**Manual de utilização**

**Rotina de Automação de Processamento de Dados do CTD com SBEDATAPROCESSING-WIN32**

**Visão Geral:**

Este manual é organizado em seções que guiarão você por todo o processo de instalação dos pré-requisitos, configuração do ambiente e execução da rotina de automação. Cada seção fornece instruções passo a passo, acompanhadas de exemplos e explicações claras.

**Objetivo:**

Este manual foi criado com o objetivo de orientá-lo na utilização dessa rotina de automação do processamento de dados brutos coletados por um instrumento oceanográfico CTD (Condutividade, Temperatura, Profundidade) utilizando o software SBEDATAPROCESSING-WIN32. A coleta e análise de dados oceanográficos desempenham um papel crucial em diversos campos, incluindo pesquisa científica, monitoramento ambiental e exploração marinha. O instrumento CTD é amplamente utilizado para medir parâmetros como temperatura, salinidade e profundidade nas águas oceânicas. O software SBEDATAPROCESSING-WIN32 é uma ferramenta poderosa para o processamento desses dados, convertendo informações brutas em conjuntos de dados utilizáveis. A rotina de automação apresentada aqui busca simplificar e otimizar a tarefa de processamento de dados, permitindo que você execute as etapas de forma eficiente e precisa.

**Principais vantagens da rotina:**

O processamento manual de dados pode ser uma tarefa complexa e demorada, sujeita a erros humanos. Compreender e dominar o software de processamento pode exigir um tempo significativo de aprendizado. Dessa forma, as principais vantagens da utilização dessa rotina são:

1. A organização: Ao realizar a rotina, entende-se que qualquer trabalho, principalmente um de cunho científico, precisa de muita organização, metodologia e padronização, afim de melhorar a consistência e a eficiência dos resultados. Pensando nisso, a rotina utiliza bibliotecas padrões do Python, como a Pathlib para criar pastas que separam partes importantes do processamento de dados:
   1. Cria, primeiramente, uma pasta-cópia dos dados brutos, para evitar trabalhar diretamente com a pasta de dados originais e, assim, evitar possíveis adulterações dos dados
   2. Cria também uma outra pasta para os dados processados. Nessa, há outras pastas:
      1. Batch\_file: uma pasta para o batch file que é utilizado durante esse determinado processamento;
      2. Arquivos\_PSA: pasta que armazenará os arquivos PSA utilizados para o processamento no software;
      3. Arquivos\_BTL: armazena os arquivos de garrafa, caso o processo Bottle Summary (09), seja realizado;
      4. Processamento\_seaplot: armazena os gráficos produzidos pelo processo Sea Plot (20);
      5. Processamento\_split: armazena os arquivos de subida e de descida extraídos a partir do processo Split (17), nessa pasta, os arquivos de subida e de descida podem ser separados em duas pastas diferentes: up e down, caso o usuário queira separar;
      6. Relatório: por fim, é criado também uma pasta em que fica registrado pontos importantes do processamento, nessa pasta há um arquivo .txt de Relatório, que explica algumas informações que foram extraídas após a rotina realizar uma breve análise dos dados brutos, além de deixar registrado as estações processadas a cada processamento, o usuário também pode fazer qualquer alteração que achar importante nesse relatório, o que facilita a comunicação e compreensão entre os pesquisadores ao lidarem com os dados. Além disso, há também outro arquivo .txt que registra a latitude, longitude, data e hora de cada estação que foi processada. Nessa pasta também será possível, se caso o usuário desejar, armazenar um arquivo kml, que permite a visualização das estações no aplicativo Google Earth.
   3. Durante o processamento, os nomes dos arquivos também são padronizados, ainda na pasta-cópia, facilitando todo o processo e permitindo a automação com a utilização de um Batch File.
   4. Todos os arquivos produzidos a partir do processamento ficam na pasta de processamento.
2. A eficiência: Além de ser uma automatização de um processo que poderia ser bastante demorado, uma vez realizado um processamento, não é necessário retomar ele do zero, pois a rotina permite retomar um processamento antigo e, para isso, é necessário apenas informar qual o caminho da pasta-cópia e o caminho da pasta de dados processados. Isso permite uma grande eficiência durante o processamento de dados.
3. Facilidade de utilizar: A rotina foi pensada para ter uma utilização simples, prática e de fácil entendimento, assim, são dadas instruções práticas na própria rotina, não sendo necessário nenhum conhecimento prévio em programação para conseguir utilizá-la.
4. Bônus: Com essa rotina, é possível, como já foi mencionado, ter relatórios organizados que explicam parte dos dados, como: se há alguma estação com arquivos incompletos, o número total de estações, quais processos o Batch\_file está realizando, além de ser possível visualizar em um só arquivo as posições, data e hora das estações. Além dessas facilidades, há a produção do arquivo KML, que permite uma melhor visualização e noção dos locais das estações e, ainda, é possível visualizar os gráficos produzidos pelo Sea Plot ao clicar em cada estação no arquivo KML.

