMO.....	xi
ABSTRACT .....	xiii
INTRODUÇÃO.....	1
MATERIAL E MÉTODOS .....	6
Local de estudo.....	6
Identificação das imagens de subida.....	8
RESULTADOS E DISCUSSÃO .....	16
CONCLUSÕES.....	29
REFERÊNCIAS .....	30
ANEXOS.....	36

## Lista de Ilustrações

- Figura 1 Foto aérea da área da Foz do Rio Doce logo após a chegada da pluma de rejeitos lamosos em Novembro de 2015, e mapa evidenciando os pontos de coleta, um com alta turbidez próxima a foz, e outro com menos turbidez mais distante. ....6
- Figura 2 Utilização do equipamento LOKI em três momentos distintos, primeiro levado para fora do navio para então ser lançado verticalmente, como mostra a imagem dois, e no ultimo estágio, o LOKI já no deck do navio, pronto para ter os dados descarregados em um computador. ....8
- Figura 3 Exemplo de perfil de variação temporal da Temperatura (°C) e Pressão (psi) da água durante um arrasto com o LOKI. Em vermelho, destaque para o momento da subida do LOKI desde próximo do fundo (máxima pressão e mínima temperatura) até próximo da superfície. ....9
- Figura 4 Fases do tratamento das imagens. A; imagem original, recortada automaticamente pelo sistema do LOKI, utilizada para definição do tom limiar do pixel. B; imagem sem fundo (fundo branco), com destaque para todas as partículas (debris+copépode). C; imagem do contorno do copépode destacado das demais partículas pela medição. ....11
- Figura 5 Etapas de tratamento de imagem utilizando o algoritmo 1, para contagem e classificação dos copépodes.....13
- Figura 6 Etapas de tratamento de imagem utilizando o algoritmo 2, para contagem e classificação dos copépodes.....13
- Figura 7 Etapas de tratamento de imagem utilizando o algoritmo 3, para contagem e classificação dos copépodes.....14

Figura 8 Etapas de tratamento de imagem utilizando o algoritmo 4, para contagem e classificação dos copépodes.....	14
Figura 9 Tratamento da imagem, primeiro a distinção entre imagens aproveitáveis (subida) e imagens não aproveitáveis (outros momentos). Após essa distinção, a correção de alguns problemas gerados em imagens adquiridas em águas turvas, como imagens não segmentadas, mal segmentadas, etc. Em seguida com essas imagens pré-selecionadas e corrigidas, testou-se 4 algoritmos de contagem e classificação do copépode pela área <sup>2</sup> , e então comparado ao tratamento realizado manualmente.....	15
Figura 10 Perfil de variação temporal da Temperatura (°C) e Pressão (mbar) da água durante um arrasto com o LOKI no ponto com alta turbidez, e baixa turbidez. Destacado em vermelho tem-se o momento em que o equipamento registrou máxima e mínima das duas variáveis, e posteriormente associada este momento às imagens.....	16
Figura 11 Diferentes dimensões que as imagens se apresentam, podendo ter tamanho máximo, ou parcial, e ter o objeto de interesse recortado perfeitamente ou mal recortado.....	18
Figura 12 Distribuição dos ROIs por tamanho da imagem, observado a maior frequência das imagens procedentes de águas mais turvas entre 1 e 10 Kb. Em baixa turbidez, a maior ocorrência das imagens foi com dimensão entre 10 e 50 Kb. ....	20
Figura 13 Comparação entre dois pontos com condições distintas de turbidez, nos tratamentos com utilização dos quatro algoritmos elaborados neste trabalho. ....	21

Figura 14 Desempenho dos 16 métodos de threshold existente no programa ImageJ, comparado com o método manual, e avaliado quanto a preservação da partícula central e remoção dos ruídos (píxels) do fundo. Ao lado tem-se a imagem original.	22
Figura 15 Dendrograma de análise de grupamento, demonstrando que o método minimum é o mais próximo ao manual, quanto limpeza do fundo e preservação da partícula central do copépode. ....	23
Figura 16 Estabelecimento do tamanho mínimo do copépode através da medição de 1000 organismos. O tamanho mínimo foi adquirido através da subtração do valor médio pelo desvio padrão. ....	24
Figura 17 Dendrograma explicando a dissimilaridade testada entre quatro algoritmos elaborados pra contagem automática de copépode, comparado com tratamento realizado manualmente. Testes realizados em dois pontos distintos, com alta turbidez e outro ponto com baixa turbidez, mais oceânico.....	27

## Lista de Tabelas

Tabela 1 Teste realizado com os 16 métodos de eliminação do fundo da imagem, com observação de presença e ausência de alguns parâmetros, excluídos ou preservados no teste.....	11
Tabela 2 Testes dos algoritmos de pré-processamento das imagens, separando dentre o total e imagens, as segmentadas, mal segmentadas e não segmentadas.	19
Tabela 3 Número de partículas obtidas individualmente por algoritmo manual e automaticamente por quatro algoritmos elaborados no programa ImageJ. Teste realizado em um conjunto aleatório de 1000 imagens. AT=Alta Turbidez, BT= Baixa Turbidez. ....	26
Tabela 4 Número de partículas obtidas individualmente por algoritmo manual e automaticamente por quatro algoritmos elaborados no programa ImageJ. Teste realizado em um conjunto de 100 imagens aleatórias, em dois pontos distintos, com alta turbidez e outro com baixa turbidez.....	26

## **RESUMO**

A quebra da Barragem do Fundão na cidade de Mariana/MG, ocorrido no dia 05 de Novembro de 2015 provocou o lançamento de rejeitos lamosos oriundos do processo de mineração para o Rio Doce, que percorre os Estados de Minas Gerais e Espírito Santo. Em cerca de algumas semanas, a lama de rejeitos percorreu o Rio Doce e alcançou a foz, na localidade de Regência. O acidente resultou em vítimas fatais e prejuízos incalculáveis a sobreviventes e demais comunidades do entorno, além de danos ambientais irreparáveis. Este acidente provocou alta turbidez da água, condição que dificultou os trabalhos realizados pelo sistema de aquisição de imagens, e inviabilizou a obtenção do ROI (região de interesse) e impossibilitou a contagem dos organismos. Ter a análise e contagem automatizada de copépoda melhora as pesquisas e viabiliza respostas rápidas a situações de risco ambiental, como o caso deste acidente. As imagens foram adquiridas pelo equipamento de aquisição de imagens o *Lightframe On-Sight Keyspecies Investigation System* (LOKI). O presente trabalho objetivou desenvolver algoritmos de tratamento dessas imagens, contando e classificando os copépodes pelo tamanho da área. Os testes foram realizados em um ponto de coleta mais próximo à foz do Rio Doce, com alta turbidez, e outro mais distante com baixa turbidez. No ponto com alta turbidez houve um acúmulo de 40.375 imagens, quantidade esta reduzida pelo algoritmo de separação das imagens aproveitáveis para 20.440 imagens. Já no ponto com baixa turbidez a aquisição foi de 21.664 imagens, quantidade reduzida para 5.832. Após a separação, as imagens foram submetidas a testes com quatro algoritmos elaborados para contagem e classificação automática comparada ao tratamento manual. Neste tratamento o primeiro algoritmo foi estatisticamente o melhor pelo teste de dissimilaridade, com margem de acerto de 96%. Os resultados obtidos neste trabalho foram bem satisfatórios com potencial melhora dos algoritmos em futuros testes, viabilizando acurácia de futuras pesquisas.

**Palavras-chave:** Rio Doce, Turbidez, Imagens, Copépode, LOKI.

## **ABSTRACT**

The failure of the Fundão Dam in the city of Mariana / MG, which occurred on November 5, 2015, led to the launching of tailings from the mining process for Rio Doce, which runs through the states of Minas Gerais and Espírito Santo. The accident resulted in fatalities and incalculable damage to survivors and other surrounding communities, as well as irreparable environmental damage. This accident caused high turbidity of the water, a condition that made the work of the image acquisition system difficult, and made it impossible to obtain ROI (region of interest) and made it impossible to count organisms.

To provide *in situ* and fast automatic analysis of Copepod populations is a challenge in some recent environmental surveys, as the leakage of the Fundão Dam in Mariana (MG). This accident caused a high turbidity water, a condition that made image acquisition difficult and quite impossible to find out ROIs (region of interest) of organisms. The goal of this work was to develop algorithms for image treatment after deployment of a Lightframe On-Sight Keyspecies Investigation System (LOKI). The algorithms focused on counting and classifying copepods by body area. The tests were carried out on a set of images from near the mouth, with high turbidity, and other with low turbidity. Under high turbidity it were obtained 40,375 images (ROIs), which were processed by the algorithms to let 20,440 (~50%) useful images. At the point with low turbidity the few images (21,664) obtained were reduced to only 5,832 useful. Therefore, all images were then submitted to tests under four algorithms elaborated for counting and automatic classification. Copepod counting provided by automatic algorithms was statistically similar to visual counting in 96% of images.

**Keywords:** Rio Doce, Turbidity, Images, Copepod, LOKI.

## INTRODUÇÃO

O zooplâncton marinho desempenha um papel fundamental na transferência de matéria orgânica dos produtores primários para níveis tróficos mais altos e, portanto, no ciclo global do carbono (Hirche *et al.*, 2012). Nesse papel, esses organismos constituem um elo fundamental na transferência de energia ao longo da teia alimentar no ambiente marinho, e sua disponibilidade é fundamental para a dinâmica do ecossistema (Harris *et al.*, 2000). O zooplâncton é um grupo diversificado de animais que habita os ecossistemas aquáticos em todo o mundo (Schmid *et al.*, 2016). Dentre essa diversidade, encontram-se os copépodes, um dos organismos mais abundantes nos oceanos, que podem representar mais de 90% da biomassa do zooplâncton marinho (Bradford-Grieve *et al.*, 1999).

Um dos principais focos de estudo na ecologia do plâncton é compreender a dinâmica espacial e temporal dos organismos e como eles interagem com seus ambientes. Nesses estudos, as redes cilíndricas tipo WP-2 são a ferramenta básica de coleta de amostras do zooplâncton (Calazans, 2011). Contudo, o tempo necessário para se analisar um grande número de amostras em microscopia impõe limitações a estudos de longo prazo ou em grandes áreas geográficas. Por essa razão, outras ferramentas vêm sendo aprimoradas, tais como os sistemas de imagens do ambiente, que fornecem medições fundamentais e observações que melhoram a compreensão dos processos físicos e biológicos (Bi *et al.*, 2015). Nas últimas três décadas, várias tecnologias capazes de registrar imagens do zooplâncton foram desenvolvidas, incluindo sistemas portáteis de laboratório, como 'ZooScan' (Grosjean *et al.*, 2004; Gorsky *et al.*, 2010), bem como sistemas *in situ*, como o 'Vídeo Plankton Recorder' (VPR) (Tang *et al.*, 1998; Davis *et al.*, 2004), o 'Underwater Vision Profiler' (UVP), o 'Zooplankton Visualization System' (ZOOVIS), o

'Lightframe On-Site Keyspecies Investigation System' (LOKI), o 'Shadow Image Particle Profiling Evaluation Recorder' (SIPPER) (Luo *et al.*, 2004), o FlowCytobot (Sosik & Olson, 2007) o Flow Cytometer e Microscope Flow CAM (Blaschko *et al.*, 2005 ; Zarauz *et al.*, 2009) e o '*In Situ* Ichthyoplankton Imaging System' (ISIIS) (Bi *et al.*, 2015). A aquisição de imagens tem se mostrado uma ferramenta importante na oceanografia biológica, possibilitando realizar a distribuição dos organismos bentônicos, o monitoramento de cardumes de peixes, e também os estudos sobre a dinâmica da distribuição do plâncton. Mesmo com os avanços técnicos em *hardware*, *software* e equipamentos, permitir a implantação de sistemas de imagens sob diversas condições ambientais inerentes a águas mais turvas, incluindo sistemas estuarinos, e extrair informações úteis a partir do grande número de imagens adquiridas continua a ser um grande desafio.

