



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

sid.inpe.br/mtc-m21b/2016/01.29.19.49-NTC

**SISTEMA DE GESTÃO E CONTROLE DE QUALIDADE  
DE DADOS (GCQD) PARA ASSIMILAÇÃO NOS  
MODELOS DO CPTEC (VERSÃO 1.0 BÁSICA)**

Luiz Fernando Sapucci  
Lucas Amarante Avango  
Nicolas Cruz Salvador  
Fábio Luiz Rodrigues Diniz  
Renato Galante  
João Gerd Zell de Mattos

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34P/3L46UF8>>

INPE  
São José dos Campos  
2015

## **PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@sid.inpe.br

## **COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):**

### **Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

### **Membros:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas  
(CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos  
(CPT)

Dr<sup>a</sup> Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

### **BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Clayton Martins Pereira - Serviço de Informação e Documentação (SID)

### **REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação  
(SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

### **EDITORAÇÃO ELETRÔNICA:**

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

sid.inpe.br/mtc-m21b/2016/01.29.19.49-NTC

**SISTEMA DE GESTÃO E CONTROLE DE QUALIDADE  
DE DADOS (GCQD) PARA ASSIMILAÇÃO NOS  
MODELOS DO CPTEC (VERSÃO 1.0 BÁSICA)**

Luiz Fernando Sapucci  
Lucas Amarante Avango  
Nicolas Cruz Salvador  
Fábio Luiz Rodrigues Diniz  
Renato Galante  
João Gerd Zell de Mattos

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34P/3L46UF8>>

INPE  
São José dos Campos  
2015



Esta obra foi licenciada sob uma Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.

## SUMÁRIO EXECUTIVO

Os dados observados de variáveis atmosféricas das mais variadas fontes observacionais chegam rotineiramente ao CPTEC através do Global Telecommunication System (GTS) e outros meios operacionais de disseminação de dados. O CPTEC através de seus acordos institucionais com os centros regionais possui uma base de dados convencionais sobre a América do Sul mais completa do que a utilizada operacionalmente em outros centros, a qual vem gradativamente crescendo com a expansão da base de dados. No que se refere aos dados de satélites a Divisão de Satélites e Sistemas Ambientais (DSA) com suas antenas e acordos internacionais, têm produzido operacionalmente uma base de dados observacional bastante densa sobre a América do Sul com o seu envolvimento no Regional ATOVS Retransmission Services (RARS) da WMO. Essa base de dados, mais completa e principalmente sua expansão, apresenta um potencial significativo para a melhoria da qualidade dos produtos de previsão numérica (PN) do CPTEC sobre a América do Sul, a qual é o propósito principal de sua missão. Mas para isso a base de dados deve ser processada, armazenada e disponibilizada adequadamente para ser utilizada no processo de assimilação de dados. O grupo de trabalho do pré-processamento da Divisão de Operações (DOP) do CPTEC é a equipe responsável por fazer operacionalmente uma triagem inicial dos dados convencionais, decodificá-los e armazená-los em um formato apropriado. Para os dados de satélites a equipe de especialistas da DSA e a estrutura operacional disponível são capazes de fornecer dados (em especial radiações e vento a superfície) com baixa latência e fluxo contínuo com forte potencial para assimilação de dados em ciclos rápidos sobre a América do Sul. No entanto, para que seja possível a assimilação desses dados nos modelos de PN do CPTEC, tanto os de abrangência global como regional, é necessário um sistema robusto de controle de qualidade usando previsão de curto prazo do modelo e a conversão em formatos de arquivos com códigos de observações utilizados nos sistemas de assimilação de dados. O presente projeto tem o propósito de desenvolver um fluxo estável para os dados recebidos (tanto no pré-processamento da DOP como na DSA), passando por um processo de preparação para a assimilação envolvendo no caso dos convencionais um controle de qualidade, o que resulta em uma base de dados com suas estatísticas de erro, necessária para o processo de assimilação na PN do CPTEC. Uma proposta no sistema de gerenciamento de projetos do CPTEC foi criada tratando do desenvolvimento de um sistema de Gestão e Controle de Qualidade de Dados para a assimilação em modelos do CPTEC (denominado pelo acrônimo GCQD). Logo na concepção dessa proposta se estudou como o desenvolvimento da mesma poderia ser particionado e se definiu o que se pode oferecer em cada uma

delas. O presente relatório técnico apresenta em detalhes a versão 1.0 do sistema GCQD, documentando todos os módulos em que o mesmo é dividido. Como é um processo em desenvolvimento outras versões mais completas deverão ser publicadas em continuidade desse desenvolvimento. A presente versão do sistema contempla uma ampliação da base de dados de satélites disponibilizando dados de radiância de outros sensores e com cobertura global dos satélites NOAA-18, NOAA-19 e MEtOp-B e dados de vento por satélite no canal do vapor d'água e no infravermelho. Com relação aos dados convencionais, todos os módulos dos sistemas de observação foram incluídos e os dados de estações à superfície foram implementados. O fator mais importante abordado nessa versão é que todo o sistema foi reestruturado para permitir a implementação compartilhada por uma maior número de colaboradores. Isso deverá contribuir significativamente para acelerar a inclusão ainda não realizada dos sistemas de observações. A publicação dessa versão é um passo para se obter uma versão onde toda a base de dados disponível sobre a América do Sul seja contemplada e os dados disponíveis para a assimilação sejam gerados com a menor latência, visando a viabilidade do ciclo rápido de atualização da condição inicial para modelagem de alta resolução. Essa tarefa de ampliação da base de dados deverá estar em constante desenvolvimento e demandará relativamente um longo período de tempo. Um desenvolvimento completo e de difícil previsão, uma vez que a disponibilidade de novos tipos de dados e modificações nos mesmos é sempre constante e requer atenção. Essa tarefa deverá fazer parte da rotina do grupo de assimilação de dados e o sistema em desenvolvimento deverá ser uma ferramenta importante para o bom êxito da mesma.

**Palavras chave:** Assimilação de dados. Fluxo de dados. GSI. Base de dados regionais. Dados convencionais.

## LISTA DE FIGURAS

	<u>Pág.</u>
1.1 Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelo do CPTEC. . . . .	2
1.2 Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelo do CPTEC. . . . .	4
2.1 Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelo do CPTEC. . . . .	9
2.2 Lista de opções do sistema. . . . .	12
3.1 Diagrama do modulo para o sistema Dummy. . . . .	19
3.2 Lista de mnemonics das tabelas NCEP. . . . .	21
3.3 Esquema da ferramenta bufrtools. . . . .	21
3.4 Esquema principal do sistema Dummy. . . . .	22
3.5 Esquema para armazenamento de arquivos BUFR. . . . .	23
3.6 Tabela de mnemônicos do NCEP para os arquivos BUFR. . . . .	24
3.7 Tabela de mnemônicos do NCEP para dados de satélites. . . . .	25
3.8 Sub-diretorios do sistema de validação Dummy. . . . .	27
3.9 Esquema de sistema de validação Dummy. . . . .	28
3.10 Saída gráfica de vento SATWND CPTEC. . . . .	30
3.11 Saída gráfica SATWND NCEP. . . . .	30
3.12 Saída gráfica GPSIPW NCEP. . . . .	31
3.13 Fluxograma do funcionamento dos processos no BFGGE. . . . .	36
3.14 Distribuição dos dados de superfície. . . . .	40
3.15 Mnemônicos equivalentes NCEP WMO (cabecalhos). . . . .	41
3.16 Mnemônicos equivalentes NCEP WMO (ângulos). . . . .	41
3.17 Mnemônicos equivalentes NCEP WMO (observações). . . . .	41
3.18 Saída gráfica de dados processados de radiância via RARS. . . . .	43
3.19 Saída gráfica de dados processados de radiância via NCEP. . . . .	43
3.20 Comparação da observação de temperatura de brilho do RARS(lado esquerdo) e do NCEP(lado direito). . . . .	44
3.21 Estrutura de diretórios para os módulos da radiâncias. . . . .	45
3.22 Estrutura de diretórios para AMSUA. . . . .	46
3.23 Análise estatística do sensor AMSUA, para os canais 5, 6, 7, 8, e 9 dos satélites NOAA-18, NOAA-19 e Metop-b. . . . .	47
3.24 Passagem dos satélites NOAA-19 e Metop-b com dados do sensor AMSUA, domínio da antena do sistema RARS. . . . .	47

3.25	Estrutura funcional para o módulo do vento por satélite. . . . .	48
3.26	Mnemônicos de vento NCEP. . . . .	52
3.27	Distribuição espacial das observações de vento por satélite. . . . .	53
3.28	Argumentos requeridos para execução do script. . . . .	57
3.29	Visualização da execução do script. . . . .	59
4.1	Detalhes para executar o PAQC. . . . .	70
4.2	Processos realizados pelo PAQC. . . . .	72



## SUMÁRIO

	<u>Pág.</u>
<b>1 Introdução</b> . . . . .	<b>1</b>
1.1 Concepção da proposta do GCQD na Assimilação de Dados . . . . .	2
1.2 Estrutura básica do pacote do GCQD . . . . .	3
1.3 Estrutura do documento . . . . .	4
<b>2 Informações gerais sobre a versão V1.0 básica do GCQD</b> . . . . .	<b>7</b>
2.1 Estrutura do sistema GCQD . . . . .	7
2.2 Obtenção/configuração/compilação e execução do pacote OBS . . . . .	9
2.2.1 Obtenção do sistema obs . . . . .	10
2.2.2 Instalação do sistema OBS . . . . .	11
2.2.3 Detalhes do script de instalação do sistema obs . . . . .	12
<b>3 Pacote BuFr GEnerator (BFGE)</b> . . . . .	<b>17</b>
3.1 Estratégia de desenvolvimento do BFGE . . . . .	17
3.1.1 Estrutura Dummy . . . . .	18
3.1.2 Modelo de Tanque de dados BUFERS para a assimilação . . . . .	22
3.1.3 Estrutura de diretórios para os dados convencionais . . . . .	23
3.1.4 Estrutura de diretórios para os dados de satélites . . . . .	24
3.1.5 Script para gerar a estrutura de subdiretórios no BufrTank . . . . .	25
3.1.6 Validação PrepBUFR Dummy . . . . .	26
3.1.7 Visualização PrepBUFR . . . . .	26
3.1.8 Decodificação . . . . .	29
3.1.9 Visualização . . . . .	29
3.1.10 Estatística básica para avaliação quantitativa do PrepBUFR . . . . .	31
3.2 Estrutura básica do BFGE e seus módulos principais . . . . .	32
3.2.1 Superfície . . . . .	36
3.2.2 Módulos de Radiância . . . . .	40
3.2.2.1 Metodologia e dados . . . . .	40
3.2.2.2 Validação dos arquivos de radiâncias gerados . . . . .	42
3.2.2.3 Resultados obtidos na assimilação . . . . .	43
3.2.2.4 Inclusão dos dados de radiância dos satélites NOAA-19 e MetOp-B no GCQD Versão básica . . . . .	45
3.2.3 Estrutura de dados SATWND (dados de vento por satélites) . . . . .	47

3.2.4	Desenvolvimento dos dados de Rádio Ocultação GNSS no sistema GCQD	53
3.3	Obtenção/ configuração/ compilação e execução do BFGE . . . . .	53
3.3.1	Detalhes sobre a instalação do BFGE . . . . .	54
3.3.2	Detalhes do script de instalação do BFGE . . . . .	54
3.3.3	Utilização do pacote BFGE . . . . .	56
3.3.4	Detalhes do script de execução do BFGE . . . . .	57
3.3.5	Passo-a-passo para a execução do BFGE . . . . .	58
<b>4</b>	<b>Pacote de Controle de Qualidade (PAQC - Package Quality Control) . . . . .</b>	<b>61</b>
4.1	Os módulos principais do pacote PAQC . . . . .	61
4.2	Obtenção/ configuração/ compilação e execução do PAQC . . . . .	65
4.2.1	Detalhes sobre a instalação do PAQC . . . . .	65
4.2.2	Detalhes do script de instalação do PAQC . . . . .	65
4.2.3	Utilização do pacote PAQC . . . . .	68
4.2.4	Detalhes do script de execução do PAQC . . . . .	69
4.2.5	Passo-a-passo para a execução do PAQC . . . . .	71
<b>5</b>	<b>Perspectivas futuras e considerações finais . . . . .</b>	<b>73</b>
5.1	Lista de versões concebidas na elaboração da proposta com o cronograma das futuras entregas . . . . .	73
5.2	Melhorias futuras identificadas para serem implementadas nas versões posteriores a V1.0 do GCQD . . . . .	74
5.3	Considerações finais . . . . .	75
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>77</b>
<b>A</b>	<b>ANEXO REVISÕES DO PROJETO . . . . .</b>	<b>79</b>

## 1 Introdução

A versão tratada nesse relatório foi implementada no sistema de gerenciamento de projetos do CPTEC, no qual um projeto foi criado. O endereço do projeto no svn é <https://svn.cptec.inpe.br/paqc> e o endereço no redmine é <https://projetos.cptec.inpe.br/projects/paqc>. Toda a estrutura do código tratado nesse documento e as páginas wikis se referem a esse repositório e considera-se que o leitor tenha acesso a ela para mais detalhes.

Todo o pacote de tratamento de dados é bastante abrangente e o desenvolvimento desse projeto deverá ser feito continuamente pelos próximos anos. Assim, está se adotando uma metodologia de desenvolvimento que envolve a definição de versões e um cronograma de entrega no qual o conteúdo das mesmas é bem detalhado e pré-definido em sua concepção. Caso haja a necessidade de modificações, implementações adicionais, essas devem ser realizadas nas futuras versões. Isso tudo salvo as correções emergenciais ou funcionalidades essenciais para a utilização do produto em desenvolvimento.

Todo o projeto com a definição dessas versões de entregas e as tarefas associadas foi planejado utilizando as ferramentas disponíveis nesse sistema. Nessa estratégia de desenvolvimento a aba de planejamento roadmap pode ser acessada durante o desenvolvimento, a qual permite o acompanhamento das tarefas e a evolução da versão em processo, bem como as tarefas associadas com as futuras versões. As páginas wikis são de grande ajuda para documentar apropriadamente todas as implementações realizadas. O conteúdo dessas páginas foi utilizado quase que integralmente para a compilação desse relatório. Problemas na eficiência desse processo foram identificados, os quais deverão ser resolvidos nas próximas versões do sistema.

## 1.1 Concepção da proposta do GCQD na Assimilação de Dados

Nessa proposta o que se pretende é organizar os dados recebidos operacionalmente pela DOP (dados convencionais) e pela DSA (dados de satélites) para a utilização na Modelagem atmosférica através da assimilação de dados. Dessa forma, o fluxo dos dados não é abordado na mesma, sendo essa a tarefa das divisões mencionadas. Dessa forma, a proposta trabalha com a premissa que os dados já estejam disponíveis e organizados em uma estrutura de diretórios que convencionou-se chamar de "Tanque", na qual os dados são separados pela data (YYYYMMDD) por tipos e subtipos obedecendo os mnemônicos usados no GSI, os quais são baseados na tabela do NCEP. O fluxo de dados serão gerenciados pela DOP no que se refere aos dados convencionais e pela DSA (dados de satélites). A figura 1.1 descreve essa ideia de forma mais ilustrativa. Durante o desenvolvimento do projeto os dados serão colocados em um tanque "Dummy" na conta gdad no scratchin, o qual simulará o processo em modo operacional. Em paralelo será trabalhado nessas divisões para que esse processo torne a ser operacionalizado seguindo o que é proposto aqui.

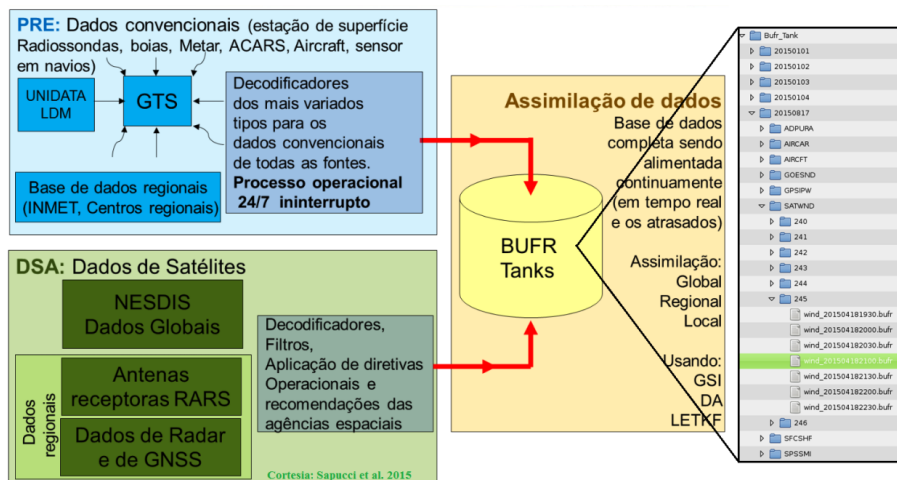


Figura 1.1 - Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelos do CPTEC.

Como é mostrado na figura 1.1 os dados convencionais das mais variadas fontes (estações de superfície, radiossondas, bóias, metar, Acars, Aircraft, sensor em navios, entre outras) espalhadas sobre todo o globo, bem como as redes de estações dos centros regionais, recebidas pelo Grupo de Pré-processamento da DOP deverão ser organizados em um depósito específico, denominado aqui de "BUFR tanque". Nesse processamento está envolvido uma árdua tarefa de decodificar as mensagens, aplicar

filtros e controle grosseiro de dados assim que as observações são recebidas, o que é feito operacionalmente e de forma ininterrupta. Para os dados de satélites o mesmo deverá ser feito pela DSA que recebendo os dados regionais via suas antenas ou mesmos os dados globais via sistema de recepção padrão, os mesmos devem ser processados e aplicados os filtros, melhorias necessárias para se ter o melhor desses dados para a assimilação. Os dados depositados no tanque, espera-se que estejam escritos em formato BUFR com as tabelas WMO, o qual é o padrão de distribuição. Como mostra a figura a direita, o tanque é organizado em diretórios por data, de forma que todos os dados de um mesmo dia sejam colocados no mesmo diretório e os mesmos estejam separados por tipo e subtipo. No exemplo é apresentado os dados de vento por satélites da DSA do dia 20150817, sendo do tipo "SATWND" e subtipo 245 que significa dados do canal do infravermelho do GOES. Essa base de dados depositada no tanque é a mesma forma de organizar sistematicamente todos os dados recebidos pelo CPTEC para a assimilação, de tal forma que ficaria disponível para futuros trabalhos envolvendo reprocessamento dos dados históricos. Além disso permitirá que os dados atrasados sejam agregados a base de dados e esses se tornarem na medida que o tempo passa uma base mais completa.

Cabe salientar que para a operacionalização dos dados para a assimilação é imprescindível uma coordenação a nível de chefias para que os grupos da DMD, DSA e DOP possam trabalhar em conjunto. O desenvolvimento das metas deverá ser feito envolvendo indiretamente os grupos mencionados na proposta a fim tornar mais eficiente sua evolução. Assim, como a necessidade de uma estrutura de fluxo de dados e sua ingestão nos modelos do CPTEC é institucional, sugere-se que esta iniciativa seja discutida em nível de coordenação do CPTEC em um momento oportuno durante a operacionalização.

## **1.2 Estrutura básica do pacote do GCQD**

O pacote é composto de um conjunto de código fonte e scripts para compilar e configurar os programas, bem como um conjunto de dados básicos para rodar o processo em modo de teste inicial. Ele está estruturado, visto na figura 1.2, em um diretório pai denominado "obs" o que já foi predefinido no projeto Regional Modeling System (RMS). Esse mesmo diretório pode ser adicionado no Global Modeling System (GMS), o qual poderá ser incluído na estrutura de processos e permitir que se tenha disponível para os usuários um processo capaz de gerar os dados para os testes na assimilação, seguindo, se for o caso, novos critérios na seleção e controle de qualidade a serem experimentados.



Figura 1.2 - Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelo do CPTEC.

A estrutura de diretório é básica contendo "bin" para os binários, "run" para os scripts de execução e o "src" contendo dois pacotes. Um pacote para o controle de qualidade dos dados convencionais (PrepBUFR) e um outro para a geração dos dados BUFRS NCEP. O primeiro será o pacote NCEP-PAQC contido dentro do pacote do sistema G3DVAR na versão inicial V0.0, o qual deverá ser substituído pelo pacote OBSPROC também do NCEP, nas versões posteriores. O segundo denominado aqui de "BUFR Generator" (BFGC) é um pacote em desenvolvimento nesse projeto que tem a função de ler os dados BUFR na tabela WMO de distribuição e escrever no formato PrepBUFR/ BUFR do NCEP para serem assimilados no GSI. O pacote BFGC vai estar focado inicialmente em gerar os dados de satélites, tais como as radiancias do RARS da DSA, dados de vento por satélites produzidos na DSA e também os dados de rádio ocultação GNSS recebidos em tempo real no CPTEC. Em uma segunda fase faz-se estudar as possibilidades de como isso deverá ser feito com os dados convencionais recebidos no pré do CPTEC. Algumas opções deverão ser estudadas, para as versões posteriores de distribuição.

### 1.3 Estrutura do documento

O presente relatório tem a função de descrever com riqueza de detalhes todas as implementações realizadas para a entrega da versão 1.0 do pacote do GCQD, descrita acima. Como a versão inicial 0.0 não foi publicada, todas as implementações realizadas desde o início do projeto serão descritas aqui desde a estruturação do ambiente de programação e a integração dos módulos em que o mesmo se divide. Em uma avaliação total desse desenvolvimento até essa versão o GCQD recebeu atenção no desenvolvimento nos seguintes aspectos:

- a) Uma metodologia de processamento dos dados para a conversão de formato de arquivos do padrão WMO para o padrão utilizado no sistema de assimilação usado atualmente no CPTEC;
- b) Uma estrutura de desenvolvimento modular para o processamento dos dados de diferentes fontes, sensores e técnicas de observação do sistema terrestre, com ferramentas de documentação e validação dos resultados;
- c) Um sistema de controle de qualidade complexo e global dos dados convencionais utilizando as previsões de curto prazo no modelo global do CPTEC.

Essa estrutura é utilizada para a inclusão de todos os tipos de dados disponíveis para a previsão de tempo do CPTEC, o qual deve ser feita de forma gradual e constante acompanhando a evolução dos sistemas de observações. Na atual versão os seguintes sistemas de observações são contemplados:

- a) Dados de radiâncias do sensor AMSU-A dos satélites GOES-18, GOES-19 e MetOp-B recebidos via RARS da DSA;
- b) Dados de vento por satélites no canal do vapor d'água e no infravermelho para os dados dos satélites GOES recebidos via RARS da DSA;
- c) Dados de estações a superfície (SINOP e METAR) recebidos via GTS no PRÉ e organizados no bando de dados MARS;

Para descrever todas essas implementações o presente documento é dividido nas seguintes seções: seção 2 apresenta a estrutura principal de todo o sistema GCQD, apresentado o fluxo dos processos, os módulos principais que o sistema está dividido como mencionado acima, bem como detalhes da obtenção, configuração, instalação e execução de todo o processo. A seção 3 descreve em detalhes o módulo de conversão de dados BUFR tabela WMO ((WORLD... , 2011)) para os dados PrepBUFR/BUFR tabela NCEP, utilizado na assimilação. Uma descrição da metodologia modular de desenvolvimento baseado em um módulo modelo é descrita, bem como outras ferramentas utilizadas para essa tarefa detalhando o passo-a-passo para a inclusão de um novo sistema de observação. Subseções são apresentadas para essas ferramentas. Outras subseções são apresentadas para descrever os quatro módulos principais de observações que são as radiâncias, vento por satélites, rádio ocultação GNSS e dados convencionais. Na seção 4 é descrito o pacote de controle de qualidade de dados usado nessa versão que é o PAQC, de forma detalhada apresentando a função de cada um dos módulos em que se divide. Na seção 5 são apresentadas as conside-

rações finais contemplando a estratégia de entrega das futuras versões do sistema e uma síntese do que deverá ser implementado nos próximos meses na continuidade da atividade.



## 2 Informações gerais sobre a versão V1.0 básica do GCQD

Data prevista para a release: 2 de dezembro de 2015.

Data que a versão foi lançada: 29 de janeiro de 2015.

Nessa seção será descrito em detalhes as informações de como foi implementado, como é o processamento e detalhes dos resultados gerados para facilitar seu uso na assimilação.