**Pré-requisitos para utilizar:**

Para utilizar essa rotina, são necessárias algumas instalações no computador do usuário. Primeiramente, é preciso ter o Python instalado, preferencialmente a versão mais recente (<https://www.python.org/>), além da instalação de algumas bibliotecas específicas para ser possível rodar a rotina. Essas bibliotecas podem ser instaladas com o código que está em “Pacote\_de\_instalação\_CTD”, arquivo que deve estar junto do manual e da rotina.

Esse código só é necessário ser rodado uma única vez se o usuário não possuir todas as bibliotecas, é recomendado sua utilização antes da primeira vez de rodar a rotina. Além disso, é necessário possuir o software SBEDataProcessing-Win32, pois é nele que o processamento é realizado. É recomendável possuir arquivos PSA para o processamento, porém pode-se construí-los do zero manualmente. Por último, é interessante baixar o aplicativo Google Earth Pro, para conseguir visualizar o arquivo KML.

**Guia de uso:**

**Passo 1: Deseja fazer um processamento novo? (s/n)**

Nessa primeira parte, deve-se escrever ‘s’ para sim e ‘n’ para não. Caso o input seja ‘s’, irá começar um processamento do zero, o que é recomendável se:

1. For seu primeiro processamento

2. Caso queira fazer um processamento diferente de outro já realizado, exemplo: com apenas uma parte dos dados, com outros dados, etc...

Caso a resposta seja ‘n’, será perguntado o caminho da pasta de dados processados: “Digite o caminho da pasta de dados processados:”, em que deve-se copiar o caminho inteiro da pasta em que estão os dados processados, do início ao fim, e, logo após, pede-se: “Digite o caminho da pasta em que os arquivos estão copiados:”, onde deve-se informar o caminho, também completo, da pasta-cópia utilizada naquele processamento em específico.

Atenção: A cópia deve ser a criada para aquele processamento, pois os nomes dos arquivos estão padronizados e precisam estar na mesma formatação. Se, por algum motivo, a resposta para essa primeira pergunta não for nem ‘s’ nem ‘n’, então será informado que a resposta não é válida: “Resposta inválida. Por favor, responda com "s" para sim ou "n" para não.” e será feita uma conferência sobre se o usuário deseja continuar ou cancelar o processamento: “Deseja cancelar o processo? (s/n)”, caso o usuário responda ‘s’, o processamento é cancelado, caso contrário, a pergunta: “Deseja fazer um processamento novo? (s/n)” é refeita.

**Paso 2: Em que pasta estão os dados brutos? (digite o caminho da pasta)**

Caso o usuário esteja realizando o seu primeiro processamento, ele deve colocar ‘s’ para a primeira pergunta. Logo após ela, será pedido para que o usuário informe o caminho (completo) da pasta em que estão os dados brutos: “Em que pasta estão os dados brutos? (digite o caminho da pasta)”. Essa parte é importante para garantir a segurança dos dados brutos

**Passo 3: Dê nome à cópia da pasta em que você irá trabalhar: (use \_ no lugar de espaços)**

Após informar a pasta de dados brutos, você deve digitar o nome da pasta com a cópia de todos os dados, nela eles serão analisados e feitas algumas alterações, como de nomenclatura, além de ser a pasta em que o batch file irá buscar os arquivos utilizados para o processamento. É recomendado que não utilize espaços no nome da pasta, para facilitar o processamento e evitar erros.

**Passo 4: Como gostaria de renomear os arquivos?**

Após informar o nome da pasta-cópia desejada, a rotina pergunta “Como gostaria de renomear os arquivos?”. Nessa parte deve ser informado o nome que será utilizado na padronização dos dados e no batch file, portanto é ideal a escolha de um nome que esteja relacionado com o programa em que os dados foram aquisitados. Além disso, não utilize espaços na nova nomenclatura.

Atenção: Caso o ‘novo nome’ dos arquivos seja exatamente igual ao nome antigo, será informado que o arquivo antigo permaneceu com a mesma nomenclatura e será recomendado a averiguação dos dados. Esse informe também será dado caso o novo nome já exista na pasta-cópia por algum motivo. Alguns motivos podem ser: repetição de arquivos de uma mesma estação, mesmo número de estações par operações diferentes numa mesma pasta, etc. O informe será: “Por favor, verifique manualmente se há mais um tipo de processamento junto aos dados e faça as alterações necessárias”.