Diversos instrumentos foram desenvolvidos ou utilizados para análises automatizadas de partículas através da detecção, contagem e medição de partículas individuais e organismos planctônicos (Christian *et al.*, 1998). A aquisição de imagens do zooplâncton com sistemas de câmeras vem acompanhada de vários problemas, dentre os quais a qualidade da imagem que deve ser suficiente para a identificação taxonômica. Para tal, as distâncias entre a câmera e o organismo devem ser curtas para reduzir a desfocagem do objeto em altas densidades de partículas e uma menor clareza de água (Schulz *et al.*, 2010). Um método eficiente para identificar com precisão e contabilizar organismos marinhos permitiria à comunidade científica acessar amostras armazenadas em coleções de plâncton, possibilitando a realização de novos projetos em larga escala espacial e temporal (Henzler *et al.*, 2010). Em adição, a classificação automatizada das imagens de plâncton impõe desafios específicos devido à proporção variável entre os grupos taxonômicos. Essa

proporção, ou biodiversidade, depende primariamente da localização geográfica, da sazonalidade, heterogeneidade morfológica, diferentes orientações, oclusão parcial e imagens únicas contendo mais do que uma única partícula (Álvarez *et al.*, 2012).

Para classificação e segmentação em imagens são utilizadas técnicas matemáticas que podem ser aplicadas às imagens obtidas (Gonzalez *et al.*, 2004). O processamento de imagens é uma tecnologia eficaz para várias aplicações na sociedade. Nessa categoria, existe processamento de imagens médicas, espaciais, subaquáticas e biológicas (Natchimuthu *et al.*, 2013).

A carência de especialistas em taxonomia no país é outro grande problema. Embora exista grande demanda de trabalhos e estudos nessa área, ainda são poucos os taxonomistas para cobrir uma diversidade grande de grupos. Hoje a classificação é feita por pessoas com diferentes graus de experiência prévia, cada um com uma *expertise* e um olhar singular ao organismo analisado. O custo dos métodos automáticos e semi-automáticos motiva qualquer pesquisador a desenvolver seu próprio método, visto que custos elevados da maioria dos equipamentos disponíveis no mercado inviabiliza a aquisição em certas regiões.

Os métodos automáticos, com seu alto potencial de aquisição, acumulam uma grande quantidade de imagens, potencializando o problema de análise e tratamento das imagens. A difícil distinção entre partículas inertes (debrís), organismos (vivos ou mortos), matéria orgânica particulada (MOP) como (fezes, carcaças) e plano de fundo (background) caracteriza o maior problema dos sistemas de captação de imagens em meio a águas turvas. Essa difícil distinção, resulta em dificuldade de recorte da região de interesse (ROI), em erros de contagem e erros de classificação. Uma imagem é formada por um conjunto de pixels, com tamanho variado, peculiar a cada imagem. Cada pixel em uma imagem em tons de cinza

("Gray Scale") recebe um valor em escala numérica que representa a tonalidade daquele pixel. O preto é representado pelo valor "0", normalmente o "*background*" ou fundo da imagem, e todas as demais partículas existentes recebem um valor entre 1 e o máximo de "255", ou branco puro.

Mesmo considerando os avanços no surgimento desses novos mecanismos de processamento de imagens e avanços em *Hardwares*, extrair informações úteis de um grande volume de imagens gerado por esses sistemas continua a ser um grande desafio. O mecanismo utilizado por diferentes sistemas para captura de informações é semelhante. No entanto, estes sistemas foram projetados para imagens adquiridas em ambientes sob condição controlada, em um laboratório ou ambiente *in situ* com águas cristalinas, onde é mais fácil obter imagens de alta qualidade. O processamento de imagens obtidas a partir de águas turvas, com alta concentração de debrís continua a ser um grande desafio (Bi *et al.* 2015). A alta turbidez da água reduz a visibilidade, o que limita a capacidade de adquirir imagens de alta qualidade. Esta condição de visibilidade torna difícil isolar de forma confiável as ROIs que contém objetos distintos do fundo. Em razão disso, o número de partículas em cada imagem resulta em superestimativa de ROIs e em inúmeros objetos segmentados erroneamente. O reconhecimento de copépodes entremeado em meio a sedimentos, debrís, dificulta o recorte da área de interesse e a segmentação correta dos organismos alvo. Assim, partículas em suspensão representam um ruído na imagem, decorrente da incerteza sobre se o pixel pertence ou não ao organismo.

As ferramentas de qualificação e quantificação de copépodes existentes ainda não têm aceitação plena. Um dos problemas está no reconhecimento do contorno irregular dos copépodes. A elaboração de algoritmos de contagem e

classificação de copépodes melhora o processo de análise dos dados e viabiliza o monitoramento ecológico e ambiental, quanto à dinâmica e densidade de organismos no ambiente marinho.

### **OBJETIVO GERAL**

Considerando-se as vantagens dos sistemas de imageamento frente às tradicionais técnicas, este estudo visou desenvolver um procedimento para processar as imagens relativamente “ruidosas” que são obtidas em águas com alta carga de partículas em suspensão. Ampliar a capacidade técnica brasileira no monitoramento em tempo real e *in situ* da dinâmica das partículas no meio marinho. Isso significa discernir a origem e a natureza das partículas em suspensão na água.

### **OBJETIVOS ESPECÍFICOS**

Desenvolver um algoritmo de separação de imagens aproveitáveis (arrasto da subida) de imagens não aproveitáveis (outros momentos). Desenvolver um algoritmo capaz de destacar as imagens segmentadas das não segmentada. Desenvolver um algoritmo de contagem e classificação automática de copépode.

## **MATERIAL E MÉTODOS**

Para elaboração dos algoritmos de processamento, foi selecionado um conjunto de imagens produzidas em áreas com distintas concentrações de sedimento em suspensão. As imagens foram obtidas em 27 de Novembro de 2015, dias após o rompimento da Barragem de Fundão em Mariana (MG), quando a pluma de rejeitos com alta carga de sedimentos em suspensão alcançou a foz do Rio Doce. Durante as operações do Navio Hidroceanográfico Vital de Oliveira (H39), da Diretoria de Hidrografia e Navegação (DHN), imagens foram registradas *in situ* por meio de um Light Frame On-Sight Key Species Investigation (Isitech LOKI). Os arrastos foram verticais com malha de 200  $\mu\text{m}$  e taxa de aquisição de imagens de 20 frames por segundo (Figura 1). Os arrastos foram realizados em dois pontos distintos ao longo do gradiente de concentração de partículas em suspensão. Um ponto situou-se mais próximo da foz do Rio Doce, com alta turbidez (>100ntu (unidades nefelométricas de turbidez)) e raso (20 metros de profundidade), e outro mais distante da foz, de águas menos turvas (<1 ntu (unidades nefelométricas de turbidez)) e mais profundo (100 metros).

### **Local de estudo**

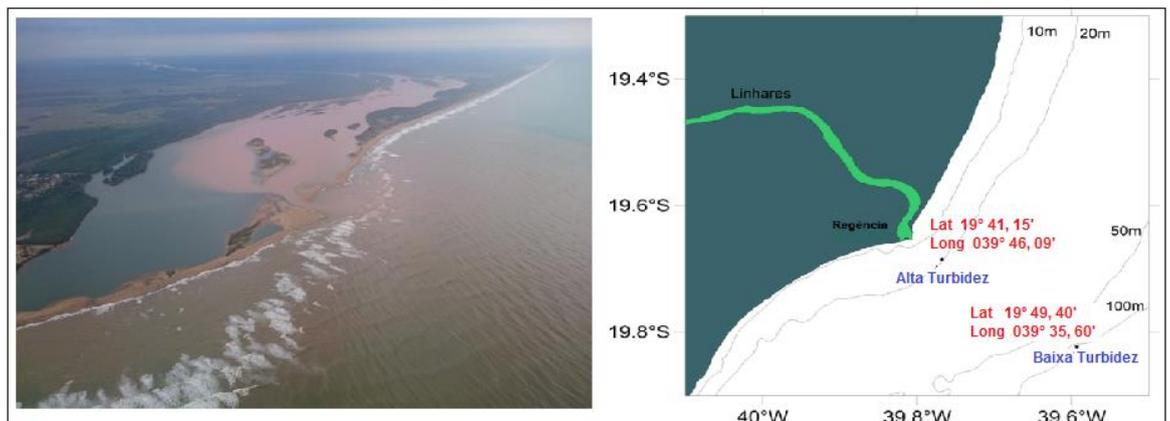


Figura 1 Foto aérea da área da Foz do Rio Doce logo após a chegada da pluma de rejeitos lamosos em Novembro de 2015, e mapa evidenciando os pontos de coleta, um com alta turbidez próxima a foz, e outro com menos turbidez mais distante.

O Lightframe On-sight Key Species Investigation (LOKI), é um sistema com 5 partes principais: (1) Uma rede de concentração, com boca de 0.28 m<sup>2</sup> e a abertura de malha de 200 µm, (2) um computador (3), um CTD com sensores de temperatura, pressão, oxigênio e fluorescência, (4) uma câmera Prosilica GC 1380H (AVT- Allied Vision Technologies, Canada) com a Pentax 2514-M e (5) uma bateria. A dimensão da imagem é de 1360 x 1024 pixels, e a resolução de 23 µm<sup>2</sup> por pixel, em 20 fotos por segundo. Possui uma unidade de LED de alta potência, sincronizado com o sinal de disparo de exposição da câmera, que permite um tempo de obturação rápida (fechamento) (55 µs) evitando o deslocamento com movimento, provocando a distorção da imagem. Em combinação com um canal de imagem de 4 mm de altura (comprimento = 31,3 mm, largura = 20,75 mm, volume = 2,6 cm<sup>3</sup>), levando a todas as partículas da imagem ficar em foco. Organismos e outras partículas são detectados e recortados do campo de visão da câmera, e posteriormente armazenados na unidade de estado sólido (SSD) do LOKI, onde mais tarde são acessados para novas análises.

O *software* de processamento das imagens do equipamento é o LOKI Browser. Assim que ligado o LOKI, como mostra a Figura 2, o software já inicia o registro das imagens instantaneamente, acumulando grande quantidade de imagens. Após esse registro, as imagens originais são pré-processadas automaticamente para recorte individualizado das partículas. Assim, são geradas imagens menores, cada uma contendo uma partícula em destaque da imagem original. Em locais de águas com alta turbidez, o software LOKI Browser encontra

dificuldades no pré-processamento dessas imagens e não executa bem a segmentação das partículas, assim como mantém em um mesmo set, imagens repetidas.



Figura 2 Utilização do equipamento LOKI em três momentos distintos, na imagem 1, sendo levado para fora do navio para então ser lançado verticalmente, como mostra a imagem 2, e no último estágio, imagem 3, o LOKI já no deck do navio, pronto para ter os dados descarregados em um computador.