Na versão V1.0 básica toda a estrutura de desenvolvimento do sistema OBS foi montada, na qual além dos ingredientes básicos contidos na versão V0.0 inicial (diga-se um pacote de dados modelo chamado Dummy, os processamentos para a conversão de dados BUFRS tabela WMO, para PrepBUFRS/ BUFRS tabela NCEP, controle de qualidade dos dados PrepBUFRS e ferramentas para validar os arquivos BUFRS gerados), essa versão apresenta a inclusão de novas fontes de dados, tais como dados de estações a superfície, dados de rádio ocultação GNSS, dados de vento por satélites do canal de infravermelho para os satélites GOES e dados de radiância dos satélites NOAA 19 e MetOp-B via RARS. Da mesma forma que a versão inicial, essa versão também é constituída de dois pacotes principais: o BuFr GEenerator (BFGE) e o PAckage of Qualit Control (PAQC) do GMAO, esse disponível no pacote do G3DVAR. Cada um desses pacotes possuem um script principal, o `*run_bfge.ksh*` e o `*run_paqc.ksh*` respectivamente, que os executam sendo ambos sequencialmente chamados pelo script principal `*run_obs.ksh*` baseado em datas iniciais de um intervalo de tempo.

### 2.1 Estrutura do sistema GCQD

O Sistema GCQD é constituído, desde sua concepção por dois processos principais, BuFr GEenerator (BFGE) e o PAckage of Qualit Control (PAQC) do GMAO. Esses pacotes têm as seguintes funções:

- BFGE: tem a função de converter o formato de arquivo BUFR tabela WMO para o formato PrepBUFR/ BUFR tabela NCEP, a qual é o formato utilizado pelo GSI. Esse pacote é um fortran composto de diversos módulos independentes, sendo cada módulo responsável por um tipo de sistema de observação. Para o caso das radiâncias os módulos são separados por tipo de sensores tendo o mesmo arquivo de escrita resultante em cada módulo. No caso dos dados convencionais diversos podem ser os módulos baseados nos diferentes tipos de sistemas observacionais, no entanto

todos devem escrever em um mesmo arquivo de saídas dos dados. Para o caso dos dados de rádio ocultação GNSS, apenas um módulo é desenvolvido o qual contempla diversas constelações de satélites em diferentes sensores, mas apenas um arquivo de dados GPS é criado. Para facilitar o desenvolvimento e mantê-lo padronizado com a colaboração de um maior numero de pessoas, nessa versão Básica\_1.0 foi implementada utilizando a estrutura Dummy, desenvolvida na versão Inicial\_0.0, e módulos vazios de todos os tipos de sistemas de observações a serem contemplados na versão Robusta\_2.0. Nessa versão são contemplados também os dados de radiâncias do sensor AMSU provenientes dos satélites NOAA-19 e MetOp-B, bem como os dados de vento por satélites na canal do infra-vermelho compondo o PrepBUFR. Mais Detalhes do BFGE e seus módulos podem ser vistos em: <https://projetos.cptec.inpe.br/projects/paqc/wiki/BFGE>.

- PAQC: esse pacote utilizado aqui é oriundo do sistema G3DVAR. Para a compilação do mesmo foi necessário a compilação da biblioteca NCEP\_shared e a GMAO\_shared. A primeira foi a mesma obtida via external do svn apontando para o repositório do G3DVAR. A GMAO\_shared foi baixada do site do GMAO e incluída no pacote. Esse pacote é responsável por ler os arquivos PrepBUFR gerados no BFGE e comparando com os dados do background do G3DVAR faz uma avaliação e modifica as flags do controle de qualidade gerando um novo arquivo PrepBUFR. Diversos processos compõem o PAQC, o qual trata os sistemas de observação contidos no PrepBUFR separadamente e por fim todos os dados em conjunto. Detalhes do PAQC e seus componentes podem ser vistos em: <https://projetos.cptec.inpe.br/projects/paqc/wiki/PAQC>.

Na figura 2.1 é apresentado um esquema ilustrativo da sequencia em que os processos são chamados em um fluxograma dos scripts responsáveis por executar cada um dos dois processos descritos acima. Nessa figura, o script principal chamado run\_obs.sh é detalhado em especial os argumentos necessários para a sua execução, bem como o lastro principal de um período de datas a serem processadas. Um destaque é feito para o fluxo de dados de cada um desses processos, os quais são utilizados setas coloridas para separar os diferentes tipos de dados: dados convencionais, dados de radiância e dados de rádio ocultação GNSS. Os dados existentes no BUFR\_Tanque, disponibilizados pelo banco de dados Meteorológicos (DOP) e de satélites (DSA) no formato WMO são convertidos pelos BFGE em dados PrepBUFR/ BUFR tabela NCEP, próprios para a assimilação de dados. Como pode ser observado, apenas os

dados convencionais são tratados no PAQC. Um destaque adicional é feito nessa figura para um das ferramentas do BFGGE que é o pacote de validação dos arquivos gerados, denominado VALIDA, o qual é responsável por validar os arquivos gerados apresentando uma distribuição espacial e uma estatística básica dos dados incluídos no arquivo PrepBUFR/ BUFR NCEP gerado.

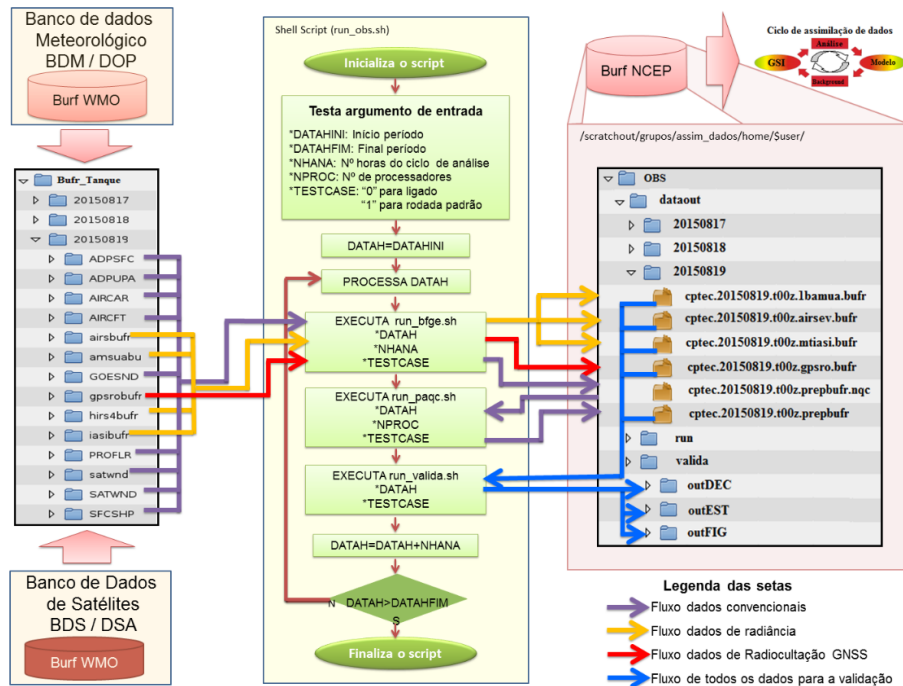


Figura 2.1 - Diagrama do sistema de assimilação de dados para o controle de qualidade nos modelos do CPTEC.

Para o desenvolvimento de todo o pacote, o mesmo foi dividido em diversas atividades para facilitar o desenvolvimento do projeto. A lista completa das atividades podem ser acessadas em gantt: <https://projetos.cptec.inpe.br/projects/paac/issues/gantt> da versão Basica\_1.0.

Para a instalação foram criados scripts de compilação/configuração e execução de todos os processos, os quais são detalhados nas próximas seções abaixo. Esses scripts praticamente não sofreram modificações com relação aos apresentados na versão Inicial\_0.0.

## 2.2 Obtenção/configuração/compilação e execução do pacote OBS

Nessa seção são apresentadas as informações para se fazer o download de todo o pacote na conta do usuário, bem como a configuração para criar a estrutura de

diretórios, a instalação dos pacotes incluindo as bibliotecas e finalmente a execução usando dados de um TestCase também disponível na distribuição.

### 2.2.1 Obtenção do sistema obs

O pacote contendo todos as rotinas e scripts de configuração/ instalação/ execução do sistema OBS está disponível no repositório SVN do CPTEC no ramo trunk, a atual versão em desenvolvimento e no tag a versão finalizada após a release da versão V0.0. Para obter essas versões do pacote desse sistema, deverão ser realizadas as seguintes etapas:

- a) Login no supercomputador TUPA:

```
ssh usuario@tupa.cptec.inpe.br -XC
```

- b) Acessar o diretório home do usuário no scratchin (/scratchin/grupos/assim\_dados/home/\$USER), o que pode ser feito via variável \$SUBMIT\_HOME:

```
cd $SUBMIT_HOME
```

- c) Uma outra possibilidade que foi viabilizada no desenvolvimento do pacote foi a instalação do OBS dentro da estrutura de diretórios do projeto RMS (Regional Modeling System) ou no futuro GMS (Global Modeling System), os quais o diretório (\$home/RMS/obs) já se encontra contemplado na estrutura dos diretórios. Se for essa a sua opção, depois de instalar o pacote RMS, acesse o diretório obs (\$home/RMS/obs):

```
cd $home/RMS/obs
```

- d) Extração do código do sistema OBS do ramo tag para o diretório escolhido usando o comando export:

```
svn export https://svn.cptec.inpe.br/paqc/tag/obs
```

- e) Para o caso onde a versão publicada teve que ter uma revisão para correção de bugs, informações sobre o número da revisão deverá estar adicionada no diretório obs no ramo tag. Ver mais detalhes sobre essas revisões no item acima. Nesse caso um erro ocorrerá ao digitar o comando acima, por indicar que a versão (obs) não exista. Use o comando list para avaliar a última versão e escolha-la com o comando anterior. Para listar as versões,

digite:

```
svn list https://svn.cptec.inpe.br/paqc/tag
```

- f) Caso o usuário queira e possua habilidade e permissões para contribuir com a versão em desenvolvimento o acesso ao pacote deve ser via comando checkout para acionar o controle das alterações na versão. Para isso deve-se acessar a versão disponível no trunk:

```
svn co https://svn.cptec.inpe.br/paqc/trunk/obs
```

**Nota 1:** O repositório do projeto obs no trunk está vinculado aos repositórios do G3DVAR para evitar a duplicação desnecessária da biblioteca NCEP\_shared que já está incluída no repositório SVN. Dessa forma, ao baixar o obs, o svn vai nos repositórios do G3DVAR e baixa os diretórios desejados para a estrutura de diretórios do sistema. A função do svn utilizada para isso é a externals.

Observe que é preciso que o usuário nessa versão trunk seja incluído no projeto G3DVAR para ter acesso ao repositório. Na versão tag esses repositórios não são linkados via external e sim copiados para a versão tag de distribuição.

**Nota 2:** Na construção da estrutura do sistema OBS foi incluída a possibilidade de que o mesmo possa ser adicionado como um pacote externo nos sistemas de assimilação de dados, como por exemplo no sistema RMS. Para isso foi pensando na utilização do comando external do svn, no qual o sistema OBS possa ser incluído na lista do sistema.

### 2.2.2 Instalação do sistema OBS

Como nos outros pacotes, o pacote de configuração/ instalação do OBS foi organizado seguindo a mesma estratégia de implementação adotada no GDAD. Nela o script config\_obs.ksh é responsável por fazer todas as etapas de configuração e instalação de todos os pacotes que envolve o Sistema de Gestão e Controle de Qualidade de Dados para a Assimilação (GCQD). A estrutura é toda montada sobre o diretório obs que é baixada no scratchin via SVN (feita na seção anterior) e ao configurar o sistema uma árvore de diretório no scratchout é criada. Isso é feita para manter a estratégia de processamento arquitetada na concepção da máquina Tupã. Nessa estratégia os códigos e scripts, mantidos em segurança no SVN sobre controle de versões, é compilado no scratchin, de onde é rodado o script principal, denominado run\_obs.ksh, que tem a função de enviar para o scratchout todos os scripts, exe-

cutáveis e demais arquivos necessários para o processamento, onde são executados envolvendo submissão de processos para os nós do Tupã. Os arquivos de resultados finais (com maior fluxo de dados) são enviados para a árvore de diretório do obs no scratchout de onde são enviados para fitas ou discos externos. Observe que com essa estratégia nada é enviado para o home do usuário, pois embora seja uma área com backup, a mesma tem restrição de espaço o que limita a sua utilização envolvendo diversos sistemas.

### 2.2.3 Detalhes do script de instalação do sistema obs

Para a instalação do sistema OBS utiliza-se o script `config_obs.ksh`, o qual executa diferentes tarefas definidas via argumento de entrada. Uma descrição detalhada de cada uma das opção segue abaixo:

#### A opção ajuda do `config_obs.ksh`

Seguindo as instruções de ajuda (ou caso nenhum argumento de entrada é fornecido), apresenta uma lista das possíveis opções. Essa lista é apresentada na figura 2.2:

```
luiz.sapucci@eslogin10:/scratchin/grupos/assim_dados/home/luiz.sapucci/obs> ./config_obs.ksh ajuda
> Opcao escolhida: ajuda

Uso.....: ./config_obs.ksh <opcao>

Opcoes...: <opcao> configurar -> configura o OBS (estrutura de pastas e arquivos)
           compilar    -> compila o bfgc + paqc (cria executaveis)
           clean_all   -> limpa executaveis
           vars_export  -> exporta variaveis
           ajuda       -> apresenta uma ajuda

Exemplos: ./config_obs.ksh configurar
           ./config_obs.ksh compilar
           ./config_obs.ksh clean_all
           ./config_obs.ksh vars_export
           ./config_obs.ksh ajuda

PS. Aconselha-se a utilizacao das opcoes em ordem descrita acima.
    No caso da opcao vars_export utilize-a apenas depois de configurar e compilar.

> Para mais informacoes sobre esta distribuicao do GCOD
> leia a wiki do projeto https://projetos.cptec.inpe.br/projects/paqc/wiki
```

Figura 2.2 - Lista de opções do sistema.

**A opção configurar do `config_obs.ksh`** A opção configurar tem a função de preparar a estrutura de diretórios no scratchout do Tupã. Exporta as variáveis chamando a opção `vars_export` (ver mais detalhes abaixo) e cria os diretórios que as variáveis estão associadas. Para que esse sistema seja versátil e possa ser instalado em qualquer outro sistema como parte do mesmo, foi implementado de forma que na configuração a função `var_export` permite que seja passado para o sistema o en-

dereço onde se quer rodar o obs e gerar os resultados. Seria a variável \$obs\_work. Por exemplo, para o caso do mesmo ser integrado ao RMS o endereço adequado deveria ser \$WORK\_HOME/RMS, no qual seria criado o diretório obs. Para o caso convencional esse endereço seria o home do usuário no scratchout (\$WORK\_HOME).

**A opção vars\_export do config\_obs.ksh** A opção vars\_export exporta uma sequência de variáveis associadas a todo o sistema, as quais são utilizadas pelos demais scripts integrados no sistema e com essa opção os mesmos conseguem carregar essa mesma estrutura de variáveis mantendo uma padronização. Esse script com essa opção pode ser até colocado no .login do usuário e se os nomes das variáveis for por ele memorizadas pode permitir uma navegação por todas as componentes do obs. Existe uma lógica na composição das avariáveis para facilitar essa memorização. A sequência principal de diretórios é:

- Do obs no scratchin:
  - subt\_obs ..... igual ao obs\_home que é o home da instalação do obs no Tupã
  - obs\_run ..... o diretório onde é executado o obs
  - obs\_bin ..... o diretório dos binários principais do obs
  - obs\_src ..... o diretório raiz dos fontes do sistema
- Do pacote BFGE no scratchin:
  - bfge\_src ..... diretório do source do bfge
  - bfge\_src\_conv ..... diretório do source do módulo de dados convencionais
  - bfge\_src\_nooa\_amsua ..... diretório do source do módulo de dados de radiância do amsua
- Do pacote PAQC no scratchin:
  - paqc\_src ..... o diretório do source root do PAQC
  - paqc\_2src ..... o diretório do source do PAQC
  - paqc\_bin ..... o diretório dos binários do PAQC

- paqc\_etc ..... o diretório das libs e outros do PAQC
- Diretórios de entrada de dados do pré TANQUE
  - TANQUE\_DUMMY ..... Diretório de entrada dos dados do TestCase (Tanque) na estrutura pré definida (setada para /scratchin/grupos/assim\_dados/home/gdad/DataFix/OBS/testcase/Bufr\_Tanks)
  - obs\_datain ..... Diretório onde os dados BUFERS de entrada devem estar disponíveis operacionalmente.
- Do obs no scratchout:
  - obs\_work\_run ..... diretório root de trabalho do obs
  - obs\_work\_run\_bfge ..... diretório de trabalho do BFGE
  - obs\_work\_run\_paqc ..... diretório de trabalho do PAQC
  - obs\_valida ..... diretório de saída dos dados da validação
  - obs\_dataout ..... diretório de saída dos dados PrepBUFERS e BUFERS para a assimilação no GSI.
- Diretórios do G3DVAR
  - work\_gsi\_datain\_bkg ..... dados do background do G3DVAR entre outros diretórios
- Utilitários
  - inctime ..... ferramenta para incremento de tempo disponível no pacote do G3DVAR (/stornext/home/\$USER/G3DVAR/util/inctime/inctime)
  - public\_gdad ..... diretório de dados públicos do user gdad (/scratchin/grupos/assim\_dados/home/gdad/public)
  - Fix\_obs ..... diretório de dados fixos para o obs também no user gdad (/scratchin/grupos/assim\_dados/home/gdad/DataFix/OBS)

**Nota 1.** Esse pacote precisa do IncTime instalado no local indicado pela variável



definida no `config_obs.ksh` apontado para a instalação no G3DVAR. Caso não tenha o sistema G3DVAR instalado um erro deverá ocorrer. Esse mesmo aplicativo está instalado no sistema RMS, e se for esse o caso deve-se ajustar o endereço. Caso não tenha a instalação desses sistemas, faça a instalação desse aplicativo baixando uma versão do svn no endereço <https://svn.cptec.inpe.br/gdad/jgerd/tags/inctime> e siga as instruções de instalação no `Readme`.

**A opção compilar do `config_obs.ksh`** A opção compilar tem a função de instalar todas as componentes necessárias para funcionar o OBS utilizando os executáveis gerados nos pacotes BFGE e PAQC. Inicialmente verifica se está no `eslogin01` (pré-requisito para a instalação dos pacotes no Tupã), carrega os módulos do Tupã usando a configuração pré-definida e testada das versões dos compiladores e bibliotecas do sistema, exporta as variáveis de ambiente (chamando a função `vars_export`), compila os módulos do BFGE que são: `bufrlib`, `mbufr-adt`, `f90lib`, o principal denominado `prepobs` e finalmente o `decode` para a validação. Depois disso verifica se os arquivos executáveis foram gerados. São eles:

- `$obs_bin/prepobs.x`
- `$obs_bin/prepbufr_decode_conv.x`
- `$obs_bin/bufr_decode_radiance.x`
- `$obs_bin/valida_dec_estat_prepbufr.x`

Após a instalação do BFGE, o script compila o PAQC. Para isso o mesmo exporta algumas variáveis de ambiente e executa o comando `"gmake install"` no diretório do source do PAQC (`$paqc_2src`). As saídas da tela com as informações da instalação são armazenadas no arquivo `$paqc_src/compilePAQC.log`. Um teste dos arquivos executáveis gerados é feito para avaliar o sucesso do processo:

- `$paqc_bin/echorc.x`
- `$paqc_bin/combfrd.x`
- `$paqc_bin/scanbuf0.x`
- `$paqc_bin/ssprepqc`
- `$paqc_bin/zeit_ci.x`

- \$paqc\_bin/zeit\_co.x
- \$paqc\_bin/zeit\_pr.x
- \$paqc\_bin/ssprevents.x
- \$paqc\_bin/cqcbufr.x
- \$paqc\_bin/raobcore.x
- \$paqc\_bin/hradcor.x
- \$paqc\_bin/cqcvad.x
- \$paqc\_bin/profcqc.x
- \$paqc\_bin/acarsqc.x
- \$paqc\_bin/oiqcbufr.x
- \$paqc\_bin/prepacqc.x

### **3 Pacote BuFr GEnerator (BFGE)**

Esse pacote tem a função de converter os dados em formato BUFR da tabela WMO para o formato PrepBUFR/ BUFR tabela NCEP, o qual é o formato utilizado pelo GSI. Embora as conversões do formato de arquivo sejam semelhantes entre os diferentes tipos de dados, bastando apenas que sejam identificados os códigos de leitura WMO e escrita na tabela NCEP, todo o processo é separado em módulos principais por tipo de dados e tratados separadamente. Isso é feito para tornar mais fácil as futuras modificações no código, para que um tipo de dado não interfira em outro. Essa separação modular, embora às vezes deixe o código pouco otimizado com algumas rotinas semelhantes, permite uma separação do fluxo e facilita a interpretação do código. As futuras versões provavelmente deverão caminhar para um código mais eficiente. A estratégia de implementação adotada foi que apenas uma rotina principal denominada `bfge.f90` seja responsável por executar todos os módulos e gerar assim os diferentes tipos de arquivos. A estratégia adotada na concepção do projeto foi que cada sistema de observação fosse tratado separadamente em módulos distintos, os quais os mais semelhantes sejam agrupados em módulos maiores, como por exemplo os sensores de radiancias em módulos distintos, mas agrupados em um módulo de processamento de radiancias.

#### **3.1 Estratégia de desenvolvimento do BFGE**

Para o desenvolvimento dos módulos específicos para cada sistema de observação no BFGE foi criado um Dummy composto basicamente de uma rotina de configuração, uma de seleção dos dados e uma terceira de escrita. O desenvolvimento de um novo sistema de observação a ser incluído no BFGE deve tratar das seguintes tarefas:

- Levantamento dos dados e estudo da estrutura de escrita do sistema de observação e sua utilização na assimilação do GSI;
- Preparação do ambiente de implementação no repositório do pacote usando o pacote Dummy;
- Organização dos dados no TestCase padrão do pacote;
- Implementação do código obedecendo a estrutura Dummy de desenvolvimento e as exigências para as entregas no SVN;
- validar os resultados usando o pacote de validação e implementar funcionalidade de o dado exigir função ainda não contempladas no pacote;

-gerar uma página wiki documentando em detalhes todo o desenvolvimento realizado relacionado com as revisões realizadas no SVN.

Para isso algumas ferramentas básicas destinadas para apoiar o desenvolvimento foram elaboradas, as quais são listadas abaixo:

- Estrutura Dummy: para facilitar o desenvolvimento e mantê-lo padronizado com a colaboração de um maior número de pessoas, foi implementada uma estrutura Dummy. Essa estrutura deve ser usada como padrão de desenvolvimento de módulos para a conversão de dados BUFRS da WMO para PrepBUFR/ BUFR do NCEP, com recursos de leitura (BUFR-WMO) processamento e escrita de dados (PrepBUFR). Essa estrutura deve ser utilizada para a inclusão de novos sistemas observacionais, a qual permite uma padronização do código mantendo-o estruturado.
- Estrutura Dummy\_bufr\_tank: para os testes iniciais de avaliação do pacote foi organizado um depósito modelo de arquivos BUFR WMO, os quais devem conter um conjunto básico de dados para os testes iniciais de todos os sistemas de observações contemplados no BFGÉ. Esse depósito foi criado no diretório DataFix do usuário GDAD. Nesse Tanque Dummy foi definida a estrutura de diretórios, os quais o tanque deve obedecer quando estiver na operação junto a DOP. Com a inclusão de novos sistema de observações, os dados do TestCase devem ser colocados nesse tanque.
- Validação do PrepBUFR/ BUFR gerados: visando a avaliação básica dos produtos gerados no pacote, também é disponibilizado um pacote de validação dos arquivos PrepBUFRS/ BUFRS gerados, no qual uma distribuição espacial das observações é apresentada em forma de figuras e valores de estatísticas básicas são tabelados em arquivos de saída.

Essas ferramentas são detalhadas nas próximas subseções.

### **3.1.1 Estrutura Dummy**

A estrutura básica do modelo Dummy consiste na criação de um pacote inicial estruturado em módulos e sub-rotinas que tem por objetivo de auxiliar o desenvolvedor na criação de novos pacotes de outros tipos de observação a partir desse pacote modelo, facilitando assim todo o esquema de estruturação do programa em geral, sendo possível com maior facilidade incluir novas funcionalidades de observações.

