

# PREPARAÇÃO DE DADOS DE OBSERVAÇÃO PARA O SISTEMA DE ASSIMILAÇÃO DE DADOS DO CPTEC

Rosangela Cintra<sup>1</sup>

Haroldo Fraga de Campos Velho<sup>2</sup>

## Abstract

Thousands of meteorological observations collected from the meteorological services around the world and satellites data are used for operating systems for Numerical Weather Prediction (NWP) employing different techniques of data assimilation. The Center for Weather Forecast and Climatic Studies of the National Institute for Space Research (CPTEC-INPE) is implementing the global analysis system called PSAS (Physical-space Statistical Analysis System) (GPSAS) to generate initial conditions for its global model CPTEC/COLA. The data assimilation process requests measures of the real state of the atmosphere, these measurements are the meteorological observations received from Global Telecommunications System (GTS)/National Institute of Meteorology (INMET), from satellite data, and alternatively from the National Environmental Satellite Data and Information Service (NESDIS), for satellite data. The analysis system gives an indispensable adjust of the model state of weather forecast in agreement with the observations, once there are not observations for all the points of the domain of the models. These observations are processed using standard codes from the World Meteorological Organization (WMO), and they are prepared in agreement with the requirements of the assimilation system. Additionally, several developments are made for pre-processing the observational data for the Global PSAS. In this paper methods used for selection and adaptation of observations will be presented, as well as a general view of the space distribution of observations received and used in CPTEC.

## Resumo

Milhares de observações meteorológicas coletadas pelos serviços meteorológicos de tempo no mundo e dados gerados por satélites são utilizados para iniciar sistemas operacionais globais de

---

<sup>1</sup> Instituto Nacional de Pesquisas Espaciais – INPE

Centro de Previsão do Tempo e Estudos Climáticos – CPTEC

Rodovia Presidente Dutra Km. 39. CEP: 12630-000 - Cachoeira Paulista, São Paulo, Brasil

<sup>2</sup> Instituto Nacional de Pesquisas Espaciais – INPE

Laboratório Associado de Computação e Matemática Aplicada – CPTEC

Av dos Astronautas, 1758. CEP: 12227-010 - São José dos Campos, São Paulo, Brasil

<sup>1</sup> [rcintra@cptec.inpe.br](mailto:rcintra@cptec.inpe.br), <sup>2</sup> [haroldo@lac.inpe.br](mailto:haroldo@lac.inpe.br)

Previsão Numérica de Tempo (PNT) utilizando diferentes técnicas de assimilação de observações. O Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) está implantando o sistema de análise objetiva global PSAS (Physical-space Statistical Analysis System) (GPSAS) para gerar condições iniciais para o seu modelo global CPTEC/COLA. O processo de assimilação de dados requer medidas do estado real da atmosfera em tempo real, estas medidas são as observações meteorológicas que são recebidas no centro através do *Global Telecommunications System (GTS)*/ Instituto Nacional de Meteorologia (INMET) e como fonte alternativa para dados de satélite o *National Environmental Satellite Data and Information Service (NESDIS)*. O sistema de análise desempenha um papel indispensável para ajustar modelos de previsão de tempo de acordo com as observações, uma vez que não há observações para todos os pontos do domínio dos modelos. Estas observações, que chegam ao centro em formatos de códigos padronizados pela Organização Mundial de Meteorologia (OMM), devem ser preparadas e colocadas em formato coerente aos requisitos do sistema de assimilação. Foram desenvolvidos/adaptados diversos códigos computacionais de tratamento e preparação dos dados no formato adequado para o sistema GPSAS. Neste trabalho serão apresentados os métodos utilizados para seleção e adequação dos dados de observação, bem como uma visão geral das observações que são recebidas e utilizadas no CPTEC e como são apresentadas após sua preparação.

## **PALAVRAS CHAVES**

Assimilação dados, observações, pré-processamento de dados, previsão numérica do tempo.