### **Identificação das imagens de subida**

Do total de imagens acumuladas desde a inicialização do equipamento, apenas aquelas do fundo à superfície representam um arrasto na coluna de água e, portanto, são de interesse para caracterização e quantificação dos organismos planctônicos. A distinção e a separação dessas imagens de interesse foram realizadas com algoritmos elaborados em linguagem Python®, a partir dos parâmetros obtidos no CTD. Em síntese, cada imagem adquirida é vinculada ao momento de sua aquisição. Assim, o algoritmo de separação das imagens de interesse identifica o intervalo entre início e final do arrasto para selecionar corretamente o subconjunto de interesse. A identificação desse intervalo parte da premissa de que a pressão aumenta e a temperatura diminui de acordo com o aumento na profundidade. Assim, o início do arrasto pode ser definido como o instante final da máxima pressão e mínima temperatura (Figura 3). De modo análogo, para identificar o final do arrasto, quando se espera que o equipamento

esteja próximo da superfície. Além dos parâmetros de menor pressão e maior temperatura foi utilizado também a densidade da água.

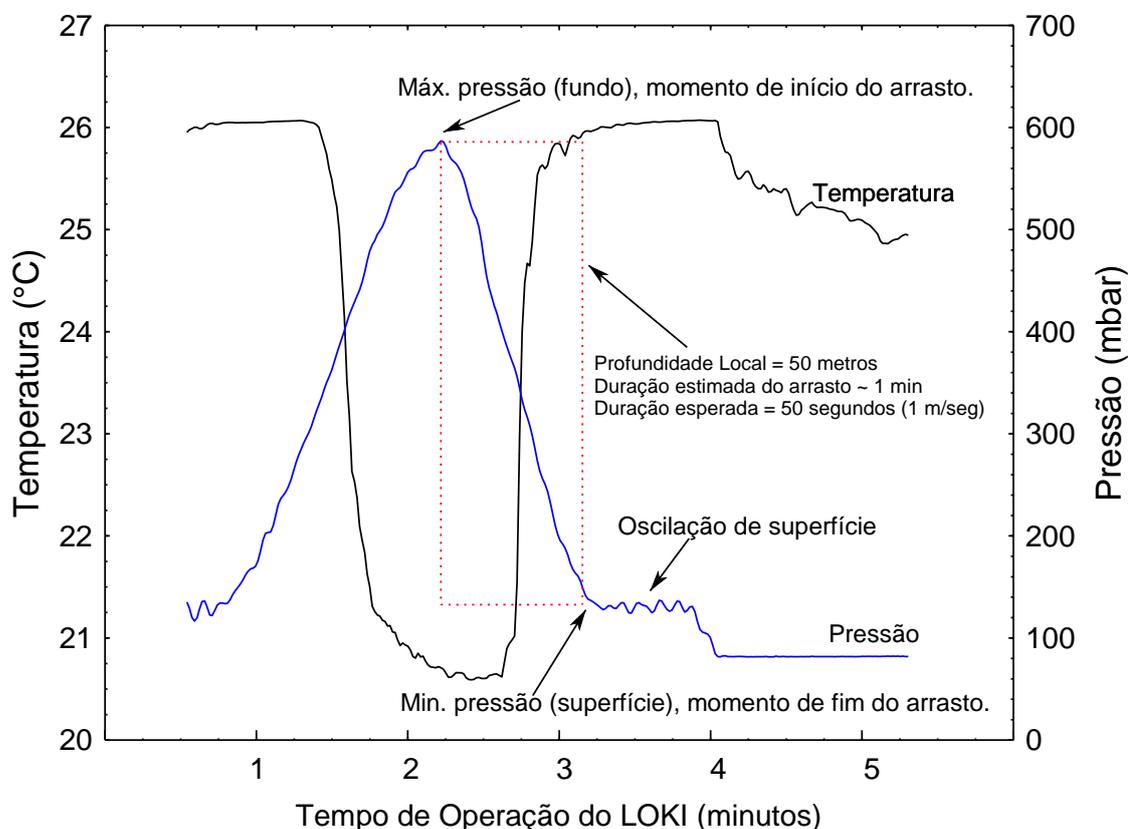


Figura 3 Exemplo de perfil de variação temporal da Temperatura (°C) e Pressão (mbar) da água durante um arrasto com o LOKI. Em vermelho, destaque para o momento da subida do LOKI desde próximo do fundo (máxima pressão e mínima temperatura) até próximo da superfície.

Uma vez separadas as imagens do intervalo de subida da rede, procedeu-se à separação de imagens de organismos e partículas isolados da imagem original (segmentadas). Foram selecionadas aleatoriamente 1000 imagens de cada ponto para teste dos algoritmos em desenvolvimento. Os algoritmos de pré-processamento foram elaborados de modo a comparar as dimensões x e y das imagens com um referencial pré-estabelecido. Uma imagem não segmentada possui um tamanho conhecido de 1360 x 1024 pixels. Uma imagem mal segmentada frequentemente possui apenas uma das dimensões, x ou y, diferente da não segmentada. Por outro

lado, imagens segmentadas são recortes de uma ROI e têm dimensões reduzidas. Imagens de dimensões inferiores ao máximo do campo visual da câmera em x e y significam necessariamente recortes de objetos, que podem representar desde distintos organismos a debris, bolhas e artefatos (ruídos).

Com intuito de realizar a distinção automatizada de imagens de organismos em meio a imagens de ruídos, 50 ROIs de copépodes foram selecionadas aleatoriamente e processadas em 16 diferentes algoritmos de remoção de fundo (Tabela 1). Como controle, as mesmas ROIs foram processadas manualmente, a partir do contorno do corpo do copépoda, feito com auxílio de uma taxonomista experiente, a pesquisadora Wanda Ribas do Instituto de Estudos do mar Almirante Paulo Moreira. O processamento para remoção de fundo foi baseado na definição do centroide do corpo do copépode, seguido da reconstrução gradual do contorno a partir das proporções na morfometria. Em outras palavras, previamente foi construído um contorno padrão para os copépodes a partir da tomada de medidas (explicado adiante) de prossoma, urossoma e antenas. Assim, a reconstrução do contorno pode ser baseada na busca de um limiar de tom do pixel representativo do centroide do corpo do animal (Figura 4) (Fase A). A matriz de reconstrução do corpo baseada nos 16 algoritmos já existentes no programa Image J, foi comparada com a matriz de medição manual (controle) por meio da distância Euclidiana. A análise de agrupamento apontou dentre os métodos aquele que resultava em medidas mais similares ao controle (manual).

Tabela 1 Teste realizado com os 16 métodos de eliminação do fundo da imagem, com observação de presença e ausência de alguns parâmetros, excluídos ou preservados no teste.

<b>Métodos</b>	<b>Copépode</b>	<b>1 Antena</b>	<b>2 Antenas</b>	<b>Cont.Cop.</b>	<b>Debrís</b>
<b>Default</b>	X	x			X
<b>Huang</b>	X		X	X	X
<b>Intermodes</b>	X	X			X
<b>IsoData</b>	X	X			X
<b>Li</b>	X		X	X	X
<b>MaxEntropy</b>	X		X	X	X
<b>Mean</b>	X		X	X	X
<b>MinError</b>	X		X	X	X
<b>Minimum</b>	X				
<b>Moments</b>	X	X			X
<b>Otsu</b>	X	X			X
<b>Percentile</b>					X
<b>RenyiEntropy</b>	X		X	X	X
<b>Shanbhag</b>	X		X	X	X
<b>Triangle</b>	X		X	X	X
<b>Yen</b>	X		X	X	X

Após o processamento de fundo, a imagem foi convertida em binária, com partículas bem destacadas (Figura 4) (Fase B) e distintos tamanhos. Por fim, foi realizada a contagem e medição das partículas, com intuito de definir o limiar de tamanho do copépode (Fase C).

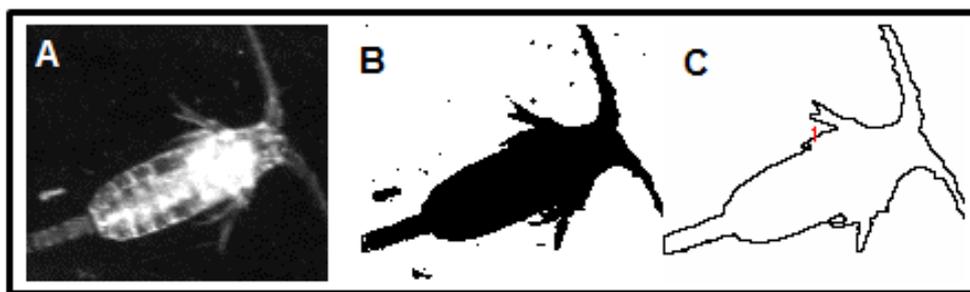


Figura 4 Fases do tratamento das imagens. A; imagem original, recortada automaticamente pelo sistema do LOKI, utilizada para definição do tom limiar do pixel. B; imagem sem fundo (fundo branco), com destaque para todas as partículas (debris+copépode). C; imagem do contorno do copépode destacado das demais partículas pela medição.

Para definição do limiar de tamanho dos copépodes e distinção dos debris de fundo, foram selecionadas aleatoriamente 100 imagens cuja medição dos parâmetros morfométricos mais significativos pudessem ser usados na definição do contorno do corpo (Tabela 2). A medição foi realizada visualmente no *software ImageJ* e a matriz de dados foi analisada no *Software Statistica 8.0*.

Para contagem e separação dos copépodes, foram elaborados 4 algoritmos com diferentes etapas sequenciais de tratamento, mas amparados igualmente na morfometria descrita acima e no processamento de fundo. Como controle, foi realizada a contagem visual por taxonomista especializado. Esses algoritmos foram elaborados com base no mesmo conjunto de 1000 imagens selecionadas acima para reduzir a variância interna dos dados. Com intuito de avaliar a constância e linearidade dos resultados obtidos em amostras pequenas (<1000 imagens), os algoritmos foram testados novamente com 3 subconjuntos de 100 imagens aleatórias de cada ponto de coleta. Diferenças no total de copépodes contados em cada subconjunto (pseudoréplicas) por cada algoritmo foram avaliadas por meio de uma Análise de Agrupamento realizada com o *Software Statistica 8.0* aplicada sobre a matriz de dissimilaridade Bray-Curtis entre os métodos. Abaixo se encontra a descrição dos algoritmos de contagem construídos em ImageJ.

**Algoritmo 1 (Sem ajuste de exposição)** – Sobre a imagem foi realizado um processamento de fundo utilizando o método automático estatisticamente mais semelhante ao manual. Em seguida, é feita a contagem direta de partículas, com limite mínimo para o tamanho dos copépodes (Figura 5).

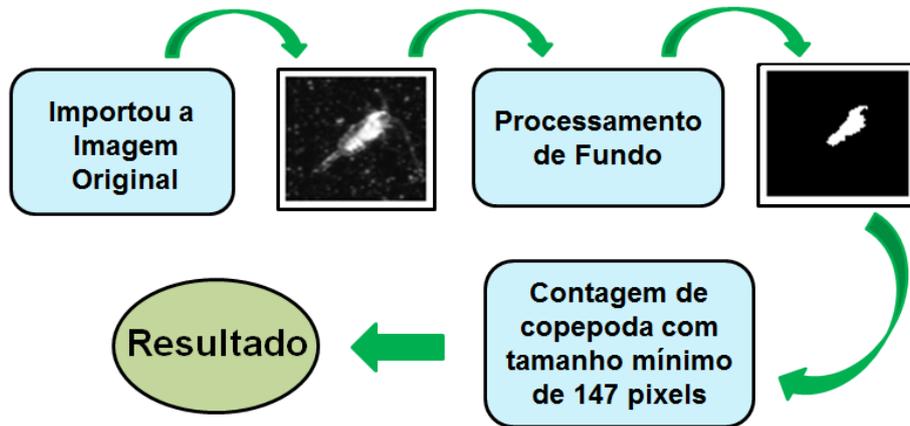


Figura 5 Etapas de tratamento de imagem utilizando o algoritmo 1, para contagem e classificação dos copépodes.