O pacote Dummy é atualmente separado em dois módulos e faz a utilização das sub-rotinas de alguns módulos independentes (módulo de variáveis globais, módulo de leitura de dados BUFR e módulo de escrita de dados PrepBUFR).

Os módulos principais do Dummy são: m\_Dummy\_select e m\_Dummy\_conf.

O módulo m\_Dummy\_conf contém a sub-rotina que faz a leitura do arquivo de configuração denominado namelist\_Dummy.PREPOBS apresentado na figura 3.1:

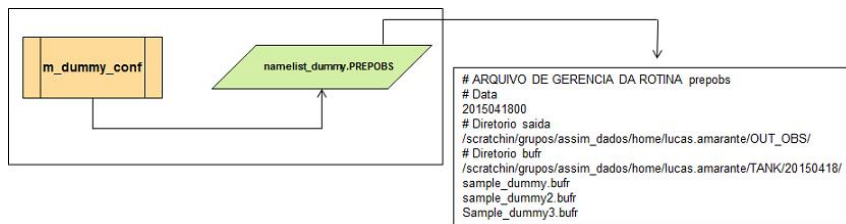


Figura 3.1 - Diagrama do modulo para o sistema Dummy.

A sub-rotina específica nesse módulo faz a leitura do arquivo de configuração e armazena o resultado em algumas variáveis que serão utilizados pelos outros módulos do programa, tais como os diretórios de entrada e saída dos arquivos BUFRS. O arquivo namelist tem uma configuração que deve ser respeitada, como inclusão de # no início de linhas não utilizadas, e também devem estar de acordo com a descrição da linha acima desses parâmetros.

O módulo principal m\_Dummy\_select consiste num módulo com uma sub-rotina responsável em executar a chamada de outras sub-rotinas localizadas nos módulos independentes do programa, que são:

- Módulo de variáveis globais: Esse módulo contém as variáveis que são necessárias utilizar ao longo das sub-rotinas de outros módulos. Foi preciso a criação de um módulo específico como esse porque havia necessidade de utilizar a mesma variável preenchida com o valor em outras sub-rotinas, visto que quando a declaração é feita dentro de uma sub-rotina, a mesma perde seu valor em outras sub-rotinas de outros módulos.
- Sub-rotina no módulo de leitura do arquivo de configuração namelist: Essa sub-rotina pertence ao pacote Dummy, conforme explicada acima faz a leitura do arquivo de configuração e armazena o resultado em algumas variáveis que serão utilizados pelos outros módulos do programa.

- Sub-rotina no módulo de leitura de dados BUFRS: A sub-rotina de leitura de dados BUFRS faz a utilização dos módulos de criação e decodificação de BUFRS da ferramenta mbuf\_r\_tools (distribuída pelo CPTEC/INPE: <http://downloads.cptec.inpe.br/publicacoes/detalhes.jsp>) e é uma adaptação de um dos programas dessa ferramenta denominado bufdump. A partir dos parâmetros informados no namelist, o dado BUFR é decodificado pelos módulos de decodificação a partir da leitura de arquivos tabela, que são um conjunto de arquivos em formato texto com a descrição dos códigos referentes a informação dentro do arquivo BUFR.

Essas tabelas são fundamentais para a correta decodificação do dado e segue um exemplo abaixo de um desses descritores (trecho retirado de um arquivo tabela):

002153	SATELLITE CHANNEL CENTRE FREQUENCY	Hz	-8	0	26	
002154	SATELLITE CHANNEL BAND WIDTH	Hz	-8	0	26	
002163	HEIGHT ASSIGNMENT METHOD	CODE TABLE	2163	0	0	4
002164	TRACER CORRELATION METHOD	CODE TABLE	2166	0	0	4
002166	RADIANCE TYPE	CODE TABLE	2166	0	0	4
002167	RADIANCE COMPUT METHOD	CODE TABLE	2167	0	0	4
004001	YEAR	YEAR	0	0	12	
004002	MONTH	MONTH	0	0	4	
004003	DAY	DAY	0	0	6	
004004	HOUR	HOUR	0	0	5	
004005	MINUTE	MINUTE	0	0	6	
004006	SECOND	SECOND	0	0	6	
004011	TIME INCREMENT	YEAR	0	-1024	11	

A decodificação desses dados são armazenados em variáveis tipo e seus valores são passados para a sub-rotina seguinte de escrita de dados PrepBUFRS.

- Sub-rotina no módulo de escrita de dados PrepBUFRS: Essa é a última sub-rotina utilizada no processo de geração de dados PrepBUFRS, sendo a mesma responsável por adicionar os valores das variáveis vindas da leitura do BUFR aos parâmetros específicos da lista de mnemônicos, os quais são descritos em algumas das tabelas NCEP que está localizada no link [http://www.emc.NCEP.noaa.gov/mmb/data\\_processing/prepbuf\\_r.doc/table\\_1.htm](http://www.emc.NCEP.noaa.gov/mmb/data_processing/prepbuf_r.doc/table_1.htm).

Na figura 3.2 pode-se ver um exemplo de alguns mnemônicos:

String name	MNEMONIC	NUMBER	DESCRIPTION
hdstr	SID	001192	STATION IDENTIFICATION
	XOB	006002	LONGITUDE
	YOB	005002	LATITUDE
	DHR	004192	OBSERVATION TIME MINUS CYCLE TIME
	TYP	001193	PREPBUFR REPORT TYPE
	ELV	010194	STATION ELEVATION
	SAID	001007	SATELLITE IDENTIFIER (SATELLITE REPORTS ONLY)
obstr	POB	007192	PRESSURE OBSERVATION
	QOB	013192	SPECIFIC HUMIDITY OBSERVATION AFTER "VIRTMP" STEP - ALWAYS RECALCULATED FROM QUALITY CONTROLLED VIRTUAL TEMPERATURE DATA)
	TOB	012192	TEMPERATURE OBSERVATION(AFTER "PREPRO" STEP - REPORTED TEMP, EITHER SENSIBLE OR VIRTUAL DEPENDING UPON DATA TYPE AFTER "VIRTMP" STEP - VIRTUAL TEMPERATURE IF MOISTURE AVAILABLE, OTHERWISE SENSIBLE)
	ZOB	010196	HEIGHT OBSERVATION
	UOB	011003	U-COMPONENT WIND OBSERVATION
	VOB	011004	V-COMPONENT WIND OBSERVATION
	PWO	013213	TOTAL PRECIPITABLE WATER OBSERVATION
	CAT	001194	PREPBUFR DATA CATEGORY
	PRSS	010195	SURFACE PRESSURE OBSERVATION

Figura 3.2 - Lista de mnemonics das tabelas NCEP.

Essa sub-rotina é uma adaptação de um dos programas da ferramenta bufrtools (disponibilizada em <http://www.dtcenter.org/com-GSI/users/downloads/index.php>).

O esquema do funcionamento do Dummy é apresentado na figura 3.3.

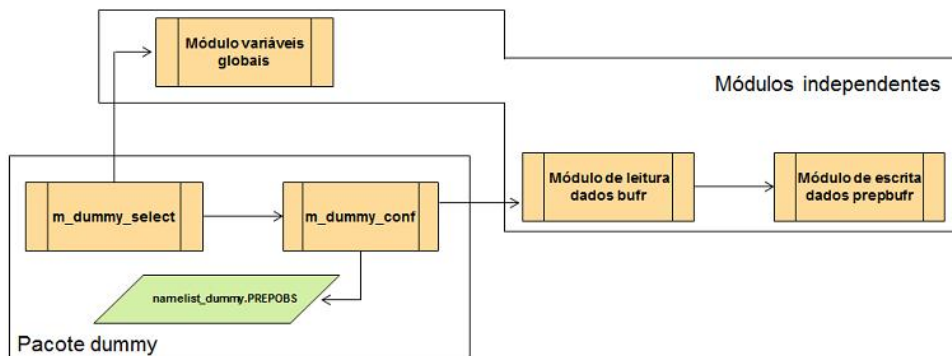


Figura 3.3 - Esquema da ferramenta bufrtools.

A ideia inicial é que toda a execução do pacote Dummy será realizada pelo programa principal que executa a sub-rotina principal do Dummy localizada no `m_Dummy_select`. E a medida que desenvolve esses procedimentos utilizando o Dummy para outros tipos de observação as sub-rotinas serão incluídas no código do programa principal para a execução das mesmas uma por vez, conforme na figura 3.4 :

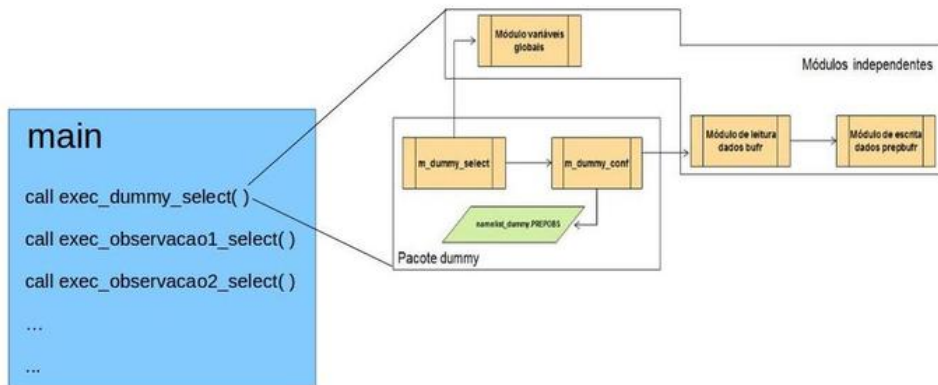


Figura 3.4 - Esquema principal do sistema Dummy.

### 3.1.2 Modelo de Tanque de dados BUFERS para a assimilação

Essa seção apresenta a criação de um diretório contendo os dados para uma rodada de teste da versão em distribuição, bem como orientar o ambiente onde os dados deverão estar disponíveis para a operacionalização de todo o pacote de processamento do GCQD. Assim foi criado um diretório chamado de BufrTank no scratchin do usuário gdat no diretório DataFix como é o procedimento padrão de desenvolvimento dos demais projetos do grupo, sendo a ideia conceitual de um repositório onde os dados disponíveis para a assimilação sejam colocados da forma como mostra a figura 3.5.



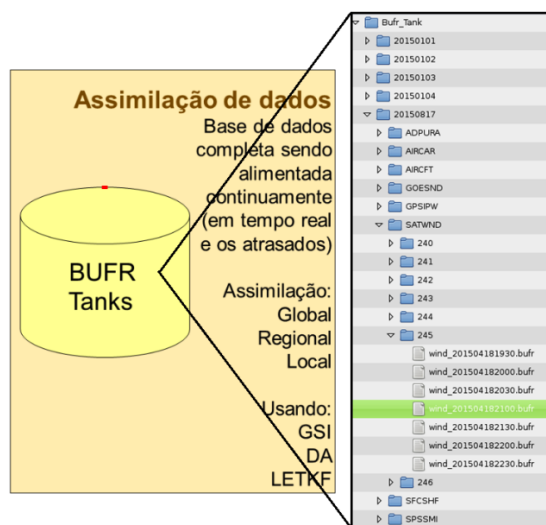


Figura 3.5 - Esquema para armazenamento de arquivos BUFR.

Para isso foi criado o diretório OBS ao lado do GSI e WRF no DataFix /scratchin/-grupos/assimdados/home/gdad/DataFix/OBS/

Nele foi criado o TestCase /scratchin/grupos/assimdados/home/gdad/Data-Fix/OBS/ testcase

No caso do projeto obs para o TestCase são necessários os dados do BufrTank /scratchin/grupos/assimdados/home/gdad/DataFix/OBS/testcase/BufrTanks

Nesse diretório, subdiretórios para cada data foram criados com o formato YYYYMMDD, nos quais subdiretórios com os mnemônicos dos dados são criados seguindo a estrutura de diretórios idealizado na coNCEPção da proposta e ilustrada na figura 3.5. Os detalhes dos mnemônicos são apresentados nas próximas sub-seções.

Nesse diretório foi colocado também um README com todas as ideias de como foi organizado os diretórios no bufrtank.

### 3.1.3 Estrutura de diretórios para os dados convencionais

A relação dos mnemônicos utilizada para a criação dos diretórios dos dados convencionais foi a lista na tabela do NCEP disponível em <http://www.emc.NCEP.noaa.gov/mmb/dataprocessing/prepbufr.doc/table1.htm> as quais são destacadas na figura 3.6.

**THE FOLLOWING ARE TABLE A ENTRIES FOR PREPBUFR MESSAGE TYPES**

<b>MNEMONIC</b>	<b>NUMBER</b>	<b>DESCRIPTION</b>
ADPUPA	A48102	UPPER-AIR (RAOB, PIBAL, RECCO, DROPS) REPORTS
AIRCAR	A48103	MDCRS ACARS AIRCRAFT REPORTS
AIRCFT	A48104	AIREP, PIREP, AMDAR, TAMDAR AIRCRAFT REPORTS
SATWND	A48105	SATELLITE-DERIVED WIND REPORTS
PROFLR	A48106	WIND PROFILER AND ACOUSTIC SOUNDER (SODAR) REPORTS
VADWND	A48107	VAD (NEXRAD) WIND REPORTS
SATEMP	A48108	POES SOUNDING, RETRIEVAL, RADIANCE DATA (TOVS)
ADPSFC	A48109	SURFACE LAND (SYNOPTIC, METAR) REPORTS
SFCSHP	A48110	SURFACE MARINE (SHIP, BUOY, C-MAN PLATFORM) REPORTS
SFCBOG	A48111	MEAN SEA-LEVEL PRESSURE BOGUS REPORTS
SPSSMI	A48112	DMSP SSM/I RETRIEVAL PRODUCTS (REPROCESSED WIND SPEED, TPW)
SYNDAT	A48113	SYNTHETIC TROPICAL CYCLONE BOGUS REPORTS
ERS1DA	A48114	ERS SCATTEROMETER WIND DATA (REPROCESSED)
GOESND	A48115	GOES SOUNDING, RETRIEVAL, RADIANCE DATA
QKSWND	A48116	QUIKSCAT SCATTEROMETER WIND DATA (REPROCESSED)
MSONET	A48117	MESONET SURFACE REPORTS
GPSIPW	A48118	GLOBAL POSITIONING SATELLITE-INTEGRATED PRECIPITABLE WATER REPORTS
RASSDA	A48119	RADIO ACOUSTIC SOUNDING SYSTEM (RASS) VIRTUAL TEMPERATURE PROFILE REPORTS
WDSATR	A48120	WINDSAT SCATTEROMETER WIND DATA (REPROCESSED)
ASCATW	A48121	ASCAT SCATTEROMETER DATA (REPROCESSED)

Figura 3.6 - Tabela de mnemônicos do NCEP para os arquivos BUFR.

Embora nem todos os tipos de dados são utilizados, a lista implementada dos subdiretórios para os dados convencionais foi a seguinte:

ADPUPA AIRCAR AIRCFT SATWND PROFLR VADWND SATEMP ADPSFC  
SFCSHP SFCBOG SPSSMI SYNDAT ERS1DA GOESND QKSWND MSONET  
GPSIPW RASSDA WDSATR ASCATW

### 3.1.4 Estrutura de diretórios para os dados de satélites

Para os dados de satélites foi utilizado os mesmos mnemônicos utilizados no GSI na ingestão de seus respectivos dados. Nessa lista é destacado o sensor ou a técnica e os satélites envolvidos podem estar separados em subdiretórios, o que poderá variar significativamente para cada um deles. A necessidade dessa subdivisão fica para ser decidida na inclusão de cada tipo de dado. Essa lista do GSI está disponível no arquivo GSIUserguide.pdf (<http://www.dtcenter.org/com-GSI/users/docs/usersguide/GSIUserGuidev3.3.pdf>) do DTC, mais especificamente na página 18, onde são apresentados todos os tipos de dados. Essa tabela é transcrita aqui na figura 3.7:

GSI Name	Content	Example file names
prepbufr	Conventional observations, including ps, t, q, pw, uv, spd, dw, sst	gdas1.t12z.prepbufr
satwnd	satellite winds observations	gdas1.t12z.satwnd.tm00.bufr_d
amsuabufr	AMSU-A 1b radiance (brightness temperatures) from satellites NOAA-15, 16, 17,18, 19 and METOP-A/B	gdas1.t12z.1bamua.tm00.bufr_d
amsubufr	AMSU-B 1b radiance (brightness temperatures) from satellites NOAA-15, 16,17	gdas1.t12z.1bamub.tm00.bufr_d
radarbufr	Radar radial velocity Level 2.5 data	ndas.t12z.radwnd.tm12.bufr_d
gpsrobufr	GPS radio occultation and bending angle observation	gdas1.t12z.gpsro.tm00.bufr_d
ssmirrbufr	Precipitation rate observations from SSM/I	gdas1.t12z.spssmi.tm00.bufr_d
tmirrbufr	Precipitation rate observations from TMI	gdas1.t12z.sptrmm.tm00.bufr_d
sbuvbufr	SBUV/2 ozone observations from satellite NOAA-16, 17, 18, 19	gdas1.t12z.osbuv8.tm00.bufr_d
hirs2bufr	HIRS2 1b radiance from satellite NOAA-14	gdas1.t12z.1bhirs2.tm00.bufr_d
hirs3bufr	HIRS3 1b radiance observations from satellite NOAA-16, 17	gdas1.t12z.1bhirs3.tm00.bufr_d
hirs4bufr	HIRS4 1b radiance observation from satellite NOAA-18, 19 and METOP-A/B	gdas1.t12z.1bhirs4.tm00.bufr_d
mnsbufr	MSU observation from satellite NOAA 14	gdas1.t12z.1bmnsu.tm00.bufr_d
airsbufr	AMSU-A and AIRS radiances from satellite AQUA	gdas1.t12z.airsev.tm00.bufr_d
mhsbufr	Microwave Humidity Sounder observation from NOAA-18, 19 and METOP-A/B	gdas1.t12z.1bmhs.tm00.bufr_d
ssmitbufr	SSM/I observation from satellite f13, f14, f15	gdas1.t12z.ssmi.tm00.bufr_d
amsrebufr	AMSR-E radiance from satellite AQUA	gdas1.t12z.amsre.tm00.bufr_d
ssmisbufr	SSMIS radiances from satellite f16	gdas1.t12z.ssmis.tm00.bufr_d
gsnd1bufr	GOES sounder radiance (sndrd1, sndrd2, sndrd3 sndrd4) from GOES-11, 12, 13, 14, 15.	gdas1.t12z.goesfv.tm00.bufr_d
l2rwbufr	NEXRAD Level 2 radial velocity	ndas.t12z.nexrad.tm12.bufr_d
gsndrbufr	GOES sounder radiance from GOES-11, 12	gdas1.t12z.goesnd.tm00.bufr_d
gimgrbufr	GOES imager radiance from GOE- 11, 12	
omibufr	Ozone Monitoring Instrument (OMI) observation NASA Aura	gdas1.t12z.orni.tm00.bufr_d
iasibufr	Infrared Atmospheric Sounding Interfero-meter sounder observations from METOP-A/B	gdas1.t12z.mtiiasi.tm00.bufr_d
gomebufr	The Global Ozone Monitoring Experiment (GOME) ozone observation from METOP-A/B	gdas1.t12z.gome.tm00.bufr_d
mlsbufr	Aura MLS stratospheric ozone data from Aura	gdas1.t12z.mlsbufr.tm00.bufr_d
tcvtil	Synthetic Tropic Cyclone-MSLP observation	gdas1.t12z.syndata.tcvitals.tm00
seviribufr	SEVIRI radiance from MET-08,09,10	gdas1.t12z.sevcsr.tm00.bufr_d
atmsbufr	ATMS radiance from Suomi NPP	gdas1.t12z.atms.tm00.bufr_d
crisbufr	CRIS radiance from Suomi NPP	gdas1.t12z.cris.tm00.bufr_d
modisbufr	MODIS aerosol total column AOD observations from AQUA and TERRA	

Figura 3.7 - Tabela de mnemônicos do NCEP para dados de satélites.

A relação completa dos mnemônicos (primeira coluna na tabela da figura 3.7) foi reduzida para uma relação mais enxuta com os dados mais importantes, a qual pode ser ampliada futuramente na medida da necessidade com a inclusão de dados não considerados nessa lista. Uma relação mais enxuta utilizada é: satwnd amsuabufr amsubufr radarbufr gpsrobufr hirs3bufr hirs4bufr airsbufr mhsbufr iasibufr atmsbufr crisbufr modisbufr

### 3.1.5 Script para gerar a estrutura de subdiretórios no BufrTank

Para facilitar a estruturação de novos períodos de dados foi desenvolvido um script para criar a estrutura de diretórios no BufrTank, que chama-se criasubdir.ksh e tem as seguintes funções:

- 1- Criar um diretório para a data;

- 2- Criar os subdiretórios para os dados convencionais e de satélites para cada data;
- 3- Criar os subdiretórios para os subtipos em que cada um dos tipos de dados são composto. No momento (24/09/2015) essa terceira tarefa não foi ainda implementada.

Os parâmetros de entrada são as data inicial e final do período desejado na forma YYYYMMDD.

### **3.1.6 Validação PrepBUFR Dummy**

O processo de validação do PrepBUFR a ser desenvolvido deverá abordar três aspectos. O primeiro uma decodificação e posteriormente visualização dos dados de forma comparativa com os dados contidos no PrepBUFR do NCEP. Uma segunda estatística básica para avaliar quantitativamente os padrões que são captados pelos dados e se valores espúrios ou muito discrepantes são observados através de cálculos estatísticos e valores de máximo e mínimo. Obviamente que diferenças deverão haver pois um tem já o controle de qualidade e o outro aqui gerado nessa fase ainda não. Uma terceira fase os dados gerados devem passar para serem assimilados e observar se os dados foram assimilados, para verificar se erros gritantes não foram adicionados no processo. As três fases serão desenvolvidas e descritas com detalhes abaixo.

### **3.1.7 Visualização PrepBUFR**

Visando avaliar a distribuição espacialmente das observações dentro dos PrepBUFRS gerados pelo pacote obs e poder compará-los com dados PrepBUFRS de outros centros de processamento foi criado uma estrutura de programas para a visualização das observações a partir de coordenadas (latitudes e longitudes). Esse pacote utiliza-se de shell script, programa em fortran e script em Grads, apresentado na figura 3.8 os diretórios correspondentes e os programas, o qual faz uma boa interação entre essas linguagens de programação e possuem certa dependência entre eles para que seja alcançado o objetivo final da figura.

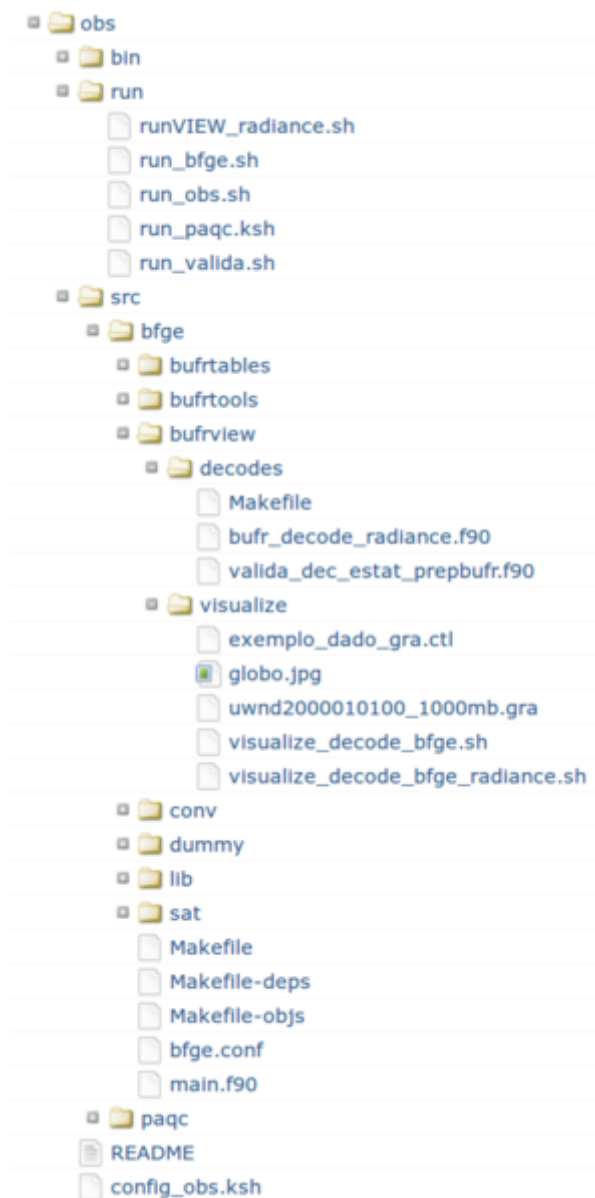


Figura 3.8 - Sub-diretórios do sistema de validação Dummy.