## **INTRODUÇÃO**

Diariamente executam-se sistemas operacionais globais de previsão numérica de tempo como o que está sendo implantado no CPTEC, utilizando mais de 500.000 itens de informação sobre a atmosfera, milhares de observações meteorológicas coletadas pelos serviços meteorológicos de tempo no mundo e dados gerados por satélites. Atualmente, preferencialmente operacional, o CPTEC executa seu sistema de PNT com condições iniciais proveniente do National Centers for Environmental Prediction (NCEP). Essas análises utilizam o Sistema de Interpolação Estatística – SSI (Parrish and Derber, 1992). Diferentes técnicas são usadas para assimilar estas observações, como a técnica variacional tri-dimensional (3D-VAR), ou Interpolação Ótima (IO); qualquer técnica utiliza como entrada de dados as observações e a primeira estimativa proveniente do modelo. Os sistemas de assimilação atuais não podem assimilar todos os tipos de dados, e nem toda a quantidade deles. Então, algoritmos de seleção de dados têm que ser usados.

Padrões de medidas para instrumentos variam amplamente. Instrumentos convencionais, como radiossondas, realizam medidas em tempos fixos e (essencialmente) de locais fixos. Instrumentos de órbita são amostras mais irregulares em espaço e tempo. A maioria observações de superfície tende a estar em áreas densamente povoadas e os padrões de instrumentos de órbita são governados através de mecânicas orbitais. Em assimilação de dados atmosféricos, é importante entender tanto quanto possível sobre cada instrumento cujas observações serão assimiladas, em particular, características dos erros dos instrumentos de observação devem ser bem entendidas.

Observações atmosféricas, atualmente, incluem: medidas diretas (*in situ*) de variáveis atmosféricas incluindo velocidade do vento, pressão, temperatura, e umidade; observações visuais de nuvens, visibilidade e tipo de precipitação; e observações de sensores remotos (medidas indiretas) de temperatura, umidade, nuvens e vento. Considera-se que observações diretas tendem a ser mais informativas que as observações de sensoriamento remoto, e em contrapartida esta última tem uma cobertura melhor (especialmente para a América do Sul). Esta combinação de observações provê uma extensa base de dados para obter condições iniciais e verificar modelos de PNT e para pesquisa de futuros entendimentos da atmosfera. A concentração das observações sinóticas convencionais está principalmente sobre no Hemisfério Norte o que limita a precisão e o alcance útil de previsões do tempo significativamente. O advento de satélites meteorológicos trouxe meios efetivos para completar a rede convencional de observações, não só provendo observações nas regiões pobres de dados dos oceanos, como no Hemisfério Sul e estratosfera, e também uma resolução horizontal mais refinada que as observações convencionais, devido a isto é muito importante a inclusão destes dados nos sistemas de assimilação.

Em Previsão Numérica de Tempo são necessários dois componentes básicos: um estado “real” da atmosfera, definido pelas observações meteorológicas, e a necessidade de um modelo atmosférico que possa ser integrado no tempo através de um procedimento numérico. Na estimação do estado inicial, a partir de um conjunto de dados de observação incompletos no espaço-tempo utiliza-se a técnica de assimilação de dados. O procedimento cíclico onde o resultado de uma integração curta, tipicamente de 6 h é usado como “estimativa inicial” (FG), combinado com as últimas observações disponíveis, formam o novo estado inicial para a próxima integração do modelo. Esta pode ser uma maneira eficiente para processar as numerosas observações da natureza, pois a previsão de tempo depende de uma representação precisa do estado inicial da atmosfera (Daley,1991).

O PSAS (*Physical-space Statistical Analysis System*) é um sistema de assimilação de dados com a versatilidade e capacidade de acomodar desenvolvimentos futuros na metodologia de assimilação de dados de uma maneira consistente com a teoria de estimação; baseado em Interpolação Ótima (OI), embora formule as equações de análise de uma maneira diferente, são

matematicamente equivalentes (Guo e da Silva, 1995), foi desenvolvido pelo Global *Modelling and Assimilation Office* (GMAO)/NASA e está sendo implantado no CPTEC.