**Algoritmo 2 (Superexposição)** – Cada valor de tom de pixel foi multiplicado por cinco, com intuito de produzir uma imagem superexposta e destacar as partículas. Em seguida, sobre a imagem foi realizado um processamento de fundo utilizando o método automático estatisticamente mais semelhante ao manual. Por fim, foi feita a contagem automática, com limite mínimo para o tamanho dos copépodes (Figura 6).

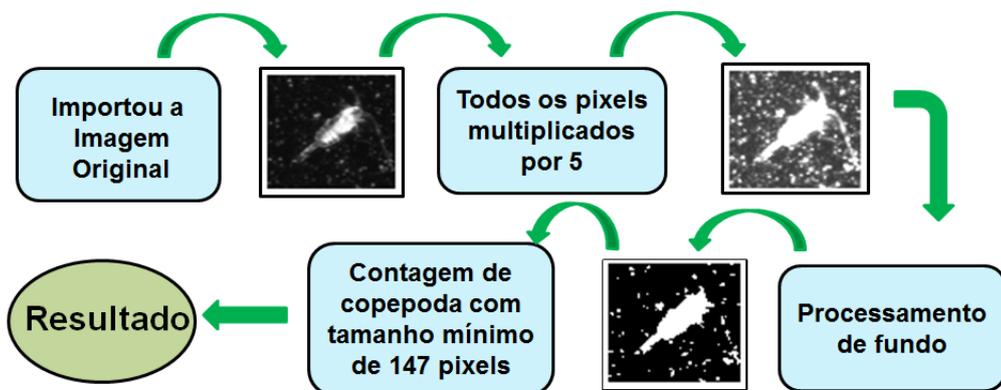


Figura 6 Etapas de tratamento de imagem utilizando o algoritmo 2, para contagem e classificação dos copépodes.

**Algoritmo 3 (Sub-exposição)** – Cada valor de tom de pixel foi subtraído um valor de 25, com intuito de produzir uma imagem subexposta e destacar apenas o copépode. Em seguida, sobre a imagem foi realizado um processamento de fundo utilizando o método automático estatisticamente mais semelhante ao manual. Por fim, foi feita a contagem automática, com limite mínimo para o tamanho do copépode (Figura 7).

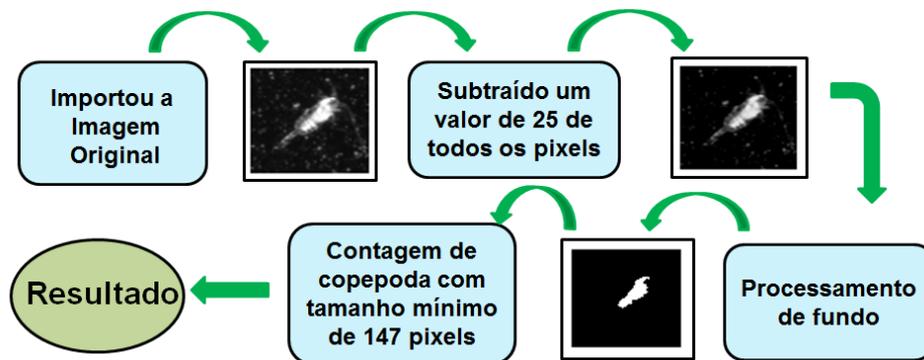


Figura 7 Etapas de tratamento de imagem utilizando o algoritmo 3, para contagem e classificação dos copépodes.

**Algoritmo 4 (Agrupamento Estatístico de Região – *Statistical Region Merging SRM*)** – Foi realizada a segmentação das partículas com base na definição de regiões (Q) de tons similares. Para reduzir a perda de partículas como artefato negativo do processamento, o número de regiões (Q) foi maximizado para 25. Em seguida, sobre a imagem foi realizado um processamento de fundo utilizando o método automático estatisticamente mais semelhante ao manual. Por fim, foi feita a contagem automática, com limite mínimo para o tamanho do copépoda (Figura 8).

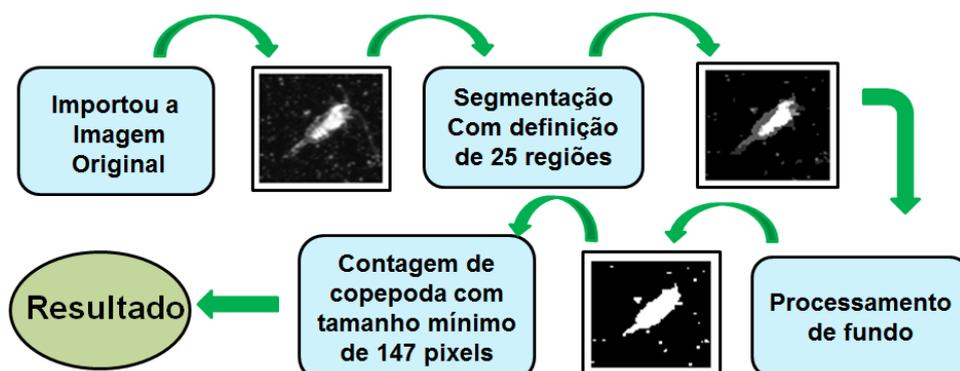


Figura 8 Etapas de tratamento de imagem utilizando o algoritmo 4, para contagem e classificação dos copépodos.

Abaixo, encontra-se uma breve descrição das etapas (Figura ).

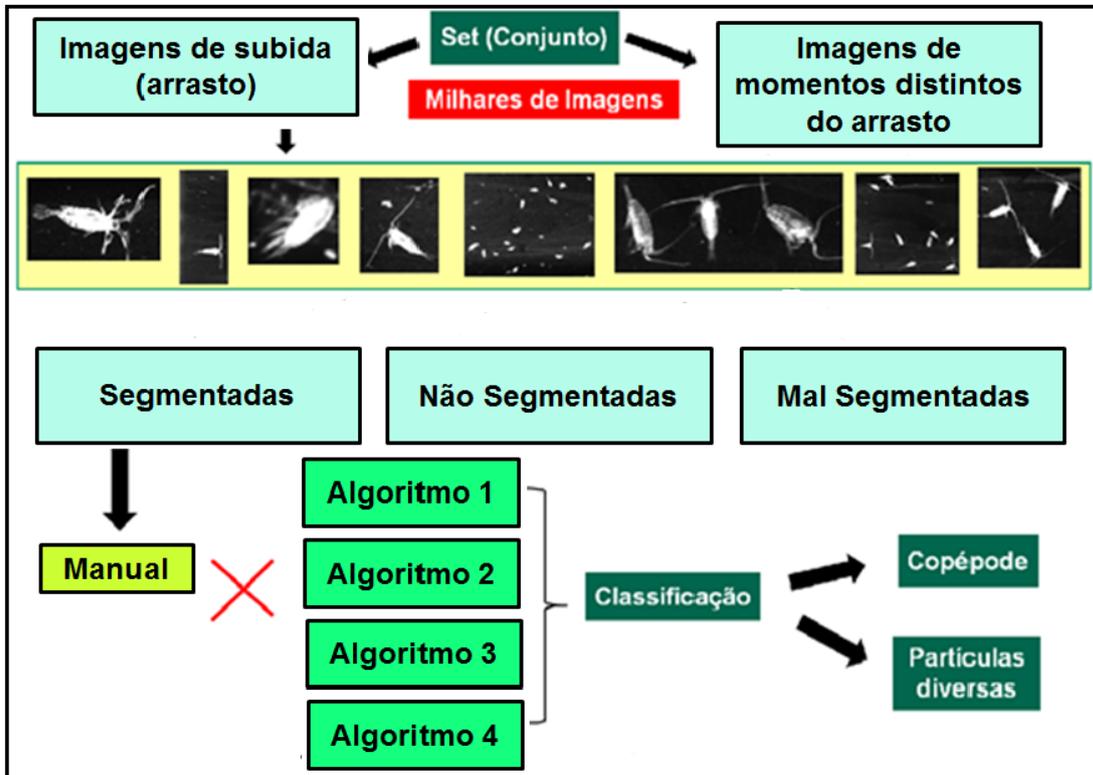


Figura 9 Tratamento da imagem, primeiro a distinção entre imagens aproveitáveis (subida) e imagens não aproveitáveis (outros momentos). Após essa distinção, a correção de alguns problemas gerados em imagens adquiridas em águas turvas, como imagens não segmentadas, mal segmentadas, etc. Em seguida com essas imagens pré-selecionadas e corrigidas, testou-se 4 algoritmos de contagem e classificação do copépode pela área<sup>2</sup>, e então comparado ao tratamento realizado manualmente.

## RESULTADOS E DISCUSSÃO

No total, foram produzidas 40.375 imagens no ponto com alta turbidez e 21.664 no ponto de baixa turbidez. Consideradas as distintas profundidades e tempos de operação do equipamento em cada ponto, mais imagens foram obtidas em águas mais rasas (20 metros e 8 minutos (2,5 m/min), ponto de alta turbidez) do que em águas mais profundas (100 metros e 5 minutos (20 m/min), ponto de baixa turbidez). A uma taxa de 20 imagens a cada segundo, seriam esperadas respectivamente 9.600 e 6.000 imagens brutas não-segmentadas nesses pontos. Padronizando-se o total de imagens pelo tempo de operação do equipamento (40.357 imagens / 8 minutos e 21.664 imagens / 5 minutos), resultaria em totais de 5044 e 4.333 imagens/min nos pontos de alta e baixa turbidez (Figura ).

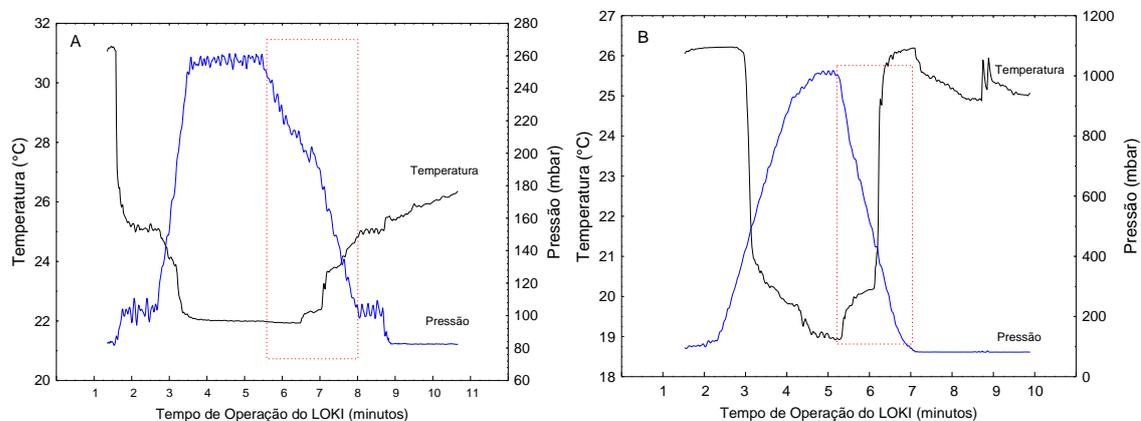


Figura 10 Perfil de variação temporal da Temperatura (°C) e Pressão (mbar) da água durante um arrasto com o LOKI no ponto com alta turbidez, e baixa turbidez. Destacado em vermelho tem-se o momento em que o equipamento registrou máxima e mínima das duas variáveis, e posteriormente associada este momento às imagens.