Todo o funcionamento é gerenciado pelo script `run_valida.sh` que se localiza no diretório `run`. Os programas de decodificação e visualização ambos se encontram no diretório `bufrview` e então é dividido em `decodes` e `visualize`. O funcionamento desses processos ocorre da seguinte forma:

O script principal cria um `namelist` de configuração inicial que será lido pelo programa de decodificação, com os diretórios dos `PrepBUFERS` criados e o local de saída dos arquivos que serão decodificados bem como a lista dos `PrepBUFERS` a serem executados. Então executado o programa de decodificação `valida_dec_es-`

tat\_prepbuf.x que é compilado a partir do programa em fortran valida\_dec\_estat\_prepbuf.f90, que está no diretório bin e src/bfge/bufrview/decodes, para gerar os arquivos texto com as informações dos PrepBUFRS. A fim de filtrar somente as observações desejadas é acrescentado em um vetor os tipos de observações que serão utilizadas para a filtragem e assim gerar novos arquivos texto temporários apenas com as observações requeridas. Após todo o processo de decodificação e filtragem dos dados é então executado o programa em shell script visualize\_decode\_bfge.sh localizado no diretório visualize, que faz os procedimentos de visualização a partir da criação e execução de script em Grads. O script de visualização tem como parâmetro passado pelo script principal o nome do arquivo PrepBUFR a ser visualizado e então prepara a estrutura adicionando valores a algumas variáveis para a criação do script em Grads que contém também a abertura de um arquivo .ctl para projeção de latitude e longitude para posteriormente executar esse script pelo Grads gerando assim as figuras com as observações distribuídas geograficamente na figura do globo terrestre utilizada nesses procedimentos.

O esquema geral do funcionamento desses processos podem ser vistos na figura 3.9

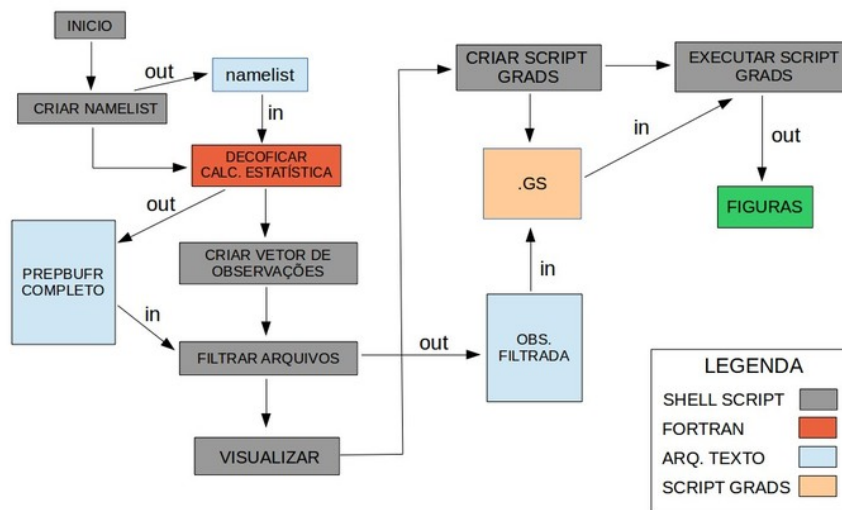


Figura 3.9 - Esquema de sistema de validação Dummy.

Todo esse procedimento é dividido em duas fases bem distintas:

### 3.1.8 Decodificação

Esse processo consiste em gerar arquivos texto com o conteúdo dos dados PrepBUFRS. Para essa etapa é utilizado o programa de decodificação da ferramenta bufrtools (especificamente denominado prep-bufr\_decode\_all encontrado no link: <http://www.dtcenter.org/com-GSI/BUFR/examples/download/convention/index.php>), e após ser adaptado conforme desejado, o mesmo após ser executado gera os arquivos no formato texto com suas respectivas informações.

### 3.1.9 Visualização

Após gerar as saídas em formato texto o procedimento a ser feito é filtrar as informações desejadas pelo tipo de observação para a sua visualização. Feito a escolha do tipo de observação requerida pode-se construir a figura com a distribuição do tipo de observação através da latitude e longitude num mapa global, e assim comparar com dados de outros centros de processamento. Abaixo segue alguns exemplos nas figuras 3.10, 3.11 e 3.12 para alguns tipos de observações como uma comparação do SATWND (observações a partir dos dados do NCEP e observações do PrepBUFR criado pelo pacote obs) e GPSIPW (observações dos dados do NCEP). Cabe ressaltar que a comparação do SATWND possuem alguns pontos que divergem nos dois PrepBUFRS, como data e os valores da janela de tempo, visto que para uma primeira versão esses testes atenderam bem a proposta solicitada.



Observações SATWND do PrepBUFR gerado pelo pacote obs (cptec):

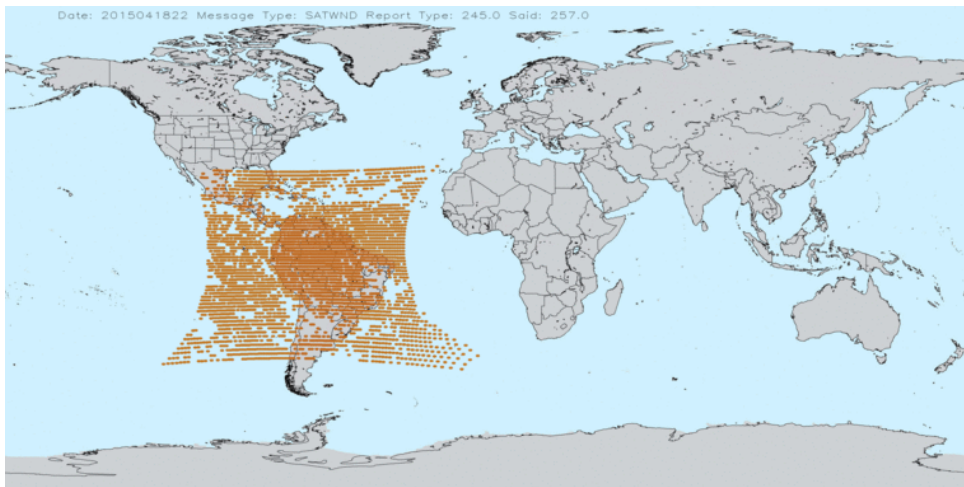


Figura 3.10 - Saída gráfica de vento SATWND CPTEC.

Observações SATWND do PrepBUFR gerado pelo NCEP:

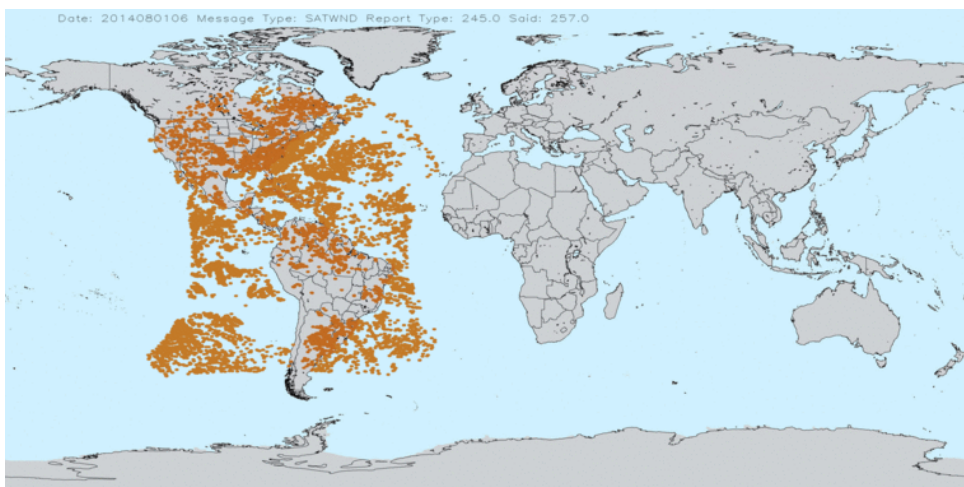


Figura 3.11 - Saída gráfica SATWND NCEP.



Observações GPSIPW do PrepBUFR gerado pelo NCEP:

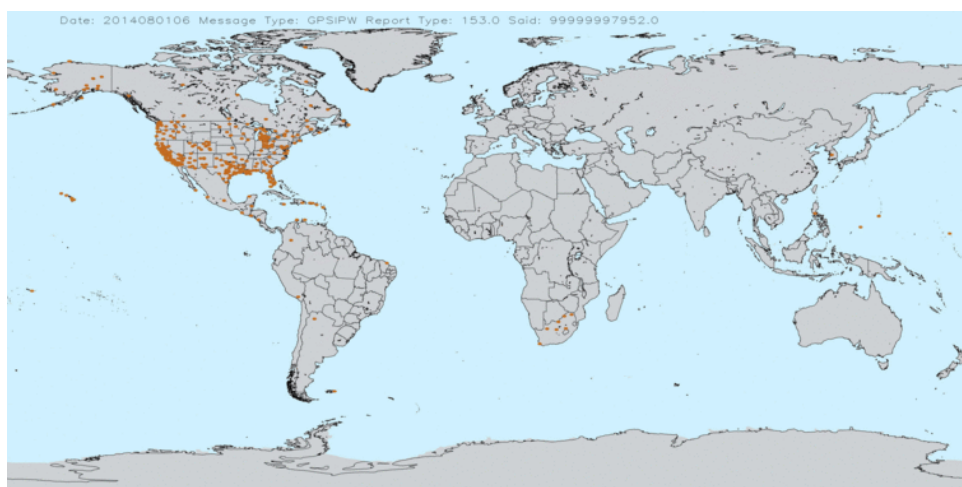


Figura 3.12 - Saída gráfica GPSIPW NCEP.

### 3.1.10 Estatística básica para avaliação quantitativa do PrepBUFR

O programa que gera valores estatísticos e valores máximos e mínimos é o mesmo programa que realiza a decodificação do PrepBUFR, o qual foi descrito nos tópicos acima. Ao passo que as mensagens são decodificadas do arquivo alguns procedimentos vão sendo realizados dentro e após o loop principal no código dos valores desejados para os tipos de observações U e V do vento. São extraídos os valores máximo, mínimo, média, desvio padrão e índice desses componentes. Todo o processo informado no tópico de decodificação é o mesmo para a geração desses valores quantitativos, pois esses cálculos estão no mesmo programa em fortran, e logo após é gerado um arquivo em formato texto com esses valores para serem comparados com outros dados PrepBUFRS.

Abaixo segue um exemplo de saída do arquivo de avaliação quantitativa.

```
%ESTATISTICA DO COMPONENTE WIND (U/V)
%TIPO MAX      MIN      MEAN      DESVIO.P IND
U      53.8000   -41.1000   1.8507    10.8737    23870
V      56.2000   -56.0000  -1.9909    8.0038     23870
```

Para as próximas versões da validação pretende-se incluir as estatísticas de outras observações que são gerados na decodificação.

### 3.2 Estrutura básica do BFGGE e seus módulos principais

Com base nessa metodologia de desenvolvimento a estrutura do pacote está organizada em função da origem dos dados, sendo uma para os dados convencionais, diretório "conv", que contempla os dados oriundos do BDM da DOP, e uma segunda para os dados de satélites, diretório "sat", que trata dos dados oriundos do BDS da DSA. Os dados convencionais compõem um módulo único com todos os diferentes sistemas de observações que escrevem seus dados em um mesmo arquivo denominado PrepBUFR. Os dados de satélites são separados em três diferentes módulos para os três principais tipos de dados importantes para a assimilação, os quais são radiâncias, dados de vento por satélites e dados de rádio ocultação dos sinais dos sistemas de navegação GNSS.

Os quatro principais módulos que dividem o pacote são detalhados abaixo:

**Convencionais (convobs):** os dados chamados convencionais são dados provenientes de estações meteorológicas em superfície, sensores em aviões, navios e boias, perfilhadores de vento e dados de lançamento de radiossondas, dropsondas e outras fontes de dados em bases terrestres ou reportados a essas. Todas essas fontes de dados são organizadas em um mesmo arquivo de leitura no GSI, o qual é denominado PrepBUFR. Inicialmente se pensou em colocar cada tipo de dado em um módulo em separado, mas a separação em módulos não traria contribuição significativa dada as suas semelhanças e por isso todos os dados são tratados em um mesmo módulo. Dessa forma as tarefas similares como configuração dos sistemas de observações e seleção dos arquivos a serem processados são feitas pela mesma rotina, mas a escrita dos dados no arquivo PrepBUFR é realizada por uma rotina específica para cada sistema de observação dada as suas particularidades. Seguindo a tabela do NCEP, há uma relação de mnemônicos para os dados convencionais e os principais, contemplados nesse projeto, são listados abaixo:

- ADPUPA: dados de ar superior na atmosfera (Radiossondas, Balão Piloto, Dropsondas);
- AIRCAR: dados de sensores meteorológicos instalados em aviões (ACARS, AIRCRAFT e MDCRS);
- AIRCFT: dados de sensores meteorológicos instalados em aviões comerciais (AIREP, PIREP, AMDAR, TAMDAR)
- PROFLR: dados de perfilhadores acústicos de vento (SODAR);

- ADPSFC: dados de estações meteorológicas a superfície terrestres (SYNOP e METAR);
- SFCSHIP: dados de estações a superfície oceânica (Navios, boias e plataformas C-MAN);
- SFCBOG: dados de pressão ao nível médio do mar (BOGUS);
- GOESND: dados de perfis atmosféricos recuperados a partir de dados dos satélites GOES (retrieval);
- GPSIPW: dados do vapor d'água integrado obtidos a partir dos dados GPS;
- ASCATW: dados de vento a superfície marinha recuperados a partir de dados de escaterômetro (ASCAT).
- SATWND: dados de vento derivados de satélite;
- VADWND: dados de vento VAD (NEXRAD);
- RASSDA: dados de sensores rádio acústico (RAAS) - perfil de temperatura virtual;

**Radiâncias (radiance):** sendo os dados de maior importância para a assimilação em modelos globais de PNT, os dados de radiância compõem um módulo principal que é subdividido em sub-módulos para cada tipo de sensor, independente dos satélites que os embarca. Para cada um desses módulos um tipo de arquivo é gerado contendo os dados do mesmo sensor provenientes de distintos satélites. Isso foi organizado dessa forma, pois é como o GSI trata a relação de sensores e satélites para os dados da radiância. Cada um desses módulos devem ser desenvolvidos obedecendo a estrutura Dummy de desenvolvimento. A relação abaixo apresenta a lista dos sub-módulos para a radiância relacionado os sensores, os satélites e os arquivos que os mesmos geram para a assimilação no GSI:

- Módulo: amsua - Satélites: NOAA-18 NOAA-19 MetOp-A e MetOp-B - Arquivo: cptec.yyyymmdd.thhz.1bamua;
- Módulo: airs - Satélites: Aqua Modis - Arquivo: cptec.yyyymmdd.thhz.airsev;

- Módulo: iasi - Satélites: MetOp-A e MetOp-B - Arquivo: cptec.yyyymmdd.thhz.mtiasi;
- Módulo: mhs - Satélites: NOAA-18 NOAA-19 MetOp-A e MetOp-B - Arquivo: cptec.yyyymmdd.thhz.1bmhs;
- Módulo: hirs4 - Satélites: NOAA-18 NOAA-19 MetOp-A e MetOp-B - Arquivo: cptec.yyyymmdd.thhz.1bhrs4;

**Vento por satélites (satwnd):** os dados de vento por satélites gerado operacionalmente na DSA/CPTEC/INPE utilizam imagens de um satélite geo-estacionário. O processo de estimativa assume a hipótese de que o deslocamento horizontal das nuvens se deve apenas ao efeito dos ventos, portanto é necessário identificar a propagação horizontal dessas nuvens em uma sequência de imagens. Uma metodologia de rastreamento de nuvens, baseada no método da Máxima Correlação Cruzada (MCC), é empregada para identificar a direção e intensidade do vento responsável por esse deslocamento. O nível de pressão que cada vetor vento representa é definido de acordo com o tipo de nuvem utilizada como traçador durante o processo de rastreamento. Esse dado é tratado no GSI como uma das fontes de dados do arquivo PrepBUFR. No entanto, como esse tipo de dado tem crescido significativamente nos últimos anos, o mesmo tem dominado o arquivo PrepBUFR, o que sugeriu a criação de um arquivo separado dos demais tipos de dados no PrepBUFR. Dessa forma, esse tipo é tratado em um módulo específico no BFGGE, uma vez que utiliza dados de satélites e conta com o apoio da DSA. A relação de satélites e tipos de canais tratados aqui são:

- Satélite GOES
  - visível
  - infravermelho próximo ( $3.9 \mu\text{m}$ )
  - vapor d'água ( $6.7 \mu\text{m}$ )
  - infravermelho janela ( $10.2 \mu\text{m}$ )
- Satélite EUMETSAT (previsto para o segundo semestre de 2016)
  - vapor d'água ( $6.2 \mu\text{m}$ )
  - infravermelho ( $10.8 \mu\text{m}$ )

**Radiocultação (rogness):** os dados de rádio ocultação GNSS são perfis de observações indiretas da atmosfera, que semelhante aos dados de radiância, carregam informação do comportamento da mesma em diferentes níveis. Os dados podem ser em três níveis de processamento, os quais podem ser listados aqui (do mais cru para o mais processado) como sendo (1) perfis do ângulo de ocultação em função do parâmetro de impacto; (2) perfis de refratividade em função da altura geométrica e (3) perfis de temperatura e umidade em função da pressão atmosférica. Atualmente o CPTEC está assimilando dados de refratividade e caminha para a assimilação do ângulo de ocultação. No GSI todos os dados de rádio ocultação independente dos sensores e constelação de satélites são tratados em um mesmo arquivo. Mas como são distintas as formas de receber os dados das diferentes constelações, os seguintes sub-módulos deverão ser criados:

- Sensor GPS, Satélites: Constelação COSMIC-I;
- Sensor GRAS, Satélites MetOp-A e MetOP-B;
- Sensor GNSS, Satélite TerraSAR;
- Sensor GPS, Satélites: Constelação COSMIC-II.

Na figura 3.13 é apresentado um esquema ilustrativo da sequência em que os processos são chamados em um fluxograma dos scripts responsáveis por executar cada um dos módulos que compõem o BFGE descritos acima. Nessa figura o script principal chamado `run_bfge.sh` é detalhado em especial os argumentos necessários para a sua execução. Um destaque é feito para o fluxo de dados de cada um desses módulos, os quais são utilizados setas coloridas para separar os diferentes tipos de dados: dados convencionais, dados de radiância e dados de rádio ocultação GNSS. Os dados existentes no BufrTank, disponibilizados pelo banco de dados Meteorológicos (DOP) e de satélites (DSA) no formato WMO são convertidos pelos BFGE em dados BUFR tabela NCEP, próprios para a assimilação de dados.

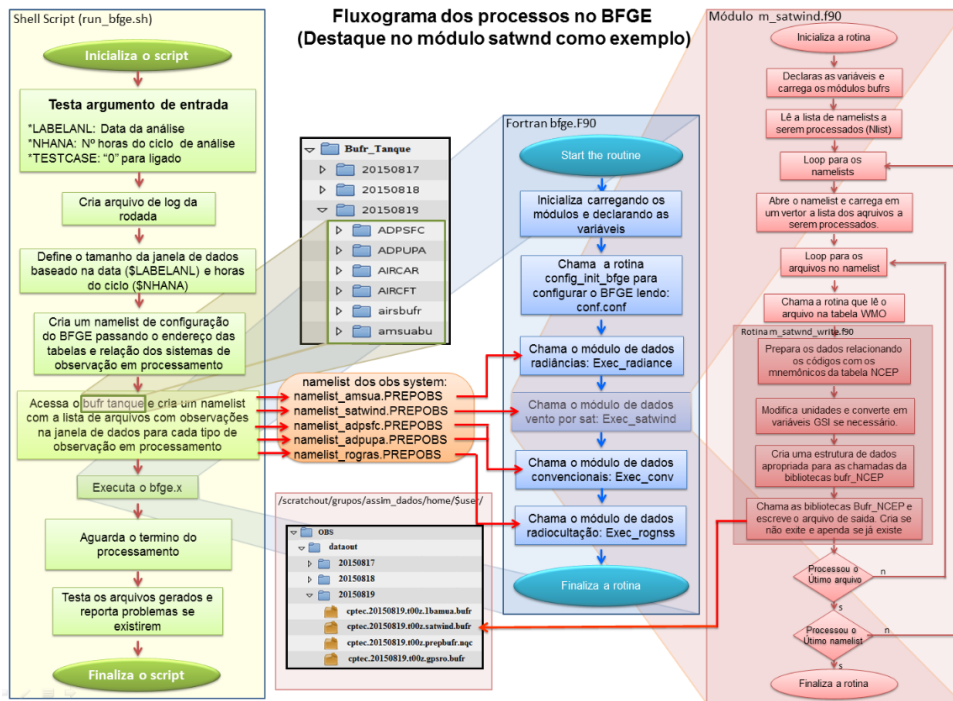


Figura 3.13 - Fluxograma do funcionamento dos processos no BFGE.

Cada um dos 4 módulos em que se divide o BFGE são detalhados nas próximas subseções.

### 3.2.1 Superfície

Seguindo a mesma metodologia da estrutura do pacote Dummy foi adicionado ao programa principal os módulos para a inclusão dos dados de superfície no Prep-BUFR. O pacote desenvolvido para essa observação utiliza-se das mesmas sub-rotinas independentes envolvidas nos outros módulos das outras observações como: módulo de leitura do BUFR e módulo de variáveis globais.

Os módulos que foram criados e que compõem a estrutura para a observação de superfície são:

- m\_adpsfc\_conf
- m\_adpsfc\_select
- m\_adpsfc\_write

O mesmo conceito é aplicado para cada novo tipo de observação que será incluído no BFGE, sendo que a estrutura é semelhante nos casos dos módulos "conf" e "select"

apenas no módulo de escrita há diferenças em sua composição, visto que alguns parâmetros são alocados para um determinado tipo de variável em uma observação e em outros casos não. Por exemplo, na leitura dos arquivos BUFRS são extraídos vários tipos de variáveis que estão localizados em posições diferentes da matriz para uma observação: nos dados de superfície a latitude está na posição 10 da matriz e nos dados de vento por satélite estão na posição 13.

Os dados de superfície aqui tratados são os advindos do MARS, com arquivos gerados sob consulta no BDM do CPTEC, sendo que os mesmos são organizados no Tank para a utilização do módulo de leitura dos BUFRS.

Os módulos descritos acima são detalhados abaixo:

**m\_adpsfc\_conf**: esse módulo compreende na leitura do arquivo de configuração namelist que é o arquivo que contém os diretórios e os nomes dos arquivos que serão executados. Um loop na relação dos arquivos a serem processados é feito no qual a quantidade deles é armazenada em uma variável global. A relação desses arquivos é gerada em um script que avalia os arquivos disponíveis dentro de uma janela de dados.

**m\_adpsfc\_select**: nesse módulo faz-se a utilização das sub-rotinas de alguns módulos independentes (módulo de variáveis globais e módulo de leitura de dados BUFR), do módulo de escrita de dados PrepBUFR e também da sub-rotina de leitura do namelist de configuração. Pode-se observar nesse módulo que a relação dos arquivos a serem processados é criada por um script que avalia no diretório tank quais os arquivos que se encontram dentro da janela de dados em cada passo do ciclo, tendo a hora da análise o centro da janela.

**m\_adpsfc\_write**: esse módulo contém a sub-rotina write\_prepbuf, responsável pela escrita dos arquivos PrepBUFRS de acordo com a leitura da observação, que no caso dessa versão inicial esta sendo feito para os dados BUFR de superfície.