**Passo 5: Conferência das informações obtidas**

Nessa etapa, serão mostradas algumas informações que foram possíveis de serem adquiridas a partir de uma breve análise dos dados brutos na pasta-cópia. Serão informados o número total de arquivos copiados, o número de arquivos do tipo HEX, BL, HDR, MRK, XMLCON e COM. Será informado o número total de estações que estão na pasta-cópia e, também, quais estações não estão completas e quais os tipos de arquivos que estão faltando em cada uma.

É muito importante revisar todas essas informações, principalmente o número de estações, pois como muitos dados brutos não estão padronizados, o código está suscetível a alguns erros na contagem das estações. Essas informações irão ser armazenados no relatório de processamento, que fica dentro da pasta de relatório, dessa forma, o usuário pode mudar e/ou adicionar qualquer informação que achar necessária.

**Paso 6: Nomeie a sua pasta de dados processados: (use \_ no lugar de espaços)**

Será pedido, dessa forma: “Nomeie a sua pasta de dados processados: (use \_ no lugar de espaços)”, para que o usuário escolha um nome para a pasta em que ficarão todos os dados provenientes do processamento. É indicado que se utilize ‘\_’ ao invés de espaços, para evitar qualquer erro durante o processamento. Recomenda-se, também que o nome seja intuitivo e esteja relacionado com a operação realizada, exemplo: “Processamento\_OP41”.

**Passo 7: Você já possui um arquivo PSA para realizar o processamento? (s/n)**

Nesse passo, você deve informar se já possui arquivos PSA ou não. Caso escreva ‘s’, é retornada uma mensagem:

“Por favor, coloque os seus arquivos PSA na pasta Arquivos\_PSA que se encontra em {nome da pasta de arquivos processados}

Confira que os arquivos PSA estão nomeados segundo a nomenclatura padrão do SBEDataProcessing

Pressione Enter após terminar...”

Caso contrário, é retornada a seguinte mensagem:

“Por favor, crie os arquivos PSA desejados e coloque-os na pasta Arquivos\_PSA que se encontra em {nome da pasta de arquivos processados}

Confira que os arquivos PSA estão nomeados segundo a nomenclatura padrão do SBEDataProcessing

Pressione Enter após terminar...”

Em ambos os casos, o usuário deve fornecer os arquivos PSA e colocá-los na pasta Arquivos\_PSA. É fundamental a conferência de que todos os arquivos PSA dos determinados processados que serão pedidos pelo Batch file estejam nessa pasta e que estejam com a nomenclatura padrão reconhecida pelo SBEDataProcessing.

As nomenclaturas são: 'AlignCTD.psa', 'BinAvg.psa', 'BottleSum.psa', 'CellTM.psa', 'DatCnv.psa', 'Derive.psa', 'Filter.psa', 'LoopEdit.psa', 'SeaPlot.psa', 'SeaPlot\_autoscale1.psa', 'SeaPlot\_ts.psa', 'Split.psa', 'WildEdit.psa', 'MarkScan.psa', 'Buoyancy.psa', 'W\_Filter.psa', 'Section.psa’, 'Strip.psa', 'Trans.psa'.

O processamento só irá funcionar caso as nomenclaturas dos PSA estejam padronizadas dessa forma, ainda assim, a rotina reconhece variações apenas de letras minúsculas ou maiúsculas desses nomes, mas é recomendado que se deixe dessa forma, ou com letras somente minúsculas ou maiúsculas nesse padrão, sem variações. A rotina só continuará depois que o usuário pressionar enter, indicando que conferiu tudo antes de passar para a próxima etapa.

**Passo 8: Você já possui um Batch File para realizar esse processamento? (s/n)**

Nessa etapa, caso você já possua um arquivo Batch File para realizar o processamento, deve digitar ‘s’ para sim. Nesse caso, você deverá fazer todas as modificações necessárias no Batch file para conseguir realizar o processamento, desde a atenção às nomenclaturas, até o informe de cada pasta em que serão armazenados os dados. Caso a resposta seja ‘n’, para não, então será feito um batch file para você, que deve informar como você quer que ele seja criado, quais processamentos quer realizar, etc.

**Paso 9: Dê nome ao seu batch file (sugestão: batch\_nome\_do\_programa):**

Caso tenha optado por realizar um Batch file novo, você deverá informar o nome que quer dar ao seu Batch file, por favor, não utilize espaços.

**Passo 10: Quais processamentos você deseja realizar?**

Nessa etapa é necessária bastante atenção, pois, ao receber a pergunta: “Quais processamentos você deseja realizar?”, deve-se escrever os números dos processos que se quer realizar, exemplo: Caso queria realizar o Data Conversion e Filter informe que quer realizar os processamentos 01, 02... Pode também informar sequências, como 01-04,...

Porém, é necessário conferir que todos os arquivos PSA de cada um dos processos que serão realizados estão na pasta de Arquivos\_PSA com a nomenclatura correta. Caso essa etapa seja negligenciada, o batch não funcionará e ocorrerá um erro na rotina, forçando o usuário a começar do zero ou a fazer o seu próprio batch manualmente, pois ele poderá retomar a algum processamento já feito apenas se tiver um Batch file na pasta Bach\_file.