Considerando os fatores tempo de operação, profundidade local e turbidez, os resultados apontam para a turbidez da água como fator mais efetivo na geração de múltiplos ROIs. Resultado similar foi observado no trabalho de Bi *et al.*, (2015),

que apontou vantagens nos métodos de tratamento de imagem de plâncton projetados para condições controladas, de águas menos turvas, com imagens mais consistentes e de alta qualidade. A velocidade de um arrasto vertical, apesar de sujeita às condições de mar e do guincho, deve ser mantida o mais próximo possível de 1 metro por segundo (Boltovskoy, 1981). Os vinte metros de profundidade do ponto mais turvo, ignorando-se a distância de segurança do fundo quando da operação do equipamento, e a taxa de aquisição de 20 frames por segundo deveriam produzir cerca de 400 imagens brutas ( $20 \text{ frames/seg} * 20 \text{ segundos}$ ) ao final da subida. Foram obtidas no ponto mais turvo 5044 imagens a cada minuto, o equivalente a 1681 imagens em 20 segundos (1/3 de minuto) de arrasto. Esse total implicou em cerca de 4,2 partículas (ROIs) extraídas de cada imagem bruta original. Por outro lado, na estação de maior profundidade e menor turbidez, respeitadas as taxas de aquisição e velocidade de arrasto, haveria 2000 imagens brutas após 100 metros de arrasto vertical. Foram obtidas 4333 imagens a cada minuto, o equivalente a 7222 em 100 segundos de arrasto, quando eram esperados 2000 ( $20 \text{ frames/seg} * 100 \text{ segundos}$ ). Esse total implicou em 3,6 partículas (ROIs) segmentadas de cada imagem bruta original.

As imagens adquiridas para a realização do presente trabalho foram adquiridas em local com alta turbidez, provocado por aporte de sedimentos lamosos, o que contribuiu para o espalhamento da luz na imagem, dificultado a tomada de decisão do programa na segmentação do organismo. Em seu trabalho, (Bi *et al.*, 2015) ressaltaram que o processamento de imagens obtidas a partir de águas turvas continua a ser um grande desafio. Em primeiro lugar, águas de alta turbidez atenua rapidamente e espalha a luz, reduzindo a visibilidade, o que limita a capacidade de

adquirir imagens de alta qualidade, por exemplo, em condições não uniforme de iluminação de fundo e de baixo contraste (Bi *et al.*, 2015).

O teste de avaliação dos algoritmos de pré-processamento revelou eficiência na identificação das imagens segmentadas (combinadas bem e mal segmentadas) e não-segmentadas. Do total de 1000 imagens de cada ponto, os algoritmos resultaram em 100% de acerto entre segmentadas e não-segmentadas (Figura 11).

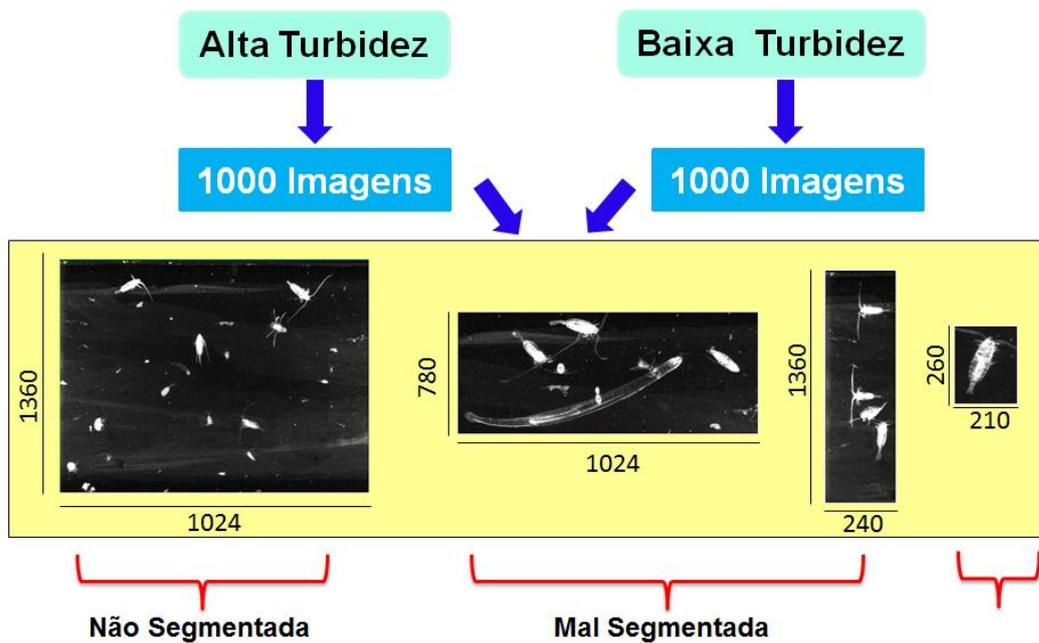


Figura 11 Diferentes dimensões que as imagens se apresentam, podendo ter tamanho máximo, ou parcial, e ter o objeto de interesse recortado perfeitamente ou mal recortado.

No entanto, houve erros significativos na separação das imagens mal e bem segmentadas. Sob condições de alta turbidez, o algoritmo superestimou a quantidade de imagens mal segmentadas e subestimou o total de imagens bem segmentadas (Tabela 2). Os resultados foram opostos sob condições de baixa turbidez.

Tabela 2 Testes dos algoritmos de pré-processamento das imagens, separando dentre o total e imagens, segmentadas, mal segmentadas e não segmentadas.

<b>Imagens</b>	<b>Contagem Manual</b>	<b>Costeira (Alta Turbidez)</b>	<b>Contagem Manual</b>	<b>Oceânica (Baixa turbidez)</b>
<b>A) Segmentadas</b>	961	961	993	993
<b>B) Mal segmentadas</b>	190	941	230	7
<b>C) Bem segmentadas (ROI)</b>	771	20	763	986
<b>D) Não segmentadas</b>	39	39	7	7

Um dos principais problemas na identificação das imagens decorreu da quantidade de partículas do fundo. Álvarez *et al.* (2012) ressaltaram que a presença de detritos nas amostras é, no entanto, uma característica comum dos dispositivos de imagem de plâncton. Quaisquer partículas na imagem representaram ruídos e as imagens do ponto de alta turbidez possuíam, a priori, um maior ruído. Esse ruído pode ser percebido na alta frequência de imagens menores que 10 Kilo bites, que representam partículas muito pequenas (Figura 12). Essas partículas predominaram entre as ROIs, com mais de 60% de todas as 1000 imagens selecionadas. Em contrapartida, a baixa turbidez representou pequena quantidade de micropartículas e, por analogia, de ruídos nas imagens. A maior frequência está nas imagens de 10Kb a 50Kb, que representam material particulado em suspensão.

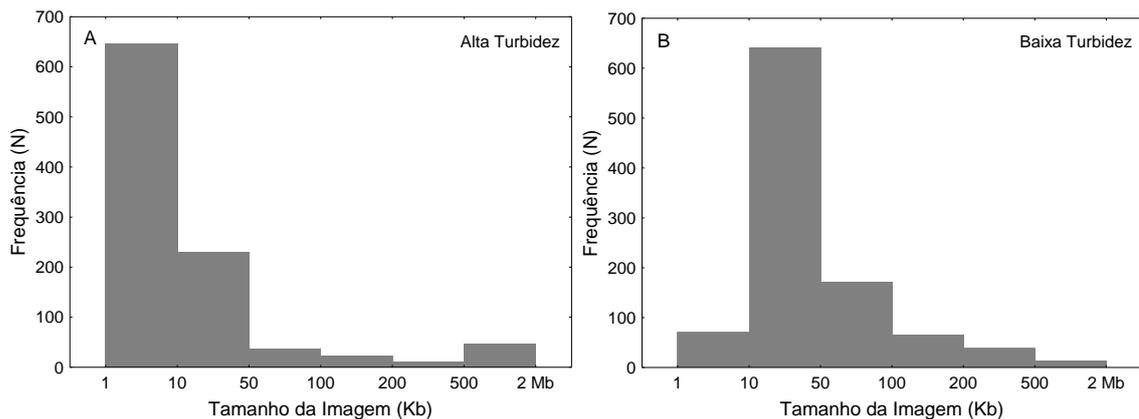


Figura 12 Distribuição dos ROIs por tamanho da imagem, observado a maior frequência das imagens procedentes de águas mais turvas entre 1 e 10 Kb. Em baixa turbidez, a maior ocorrência das imagens foi com dimensão entre 10 e 50 Kb.

Schulz *et al.* (2010) destacaram problemas na aquisição de imagens de zooplâncton com sistemas de câmeras, dentre os quais a baixa qualidade da imagem utilizada para identificação taxonômica. A presença de quantidade de material disperso na água que não seja objeto alvo potencializa a dificuldade na distinção das partículas de interesse (Bi *et al.*, 2014).

Os objetos dentro de cada imagem foram separados do fundo através de um processo chamado segmentação (Figura 13). Surovy *et al.* (2014) descreveram a dificuldade encontrada na segmentação de imagens de vegetais, com baixo contraste, podendo ter as raízes dos vegetais fragmentadas e terem a medição e a contagem subestimados.

No presente estudo, puderam-se observar problemas de segmentação no pré-tratamento das imagens de pontos com alta turbidez. Nesses pontos, a segmentação de um único copépode em meio a partículas do sedimento em suspensão foi concluída com êxito parcial pelo programa do LOKI. A segmentação é a parte fundamental no processamento de imagens (Natchimuthu *et al.*, 2013). Elaborar um algoritmo que separa o objeto de interesse do fundo facilitou os testes

com os algoritmos de contagem dos organismos por tamanho da área, e foi fundamental para a realização deste trabalho.

No entanto Castro *et al.* (2007) e Escribano *et al.* (2009) procederam à análise detalhada da distribuição vertical dos organismos por meio da utilização do LOKI, sem os problemas de segmentação. Segundo Sieracki *et al.* (2007), os sistemas de imagens são capazes de registrar a distribuição do plâncton em escala fina. No entanto, as falhas apontadas no presente trabalho indicam a necessidade de algoritmos adicionais de aperfeiçoamento dos sistemas atuais quanto a contagem e classificação de copépodes em ambientes de alta turbidez.

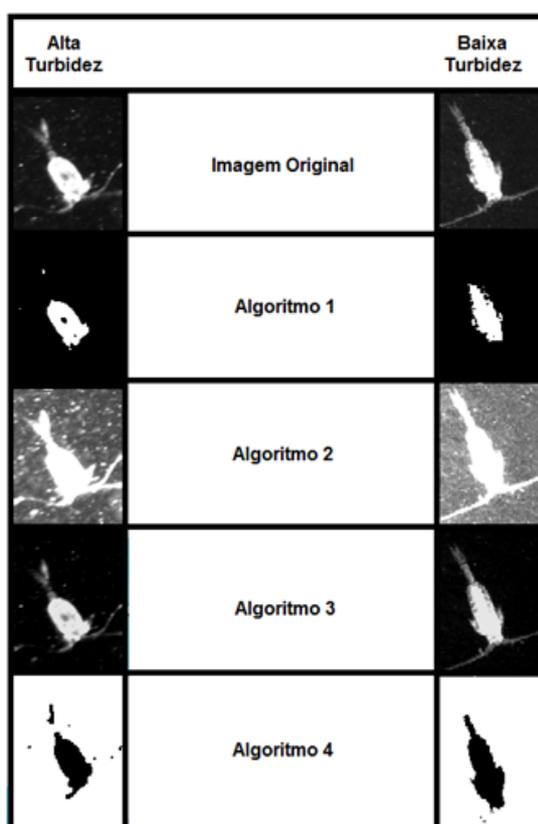


Figura 13 Comparação entre dois pontos com condições distintas de turbidez, nos tratamentos com utilização dos quatro algoritmos elaborados neste trabalho.