Por enquanto está organizado a escrita no PrepBUFR apenas para os tipos "Surface data land", encontrados na leitura do BUFR que pode ser demonstrado no trecho abaixo do arquivo:

```
:BUFR:                # EDITION = 3
0                    # MBUFR Error code
:SEC1:
```

```

0          # BUFR MASTER TABLE
214        # ORIGINATING CENTER: Madrid
0          # ORIGINATING SUBCENTER
0          # UPDATE SEQUENCE NUMBER
0          # DATA CATEGORY: Surface data ? land
0          # DATA SUBCATEGORY
0          # LOCAL DATA SUBCATEGORY
14         # BUFR MASTER TABLE VERSION NUMBER
0          # LOCAL TABLE VERSION NUMBER
2015      # YEAR
11        # MONTH
24        # DAY
7         # HOUR
0         # MINUTE
:SEC3:
7         # Num.subsets
1         # Num.descriptors
0 # Flag for Compressed data (1=compressed 0=uncompressed)
307080
:SEC4:
130      # N. VARIABLES !!!
:SUBSET 00001:
8.00000 # 1) *001001-WMO block number (Numeric)
171.00000 # 2) 001002-WMO station number(Numeric)
"LLEIDA "# 3) 001015-Station or site name(CCITT IA5)
0.00000 # 4) 002001-Type of station (Code table)
2015.00000 # 5) *004001-Year (Year)
11.00000 # 6) 004002-Month (Month)
24.00000 # 7) 004003-Day (Day)
7.00000 # 8) *004004-Hour (Hour)
0.00000 # 9) 004005-Minute (Minite)
41.63000 # 10) *005001-Latitude (high accuracy) (degree)
0.60000 # 11) 006001-Longitude (high accuracy) (degree)
185.00000 # 12) 007030-Height of station ground above mean sea level (see (m))
599.00000 # 13) 007031-Height of barometer above mean sea level (see Note (m))
100180.00000 # 14) *010004-Pressure (Pa)
102730.00000 # 15) 010051-Pressure reduced to mean sea level (Pa)

```



90.00000 # 16) 010061-3-hour pressure change (Pa)

3.00000 # 17) 010063-Characteristic of pressure tendency (Code table)

Para esses dados de superfície é utilizada a mesma estrutura de mnemônicos utilizado nos outros módulos das outras observações convencionais, porém alguns deles tiveram que ser decididos e estruturados de acordo com algumas descrições encontradas nos conteúdos das tabelas do NCEP. Como por exemplo no caso do mnemônico SID, que foi estruturado de acordo com as posições 1 e 2, visto acima no trecho do subset do BUFR da WMO descrito como WMO block number e WMO station number, que pode ser montado como: 8171 para essa estação. Outro mnemônico que foi determinado na escrita do PrepBUFR foi o TYP, significando Report Type, valor especificado para o tipo de observação da tabela NCEP (visto no link: <http://www.emc.NCEP.noaa.gov/mmb/dataprocessing/prepbufr.doc/table2.htm#c.1>), que para esse dado de superfície foi resolvido o valor 181.

A relação dos tipos e subtipos da WMO associados com os mnemônicos do NCEP são encontrados em <http://www.nco.NCEP.noaa.gov/pmb/codes/nwprod/obsprocshared/bufrdumplist.v1.0.0/fix/bufrdumplist>

Para a superfície temos: adpsfc nem 000000 000001 000002 000007, sendo os três primeiros algarismos o tipo e os 3 últimos o subtipo, que são as categorias da tabela WMO.

Mais detalhes de como o código é estruturado são vistos no link: <https://projetos.cptec.inpe.br/projects/paqc/repository/show/trunk/obs/src/bfge/conv/adpsfc>

A escrita dos dados de superfície puderam ser visualizadas utilizando o script de validação do BFGE, e na figura 3.14 pôde-se verificar a posição das observações de superfície no globo terrestre plotados por latitude e longitude.

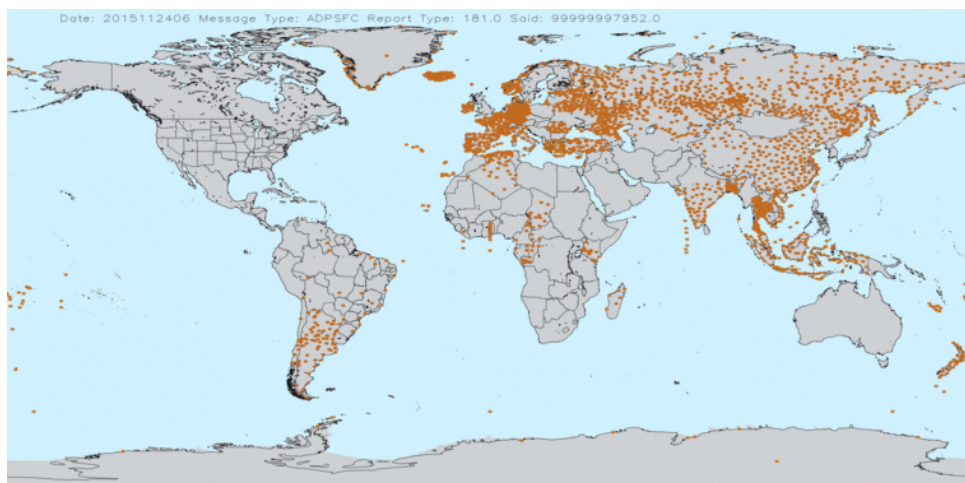


Figura 3.14 - Distribuição dos dados de superfície.

### 3.2.2 Módulos de Radiância

Nessa seção é descrito toda a implementação dos módulos de radiâncias, que contempla utilizar e converter as observações da temperatura de brilho obtidas a partir do sistema RARS (Regional ATOVS Retransmission Services) que são decifrados em vários níveis de controle em formato usado pela WMO para o formato do NCEP utilizado no sistema de assimilação GSI. O sensor usado para a análise é o microondas AMSU-A que contém 15 canais e esta baseado no parâmetro de umidade do vapor d'água. Este sensor encontra-se nas unidades dos satélites NOAA-18, NOAA-19 e METOP-B, e são recepcionados pelas antenas da divisão de satélites ambientais do CPTEC-INPE.

#### 3.2.2.1 Metodologia e dados

Inicialmente foram utilizados os dados dos sensores AMSUA que compõem o sistema ATOVS (Advanced Tiros Operational Vertical Sounder). Os dados utilizados possuem duas fontes: dados regionais ingestados pelas antenas da DSA, e dados globais (incluindo dados de outros centros de recepção) através do projeto RARS. O RARS é um projeto da Organização Mundial de Meteorologia, e consiste em um sistema de aquisição em tempo real de dados de satélite de órbita polar e de disponibilização para a comunidade através dos centros regionais de processamento. O objetivo principal do RARS é disponibilizar dados ATOVS globais em menos de 30 minutos desde a aquisição dos dados na estação. Estes dados já foram requisitados pelos centros responsáveis do RARS e são recebidos na DSA.

Em relação ao conjunto de códigos da tabela WMO e os mnemônicos da tabela NCEP, conforme citados nos tópicos anteriores, ambos precisam ser equivalentes. A tabela da WMO A,B e C descrevem os parâmetros, variáveis e dados observados com formato baseado nos índices numéricos, já as tabelas do NCEP inscritas em um determinado arquivo denominado radiance.bufrrtables descrevem os mesmos parâmetros, variáveis e dados observados em formato alfanumérico, tendo um valor numérico associado a um valor carácter (mnemônico) como mostrado na figura 3.15. Na figura 3.16 são descritos os valores correspondentes aos ângulos de observação do satélite em relação ao sol e a terra (azimutal e do zenith), e por último na figura 3.17 encontra-se o parâmetro correspondente ao dado observado, neste caso a temperatura de brilho do sensor AMSUA com 15 canais.

Tabela NCEP		Tabela WMO	Variável
SAID	corresponde a	001007	-SATELLITE IDENTIFIER (CODE TABLE 1007)
FOV	corresponde a	005043	-FIELD OF VIEW NUMBER (NUMERIC)
YEAR	corresponde a	004001	-YEAR (YEAR)
MONTH	corresponde a	004002	-MONTH (MONTH)
DAYS	corresponde a	004003	-DAY (DAY)
HOURS	corresponde a	004004	-HOUR (HOUR)
MINU	corresponde a	004005	-MINUTE (MINUTE)
SECO	corresponde a	004006	-SECOND (SECOND)
CLAT	corresponde a	005002	-LATITUDE (DEGREE)
CLON	corresponde a	006002	-LONGITUDE (DEGREE)
CLATH	corresponde a	005001	-LATITUDE (HIGH ACCURACY) (DEGREE)
CLONH	corresponde a	006001	-LONGITUDE (HIGH ACCURACY) (DEGREE)
HOLS	corresponde a	007001	-HEIGHT OF STATION (M)

Figura 3.15 - Mnemônicos equivalentes NCEP WMO (cabecalhos).

Tabela NCEP		Tabela WMO	Variavel
SAZA	corresponde a	007024	-SATELLITE ZENITH ANGLE (DEGREE)
SOZA	corresponde a	007025	-SOLAR ZENITH ANGLE (DEGREE)
BEARAZ	corresponde a	007021	-BEARING OR AZIMUTH (DEGREE TRUE)
SOLAZI	corresponde a	005022	-SOLAR AZIMUTH (DEGREE TRUE)

Figura 3.16 - Mnemônicos equivalentes NCEP WMO (ângulos).

Tabela NCEP		Tabela WMO	Variavel
TMBR	corresponde a	012063	-BRIGHTNESS TEMPERATURE (K) (15) canais

Figura 3.17 - Mnemônicos equivalentes NCEP WMO (observações).

Existe dentro da sub-rotina wrt\_amsua do módulo write\_bfr uma definição para a

variável denominada SUBSET, que distingue o tipo de sensor que esta sendo lido pelo programa.

```
vdssubset= NC021023 | A61223 | MTYP 021-023 PROC AMSU-A 1B Tb  
DATA(NOAA-15-19,METOP-2)
```

Neste caso, o valor alfanumérico corresponde ao sensor AMSUA do satélite NOAA18.

### **3.2.2.2 Validação dos arquivos de radiâncias gerados**

O sistema de validação de dados para radiância é realizado a partir do módulo `bufr_decode_radiance.f90`, que decodifica os dados de temperatura de brilho e gera um arquivo de interface para poder ver as variáveis graficamente e poder comparar os dados do satélite NOAA 18 distribuídas por ambos centros. O código segue a mesma lógica do módulo de validação Dummy. Principalmente é decodificado o sensor AMSUA pelo SUBSET NC021023, escrito logo em um arquivo os dados `hdr1` correspondentes a latitude, longitude, tipo de satélite correspondente, etc. Para identificar o tipo de satélite, um IF lógico (filtro) separa os dados correspondentes ao NOAA-18 escrevendo somente aqueles que contém o `SATID = 209`, e o resto das informações não se incluem no arquivo. As saídas gráficas são geradas pelo script `visualize_decode_bfge_radiance.sh` que utiliza a ferramenta gráfica (GRADS) e gera as saídas como as figuras 3.18 e 3.19 mostradas a seguir:

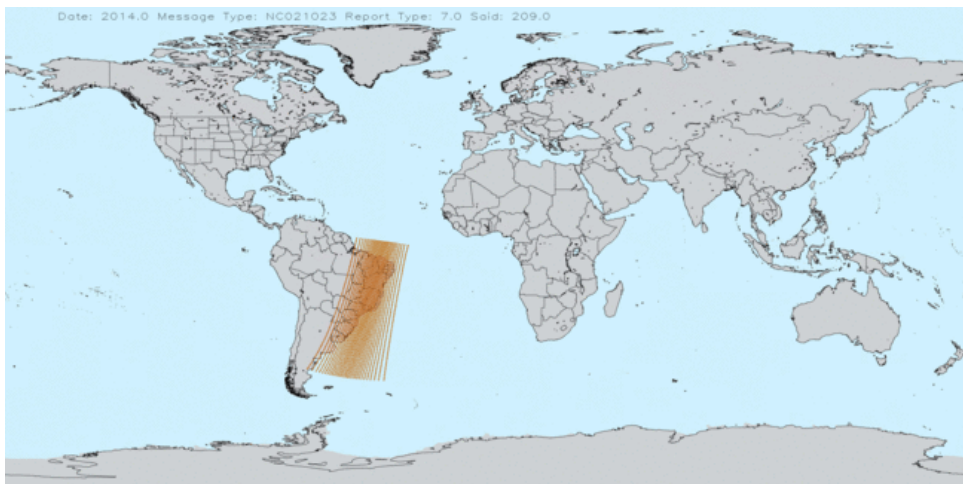


Figura 3.18 - Saída gráfica de dados processados de radiância via RARS.

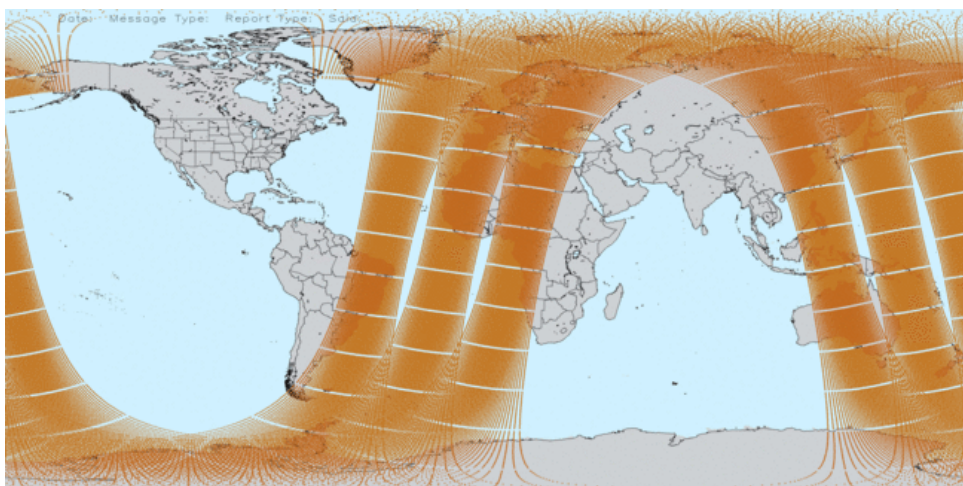


Figura 3.19 - Saída gráfica de dados processados de radiância via NCEP.

Na figura 3.18 é mostrado a saída dos dados recepcionados pelo sistema RARS, sendo recepcionados pela antena do CPTEC as 6:57 UTC do dia 24 de Agosto de 2014. A figura 3.19 corresponde aos dados emitidos pelo NCEP e correspondem as 3 passagens do NOAA-18 compreendidas entre as 00 e 7:00 UTC do mesmo período.

### 3.2.2.3 Resultados obtidos na assimilação

Os resultados da assimilação dos dados observados de radiâncias de satélites são mostrados na figura 3.20, que correspondem aos canais 1, 2 e 3, e que puderam ser comparados com os que são proporcionados pelo NCEP. Os dados pertencem ao

sensor de micro-ondas AMSUA que estão a bordo de satélites NOAA 18. No lado esquerdo são apresentados os valores da temperatura de brilho correspondentes ao sistema RARS e no lado direito as observações de temperatura de brilho obtidas via ftp pelo NCEP. Nota-se que há diferença mínima entre elas e por outro lado sabemos que a informação obtida pelo sistema RARS chega a ser decodificada em 15 minutos aproximadamente, já os dados do NCEP são disponibilizados via ftp com uma frequência maior de 3 horas. Essa comparação foi realizada para o dia 24 de Agosto de 2014 as 06:57 UTC.

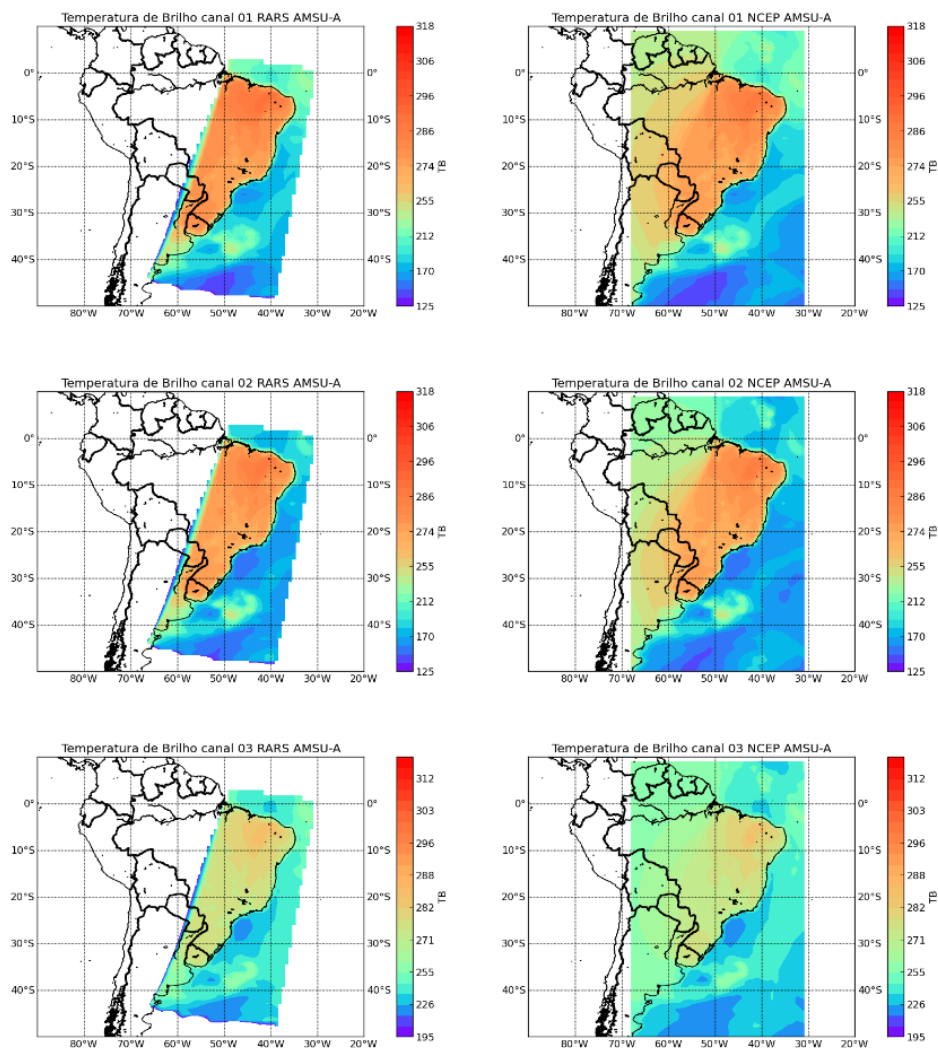


Figura 3.20 - Comparação da observação de temperatura de brilho do RARS(lado esquerdo) e do NCEP(lado direito).



### 3.2.2.4 Inclusão dos dados de radiância dos satélites NOAA-19 e MetOp-B no GCQD Versão básica

Utilizando a estrutura de desenvolvimento Dummy, e o que já foi feito para o satélite NOAA-18, foi desenvolvido um módulo de processamento dos dados de radiância dos satélites NOAA-19 e METOP-B, gerando um arquivo de dados para a assimilação no GSI. O critério para discriminar os diretórios foi chamar por tipo de sensor e neles aninhar os satélites que contém os mesmos tipos de sensores, como para este primeiro, todo foi baseado no sensor AMSU-A, a família de satélites que contém esses sensores no sistema RARS foram o NOAA-18, NOAA-19 e METOP-B, outros satélites como Aqua e o Terra que possuem esses sensores poderiam ser incorporados, mais adiante. No mesmo sentido, sensores como HIRS, AMSU-B e MHS terão as mesmas estruturas na versão robusta, como na figura 3.21.

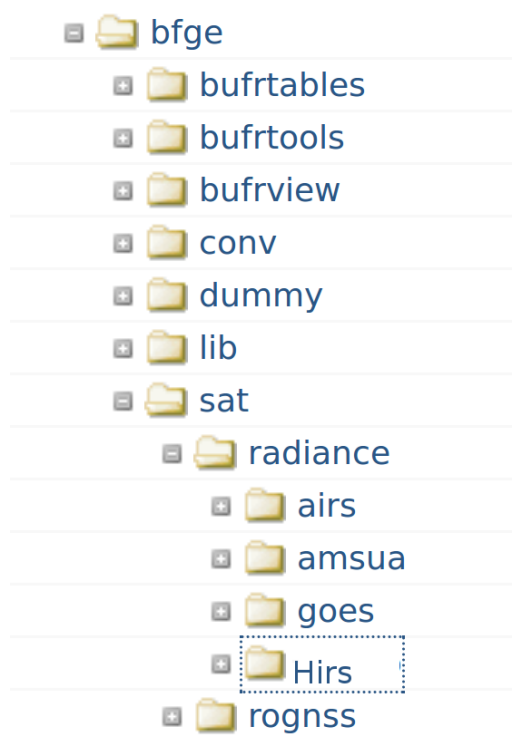


Figura 3.21 - Estrutura de diretórios para os módulos da radiâncias.

Um único módulo de leitura e escrita para o sensor AMSUA trabalha simultaneamente os dados observados pelos satélites NOAA-18, NOAA-19 e METOP-B. Este módulo se encontra no subdiretório noaa-metop, visto na figura 3.22, que gera o arquivo BUFR de saída (cptec.20140802.t06z.prepbuf.amsua), em que a cada 6 ho-

ras os dados observados pelo satélite do sistema RARS são coletados, centrados em 00:00, 06:00, 12:00, 18:00 horários UTC.

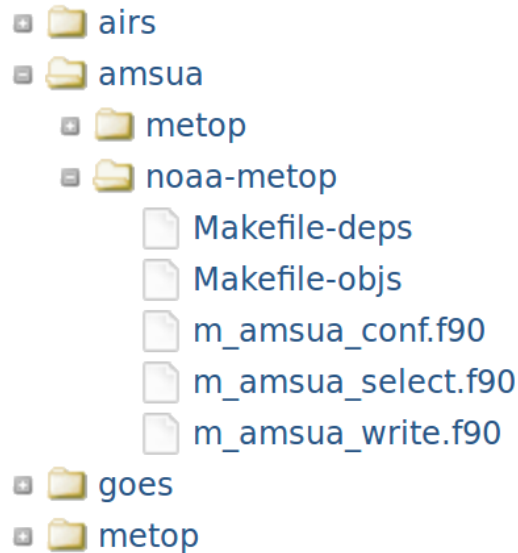


Figura 3.22 - Estrutura de diretórios para AMSUA.

Após toda a implementação do módulo foi possível gerar várias saídas de diferentes formatos para a validação dos dados. Foi implementada uma rotina que estabelece saídas dos dados em formatos ASCII para esses dados de temperatura de brilho nos 15 canais, agrupados sistematicamente em ordem da passagem dos satélites. Um segundo módulo faz a parte estatística, apresentado os valores do máximo e mínimo, média e desvio padrão, tomados para uma observação horária do satélite. Estas medidas estatísticas são dadas para os canais 5, 6, 7, 8 e 9 do espectro do vapor de água do AMSUA como mostrado são apresentados na figura 3.23.



```

%ESTATISTICA DOS CANAIS 5 e 6 NOAA-18
Nao existem dados para NOAA-18
%ESTATISTICA DOS CANAIS 5,6,8 e 9 NOAA-19
%TIPO      MAX      MIN      MEAN      DESVIO.P      IND
CH5        258.61    235.28    249.91     4.73          2310
CH6        242.02    222.89    234.26     4.27          2310
CH7        228.30    215.01    223.64     2.81          2310
CH8        223.23    210.93    218.11     2.03          2310
CH9        218.61    206.60    213.52     3.44          2310
%ESTATISTICA DOS CANAIS 5,6,7,8 e 9 METOP-1
%TIPO      MAX      MIN      MEAN      DESVIO.P      IND
CH5        263.11    240.22    254.51     4.92          2910
CH6        242.94    225.89    236.44     4.44          2910
CH7        228.55    215.48    223.91     3.43          2910
CH8        220.85    209.19    215.85     2.57          2910
CH9        217.51    206.11    210.11     3.19          2910

```

Figura 3.23 - Análise estatística do sensor AMSUA, para os canais 5, 6, 7, 8, e 9 dos satélites NOAA-18, NOAA-19 e Metop-b.

Uma outra saída em formato gráfico pôde-se visualizar a passagem dos satélites sobre o território brasileiro em diferentes escalas de cores, conforme visto na figura 3.24.