**Passo 11: Conferência para realização do processamento**

Nessa parte, é informado ao usuário para conferir se todas as informações estão certas e se todos os PSA estão adequados com o Batch file. Pressione enter após realizar todas as conferências necessárias. É muito importante realizar as conferências para evitar qualquer tipo de erro durante o processamento.

**Passo 12: Informar quais estações serão processadas**

Nessa etapa, será feita a seguinte pergunta:

“Em quais estações irá processar os dados? (Escreva uma lista com 3 dígitos. Ex.: 001-010, 020

Caso a estação tenha letras após o número, por favor, escreva, exemplo: 001ABC)

Caso queira processar todas as estações, escreva "tudo".”

O usuário deverá fornecer todas as estações que quer realizar para esse processamento. Pode fornecer uma lista de 3 dígitos, informando os números das estações, a lista pode incluir sequências, como 001-010. Além disso, se caso houver variações de uma mesma estação, por exemplo: 002A, 002c,... e se deseja processar elas, é necessário escrever o nome/letra após os 3 dígitos também. E, por último, caso deseje processar todas as estações, apenas escreva ‘tudo’.

Caso você digite uma estação que não existe, será informado que a estação não existe e a pergunta sobre quais estações deverão ser processadas será refeita. Porém, caso você digite algum termo inválido (que não obedeça ao padrão de 3 dígitos para o número da estação) será informado que sua resposta é inválida e pergunta também será repetida.

Nessa etapa serão realizados os processamentos das estações que foram pedidas e que são válidas. Além disso, a tabela de posição é feita e o relatório é atualizado.

**Passo 13: Deseja separar os arquivos de subida e descida? (s/n)**

Após a realização de todo o processamento, será perguntado se o usuário deseja separar os arquivos de descida e de subida, que estão na pasta de “Processamento\_split”, caso o usuário deseje fazer isso, será criado duas pastas ‘up’ e ‘down’, cada uma para armazenar ou arquivos de subida ou de descida. Caso a resposta seja ‘n’, para não, ele apenas não separará esses arquivos. Essa pergunta é refeita a cada processamento realizado.

**Passo 14: Deseja criar um arquivo KML? (s/n)**

Nessa etapa, o usuário digitará ‘s’ para caso queria produzir um arquivo KML dos dados que foram processados. Nesse arquivo haverá marcações indicando as posições de cada estação já processada e, caso tenha sido feio o SeaPlot da estação, haverá o gráfico produzido no processamento sendo representado em cada marcação.

Atenção: O KML produzido utiliza as informações de todas as estações que estão na tabela de posição das estações processadas, caso não queria de todas as estações, faça outro processamento ou apague as informações da tabela. Caso o usuário queria o arquivo KML, ele deverá informar o nome do seu KML, que estará na pasta ‘Relatório’.

**Passo 15: Deseja terminar o processamento? (s/n)**

Por fim, será perguntado se o usuário deseja terminar o processamento. Caso digite ‘s’, o processamento é concluído. Caso digite ‘n’, a rotina volta para o passo 12, para processar mais estações.

**Avisos:**

Essa rotina processa apenas os seguintes processos:

'01': 'Data Conversion',

'02': 'Filter',

'03': 'Align CTD',

'04': 'Cell Thermal Mass',

'05': 'Loop Edit',

'06': 'Derive',

'07': 'Derive TEOS-10',

'08': 'Bin Average',

'09': 'Bottle Summary',

'10': 'Mark Scan',

'11': 'Buoyancy',

'12': 'Wild Edit',

'13': 'Window Filter',

'17': 'Split',

'18': 'Strip',

'19': 'Translate',

'20': 'Sea Plot'.

Ou seja, os processos 14, 15, 16 e 21 não serão executados de forma alguma nessa rotina, mas, caso o usuário deseje modificar o Batch manualmente e colocar os arquivos PSA necessários para realizar esses processos, pode fazê-lo.

Os nomes dos processos que não são realizados:

'14': 'ASCII In',

'15': 'ASCII Out',

'16': 'Section',

'21': 'SeaCalc III'.

Caso o usuário deseje realizar o ‘Derive’ e tenha esse arquivo PSA, será feita uma conferência automática para, caso também tenha um arquivo PSA do 'Derive TEOS-10', esse também será realizado. Essa preferência foi colocada por ser uma versão mais recente do processo ‘Derive’.

**Agradecimentos:**

Agradecemos a preferência e esperamos que este manual e a rotina de automação simplifiquem suas tarefas de processamento de dados. Aproveite ao máximo esta ferramenta para melhorar a eficiência de suas análises e contribuir para o avanço do conhecimento oceanográfico.