A escolha do algoritmo de redução de ruído automático menos impactante na estimativa do centroide do copépode revelou que a busca do limiar foi mais eficiente quando a distribuição da frequência dos tons de pixels, após cálculo da média móvel

de janela equivalente a 3 pixels, resultou bimodal (algoritmo do mínimo). O limiar aceito se situou no ponto mínimo entre as duas modas e destacou quase que exclusivamente o prossoma (Figura 14)

Schmid *et al.* (2016) ressaltaram que na taxonomia tradicional de copépodes, o comprimento do prossoma é frequentemente usado para identificar espécies em vez do comprimento total.

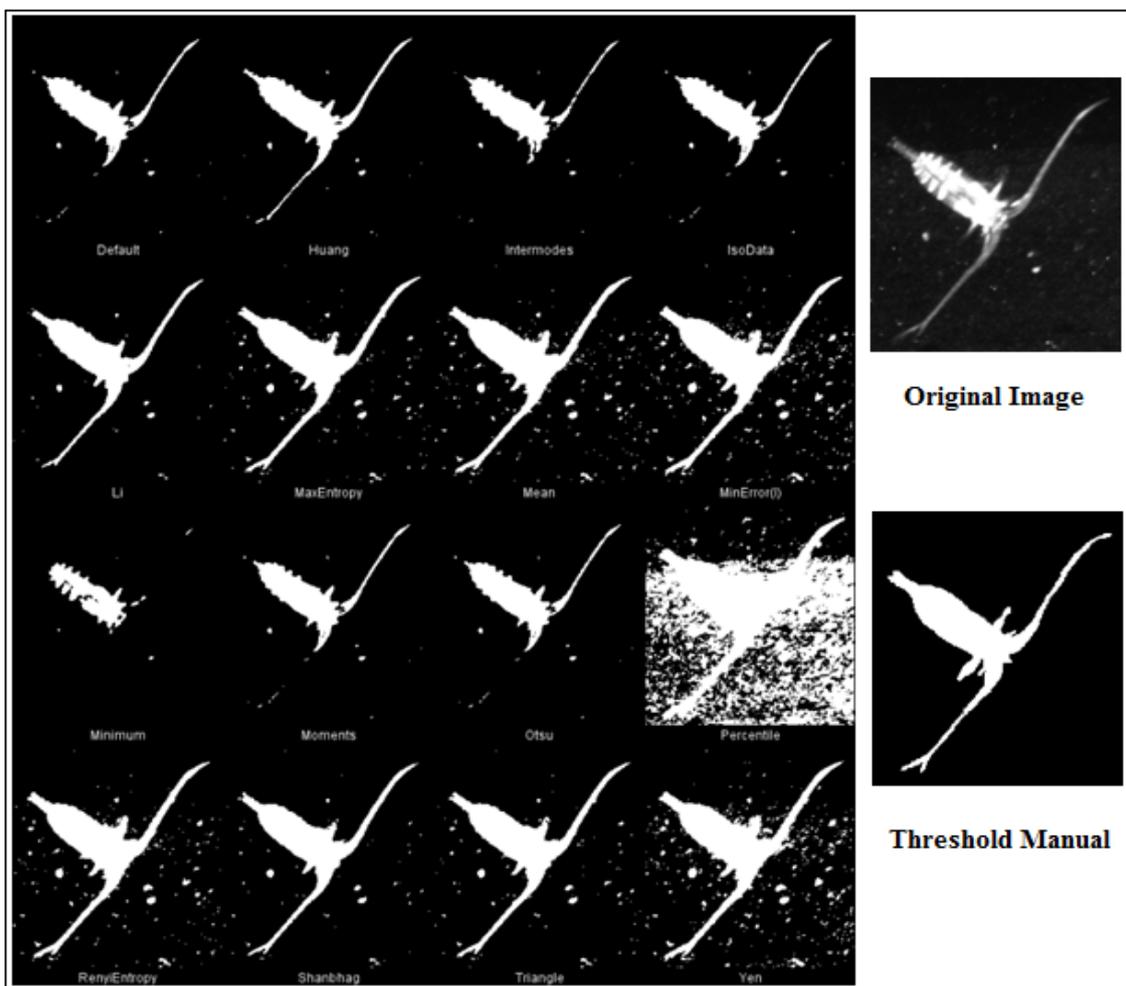


Figura 14 Desempenho dos 16 métodos de threshold existente no programa ImageJ, comparado com o método manual, e avaliado quanto a preservação da partícula central e remoção dos ruídos (píxels) do fundo. Ao lado tem-se a imagem original.

O método de Otsu é descrito por Sahoo *et al.* (1988) como o melhor método de *threshold* usado como limiar global para separação de partículas ou objetos do

fundo, descrito também nos trabalhos de Hongmin *et al.* (2014) e Ping-Sung *et al.* (2001). Entretanto, para o tratamento de imagens de copépodes no presente estudo, a escolha do método de *threshold* 'minimum' atendeu melhor às condições definidas no método usado no trabalho quanto à preservação da partícula central e limpeza dos pixels ruidosos de fundo. Imagens processadas por esse método resultaram em medidas morfométricas mais similares àquelas feitas manualmente e têm maior potencial de reconstrução de contorno com menor ruído. Surovy *et al.* (2014) concluíram que o grau de precisão na estimativa de tamanho de área de um objeto analisado, em particular nas imagens de estrutura vegetal, como o trabalho dele, se deve ao acerto na escolha do *threshold* (Figura 15).

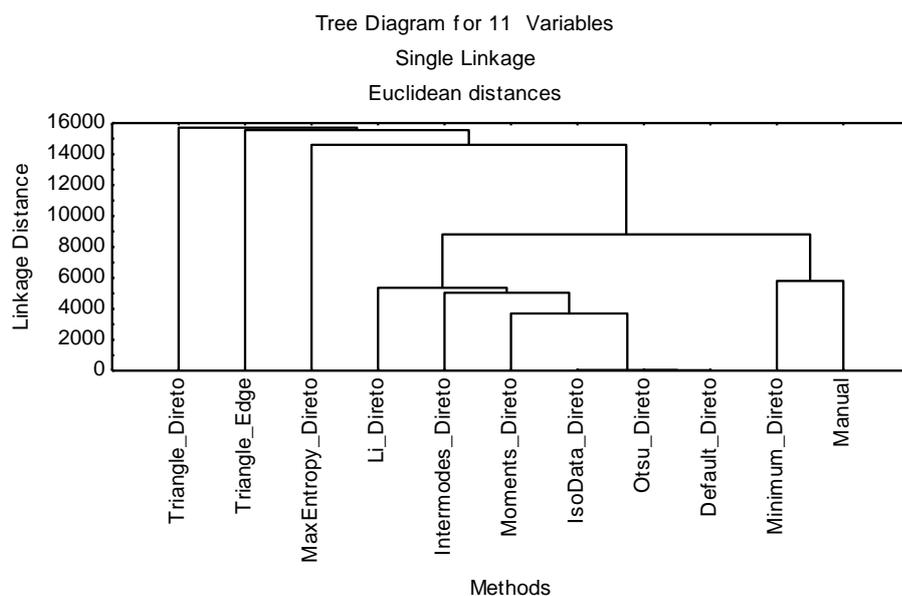


Figura 15 Dendrograma de análise de grupamento, demonstrando que o método minimum é o mais próximo ao manual, quanto limpeza do fundo e preservação da partícula central do copépode.

De acordo com Bi *et al.* (2015), é difícil usar um único valor limiar global para converter as imagens de escala de cinza em imagens binárias. Devido à iluminação não uniforme sobre o campo de visão, é difícil impor um limiar acima do qual todos os pixels tornam-se brancos e abaixo do qual tornam-se pretos.

O valor médio para a área corporal de um copépode, excluídas cerdas, equivaleu a  $187.795 \mu\text{m}^2$  ( $\approx 355 \text{ pixels}^2$ ). O alto desvio padrão ( $109.503 \mu\text{m}^2$ ) demonstrou a ampla variação decorrente do desenvolvimento ontogenético. No âmbito do presente estudo, o valor mínimo da área corporal do copépode foi estabelecido estatisticamente em  $77.763 \mu\text{m}^2$  ( $147 \text{ pixels}^2$ ) (Figura 16). Qualquer partícula acima desse limiar tem potencial para ser um copépode e não pode ser descartado sem passar pelo algoritmo de reconhecimento.

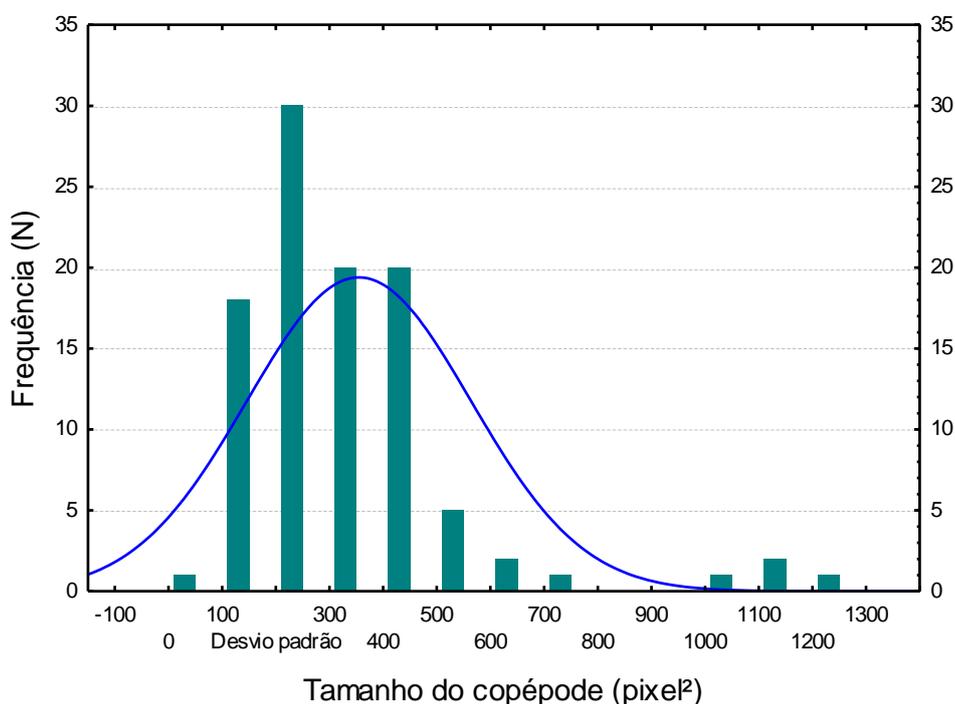


Figura 5 Estabelecimento do tamanho mínimo do copépode através da medição de 1000 organismos. O tamanho mínimo foi adquirido através da subtração do valor médio pelo desvio padrão.