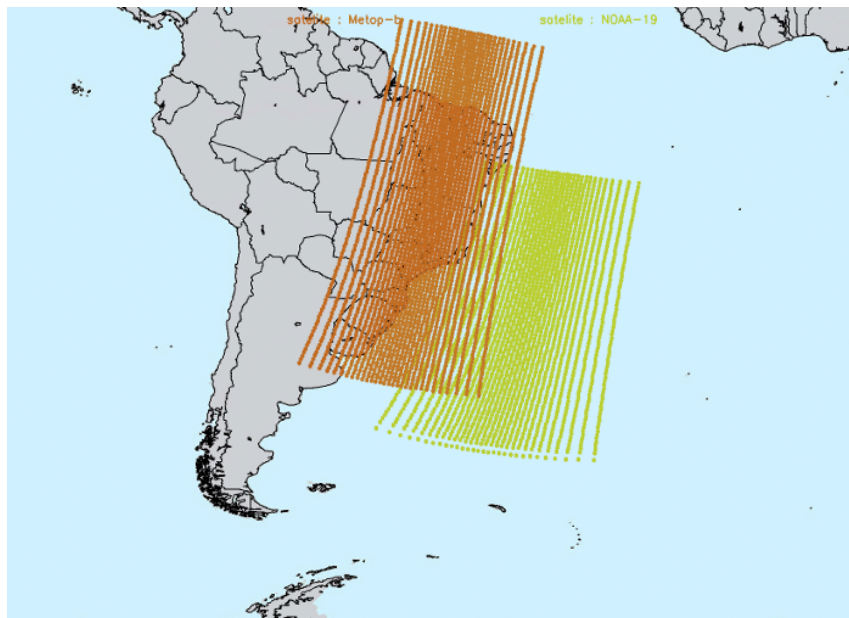


Figura 3.24 - Passagem dos satélites NOAA-19 e Metop-b com dados do sensor AMSUA, domínio da antena do sistema RARS.

### 3.2.3 Estrutura de dados SATWND (dados de vento por satélites)

Esta seção trata do processamento dos dados referentes ao vento em diferentes níveis da troposfera estimado a partir de satélites geo-estacionários.

O objetivo atual é assimilar os AMVs gerados operacionalmente na DSA/CPTEC. Inicialmente, assimilaremos apenas os campos de vento estimados a partir das imagens dos canais 6.2  $\hat{\mu}$  m (banda de absorção pelo vapor d'água, WV) e 10.8  $\hat{\mu}$  m

(infravermelho janela, IR).

A estrutura para os dados convencionais, realizada primeiramente para os dados de vento derivados de satélites geo-estacionários disponibilizados pela DSA, foi desenvolvida seguindo a mesma metodologia do pacote Dummy.

Para o pacote de dados de vento são utilizados os mesmos tipos de módulos e sub-rotinas do Dummy, que podem ser vistos no esquema da figura 3.25:

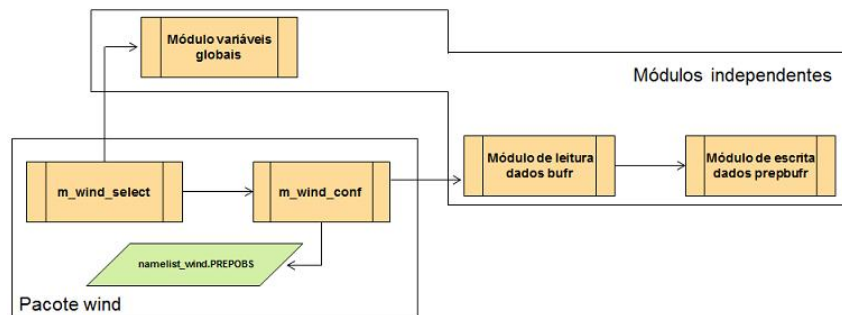


Figura 3.25 - Estrutura funcional para o módulo do vento por satélite.

De forma semelhante ao pacote Dummy, os módulos principais para esse tipo de observação são: `m_wind_select` e `m_wind_conf`.

O módulo principal do pacote é o `m_wind_select`, que é responsável pela execução de quatro sub-rotinas localizadas nos módulos independentes do pacote OBS. Tais sub-rotinas são:

- Módulo de variáveis globais;
- Sub-rotina no módulo de leitura do arquivo de configuração `namelist`;
- Sub-rotina no módulo de leitura de dados BUFRS;
- Sub-rotina no módulo de escrita de dados PrepBUFRS;

Inicialmente é configurado um arquivo texto denominado `namelist` contendo as informações necessárias para a geração dos arquivos PrepBUFR com os dados de vento. Nesse arquivo são determinados a data, o nome e o diretório de saída dos arquivos PrepBUFRS que serão gerados além da lista dos arquivos BUFRS de vento que serão lidos ao longo do processo para a janela de tempo utilizada.

Um arquivo PreBUFR está organizado e separado em horários sinóticos (0, 6, 12 e 18 horas) ao longo de um dia. Para alguns tipos de observações, possuem uma janela

de tempo (tempo da observação menos o tempo do ciclo) para o respectivo horário determinado no ciclo do arquivo. O DHR (observation time minus cycle time) é o parâmetro que indica a quantidade de horas anterior e posterior à observação em relação ao tempo do ciclo de assimilação, por exemplo, para as observações de vento do dia 01 às 00:30h será adicionado ao DHR o valor 0.5 em um arquivo PrepBUFR referente a 00:00h. Dependendo da observação, será específico uma janela de tempo, que no caso dos ventos é de -2.5 a 3.0 significando a hora do ciclo menos 02:30h e ciclo mais 3h.

A partir da leitura do namelist pela sub-rotina `m_wind_conf`, alguns desses parâmetros são organizados em variáveis que serão utilizadas ao longo do processamento para a leitura dos arquivos BUFR e escrita dos arquivos PrepBUFRS.

O gerenciamento da execução do `m_wind_conf` e da leitura dos arquivos BUFR, que utiliza-se da metodologia de (FERREIRA, 2007), e escrita dos PrepBUFR é feita pela sub-rotina `m_wind_select`. Se o arquivo PrepBUFR não existir, este será criado, senão, os dados de vento contidos no arquivo BUFR do horário em questão serão acrescentados ao arquivo PrepBUFR existente, cujo processo se chama `append`.

Ao executar a leitura dos dados de vento em formato BUFR, é possível averiguar as informações escrevendo-as em um arquivo texto. O arquivo texto com essas informações é escrito da seguinte forma:

```
:BUFR: # EDITION = 3
0 # MBUFR Error code
:SEC1:
0 # BUFR MASTER TABLE
46 # ORIGINATING CENTER: Brazilian Space Agency INPE
0 # ORIGINATING SUBCENTER
1 # UPDATE SEQUENCE NUMBER
5 # DATA CATEGORY: Single level upper-air data (satellite)
0 # DATA SUBCATEGORY
0 # LOCAL DATA SUBCATEGORY
12 # BUFR MASTER TABLE VERSION NUMBER
0 # LOCAL TABLE VERSION NUMBER
2015 # YEAR
4 # MONTH
18 # DAY
22 # HOUR
```

30 # MINUTE  
 :SEC3:  
 3532 # Num.subsets  
 24 # Num.descriptors  
 1 # Flag for Compressed data (1=compressed 0=uncompressed)  
 001007  
 002020  
 101015  
 031031  
 001031  
 001032  
 033007  
 :SEC4:  
 37 # N. VARIABLES !!!  
 :SUBSET 00001:  
 257.00000 # 1) 001007-SATELLITE IDENTIFIER (CODE TABLE 1007)  
 1.00000 # 2) 002020-SATELLITE CLASSIFICATION (CODE TABLE 2020)  
 2015.00000 # 3) 004001-YEAR (YEAR)  
 4.00000 # 4) 004002-MONTH (MONTH)  
 18.00000 # 5) 004003-DAY (DAY)  
 22.00000 # 6) 004004-HOUR (HOUR)  
 30.00000 # 7) 004005-MINUTE (MINUTE)  
 0.00000 # 8) 004006-SECOND (SECOND)  
 28.62000 # 9) 005001-LATITUDE (HIGH ACCURACY) (DEGREE)  
 -102.01999 # 10) 006001-LONGITUDE (HIGH ACCURACY) (DEGREE)  
 9.00000 # 11) 002152-SAT INST USED IN DATA PROCESS(6) (FLAG TAB 2152)  
 1.00000 # 12) 002023-SAT DERIVED WIND COMP METHOD (CODE TAB 2023)  
 66700.00000 # 13) 007004-PRESSURE (PA)  
 350.00000 # 14) 011001-WIND DIRECTION (DEGREE TRUE)  
 3.10000 # 15) 011002-WIND SPEED (M/S)  
 Null # 16) 222000  
 Null # 17) 236000  
 Null # 18) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 19) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 20) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 21) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 22) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)

Null # 23) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 24) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 25) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 26) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 27) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 28) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 Null # 29) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 0.00000 # 30) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 0.00000 # 31) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 0.00000 # 32) 031031-DATA PRESENT INDICATOR (FLAG TABLE 31031)  
 46.00000 # 33) 001031-IDENT OF ORIG/GEN CENTRE (CODE TAB 1031)  
 1.00000 # 34) 001032-GENERATING APPLICATION (CODE TABLE 1032)  
 18.00000 # 35) 033007-PER CENT CONFIDENCE (%)  
 18.00000 # 36) 033007-PER CENT CONFIDENCE (%)  
 18.00000 # 37) 033007-PER CENT CONFIDENCE (%)  
 :SUBSET 00002:  
 257.00000 # 1) 001007-SATELLITE IDENTIFIER (CODE TABLE 1007)

Os valores da CODE TABLE (listado em algumas das variáveis vista acima no exemplo da leitura do BUFR) podem ser vistos nas tabelas da WMO (versão 15). Esses valores descritos de acordo com as tabelas da WMO devem ser ajustados para a escrita do PrepBUFR de acordo com a tabela do NCEP.

Alguns dos parâmetros presentes nesse BUFR mostram a quantidade de subsets e variáveis, facilitando assim a interação dessas mensagens dentro dos loops no módulo de escrita do arquivo PrepBUFR.

Todas as mensagens extraídas do arquivo BUFR são armazenadas em um conjunto de variáveis tipo, que são organizadas como valores de entrada para a escrita de um arquivo no formato PrepBUFR, o qual é necessário também que haja verificações, visando confirmar se todo o procedimento de escrita nesse formato está sendo feito corretamente.

Para escrever um PrepBUFR alguns parâmetros denominados mnemônicos são preenchidos de acordo com os valores coletados dos arquivos BUFRS. Essa lista de mnemônicos pode ser vista em diversas tabelas do NCEP que estão no link: <http://www.emc.NCEP.noaa.gov/mmb/dataprocessing/prepbufr.doc/table1.htm>. Esses mnemônicos, conforme dito acima, são preenchidos com os valores lidos dos ar-

quivos BUFRS e alguns deles são preenchidos na própria sub-rotina, como por exemplo, o DHR (observation time minus cycle time), que a partir da hora minuto extraída do BUFR é preenchido a variável diff time, e no caso também do Report Type, que embora não contido no arquivo BUFR foi determinado pelo valor 245, significando SATWND - NESDIS IR (LONG-WAVE) CLOUD DRIFT (ALL LEVELS) (GOES) - u, v, o qual é visto com os outros tipos de observações convencionais no link: <http://www.emc.NCEP.noaa.gov/mmb/dataprocessing/prepbufr.doc/table2.htm>, e conforme um exemplo na figura 3.26:

String name	MNEMONIC	NUMBER	DESCRIPTION
hdstr	SID	001192	STATION IDENTIFICATION
	XOB	006002	LONGITUDE
	YOB	005002	LATITUDE
	DHR	004192	OBSERVATION TIME MINUS CYCLE TIME
	TYP	001193	PREPBUFR REPORT TYPE
	ELV	010194	STATION ELEVATION
	SAID	001007	SATELLITE IDENTIFIER (SATELLITE REPORTS ONLY)
obstr	POB	007192	PRESSURE OBSERVATION
	QOB	013192	SPECIFIC HUMIDITY OBSERVATION AFTER "VIRTMP" STEP - ALWAYS RECALCULATED FROM QUALITY CONTROLLED VIRTUAL TEMPERATURE DATA)
	TOB	012192	TEMPERATURE OBSERVATION(AFTER "PREPRO" STEP - REPORTED TEMP, EITHER SENSIBLE OR VIRTUAL DEPENDING UPON DATA TYPE AFTER "VIRTMP" STEP - VIRTUAL TEMPERATURE IF MOISTURE AVAILABLE, OTHERWISE SENSIBLE)
	ZOB	010196	HEIGHT OBSERVATION
	UOB	011003	U-COMPONENT WIND OBSERVATION
	VOB	011004	V-COMPONENT WIND OBSERVATION
	PWO	013213	TOTAL PRECIPITABLE WATER OBSERVATION
	CAT	001194	PREPBUFR DATA CATEGORY
	PRSS	010195	SURFACE PRESSURE OBSERVATION

Figura 3.26 - Mnemônicos de vento NCEP.

Nessa mesma sub-rotina é feita a verificação se já existe ou não o PrepBUFR e então é escolhido uma das duas formas de geração de PrepBUFR: a criação a partir de zero observações ou o append a partir do PrepBUFR já existente irá acrescentando as observações.

Após toda a execução do loop de arquivos é gerado então o PrepBUFR com as observações de vento por satélite. E como uma forma de representar essas observações visualmente foi utilizado a estrutura de visualização dos dados PrepBUFRS gerados, apresentada nos tópicos anteriores, visando avaliar a distribuição das observações, podendo assim comparar os mesmos tipos com outros PrepBUFRS de outros centros de processamentos (ex: dados do NCEP).

O PrepBUFR criado com essas observações pode ser visualizado e é apresentado na figura 3.27.

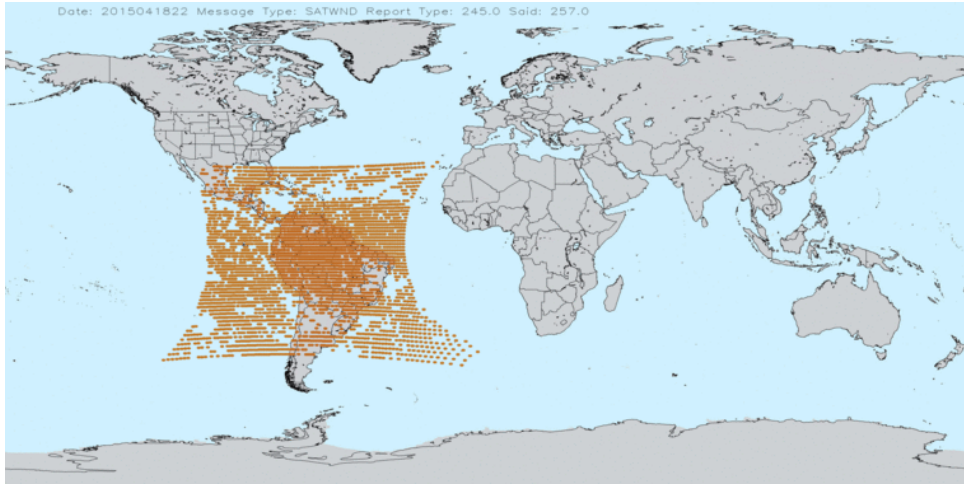


Figura 3.27 - Distribuição espacial das observações de vento por satélite.

### 3.2.4 Desenvolvimento dos dados de Rádio Ocultação GNSS no sistema GCQD

Os dados de rádio ocultação são relativamente mais complexos de se adicionar no BUFR tabela NCEP, pois possuem um cabeçalho de informações para cada dado de um mesmo perfil. Para isso são criados diferentes sequencias de informações denominadas ROSEQ01, ROSEQ02 e ROSEQ03.

Essa informações podem ser acessadas em <http://www.nco.ncep.noaa.gov/pmb/codes/nwprod/fix/bufrtab.003>, onde são encontrados quais os mnemônicos utilizados nas sequencias ROSEQ01 ROSEQ02 e ROSEQ03.

Esse módulo está em desenvolvimento e deverá ser entregue a primeira versão nas próximas entregas do GCQD.

### 3.3 Obtenção/ configuração/ compilação e execução do BFGE

Nessa seção são apresentadas as informações relevantes do processo de configuração, instalação e execução do pacote BFGE. É descrito o download de todo o pacote na conta do usuário, bem como a configuração para criar a estrutura de diretórios, a instalação dos pacotes incluindo as bibliotecas e finalmente a execução usando dados de um TestCase também disponível na distribuição. Embora o BFGE possa

executar independente do GCQD, ajustes são necessários para isso, os quais não são tratados aqui, pois as informações apresentadas pressupõe a versão integrada.

### 3.3.1 Detalhes sobre a instalação do BFGE

Como nos outros pacotes, o pacote de configuração/ instalação do OBS foi organizado seguindo a mesma estratégia de implementação adotada no GDAD. Nela o script `config_obs.ksh` é responsável por fazer todas as etapas de configuração e instalação de todos os pacotes que envolve o GCQD. A estrutura é toda montada sobre o diretório `obs` que é baixada no `scratchin` via SVN (apresentada em uma das seções anteriores) e ao configurar o sistema, uma árvore de diretórios no `scratchout` é criada. Isso é feito para manter a estratégia de processamento arquitetada na concepção da máquina Tupã. Nessa estratégia, os códigos e scripts, mantidos em segurança no SVN sobre controle de versões, é compilado no `scratchin`, de onde é rodado o script principal, denominado `run_obs.ksh`, que tem a função de enviar para o `scratchout` todos os scripts, executáveis e demais arquivos necessários para o processamento, onde são executados envolvendo submissão de processos para os nós do Tupã. Os arquivos de resultados finais (com maior fluxo de dados) são enviados para a árvore de diretório do `obs` no `scratchout` de onde são enviados para fitas ou discos externos. Observe que com essa estratégia nada é enviado para o home do usuário, pois embora seja uma área com backup, ela tem restrição de espaço, o que limita a sua utilização envolvendo diversos sistemas.

### 3.3.2 Detalhes do script de instalação do BFGE

Com a utilização do script `config_obs.ksh`, com a opção configurar selecionada é criado a seguinte lista de variáveis associadas com o BFGE:

```
export bfgesrc=$obssrc/bfge
export mbufrlib=$bfgesrc/lib/bufrlib
export mbufradt=$bfgesrc/lib/mbufr-adt
export f90lib=$bfgesrc/lib/f90lib
export bfgetables=$bfgesrc/bufrtables
export bfgetablesWMO=$bfgetables/wmo
export bfgetablesNCEP=$bfgetables/NCEP
export bfgetablesNCEPCON=$bfgetables/NCEP/prepbufr.table
export bfgetablesNCEPRAD=$bfgetables/NCEP/radiance.bufrtable
export bfgetools=$bfgesrc/bufrtools
export bfgebuferview=$bfgesrc/bufrview
```



```
export bfgedecode=$bfgesrc/bufrview/decodes
export bfgevisualize=$bfgesrc/bufrview/visualize
export obsworkrunbfge=$obsworkrun/bfge
export obsworkrunbfgegelog=$obsworkrunbfge/logfiles
```

Com a utilização do script `config_obs.ksh`, com a opção `compilar` selecionada é compilado os módulos do BFGÉ que são apresentados abaixo com os respectivos comandos para a compilação:

**bufrlib:** `$bfgesrc/lib/bufrlib` que é diretório das bibliotecas BUFRS;

```
cd $mbufrlib
export FC=ftn
export CC=gcc
$CC -c -DUNDERSCORE `./preproc.sh` *.c
$FC -c modv*.f moda*.f `ls -1 *.f | grep -v "mod[av]"`
ar crv libbufr.a *.o
```

**mbufr-adt:** `$bfgesrc/lib/mbufr-adt` que é diretório das bibliotecas mbufr-adt;

```
cd $mbufradt; make
```

**f90lib:** `$bfgesrc/lib/f90lib` que é diretório das bibliotecas f90lib;

```
cd $f90lib; make
```

**prepobs:** `$obsworkrun/bfge` diretório do código fonte do BFGÉ, com todos seus módulos para as diferentes fontes de dados;

```
cd $bfgesrc; make
```

Após toda a compilação é verificado se o seguinte arquivo executável foi gerado:

`$obsworkrun/prepobs.x`: programa principal do BFGÉ contendo todos os módulos integrados;

Com a utilização do script `config_obs.ksh`, com a opção `clean_all` selecionada é feito a limpeza de todos os diretórios dos pacotes de instalação do BFGGE apagando os arquivos temporários usados durante esse processo nos seguintes diretórios:

```
$mbufrlib; make clean
```

```
$mbufradt; make clean
```

```
$f90lib; make clean
```

```
$bfgesrc; make clean
```

### 3.3.3 Utilização do pacote BFGGE

Os dados necessários para esse TestCase do sistema estão disponíveis no scratchin em um diretório comum a todos os usuários do sistema no user "gdad", de forma que não se precisa replicá-los e nem ocupar espaço no SVN. A execução de todo o sistema é realizado ao rodar o script `run_obs.sh` que tem a função de incrementar uma variável com a data e hora das rodadas para o período desejado, rodando sequencialmente os processo envolvidos, chamando o script `run_bfge.sh` que executa o BFGGE para gerar os dados PrepBUFR tabela NCEP.

Em uma descrição geral do script `run_bfge.sh`, nota-se que o mesmo:

- Recebe via argumento de entrada a data para o processamento dos dados como sendo a data do centro da janela de dados. Um segundo argumento é o passo da análise para se definir o tamanho da janela de dados sendo metade desse passo em horas o intervalo considerado antes e depois do centro da janela. Um terceiro argumento é a flag do TestCase, onde se define o diretório que os dados devem ser procurados para o processamento;

- Esse script cria um namelist para o programa principal, chamado `bfge.conf` contendo os diretórios das tabelas e uma lista dos namelists dos módulos contemplados na rodada;

- Cria um arquivo de log da rodada com a data informada e a hora em que foi executado o script, do tipo `$obsworkrunbfge/$(date +%Y%m%d_%H%M%S).log`;

- Trata o tamanho da janela de forma flexível, podendo ser desde arquivos a cada 6 horas como a cada 1 hora para a assimilação no ciclo rápido;

- Cria para cada sistema de observação um namelist contendo uma lista dos arquivos

disponíveis para o processamento;

-Executa o programa principal processando todos os tipos de dados considerados e gera os arquivos de dados PrepBUFR;

-Verifica se todos os arquivos foram gerados com êxito.

### 3.3.4 Detalhes do script de execução do BFGE

Para rodar todo o processo BFGE basta acionar o script `run_bfge.sh` passando os parâmetros seguindo as recomendações da opção ajuda, que é apresentada quando nenhum argumento é dado. Essa ajuda é apresentada na figura 3.28 para facilitar a descrição de cada um dos argumentos requeridos para a execução do script.

```
luiz.sapucci@eslogin07:/scratchin/grupos/assim_dados/home/luiz.sapucci/obs/run> ./run_bfge.sh

-----
>>>>>      Execução do BFGE no Sistema de GCQD      <<<<<<
-----

==> Para rodar o processo BFGE digite:

> ./run_bfge.sh LABELANL NHANA TESTCASE

* LABELANL      : Data da analise ou do centro da janela de dados (YYYYMMDDHH) ex. 2015010100
* NHANA         : Passo do ciclo de analise para Janela de dados (tempo (+) e (-) relacao a hora sinotica). Ex: 6
* TESTCASE      : Flag para opcao testCase sendo:
                  =0 (Not testCase) Nao se deseja rodar o testeCase usando os dados bufr tabela
                   wmo do tanque predefinido na variavel obs datain no config_obs.ksh.
                   Configuracao inicial e o mesmo diretorio do Tanque Dummy que deve ser modificado pelo usuario
                  =1 utiliza dados bufr da tabela wmo disponiveis no diretorio definido pela
                   variavel TANQUE_DUMMY no config_obs.ksh. Esse repositorio dos dados do testcase esta disponivel em
                   /scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufr_Tanks

==> Para testCase basico do BFGE digite:

> ./run_bfge.sh 2014080206 6 1

==> A opcao ajuda mostra as informacoes sobre os argumentos de entrada. Para isso digite:

> ./run_bfge.sh ajuda

Para mais informacoes sobre o pacote leia os ../doc/README_GCQD_from_wiki_in_*.pdf usando o evince:
> evince ../doc/README_GCQD_from_wiki_in_*.pdf

Creditos: GDAD/DMD/CPTEC/INPE
```

Figura 3.28 - Argumentos requeridos para execução do script.