As aproximações locais e a seleção de dados do esquema OI são eliminadas. O sistema é similar aos sistemas de análise espectral variacional de Centros como NCEP e ECMWF, porém trabalha diretamente no espaço físico. O PSAS desempenha uma grande parte de seus cálculos no espaço da observação, o que resulta em ganhos computacionais, pois a dimensão do espaço de observação é uma ordem de magnitude menor que o da previsão do estado. O GPSAS é o sistema de análise para o modelo global CPTEC/COLA que está sendo implantado (Herdies, 2002; Cintra, 2002) e o RPSAS é o sistema de análise para o modelo Regional Eta, que está em operacionalização no Centro. (Espinoza et al., 2002; Cintra et al., 2001).

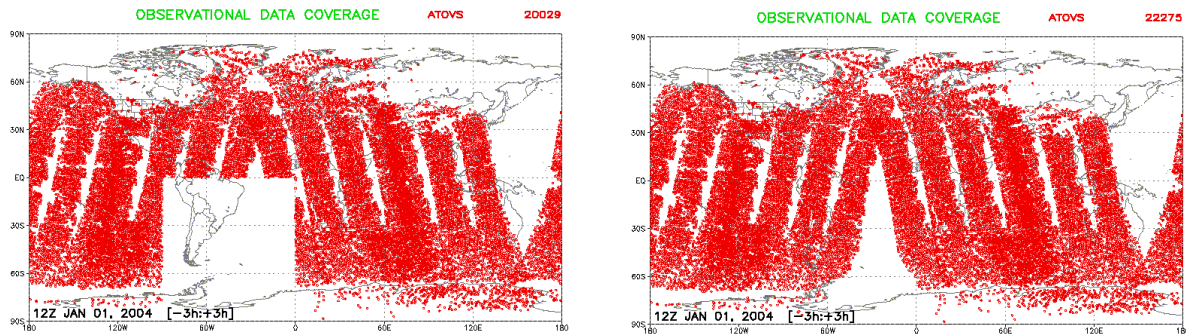
## AQUISIÇÃO DE DADOS

A WWW (World Weather Watch) mantida pela Organização Mundial de Meteorologia (OMM) em níveis: *global, regional, e nacional*, possui elementos como o Sistema de Observação Global (GOS) e o Sistema de Telecomunicação Global (GTS). O GOS provê dados de observação para agrometeorologia, meteorologia aeronáutica, e climatologia, inclusive o estudo de clima e mudança global. Todas as observações coletadas pelo GOS trafegam pelo GTS, composto de uma rede automática, para a distribuição de dados de observação. No Brasil, o órgão da OMM é o Instituto Nacional de Meteorologia (INMET) em Brasília/DF, órgão este responsável por repassar as mensagens coletadas no mundo, através do GTS. E este órgão repassa ao CPTEC os dados de observação globais, disponíveis, fazem parte da rotina operacional de aquisição de dados do CPTEC

O CPTEC recebe outros dados não distribuídos pelo GOS, nesta categoria encontram-se: as Plataformas de Coleta de Dados (PCD), dados reconstituídos de satélite (ATOVS), provenientes do National Environmental Satellite Data and Information Service (NESDIS), e provenientes da Divisão de Satélites Ambientais (D.S.A) do CPTEC/INPE, e ainda recebe dados de pesquisas.

Radiômetros de infravermelho e microondas a bordo de satélites de órbita polar NOAA (Administração Nacional do Oceano e da Atmosfera/EUA), são usados para observar radiâncias. Os dados de radiâncias são convertidos para perfis de temperatura usando um procedimento de recuperação baseado em física, estes dados reconstituídos ou “*retrievals*” são conhecidos como ATOVS (Advanced TOVS). A cobertura global é alcançada por cada satélite a cada 12 horas. Os dados produzidos são temperaturas e umidade em níveis de pressão, são assínóticos e com resolução horizontal de 80 km ou maior. Estes dados são muito importantes para completar a informação da atmosfera em todo o globo. A transmissão de dados ATOVS, via GTS/INMET,

apresenta, na distribuição espacial global, falta de dados no quadrante (0° 90S – 0° 90W), quadrante onde está a América do Sul. Como fonte alternativa do recebimento destes dados de satélite ATOVS com a cobertura global completa, o CPTEC, através de participantes do grupo ITSC (*Internacional TOVS Studies Conference*), obteve uma fonte alternativa de recepção de dados. Através de acordo junto ao NESDIS, copiam-se os dados ATOVS do “site” [www.cemsc.noaa.gov](http://www.cemsc.noaa.gov). Estes dados teoricamente seriam os mesmos transmitidos para o GOS/GTS/INMET, são adquiridos no CPTEC, também pela rotina operacional de aquisição de dados.