O teste de contagem de copépodes com base nos quatro algoritmos desenvolvidos revelou percentuais de acerto entre 90% e 507% (Tabela 3). Natchimuthu *et al.* (2013) encontraram 97% de precisão na contagem e classificação de microalgas pelo tamanho e forma. O presente trabalho elaborou e testou quatro algoritmos que realizasse a contagem e separação dos organismos levando em consideração o tamanho da área, e foi encontrado em um algoritmo elaborado para contagem e classificação automática uma similaridade de 96% com o algoritmo

manual, quanto a contagem dos copépodes em dois conjuntos de imagens, um conjunto em local de águas com alta turbidez e outro local de lugar com baixa turbidez. Schmid *et al.* (2016) obtiveram em seus testes perda dos copépodes menores, mais translúcidos, provavelmente devido ao algoritmo que detecta a partícula ao passar pelo fluxo da câmera.

Os maiores percentuais de acerto ocorreram nas imagens de alta turbidez e sugerem que métodos de subexposição têm potencial para reduzir os ruídos que as partículas de sedimento podem representar. O método mais eficiente em nossos resultados envolveu a erosão da imagem do copépode ao mínimo necessário para se obter uma distribuição bimodal, conforme resultados apontados acima. Forest *et al.* (2012) encontraram uma diferença notável entre o biovolume estimado nas análises do ZooScan e de estimativas morfométricas das contagens manuais.

Apenas partículas de tamanho equivalente a um copépode pequeno foram capazes de permanecer na imagem após o processamento pela média móvel e interferir na contagem. Similarmente, o algoritmo de subexposição, com subtração de valor fixo (25) no tom do pixel, resultou em contagens próximas, embora subestimadas. Todos os métodos aplicados para as imagens com pequenos ruídos, da estação de baixa turbidez, resultaram em superestimativa de contagem. Os algoritmos 2 e 4, que partem da premissa de aumento na exposição da imagem, tenderam a destacar as partículas de fundo, particularmente em águas de alta turbidez. Schulz *et al.* (2010) encontraram em seus resultados de contagem de partículas, cerca de 15% de detritos e partículas indesejáveis.

Tabela 3 Número de partículas obtidas individualmente por algoritmo manual e automaticamente por quatro algoritmos elaborados no programa ImageJ. Teste realizado em um conjunto aleatório de 1000 imagens. Alta Turbidez, Baixa Turbidez.

<b>Contagem de Copépodes</b>				
<b>Algoritmo</b>	Alta Turbidez	Controle	Baixa Turbidez	Controle
<b>Manual</b>	678	100%	1038	100%
<b>Algoritmo 1</b>	644	95%	1176	113%
<b>Algoritmo 2</b>	3441	507%	1782	171%
<b>Algoritmo 3</b>	612	90%	1178	113%
<b>Algoritmo 4</b>	2423	357%	1937	186%

Como esperado, os algoritmos erosivos resultaram em contagens mais próximas do controle para amostras de alta turbidez. No entanto, esses mesmos métodos tenderam a superestimar de forma significativa o total de copépodes. Esse “erro” foi particularmente percebido em imagens de baixa turbidez e sugere que seja dependente da quantidade de ruído nas imagens (Tabela 4). O número de organismos como copépodes classificados pelo procedimento semi-automatizado tende a ser menor do que as contagens manuais (~10 mais baixo) (Bi *et al* 2014). O exoesqueleto de quitina desses crustáceos resulta em contorno bem identificável do corpo, porém os detalhes geralmente não são suficientemente claros, provavelmente devido ao tamanho pequeno ou à velocidade do obturador (Schulz *et al.*, 2010).

Tabela 4 Número de partículas obtidas individualmente por algoritmo manual e automaticamente por quatro algoritmos elaborados no programa ImageJ. Teste realizado em um conjunto de 100 imagens aleatórias, em dois pontos distintos, com AT = alta turbidez e outro com BT = baixa turbidez.

<b>Teste</b>	<b>AT_1</b>	<b>AT_2</b>	<b>AT_3</b>	<b>BT_1</b>	<b>BT_2</b>	<b>BT_3</b>
<b>Manual</b>	48	89	71	121	128	137
<b>Algoritmo 1</b>	78	91	71	127	135	148
<b>Algoritmo 2</b>	402	651	703	192	191	215
<b>Algoritmo 3</b>	75	89	67	127	135	148
<b>Algoritmo 4</b>	375	425	306	201	209	211

A avaliação de constância e linearidade para amostras pequenas resultou em percentuais de similaridade de 30% a 95% entre os métodos.

Para amostras do ponto com alta turbidez, foram feitas análises de agrupamento utilizando a dissimilaridade com coeficiente de Bray-curtis, os algoritmos 1 e 3 foram considerados estatisticamente semelhantes entre si por análise de cluster. Os algoritmos mais próximo ao tratamento manual, com grau de similaridade de mais de 95%, e os algoritmos 2 e 4 apresentaram semelhança entre si e similaridade de 30% com o tratamento controle (manual), com erro de 70% na contagem dos organismos. Nos mesmos testes realizados no ponto com baixa turbidez, o grau de dissimilaridade entre os testes foram de quase 20%, considerando uma semelhança de todos os algoritmos de quase 80% do tratamento controle, porém como observado no ponto com alta turbidez, o desempenho dos algoritmos 1 e 3 foram considerados mais próximos do controle, com grau de semelhança de mais de 95%, como mostram os dendrogramas da Figura 17.

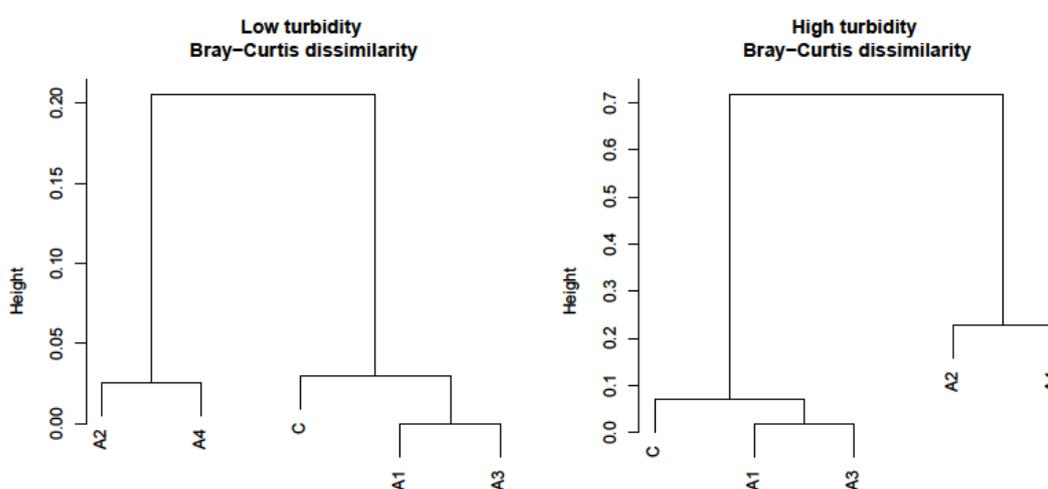


Figura 6 Dendrogramas de dissimilaridade testada entre quatro algoritmos elaborados pra contagem automática de copépode, comparado com tratamento realizado manualmente. Testes realizados em dois pontos distintos, com alta turbidez e outro ponto com baixa turbidez, mais oceânico.

Sistemas automatizados e semi-automatizados de imageamento do zooplâncton têm sido procurados há muito tempo como parte de uma abordagem

moderna para monitorar o meio ambiente marinho (Basedow *et al.*, 2013). A necessidade de sensores capazes de fornecer dados de abundância e biomassa com um espaço de tempo e alta resolução gerou um esforço crescente para viabilizar o fosso entre diferentes métodos de amostragem contemporânea em ciências do mar (Basedow *et al.*, 2013).

Os resultados obtidos no presente estudo ressaltam a necessidade de desenvolvimento de novas tecnologias de processamento das imagens de plâncton, assim como aperfeiçoamento dos algoritmos existentes.

Hu e Davis (2005) apontaram a tendência crescente de sistemas de imageamento óptico de biota marinha. Esses novos métodos de identificação automática são necessários para melhorar a precisão da classificação.

Os resultados estabelecidos por Alvarez *et al.* (2012) demonstraram que é possível automatizar o processo de classificação das partículas, eliminando as sem interesse.

Elaborar um sistema semi-automatizado de contagem e classificação de copépodes tem sido objetivo de muitos pesquisadores e desenvolvedores. Correlacionar as partículas presentes em tais ambientes com os parâmetros dispostos na água viabiliza o monitoramento de alterações nas comunidades biológicas.

## **CONCLUSÕES**

O presente trabalho objetivou desenvolver algoritmos de tratamento de imagens de copépodes, separando dentre o conjunto de imagens adquiridas, apenas as imagens aproveitáveis, ou seja, imagens do fundo à superfície. Os algoritmos elaborados para separação das imagens aproveitáveis foram satisfatórios, eliminando as imagens de momentos distintos do fundo à superfície no arrasto vertical. Os algoritmos elaborados para separação das imagens segmentadas e não segmentadas obtiveram grande porcentagem de acerto. Quanto à separação das imagens mal segmentadas, o algoritmo necessita de ajustes adicionais. Na utilização do algoritmo de contagem e separação dos copépodes, obteve-se um ótimo resultado, comprovando estatisticamente a semelhança entre o elaborado e o esperado realizado de forma manual. Com os resultados obtidos neste estudo, pode-se destacar a importância dos esforços destinados ao desenvolvimento e aperfeiçoamento destes algoritmos, pois a acurácia dos seus resultados, aumentam a confiabilidade das pesquisas e viabiliza diversos trabalhos de monitoramento ambiental e ecológicos. Neste escopo foi concluído com este estudo, que a elaboração de algoritmos para tratamento de imagens permite ter agilidade nos resultados. Os resultados deste trabalho ressaltaram a importância das ferramentas automáticas que realizam a contagem e classificação dos copépodes no ambiente *in situ*. Na eliminação automática de imagens não aproveitáveis foram obtidos sucesso e economia de tempo, assim também como na separação das imagens segmentadas e não segmentadas. Na classificação dos copépodes por tamanho da área, o resultado obtido pelo algoritmo elaborado foi positivo e expressivo, demonstrando a eficácia da ferramenta e evidenciando a importância de sua elaboração e futuras correções para aperfeiçoamento da ferramenta.

## **REFERÊNCIAS**

Álvarez, E., López-Ururutia, Á., & Nogueira, E. (2012). Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: Application to FlowCAM. *Journal of Plankton Research*, 34(6), 454-469. <https://doi.org/10.1093/plankt/fbs017>

Bachiller, E., Fernandes, J.A., (2011). Zooplankton Image Analysis Manual: Automated identification by means of scanner and digital camera as imaging devices. 18(2): 16-37. *Revista de Investigación Marina*.

Basedow, S. L., Tande, K. S., Norrbin, M. F., & Kristiansen, S. (2013). Capturing quantitative zooplankton information in the sea: Performance test of laser optical plankton counter and video plankton recorder in *Calanus finmarchicus* dominated summer situation. *Progress in Oceanography*. 108(April 2016), 72-80. <https://doi.org/10.1016/j.pocean..2012.10.005>

Benfield, M.C., Grosjean, P., Culverhouse, P.F., Irigoien, X., Sieracki, M.E., Lopez-Urrutia, A., Dam, H.G., Hu, Q., Davis, C.S. & Hansen, A. (2007). RAPID: research on automated plankton identification. *Oceanography*, 20: 172-187

Bi, H., Guo, Z., Benfield, M. C., Fan, C., Ford, M., Shahrestani, S., & Sieracki, J. M., (2015). A semi-automated image analysis procedure for in Situ plankton imaging systems. *PLoS ONE*, 10(5), 1-17. <https://doi.org/10.1371/journal.pone.0127121>

Blaschko, M. B., Holness, G., Mattar, M. A. *et al.* (2005) Automatic in situ identification of plankton. Proceedings of the Seventh IEEE Workshops on Application of Computer Vision (WACV/MOTION'05), 1, 79–86.