Como descrito na figura, o `run_bfge.sh` possui duas funções acionadas pelo primeiro argumento ao executar o script, sendo:

**Função ajuda:** apresenta o mesmo conteúdo apresentado na figura 3.28, o que é o mesmo se nenhum argumento for fornecido.

**Função running:** a função de execução do processo é acionada se os argumentos são suficientes para a rodada. Um total de 3 argumentos são necessários para executar o script, os quais são apresentados com mais detalhes:

- LABELANL: data para o processamento da janela de dados (sendo do tipo YYYYMMDDHH);
- NHANA: passo do ciclo de análise para a definição da janela de dados usada na criação dos arquivos PrepBUFRS;
- TestCase: é uma flag para indicar onde o sistema deve procurar os dados necessários para a execução.

**para a opção 0:** que representa "NoTestCase", não se deseja rodar o TestCase usando os dados BUFR da tabela WMO do tanque pré-definido na variável obsdatain no config\_obs.ksh. O diretório do TanqueDummy deve ser modificado pelo usuário;

**Para a opção 1:** utiliza os dados BUFR da tabela WMO disponíveis no diretório definido pela variável TanqueDummy no config\_obs.ksh. Esse repositório dos dados do TestCase esta disponível em: /scratchin/grupos/assimdados/home/gdad/Data-Fix/OBS/testcase/BufrTanks.

### 3.3.5 Passo-a-passo para a execução do BFGE

-Login no supercomputador CRAY XE6:

```
ssh usuário@tupa.cptec.inpe.br -XC
```

-Acessar ao diretório scratchin (variável \$SUBMIT\_HOME):

```
cd $SUBMIT_HOME
```

-Entrar no diretório dos scripts do obs, que caso tenha sido feito a instalação padrão, o acesso para esse diretório é:

```
cd obs/run
```

-Executar inicialmente o run\_bfge.sh sem argumento para rever as informações descritas acima, certificando que não há outras opções implementadas e não tratadas ainda:

```
./run_bfge.sh ajuda
```

-A saída deverá ser a mesma apresentada na figura 3.28. Depois disso execute o TestCase seguindo os argumentos apresentados abaixo:

```
./runbfge.sh 2014080206 6 1
```

Na rodada TestCase observe a sequencia dos processos realizados pelo script na figura 3.29, onde algumas informações entre os módulos foram suprimidas na figura

para simplificar o entendimento da sequencia dos processos:

```
luiz.sapucci@login07:/scratchin/grupos/assim_dados/home/luiz.sapucci/obs/run> ./run_bfge.sh 2014080206 6 1

-----
>>>> Running BFGE package for data assimilation ... <<<<<
-----

> Analise para      : 2014080206
> Passo do ciclo de analise : 6
> TestCase         : 1

>>> Diretório dos bufns file: /scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks

Janela de tempo: 201408020301 a 201408020900

> Criando a lista de arquivos disponiveis de dados: DUMMY
Escaner dos arquivos:
-----
#ARQUIVO DE GERENCIA DA ROTINA prepobs
#Data do prepbufn
2014080206
#Diretorio saida prepbufn
/scratchout/grupos/assim_dados/home/luiz.sapucci/obs/dataout/20140802/
#Nome prepbufn
dummy.20140802.t06z.prepbufn.nqc
#Arquivos bufns janela tempo
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020330.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020400.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020430.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020500.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020530.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020600.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020630.bufn
/scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/Bufn_Tanks/20140802/DUMMY/245/S11235861_201408020700.bufn

> Criando a lista de arquivos disponiveis de dados: ADPSFC
Escaner dos arquivos:
-----
#ARQUIVO DE GERENCIA DA ROTINA prepobs
#Data do prepbufn
2014080206
#Diretorio saida prepbufn
/scratchout/grupos/assim_dados/home/luiz.sapucci/obs/dataout/20140802/
#Nome prepbufn
cptec.20140802.t06z.prepbufn.nqc
#Arquivos bufns janela tempo
-----

> Executando prepobs.x .. Aguarde ... |

-----
TDA BFGE executado com sucesso. Arquivos bufns gerados:
-----
Arquivos gerados:
cptec.20140802.t06z.prepbufn.amsua
cptec.20140802.t06z.prepbufn.nqc
dummy.20140802.t06z.prepbufn.nqc
```

Figura 3.29 - Visualização da execução do script.

Os arquivos de saída do processo dos dados em PrepBUFR tabela NCEP desse Test-Case podem ser acessados em `$WORK_HOME/obs/dataout/20140802`. O arquivo de log da rodada efetuada pode ser acessado no diretório `$WORK_HOME/obs/run/bfge/logfiles`, no qual a data e hora da execução ficam registradas no nome do arquivo.

## 4 Pacote de Controle de Qualidade (PAQC - Package Quality Control)

O PAQC é o programa do NCEP utilizado para realizar um controle de qualidade de dados observacionais, o qual é composto de um processo complexo de verificações encadeadas, as quais empregam análise multi-variada para verificar todas as observações contra as vizinhas mais próximas. Módulos especiais são aplicados de ante mão e separadamente para alguns sistemas de observações e faz uma varredura avaliando em separado as observações desse sistema aplicando metodologias específicas. Esse é o caso dos dados de radiossondas (ADPUPA), dados de sensores em aviões (ASCAR e AIRCFT), perfilhadores de vento doppler (PROFLR e VADWND). Esses dados são tratados em módulos anteriores ao módulo principal, os quais são tratados em detalhes na próxima subseção. Várias interações do esquema básico são necessárias para completar o processo. Durante cada interação cada uma das observações recebe uma flag que indica se passou ou falhou no processo. Se uma observação qualquer passa, significa que ela pode ser utilizada para verificar as demais em uma subsequente interação, mas se falhar ela é excluída do processo sendo ignorada nas próximas verificações. Esse processo iterativo é desenvolvido até que o número de observação excluída tende a zero.

### 4.1 Os módulos principais do pacote PAQC

No pacote do PAQC além dos módulos de controle de qualidade existem alguns processos destinados a preparação dos dados para o processo. Os principais módulos de preparação são descritos abaixo:

- **block-unblock:** Converte o fluxo de bytes BUFR em arquivos binários para um formato em blocos Fortran e vice-versa. O código fonte está tanto para colocar em 'unblock' e a nova rotina "cwordsh" requer que seja linkado a atual biblioteca NCEP BUFRLIB. Esta versão do 'block' evita dados incorretos nos arquivos de entrada, como cabeçalhos de aviso, e até mesmo palavras de comando Fortran, para que os arquivos já bloqueados sejam novamente bloqueados sem gerar inconsistências.
- **prepobs\_prevents:** É executado em dois modos: "PREVENT" e "POSTEVENTS". No modo PREVENT, prepara relatórios observacionais PrepBUFR para um subsequente controle de qualidade e programas de análise. Isto é feito ao interpolar os campos do first-guess (em formato espectral) para a localização das observações no PrepBUFR com a codificação dos valores do first-guess em registros PrepBUFR codificando os "prevent" e/ou

eventos "virtmp" em registros de PrepBUFR e codifica os erros de observação do arquivo de especificação de erro em registros de PrepBUFR. Nesse modo também, pode-se modificar o marcador da flag de qualidade dos dados para não ser utilizado em um subsequente passo do controle de qualidade e processo de análise, em um processo de filtragem. No modo "POSTEVENTS", depois de executado todo o controle de qualidade e programa de análise, é feita a interpolação da análise global (em coordenadas sigma espectral) para a localização das observações e codifica esses valores da análise para o arquivo PrepBUFR. Para mais informação sobre "prevent" e eventos de "virtmp", veja o bloco de notas da rotina w3lib "gblevents". Esse programa chama a rotina gblevents, a qual é executada aqui no modo "prevent". A rotina w3lib gblevents faz o empilhamento do trabalho e depois cada registro é atualizado pela rotina gblevents, os quais são escritos em uma versão pré-processada do arquivo de PrepBUFR.

Os módulos de controle de qualidade são descritos aqui de forma sucinta e apresentados na sequência que os mesmos são aplicados na execução do pacote:

- **adcor (Radiossondas):** esse módulo prepara correções para observações de altura e temperatura das radiossondas a partir das diferenças entre as radiossondas (principalmente o efeito radiativo nos instrumentos). O ângulo solar e o ângulo de elevação solar são calculados. Sub-rotinas contendo tabelas de correções são chamadas e as correções são aplicadas diretamente nos níveis obrigatórios e correções de temperatura são interpoladas para a superfície e níveis significativos na troposfera.
- **prepobs\_cqcbufr.fd (Radiossondas):** esse módulo realiza um complexo controle de qualidade de observações de altura e temperatura realizadas por radiossondas. Erros são detectados e muito corrigidos baseados nas informações do perfil. Algumas verificações utilizadas são: hidrostática, incremento, estatística horizontal, estatística vertical e temporal, linha de base e taxa de queda vertical. Também são calculadas a radiação (intersondas) e correções de vieses. Adicionalmente é aplicada uma homogeneização das séries de temperatura usando estatísticas de inovação.
- **prepobs\_profqc.fd (Perfilhadores de vento Doppler):** realiza um complexo controle de qualidade de perfilhadores de vento Doppler. Os dados de entrada e saída são arquivos no formato PrepBUFR. As verificações realizadas são: incrementos, estatística vertical e temporal e uma combi-



nação de estatísticas vertical e temporal aplicadas simultaneamente.

- **prepobs\_prepacqc.fd (Dados de sensores em aviões AIRCFT):** lê um arquivo PrepBUFR contendo todos os tipos de dados pré-processados e seleciona as mensagens que se referem a "AIRCFT" organizando pelos identificadores das estações. Realiza uma verificação rastreada e identifica observações pela posição chamando uma pilha. Aplica um controle de qualidade fazendo uma verificação nos voos, removendo duplicatas, comparando a localização das observações e se preciso, produz superobs de dados de ventos, e posteriormente passando por uma verificação de qualidade. Uma série de novas flags de controle de qualidade PrepBUFR são incluídas em cada observação. Finalmente escreve os eventos empilhados (considerando as marcas de qualidade atualizadas no PrepBUFR) dentro dos dados PrepBUFR existentes. Em todos os casos, o novo arquivo contém todas as observações originais menos as duplicadas e adição das superobs. Observações usadas para gerar essas superobs são marcadas para assegurar que as mesmas serão omitidas no esquema de análise.
- **prepobs\_acarsqc.fd (Dados de sensores em aviões AIRCAR):** lê um arquivo PrepBUFR contendo todos os tipos de dados pré-processados e seleciona as mensagens que se referem a "AIRCAR". Realiza-se rudimentares verificações de qualidade nos dados, tais como erros grosseiros e verificação de sanidade. Uma série de novas flags de controle de qualidade PrepBUFR são incluídas em cada observação. Finalmente escreve os eventos empilhados (considerando as marcas de qualidade atualizadas no PrepBUFR) dentro dos dados PrepBUFR existentes. Em todos os casos, o novo arquivo contém todas as observações originais menos as duplicadas dentro da janela requerida de dados.
- **prepobs\_cqcvad.fd (Dados de vento por radar):** Esse módulo realiza um complexo controle de qualidade de ventos de VAD proveniente de radares WSR-88D.
- **oiqc (Complexo controle de qualidade):** Esse é o módulo principal utilizado para a realização do controle de qualidade de dados observações de forma global, o qual é composto de processo complexo empregando análise multivariada para verificar todas as observações contra as vizinhas mais próximas. Várias interações do esquema básico são necessárias para completar o processo. Durante cada interação todas as observações são

sujeitas a 5 interpolação de verificação:

- a) **Verificação 1:** interpola os valores de comparação dos dados de temperatura mais próximos;
- b) **Verificação 2:** avalia a consistência dos valores obtidos na verificação 1;
- c) **Verificação 3:** Interpola os valores de comparação da componente zonal e meridional do vento.
- d) **Verificação 4:** Interpolação das observações mais próxima em um mesmo perfil em avaliação. Uma combinação ponderada dessas quatro verificações é feita com base em fatores, os quais determinam, o quanto da interpolação pode ser esperado por ser mais relevante do que outras. Por exemplo, a quantidade e localização espacial e tempo de observações utilizadas em uma dada verificação, tanto quanto a quantidade relativa das observações envolvidas é usada para indicar o quanto dos resultados devem ser ponderadamente confiáveis do que tomar uma decisão em um caso particular.
- e) **Verificação 5:** Interpola um valor de um conjunto multivariado de observações na verificação 1 e 3, as quais são usadas para medir como a análise deve esboçar (ou não esboçar) a informação contidas nas observações em avaliação. A tolerância permitida para o desvio de um esperado resultado nas verificações é proporcional a medida desse esboço calculado na verificação 5. Durante cada interação cada uma das observações recebe uma flag que indica se passou ou falhou nas 5 verificações descritas acima. Se uma observação qualquer passa, significa que ela pode ser utilizada para verificar as demais em uma subsequente interação, mas se falhar ela é excluída do processo sendo ignorada nas próximas verificações. Esse processo iterativo é desenvolvido até que o número total de observações excluídas é igual em dois ciclos em um total mínimo de 4 ciclos, no qual de forma preventiva para a convergência do sistema um procedimento arbitrário de parada é envolvido para resolver alguns casos ambíguos.

## 4.2 Obtenção/ configuração/ compilação e execução do PAQC

Nessa seção são apresentadas as informações relevantes do processo de configuração, instalação e execução do pacote PAQC. Da mesma forma que a seção do pacote BFGÉ, pressupõem nessa seção que todas as tarefas a respeito da instalação do sistema GCQD descrita em sua seção tenham sido executadas. Nela é descrita o download de todo o pacote na conta do usuário, bem como a configuração para criar a estrutura de diretórios, a instalação dos pacotes incluindo as bibliotecas e finalmente a execução usando dados de um TestCase, também disponível na distribuição. Embora o PAQC possa executar independente do GCQD, ajustes são necessários para isso, os quais não são tratados aqui, pois as informações apresentadas pressupõe a versão integrada. Logo a obtenção do código é suprimida aqui, pois não há detalhes específicos do PAQC para serem reportados além dos relatados na obtenção do código do GCQD.

### 4.2.1 Detalhes sobre a instalação do PAQC

Como nos outros pacotes, o pacote de configuração/ instalação do OBS foi organizado seguindo a mesma estratégia de implementação adotada no GDAD. Nela o script `config_obs.ksh` é responsável por fazer todas as etapas de configuração e instalação de todos os pacotes que envolve o sistema. A estrutura é toda montada sobre o diretório `obs` que é baixada no `scratchin` via SVN (feita na seção anterior) e ao configurar o sistema uma árvore de diretório no `scratchout` é criada. Isso é feita para manter a estratégia de processamento arquitetada na concepção da máquina Tupã. Nessa estratégia os códigos e scripts, mantidos em segurança no SVN sobre controle de versões, é compilado no `scratchin`, de onde é rodado o script principal, denominado `run_obs.ksh`, que tem a função de enviar para o `scratchout` todos os scripts, executáveis e demais arquivos necessários para o processamento, onde são executados envolvendo submissão de processos para os nós do Tupã. Os arquivos de resultados finais (com maior fluxo de dados) são enviados para a árvore de diretório do `obs` no `scratchout` de onde são enviados para fitas ou discos externos. Na instalação do GCQD todos os módulos do PAQC são contemplados e a estrutura de diretórios necessária é criada e as bibliotecas, pré-requisitos do pacote, são instaladas. Os detalhes específicos para esse pacote são destacados abaixo.

### 4.2.2 Detalhes do script de instalação do PAQC

Com a utilização do script `config_obs.ksh`, com a opção configurar selecionada é criado a seguinte lista de variáveis associadas com o PAQC e as exporta no `scratchin`:

```
export paqc_src=$obs_src/paqc ..... Diretório raiz do PAQC
export paqc_2src=$paqc_src/src ..... Diretório do código fonte do PAQC
export paqc_bin=$paqc_src/Linux/bin ..... Diretório dos arquivos binários a serem
gerados na compilação
export paqc_etc=$paqc_src/Linux/etc ..... Diretório ETC do pacote PAQC
```

No scratchout: export obs\_work\_run\_paqc=\$obs\_work\_run/paqc ..... Diretório onde o PAQC é executado

```
export obs_work_run_paqc_log=$obs_work_run_paqc/logfiles ..... Diretório de saída
dos arquivos de log da rodada
```

```
export obs_dataout=$obs_work/dataout ..... Diretório de entrada e saída dos dados
do PAQC.
```

Com a utilização do script config\_obs.ksh, com a opção compilar selecionada é compilado toda a árvore de diretório do PAQC, a qual está na mesma estratégia de compilação utilizada no NCEP, com uma cascata de Makefile em acionamento. Esse processo é feito através do config\_obs.ksh através dos seguintes passos:

Carregando os modules do Tupã setando o NetCDF e versão do compilador pgi:

```
. /opt/modules/default/etc/modules.sh
module load netcdf
module swap pgi pgi/12.2.0
```

Entra do diretório raiz dos código do PAQC:

```
cd $paqc_2src
```

Exporta as variáveis \$ESMA\_FC do compilador padrão e diretório base na \$BASE-  
DIR:

```
export ESMA_FC=ftn export BASEDIR=$paqc_src/dummybasedir
```

Executa o gmake com a opção "install" no diretório raiz guardando nele um arquivo de log da instalação chamado compilePAQC.log:

```
gmake install ESMADIR=$paqc_src 2>&1 | tee $paqc_src/compilePAQC.log
```

Verificando o sucesso do processo avaliando se uma lista de arquivos executáveis foi criada:

```
$paqc_bin/echorc.x
$paqc_bin/combfrd.x
$paqc_bin/scanbuf0.x
$paqc_bin/ssprepqc
$paqc_bin/zeit_ci.x
$paqc_bin/zeit_co.x
$paqc_bin/zeit_pr.x
$paqc_bin/ssprevents.x
$paqc_bin/cqcbufr.x
$paqc_bin/raobcore.x
$paqc_bin/hradcor.x
$paqc_bin/cqcvad.x
$paqc_bin/profcqc.x
$paqc_bin/acarsqc.x
$paqc_bin/oiqcbufr.x
$paqc_bin/prepacqc.x
```

Caso alguns arquivos da lista não tenham sido criados uma mensagem de erro indica o arquivo. A instalação do pacote sendo como é feita pelo script `config_obs.ksh`, no entanto, pode-se instalar unicamente o PAQC, utilizando o script <https://svn.cptec.inpe.br/paqc/trunk/obs/src/paqc/install>, no qual todas as etapas acima são realizadas.

Com a utilização do script `config_obs.ksh`, com a opção `clean_all` selecionada limpa todos os diretórios dos pacotes de instalação do PAQC apagando os arquivos temporários usados durante esse processo em toda a árvore de diretórios do pacote, obedecendo as diretivas do Makefile com o comando `distclean`, da forma:

```
cd $paqc_2src
gmake distclean ESMADIR=$paqc_src
```

### 4.2.3 Utilização do pacote PAQC

Baseado nos processos já descritos sobre a utilização do sistema GCQD com a execução básica usando os dados do TestCase disponíveis, destaca-se aqui os detalhes mais relevantes da execução do PAQC.

Os dados necessários para esse TestCase do sistema estão disponíveis no scratchin em um diretório comum a todos os usuários do sistema no user "gdad", de forma que não se precisa replicá-los e nem ocupar espaço no SVN. A execução de todo o sistema é realizado ao rodar o script run\_obs.sh que tem a função de incrementar uma variável com a data e hora das rodadas para o período desejado, rodando sequencialmente os processo envolvidos, chamando o script run\_bfge.sh para gerar, entre outros arquivos, o arquivo PrepBUFR de entrada do PAQC e posteriormente chama o run\_paqc.sh responsável por chamar o pacote PAQC e fazer o controle de qualidade nos dados contidos nesse arquivo.

Em uma descrição geral do script run\_paqc.sh, nota-se que o mesmo:

- a) Recebe via argumento de entrada a data da análise que se refere o arquivo em processamento. Um segundo argumento é o numero de processadores a ser utilizado. Um terceiro argumento é a flag do TestCase onde se define o diretório dos dados a serem utilizados no processamento;
- b) Testa os argumentos de entrada e seta as variáveis apropriadamente, em especial para o tipo de TestCase, selecionando os arquivos PrepBUFRS e o Background do modelo;
- c) Esse script cria um diretório de execução e apaga o antigo, caso exista;
- d) Cria variáveis FV e adiciona o diretório criado no path do usuário;
- e) Cria um subdiretório bin e copia para o mesmo todos os executáveis necessários;
- f) cria um subdiretório etc e copia para o mesmo todos os arquivos de configuração necessários;
- g) Testa o arquivo de entrada e carrega em variáveis;
- h) Linka os arquivos de Background;

- i) Organiza o arquivo em blocos usando a ferramenta block do PAQC;
- j) Lista os arquivos e combina em apenas um;
- k) Scaneia o PrepBUFR e lista todos os tipos de sistemas de observação nele existente;
- l) Faz um teste específico para os tipos AIRCFT, PROFLR e AIRCAR e seta variáveis avisando a existência;
- m) Executa o PAQC rodando o script ssprepqc chamando todos os processos em cadeia;
- n) Espera o processo terminar e lista os arquivos no diretório de execução;
- o) Renomeia o arquivo de log gerado no processo com os dados da data e hora da execução, do tipo \$obs\_work\_run\_paqc\_log/paqc\_\$LABELANL\_\$RUNTMqc.log;
- p) Busca o arquivo de saída com o controle de qualidade adicionado e apresenta uma mensagem de sucesso ou fracasso.

#### **4.2.4 Detalhes do script de execução do PAQC**

Para executar todo o processo PAQC basta acionar o script run\_paqc.sh passando os argumentos de entrada seguindo as recomendações da opção ajuda que é apresentada quando nenhum argumento é dado. Essa ajuda é apresentada na figura 4.1 para facilitar a descrição de cada um dos argumentos requeridos para a execução do script.

```

Luiz.sapucci@eslogin07:/scratchin/grupos/assim_dados/home/Luiz.sapucci/obs/run> ./run_paqc.ksh

-----
>>>>> Execucão do PAQC no Sistema de GCQD <<<<<<
-----

==> Para rodar o processo PAQC digite:

> ./run_paqc.ksh LABELANA NPROC TESTCASE

* LABELANA      : Data da analise ou do centro da janela de dados (YYYYMMDDHH) ex. 2015010100
* NPROC         : Numero de processadores a ser utilizado (no maximo 4)
* TESTCASE      : Flag para opcao testCase sendo 0 para NotTestCase e 1 ou 2 ou 3 para as opcoes:
                  =0 Nao se deseja rodar o testeCase usando os dados bufr gerados no pacote do OBS
                  bfrg lendo os dados em /scratchout/grupos/assim_dados/home/Luiz.sapucci/obs/dataout
                  Para isso rode o pacote bfrg com dados convencionais antes de chamar o paqc
                  com atencao para as datas selecionadas. Exemplo:
                  ./run_paqc.ksh YYYYMMDDHH 4 0
                  =1 utiliza dados bufr gerados no pacote bfrg e em conjunto com o PAQC, essa opcao
                  e para teste conjunto do paqc e bfrg. Exemplo:
                  ./run_paqc.ksh 2014080106 4 1
                  =2 utiliza dados bufr operacional do NCEP do tipo GDAS disponivel para teste na
                  data 2014080206. Exemplo:
                  ./run_paqc.ksh 2014080206 4 2
                  =3 utiliza dados bufr gerados previamente no pacote do OBS bfrg e disponiveis para
                  o testcase do PAQC na data 2014080106 a 2014080318. Exemplo:
                  ./run_paqc.ksh 2014080106 4 3

PS. Dados background sao oriundos do G3DVAR, para opcao testCase os dados sao lidos no DataFix do Gdad
Para rodar normalmente os dados sao lidos /scratchout/grupos/assim_dados/home/Luiz.sapucci/G3DVAR/datainout/gsi/

==> Para testCase basico do Paqc digite:

> ./run_paqc.ksh 2014080206 4 2

==> A opcao ajuda mostra as informacoes sobre os argumentos de entrada. Para isso digite:

> ./run_paqc.ksh ajuda

Para mais informacoes sobre o pacote leia os ../doc/README_GCQD_from_wiki_in_*.pdf usando o evince:
> evince ../doc/README_GCQD_from_wiki_in_*.pdf

Creditos: GDAD/DMD/CPTEC/INPE

```

Figura 4.1 - Detalhes para executar o PAQC.

Como descrito na figura, o run\_paqc.sh possui duas funções acionadas pelo primeiro argumento ao executar o script, sendo:

- **Função ajuda:** apresenta o mesmo conteúdo apresentado na figura 4.1, o que é o mesmo se nenhum argumento for fornecido;
- **Função running:** a função de execução do processo é acionada se os argumentos são suficientes para a rodada. Um total de 3 argumentos são necessários para executar o script, os quais são listados com mais detalhes:

LABELANL: Data para o processamento da janela de dados (sendo do tipo YYYYMMDDHH);

NPROC: Número de processadores a serem utilizados (no máximo 4);

TestCase: é uma flag para indicar onde o sistema deve procurar os dados necessários para a execução, nos quais se incluem os dados PrepBUFRS



de entrada, e os campos de Background do modelo global.

**para a opção 0:** representa "NoTestCase", não se deseja rodar o TestCase usando os dados PrepBUFR gerados no pacote do BFGE. Os dados são oriundos de outras fontes;

**Para a opção 1:** utiliza os dados PrepBUFRS gerados no pacote BFGE e em conjunto com o PAQC, essa opção é para o teste conjunto do PAQC e BFGE;

**Para a opção 2:** utiliza os dados PrepBUFRS operacionais do NCEP do tipo GDAS, disponível para teste na data 2014080206;

**Para a opção 3:** utiliza os dados PrepBUFRS gerados previamente no pacote do BFGE e disponíveis para o TestCase do PAQC na data 2014080106 a 2014080318;

**Nota 1.** Os dados do Background são oriundos do G3DVAR, para opção TestCase os dados são lidos no DataFix do Gdad. Para rodar normalmente os dados do Background devem estar disponíveis no diretório da variável \$work\_gsi\_datain\_bkg.

#### 4.2.5 Passo-a-passo para a execução do PAQC

- a) Login no supercomputador Tupã:
  - ssh usuário@tupa.cptec.inpe.br -XC
- b) Acessar o diretório scratchin (variável \$SUBMIT\_dataoutHOME):
  - cd \$SUBMIT\_HOME
- c) Acessar o diretório dos scripts do obs, que caso tenha sido feito a instalação padrão o acesso para esse diretório é:
  - cd obs/run
- d) Executar inicialmente o run\_paqc.sh sem argumento para rever as informa-

ções descritas acima certificando que não há outras opções implementadas e não tratadas nessa página:

- e) `./run_paqc.sh` ajuda
- f) A saída deverá ser a mesma apresentada na figura 4.1. Depois disso execute o TestCase seguindo os argumentos apresentados abaixo:
- g) `./run_bfge.sh 2014080206 4 2`

Observe a sequencia dos processos realizados pelo script na figura 4.2, onde algumas informações entre os módulos foram suprimidas na figura para simplificar o entendimento da sequencia dos mesmos:

```
Luiz.sapucci@eslogin07:/scratchin/grupos/assim_dados/home/Luiz.sapucci/obs/run> ./run_paqc.ksh 2014080206 4 2
-----
====> Running PAQC package for data assimilation ... <====
-----
> Analise para      : 2014080206
> Numero de processadores : 4
> TestCase        : 2

>>> Diretorio background : /scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/BackGround
>>> Arquivo buf de entrada: /scratchin/grupos/assim_dados/home/gdad/DataFix/OBS/testcase/PrepBufR_bfge/20140802/gdas1.20140802.t06z.prepbuf.r
20140802.t06z.prepbuf.r

***WARNING*** G3DVAR will update quality markers from NCEP PrepBUFR
blocking from: gdas1.20140802.t06z.prepbuf
to: gdas1.20140802.t06z.prepbuf.blk
FORTRAN STOP
FORTRAN STOP

Escaner do arquivo buf: gdas1.20140802.t06z.prepbuf.blk
-----
date: 2014080206      0
ADPUPA 164
AIRFCFT 10215
SATWND 134917
PROFLR 35
VADWIND 1718
ADPSFC 126648
SFCSHP 21070
GPSIPW 1903
RASSDA 249
ASCATW 231206
-----
> Executando tpreqc chamando todos os processos em cadeia do PAQC .. Aguarde ... |
> Escrevendo o arquivo prepbuf de saída.
-----
<<<< PAQC - running data assimilation - arquivos de saída com log
-----
/scratchout/grupos/assim_dados/home/Luiz.sapucci/obs/dataout/20140802/gdas1.20140802.t06z.prepbuf.blk
```

Figura 4.2 - Processos realizados pelo PAQC.

Os arquivos de saída do processo dos dados em PrepBUFR formato tabela do NCEP com as flags de controle de qualidade são gerados a partir do arquivo PrepBUFR gerado no BFGE. Observe que a diferença desses arquivos é a extensão do nome do arquivo onde os dados sem controle de qualidade possuem a extensão "nqc" que significa "Not Quality Control". Os arquivos de entrada e saída podem ser acessados em `$WORK_HOME/obs/dataout/20140802`. O arquivo de log da rodada efetuada pode ser acessado no diretório `$WORK_HOME/obs/run/paqc/logfiles`, no qual a data e hora da execução ficam registradas no nome do arquivo.

## 5 Perspectivas futuras e considerações finais

Na continuidade do desenvolvimento do projeto principal que norteia essa proposta espera-se que uma maior quantidade de bases de dados seja incluída e que principalmente a estrutura de desenvolvimento seja mantida, para assim garantir que a melhor base de dados esteja sendo ingerida pela modelagem atmosférica do CPTEC, e que essa esteja sendo feita de forma robusta e organizada. Logo na concepção do projeto se estudou como o desenvolvimento da proposta pode ser particionada em versões de entrega para a operação do CPTEC e se definiu o que se pode oferecer em cada uma delas. Para descrever a evolução da proposta é detalhada na próxima subseção a lista de versões preliminarmente definida.

### 5.1 Lista de versões concebidas na elaboração da proposta com o cronograma das futuras entregas

Essa lista é muito útil para descrever o que se pode esperar desse projeto como perspectivas futuras e de forma geral aborda um cronograma de entregas para as próximas versões. Como esse é um processo dinâmico outras versões intermediárias podem ser idealizadas. A lista de versões é descrita abaixo:

- **Versão Inicial V0.0:** essa versão teve sua entrega no dia 29 de outubro de 2015. Nessa versão denominada "inicial" toda a estrutura de desenvolvimento do projeto foi montada, contemplando todos os ingredientes básicos iniciais para uma primeira versão disponível para a operação. Um pacote de conversão de dados modelo chamado Dummy foi desenvolvido, bem como ferramentas para validar os arquivos BUFRS gerados. Essa versão contemplou os dados de vento por satélite como arquivo de dados convencionais e dados de radiância do satélite NOAA-18 recebidos via RARS da DSA.
- **Versão básica V1.0:** essa versão é a que se está sendo documentada com a publicação do presente relatório. A mesma contempla uma ampliação da base de dados de satélites disponibilizando dados de radiância de outros sensores e com cobertura global dos satélites NOAA-19 e MEtOP-B. Dados de vento por satélite do canal infra vermelho também foi incluído. Os dados de rádio ocultação GNSS estavam previstos para serem disponibilizados nessa versão, mas tiveram que ser postergados para a próxima, por problemas nas tabelas NCEP de escrita dos arquivos BUFRS. Com relação aos dados convencionais, todos os módulos dos sistemas de observações incluídos foram estruturados e os dados de estações a superfície foram im-

plementados. O fator mais importante abordado nessa versão é que todo o sistema foi reestruturado para permitir a implementação compartilhada de uma maior número de colaboradores. Isso deverá contribuir significativamente para a aceleração da inclusão dos sistemas de observações ainda não incluídos.

- **Versão robusta V2.0:** nessa versão prevista para junho de 2016 deverá ser incluído o pacote de controle de qualidade OBSPROC do NCEP e o conjunto de dados mais completo tanto de satélites como de dados convencionais. Para os dados convencionais as mais variadas fontes de dados deverão ser incluídas com especial ênfase para a América do Sul. No que diz respeito aos dados de satélites espera-se que os dados de radiância dos sensores multi-espectrais sejam incluídos sobre todo o globo e a melhor base de dados de rádio ocultação seja utilizada, com o fluxo via DSA totalmente explorado.
- **Versão completa V3.0:** uma versão futura ainda não prevista e com data de entrega não definida, deverá ser viabilizada na qual a melhor base de dados sobre a América do Sul seja contemplada e os dados disponíveis para a assimilação sejam gerados com a menor latência, visando a viabilidade do ciclo rápido de atualização da condição inicial para modelagem de alta resolução. Essa versão espera-se que esteja disponível no final de 2016, momento que poderá assegurar que a base de dados disponível para a assimilação seja a melhor base de dados disponível no CPTEC, a qual espera-se que seja muito superior ao que é disponível no NCEP sobre a América do Sul e semelhante aos dados usados nesse referido centro no resto do globo.

## 5.2 Melhorias futuras identificadas para serem implementadas nas versões posteriores a V1.0 do GCQD

A presente versão, não sendo ainda uma mais robusta, tem em sua elaboração diversas bases de dados ainda não implementadas. Além disso, no desenvolvimento da proposta foram identificados diversos pontos que poderiam ser melhores desenvolvidos, mas que não o foram realizados para não prejudicar o andamento de outras atividades que não são dependentes dessas melhorias. Nessa seção será feita uma listagem dos pontos que foram identificados e uma previsão em termos de versões futuras em que essas melhorias deverão estar disponíveis. As principais que merecem

destaque são:

-Correção de um bug na biblioteca NCEP BUFR para o PAQC de dados de Profiler e de AIRCFT, pois na execução inicial do PAQC foi identificado um bug no controle de qualidade dos perfilhadores e para os dados de AIRCFT. Ao testar o PAQC numa determinada tarefa usando os dados do BUFR GDAD do NCEP, um problema foi identificado na biblioteca NCEP BURF ao fazer o controle de qualidade desses dados. Esse problema foi postergado para ser resolvido na próxima versão permitindo a publicação da versão V0.0. Pretende-se investir nessa tarefa para a versão V2.0 robusta, a ser disponível em junho de 2016.

-Um estudo para verificar a possibilidade de substituir o processo PAQC pelo novo sistema de controle de qualidade do NCEP, que é o ObsProc, o qual está em implementação em uma versão futura do GCQD. Toda a estrutura do ObsProc já está em processo de implementação em tarefas e sub tarefas. Aliada a essa tarefa tem-se a utilização das saídas do novo modelo do CPTEC como Background para realização do controle de qualidade dos dados aqui tratados. Modificações e adaptações deverão ser necessárias para o êxito dessa tarefa, o que demandará tempo e dedicação da equipe executora. A previsão dessa implementação é para versões posteriores a versão 2.0 Robusta, que é a última versão prevista nesse projeto até a publicação dessa versão.

-Aprimorar a ferramenta que cria o arquivo PrepBUFR abordando outros tipos de dados tais como radiossondas, dados de aviões, navios, boias, entre outras bases de dados. A versão Básica 1.0 contempla apenas os dados de estações SYNOP.

-E para o caso dos dados de satélites, outros sensores de radiância dos satélites já utilizados e recebidos na DSA e DOP devem ser adicionados, em especial dos sensores hiperespectrais. Para o caso dos dados de rádio ocultação, depois de disponibilizar uma versão inicial com os dados do MetOp, outras constelações deverão ser adicionadas posteriormente, tais como os dados do COSMIC e TerraSar e outras futuras como o COSMIC-2, entre outras.

### **5.3 Considerações finais**

O presente relatório, em um primeiro de uma série, apresenta em detalhes toda a filosofia de processamento dos dados empregada no sistema de Gestão e Controle de Qualidade de dados (GCQD) para a assimilação em modelos de PNT do CPTEC.

Esse sistema não apenas se apresenta como uma ferramenta para o tratamento dos dados, mas como um ambiente organizado onde diferentes tipos de dados, das mais variadas fontes e meios de recepção podem ser tratados de forma integrada, organizada e clara. Os processos são organizados e os diferentes sistemas observacionais são implementados em módulos de um sistema único. Uma metodologia de inclusão de novas fontes de dados é sugerida, para a qual ferramentas e versões de modelos são disponibilizados, facilitando o desenvolvimento e permitindo a validação e documentação. As tarefas associadas com a ampliação da base de dados utilizadas deverão estar em constante desenvolvimento e demandará relativamente um longo período de tempo. Um desenvolvimento completo e de difícil previsão uma vez que a disponibilidade de novos tipos de dados e modificações nos mesmos é sempre constante e requer atenção. Essa tarefa deverá fazer parte da rotina do grupo de assimilação de dados e o sistema em desenvolvimento deverá ser uma ferramenta importante para o bom êxito da mesma. Na Previsão Numérica de Tempo, independência, eficiência e qualidade são características importantes do processo de obtenção dos resultados. Como essas características se completam, o desenvolvimento de processos em que essas três características são atendidas deve ser a meta principal de todo centro de previsão operacional. O desenvolvimento do sistema GCQD é um passo extremamente importante para a obtenção dessas três características no CPTEC. Com um processo próprio de gestão de toda a base de dados disponível para a assimilação na modelagem numérica, faz o centro ser independente de condições iniciais geradas em outros centros. O fato de não ter que esperar processos externos serem finalizados e dominando todo o processo de gestão dos dados possibilita que os resultados sejam gerados de forma eficiente, com disponibilidade antecipada. Além disso, a melhor base de dados para a América da Sul pode ser garantida o que impacta diretamente na melhor qualidade dos resultados sobre esse domínio, o que é a missão do CPTEC.

## REFERÊNCIAS BIBLIOGRÁFICAS

FERREIRA, S. H. S. **MBUFR: módulo para codificação e decodificação de dados em FM94-BUFR (Versão 1.5)**. São José dos Campos: INPE, 2007. 90 p. Programa de Computador. Disponível em: <http://urlib.net/sid.inpe.br/mtc-m17@80/2007/02.28.14.37>. Acesso em: 01 set 2016. 49

WORLD METEOROLOGICAL ORGANIZATION (WMO). Manual on codes - international codes. 2011. 5





## A ANEXO REVISÕES DO PROJETO

GCQD - Gestão e Controle de Qualidade de dados para AD													
Visão geral	Atividade	Planejamento	Tarefas	Nova tarefa	Gantt	Calendário	Notícias	Documentos	Wiki	Arquivos	Repositório	Revisões de código	Configurações
<b>Revisões</b>													
#	Data	Autor											
78	29 January 2016 14:45	Fernando Sapucci, Luiz	Copiando a versão obs_V1.0 do branch para o ramo tag para publicação da versão.										
77	29 January 2016 13:45	Amarante, Lucas	Foi comentado o namelist referente ao rognss do script run_bfge e foi deslocado um print que mostra o tipo de observação utilizada no módulo para dentro de um if (mudança que não impacta fortemente nos módulos.)										
76	29 January 2016 09:19	Amarante, Lucas	Organizando a estrutura de diretórios do bfge, exclusão dos diretórios conv e sat que estão atualmente no data/										
75	28 January 2016 15:03	Fernando Sapucci, Luiz	Modificando o diretório raiz no scratchin e scratchout de obs para obs_V1.0										
74	28 January 2016 11:00	Fernando Sapucci, Luiz	Revisão para retirar os external dos links aos projetos. Os externals retirados e os respectivos diretórios no obs foram: src/paqc/Linux <a href="https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSIsa/Linux/src/paqc/src/NCEP_Shared">https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSIsa/Linux/src/paqc/src/NCEP_Shared</a> <a href="https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSIsa/src/NCEP_Shared...">https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSIsa/src/NCEP_Shared...</a>										
73	28 January 2016 10:27	Fernando Sapucci, Luiz	Cópia do trunk para o ramo branch para o check final antes do lançamento da versão V1.0 básica do pacote de processos desenvolvido no projeto do Sistema de Gestão e Controle de Qualidade de Dados para a Assimilação (GCQD). A versão copiada é a da revisão 72 onde...										
72	25 January 2016 15:34	Amarante, Lucas	Foi reestruturado novamente essa revisão. O programa principal chama bfge.F90 e foi dividido o módulo select que antes estava somente em um módulo para os 4 módulos dos tipos de observação (convencionais, radiancia, rognss e satwind). Foi criado o diretório data para melhor...										
71	22 January 2016 15:27	Amarante, Lucas	Mudança na estrutura do bfge: todo o controle está sendo feito agora por apenas 1 select que recebe o vetor de observações e executa a leitura dos buffers e escrita correspondente a cada observação da lista. Modificações foram feitas no script run_bfge, main, e nos Makefiles.										
70	20 January 2016 09:31	Salvador, Nicolas	Adicionando gráficos individuais para radiancias RARS por tipo de satélite no módulo de validação, também incremento de 5 saídas estatísticas para canais do METOP-B, NOAA-18 e NOAA-19.										
69	17 December 2015 10:02	Salvador, Nicolas	implementação da validação do sensor AMSUA para o satélite METOP-1 da meteosat, saídas gráficas e estatísticas dos canais 5-9.										
68	15 December 2015 15:00	Amarante, Lucas	Inclusão da verificação do (002023-SATELLITE DERIVED WIND COMPUTATION METHOD) no módulo de escrita do prepbufr para os dados de vento por satélite. Essa verificação faz parte do nome para o mnemonic SID do prepbufr. Verificação: C SATELLITE DERIVED WIND METHOD AS READ IN - ISATIDS:...										
67	09 December 2015 10:15	Amarante, Lucas	Ajuste no módulo da observação de dados de superfície, ajustados para o funcionamento e criação do prepbufr para os dados de superfície na categoria surface data land. Ajustes em outros módulos como leitura bufcr acrescimo de um vetor para data category.										
66	07 December 2015 15:53	Salvador, Nicolas	Atualização no script conhg_obs.ksh, a compilação do paqc está sendo...										
65	30 November 2015 07:46	Amarante, Lucas	Implementação de módulo noaa-metop para leitura de dados do sensor AMSUA que consegue ler simultaneamente os dados observados pelos satélites NOAA-18 NOAA-19 decodificador do NOAA-19 e as estatísticas dos canais 5,6,7,8 e 9. A...										
64	27 November 2015 16:15	Amarante, Lucas	Modificação no run_bfge: ajuste na função de criação do namelist para a observação de superfície. Foram realizadas as atualizações principalmente no write dos módulos de observação: Exclusão da rotina criada gera_dhr para utilização da subrotina da biblioteca w3lib (w3difdat) para o cálculo preciso do DHR (diff time) Atualização do Makefile criação de variáveis para os diretórios ...										
63	27 November 2015 14:21	Fernando Sapucci, Luiz	Copiando a versão da revisão r0.1 para a tag principal com os external removidos e os conteúdos replicados na tag.										

61	○ ○	27 November 2015 14:02	Fernando Sapucci, Luiz	<a href="https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSisa/Linux...">https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSisa/Linux...</a> Para retirar foi usado anteriormente um arquivo vazio para sobrescrever o arquivo de externals. Nessa revisão foi feita uma ação mais eficiente usando um comando de propriedades do SVN. O comando foi <code>svn propset svn:externals para mais detalhes sobre manipulação de external ver paginas 61 em...</code> retirando os external dos links aos projetos. Os externals retirados e os respectivos diretórios no obs foram: <code>src/paqaLinux</code> <code>https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSisa/Linux</code> <code>src/paqaNCEP_Shared</code>
60	○ ○	27 November 2015 13:45	Fernando Sapucci, Luiz	<code>https://svn.cptec.inpe.br/g3dvar/trunk/G3DVAR/cptec/gsi/SRC/GSisa/Linux</code> <code>src/paqaNCEP_Shared</code> Criando um branch para resolver um bug na versão V0.0 que é a existência de diretórios ou parte do código linkados a outros projetos via comando externals, o qual é destinado para o desenvolvimento, mas não para a distribuição das versões.
59	○ ○	27 November 2015 11:01	Fernando Sapucci, Luiz	Nessa revisão todo o cabeçalho da estrutura Dummy foi revisado e reescrito. Informações inúteis permaneceram e outras foram removidas para a wiki da estrutura dummy. Uma cabeçalho padrão do Dummy permaneceu nos módulos e informações específicas dos diferentes tipos de...
58	○ ○	26 November 2015 15:53	Fernando Sapucci, Luiz	Foram realizados as mudanças e ajustes nas funções do run_bfge e acrescimo da funcao para criar o namelist do modulo de observacao de superficie ADPSC. Foi criado o modulo de observacao da superficie a partir do modulo do dummy.
57	○ ○	25 November 2015 16:06	Amarante, Lucas	Ajustado o calculo do difftime da funcao dos modulos de escrita do... Revisão para incluir modificações na janela de dados dos dados de Radioculação GRS5 no script que roda o BFGE. Ainda preciso de ajustes com relação ao nome dos arquivos.
56	○ ○	25 November 2015 10:13	Fernando Sapucci, Luiz	Adicionando analise a partir do sensor AMSU-A para escrita e appendando segundo o tipo de satellite (NOAA18, NOAA19 e METOP-B)
55	○ ○	25 November 2015 08:27	Salvador, Nicolas	Nessa revisão esta se fazendo (1) A estruturação dos dados de radio ocultação GRS5, baseada nas informações da tarefa #1544, (2) A adição do modulo de dados de Radio Ocultação GRS5 no obs usando a estrutura Dummy (tarefa 1546) adicionando o diretório <code>src/bfge/sat/rogrss/metop/</code> mas ainda...
54	○ ○	24 November 2015 11:12	Fernando Sapucci, Luiz	Foi criado a Subrotina para gerar o valor do mnemonic dhr (diff time da observacao) a partir da entrada da variavel de tempo lida no buff. Faz varias verificacoes de acordo com a hora minuto para gerar o valor correspondente ao DHR em relacao a hora sinotica. Foi visto num... Incorporando Modulo metop para radiâncias para leitura e escrita buff do sensor metop-B.
53	○ ○	23 November 2015 16:04	Amarante, Lucas	Incorporando leitura de dados NOAA-19 para o modulo NOAA e também estatísticas básicas para canais S e D do satellite NOAA-19.
52	○ ○	23 November 2015 09:39	Salvador, Nicolas	Nesta versao foi realizado ajustes na estrutura da janela de dados para ser incluída junto das observacoes no prebuffer. Foram ajustados o script run_bfge utilizando a variavel NANS para a janela de tempo de dados. Modificados os programas em fortran do pacote dummy e wind...
51	○ ○	23 November 2015 09:02	Salvador, Nicolas	Atualizacoes realizadas: -ajustes na estrutura do bfge no caso da escrita do prebuffer em relacao as observacoes. Na ultima versao o modulo de escrita prebuffer era um modulo generico que estava em buffers, apos reavaliarmos alguns pontos como a preparacao dos vetores e matrizes que serao lidos nos...
50	○ ○	17 November 2015 16:23	Amarante, Lucas	Mudancas realizadas: -documentacao inicial dos programas em fortran fazendo breve descricao dos mesmos; -inclusao do programa fortran de leitura dos dados GPSRO (radio ocultacao) - tabela NCEP) para inicio de desenvolvimento dessa observacao no bfge; -acrescimo na funcao de compilacao do config_obs.ksh
49	○ ○	13 November 2015 14:23	Amarante, Lucas	Copando a versao obs_V0.0 inicial do Branch para a tag, para a publicação da primeira versao do pacote desenvolvido no projeto de Gestao e Controle de Qualidade de Dados para a Assimilacao (GCQD). Para mais informacoes sobre a versao em publicacao acesse a wiki... Renomeando o diretorio raiz com as informacoes da versao em release.
48	○ ○	12 November 2015 10:09	Amarante, Lucas	Removendo o diretorio criado erroneamente.
47	○ ○	29 October 2015 18:06	Fernando Sapucci, Luiz	Copia do track para o ramo branch para o check finais da versao V0.0 inicial do pacote de processos desenvolvido no projeto do Sistema de Gestao e Controle de Qualidade de Dados para a Assimilacao (GCQD). Esse diretorio sera renomeado para obs_V0.0 na proxima revisao para manter o padrao de desenvolvimento de... Branch de check finais da versao V0.0 inicial do pacote de processos desenvolvido no projeto do Sistema de Gestao e Controle de Qualidade de Dados para a Assimilacao (GCQD)
46	○ ○	29 October 2015 18:00	Fernando Sapucci, Luiz	
45	○ ○	29 October 2015 17:57	Fernando Sapucci, Luiz	
44	○ ○	29 October 2015 17:56	Fernando Sapucci, Luiz	
43	○ ○	29 October 2015 17:47	Fernando Sapucci, Luiz	