Bradford-Grieve, J. M.; Markhaseva, E.I.; Rocha, C. E. F., (1999). Copepoda. In: Boltovskoy, D. (Ed.). South Atlantic Zooplankton. Leiden. Backhuys Publishers, V. 2, p. 869-1098.

Calazans D., Muelbert J. H., e Muxagata E. (2011).Estudos Oceanográficos:do instrumental ao prático. capítulo 9. página 200-274.

Castro, L.R., Troncoso, V.A., Figueroa, D.R., (2007). Fine-scale vertical distribution of coastal copepods in the Golfo de Arauco, central Chile, during the upwelling season. Prog. Oceanogr. 75, 486-500.

Davis, C. S., Hu, Q., Gallager, S. M. *et al.* (2004) Real-time observation of taxa-specific plankton distributions: an optical sampling method. Mar. Ecol. Prog. Ser., 284, 77–96.

Dietrich A. and Uhlig G., Crustaceana. Supplement No. 7, Studies on Copepoda II (Proceedings of the first International Conference on copepoda, Amsterdam, the Netherlands, 24-28 August 1981) (1984), pp. 159-165. Published by: Brill, Stable URL: <http://www.jstor.org/stable/25027550>

Forest A., Stemmann L., Picheral M., Burdof L., Robert D., Fortier L., Babin M. (2012). Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort Sea: combined results from an Underwater Vision Profiler and vertical net tows. *Biogeosciences*, 9, 1301-1320.

Frontier, S. Atlas del Zooplancton del Atlántico Sudoccidental y métodos de trabajo con el zooplancton marino. Mar del Plata: INIDEP, (1981).

Gonzalez, R. C., Woods R. E., and Eddins S. L. (2004). Digital Image Processing Using MATLAB. Prentice Hall. pag. 72.

Gorsky, G., Ohman, M. D., Picheral, M. *et al.* (2010) Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.*, 32, 285–303.

Grosjean, P., Picheral, M., Warembourg, C. *et al.* (2004) Enumeration, measurement and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J. Mar. Sci.*, 61, 518–525.

Henzler, C.M., Hoaglund E.A., Gaines S. D., (2010). FISH-CS – a rapid method for counting and sorting species of marine zooplankton. *Marine Ecology Progress Series*, 410, 1-11. Retrieved from <http://www.jstor.org/stable/24873996>

Hirche, H. J., Schulz, J., & Hanken, T. (2012). A modular imaging system for collection and analysis of live and preserved zooplankton samples. *Program Book –*

OCEANS 2012 MTS/IEEE Yeosu: The Living Ocean Coast – Diversity of Resources and Sustainable Activities. <https://doi.org/10.1109/OCEANS-Yeosu.2012.6263514>

Hongmin C., Zhong Y., Xinhua C., Weiming X., Xiaoyin X. (2014). A New Iterative Triclass Thresholding Technique in Image Segmentation. IEE TRANSACTION ON IMAGE PROCESSING, Vol. 23, NO. 3. [http: redpel.com+917620593389](http://redpel.com+917620593389)

Hu Q., e Davis C., (2005). Automatic plankton image recognition with co-occurrence matrices and Support Vector Machine. MARINE ECOLOGY PROGRESS SERIES. June 23, vol. 295:21-31,2005. Inter Research. <http://www.int-res.com>

Luo, T., Kramer, K., Goldgof, D. B. *et al.* (2004) Recognizing plankton images from the shadow image particle profiling evaluation recorder. IEEE Trans. Syst. Man Cyber. B, 34, 1753–1762.

Natchimuthu S., Chinnaraj P., Parthasarathy S., Senthil K., (2013). Automatic Identification of Algal Community from Microscopic Images. Bioinformatis and Biology Insights. Volume 7. October 10, 2013. <http://doi.org/10.4137/BBI.S12844>

Ping-Sung L., Tse-Sheng C., Pau-Choo C. (2001). A Fast Algorithm for Multilevel Thresholding. Journal of Information Science and Engineering. 713-727 pag. December 30. [http://www.iis.sinica.edu.tw/JISE/2001/200109\\_01](http://www.iis.sinica.edu.tw/JISE/2001/200109_01)

Sahoo, P.K., Soltani, S., Wong, A.K., Chan, Y.C., (1988). A survey of threshold techniques. *Computer Vision, Graphics and Image Processing* 41, 233-260.

Schmid, M. S., Aubry, C., Grigor, J., Fortier, L. (2016). The LOKI underwater imaging system and an automatic identification model for the detection of zooplankton taxa in the Arctic Ocean. *Methods in Oceanography*. DOI: 10.1016/j.mio.2016.03.003

Schulz, J., Barz, K., Ayon, P., Ludtke, A., Zielinski, O., Menedoht, D., Hirche, H. J., (2010). Imaging of plankton specimens with the lightframe on-sight keystone investigation (LOKI) system. *Journal of the European Optical Society – Rapid Publications* 5, 10017s.

Sosik, H. M. and Olson, R. J. (2007) Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Meth*, 5, 204–216.

Sieracki M. E. *et al* (2007)., *Research on Automated Plankton Identification*. *Oceanography*. Vol. 20, N° 2. Published June 2007. DOI: 10.5670/oceanog.2007.63 - Source: OAI

Surovy P., Dinis C., Marusak R., Ribeiro N. de A., (2014). Importance of automatic threshold for image segmentation for accurate measurement of fine roots

of woody plants. *Lesnícky Casopis Forestry Journal*. 60 244-249.  
<http://www.nlcsk.sk/fj/>

Tang, X., Stewart, W. K., Vincent, L. *et al.* (1998) Automatic plankton image recognition. *Artif. Intell. Rev.*, 12, 177–199.

Wiebe, P.H. & Benfield, M.C. (2003). From the Hensen net toward four dimensional biological oceanography. *Progress in Oceanography*, 56:7-136

Zarauz, L., Irigoien, X. and Fernandes, J. A. (2009) Changes in plankton size structure and composition, during the generation of a phytoplankton bloom, in the central Cantabrian Sea. *J. Plankton Res.*, 31, 193–207.

## **ANEXOS**

**Tabela 1.** Tamanho médio de um copépode obtido pela medição manual da área<sup>2</sup> realizado individualmente em 100 copépodes. Seguintes valores extraídos, a média de tamanho de um copépode foi de 354,95 pixels<sup>2</sup>, Desvio padrão foi 207,21 pixels<sup>2</sup>. Em cima destes dados foi definido o menor copépode subtraindo do desvio padrão, e o maior somado ao desvio padrão.

Imagem	Área	Mean	StdDev	Mode	Min	Max
1	238	246.534	16.319	255	172	255
2	444	250.592	11.587	255	196	255
3	243	222.379	33.543	255	154	255
4	189	247.339	22.227	255	160	255
5	238	242.538	24.208	255	154	255
6	215	242.158	26.381	255	154	255
7	234	239.637	29.253	255	154	255
8	425	238.355	28.486	255	154	255
9	178	243.315	25.867	255	154	255
10	186	239.473	22.300	255	184	255
11	349	246.805	17.962	255	170	255
12	401	246.426	17.470	255	185	255
13	369	210.528	46.795	255	106	255
14	191	232.660	38.943	255	124	255
15	624	167.176	71.293	255	36	255
16	181	208.365	55.875	255	96	255
17	357	253.843	4.844	255	206	255
18	203	252.714	6.830	255	213	255
19	286	254.483	3.164	255	226	255
20	469	218.420	44.723	255	117	255
21	273	247.462	16.068	255	184	255
22	467	250.779	12.917	255	186	255
23	274	242.631	21.086	255	160	255
24	256	246.891	15.745	255	192	255
25	185	244.324	17.811	255	192	255
26	229	250.205	11.080	255	205	255
27	141	248.922	14.283	255	196	255
28	415	224.692	31.618	255	156	255
29	202	238.837	26.706	255	159	255
30	215	239.321	24.830	255	156	255
31	503	215.620	47.535	255	114	255
32	412	239.311	29.584	255	141	255
33	456	157.366	68.437	255	52	255
34	581	161.530	68.613	255	37	255
35	411	226.954	48.517	255	93	255
36	580	108.888	69.104	255	30	255
37	460	190.580	63.738	255	53	255
38	357	198.132	47.242	255	105	255
39	646	236.909	37.013	255	98	255
40	516	233.221	40.155	255	121	255
41	173	227.965	42.505	255	122	255
42	366	189.003	67.551	255	63	255
43	1143	189.202	59.412	255	79	255
44	159	211.673	54.819	255	90	255
45	465	226.553	51.734	255	90	255
46	286	221.101	52.623	255	91	255
47	321	218.302	53.190	255	91	255
48	127	220.087	51.145	255	78	255
49	475	186.021	54.356	255	56	255

50	461	231.191	31.710	255	140	255
51	181	242.823	26.865	255	148	255
52	1080	236.949	38.399	255	118	255
53	486	228.720	42.604	255	118	255
54	284	136.063	59.000	255	45	255
55	1208	230.642	41.933	255	106	255
56	334	233.317	40.125	255	118	255
57	376	207.904	48.722	255	114	255
58	741	192.177	65.301	255	59	255
59	166	151.193	45.640	158	59	236
60	469	239.670	34.673	255	123	255
61	184	226.370	40.046	255	121	255
62	323	146.693	57.271	255	63	255
63	303	238.937	29.202	255	155	255
64	315	228.752	33.568	255	154	255
65	291	231.405	34.736	255	133	255
66	343	197.653	54.701	255	73	255
67	231	177.004	52.121	255	82	255
68	271	170.517	66.022	255	60	255
69	460	230.293	40.579	255	112	255
70	290	191.917	51.109	255	78	255
71	408	235.174	40.586	255	112	255
72	395	224.658	42.201	255	105	255
73	315	225.438	44.649	255	112	255
74	557	172.706	57.449	255	75	255
75	410	244.702	18.139	255	191	255
76	283	245.216	18.519	255	160	255
77	469	249.354	15.271	255	189	255
78	205	172.298	60.471	255	78	255
79	332	203.361	60.658	255	85	255
80	317	203.640	54.837	255	86	255
81	281	201.779	55.108	255	85	255
82	305	192.279	64.380	255	85	255
83	270	217.067	46.245	255	87	255
84	213	218.761	55.374	255	85	255
85	360	206.211	58.508	255	85	255
86	395	148.732	55.598	255	53	255
87	156	197.974	31.748	255	129	255
88	378	131.833	72.445	255	31	255
89	128	239.117	22.290	255	176	255
90	92	222.446	22.812	255	179	255
91	154	246.227	19.705	255	186	255
92	207	239.696	19.929	255	187	255
93	481	183.605	70.548	255	66	255
94	1144	154.700	64.603	255	30	255
95	234	152.769	47.199	255	80	255
96	238	184.155	45.356	255	109	255
97	227	244.471	21.509	255	173	255
98	237	243.262	22.775	255	173	255
99	191	234.272	26.731	255	164	255
100	202	235.035	23.480	255	173	255

**Média de tamanho de um copéode** 354,95

**Desvio padrão** 207,2134

**Menor = Menor - Desvio Padrão** 147,7366

**Maior = Maior + Desvio Padrão**

562,1634

---