a)

b)

Figura 1 – Dados ATOVS recebidos: a) via INMET b) via NESDIS.

A Divisão de Satélites Ambientais (DSA) do INPE recebe radiâncias do satélite NOAA-15, NOAA-16 e NOAA-17, e utiliza o software ICI (*Inversion Coupled with Imager*) desenvolvido na Météofrance ([www.meteorologie.eu.org/ici/](http://www.meteorologie.eu.org/ici/)), para recuperar perfis de temperatura e umidade. Estes perfis são transformados para perfis de temperatura virtual em 15 níveis e são repassados diretamente para serem lidos pela interface de entrada do sistema de assimilação de dados para o modelo Regional do CPTEC, completando as observações necessárias dentro do domínio do modelo Eta. Estas observações não passam pelo sistema de pré-processamento de dados. Pode-se obter estas informações das variáveis de temperatura e razão de mistura como são recuperadas pelo software ICI, em 40 níveis, para ser codificada a exemplo do arquivo recebido do GTS e do NESDIS, e serem incorporadas no sistema de assimilação global GPSAS.

E ainda, quando acontecem experimentos em campo para coletar medidas de determinadas regiões do Brasil ou América do Sul, com a finalidade de pesquisa e estudos determinados, alguns destes dados ficam disponíveis no CPTEC, por exemplo, radiossondas do projeto LBA (Experimento de Grande Escala da Interação Atmosfera-Biosfera na Amazônia). Houve uma campanha de campo WETAMC (*Wet Season Atmospheric Mesoscale Campaign*) componente do LBA e do TRMM (*Tropical Rainfall Measuring Mission*) onde foram coletadas radiossondas para estudo da atmosfera, no sudoeste da Amazônia durante os meses de cheia: janeiro e fevereiro de

1999, e outro experimento, SALLJEX (*South American Low-level Jet Experiment*), onde foram coletadas informações de radiossondas em janeiro e fevereiro de 2003. Para a utilização em pesquisas de previsão de tempo e assimilação de dados, estas observações foram adquiridas através de pesquisadores e preparadas para inclusão nas análises dos modelos regional e global.

## **O SISTEMA DE PRÉ-PROCESSAMENTO DE DADOS.**

As observações recebidas no CPTEC são lidas e decodificadas através de um sistema de aquisição e pré-processamento de dados, que é o software de decodificação/codificação das mensagens adquiridas no centro, o sistema foi desenvolvido pelo Centro Europeu de Previsão de Tempo de Médio Prazo (ECMWF), onde ainda hoje é usado operacionalmente. O sistema foi repassado ao CPTEC em 1992 e então, adaptado para ser utilizado; diversas modificações foram feitas no sentido de aprimorá-lo e adaptá-lo às necessidades e sistemas computacionais. O sistema de aquisição de dados foi desenvolvido no CPTEC e apresenta características diversas do ECMWF. Em 1995, este software entrou em operação gerando observações para alimentar a análise objetiva/JMA (*Japan Meteorological Agency*) do modelo global do CPTEC. Em 1999, os arquivos de mensagens codificadas pelo pré-processamento de dados passaram a alimentar o Banco de Dados Meteorológico (BDM) do CPTEC implantado naquela data.

O sistema de pré-processamento atual com a aquisição de dados do CPTEC recebe diversos tipos de observações pré-definidas, definição esta acompanhada pela necessidade de obtenção de variáveis de estado para a assimilação de dados. São as pré-processadas mensagens de superfície, de ar superior, de satélite. O sistema de aquisição de dados recebe os diversos arquivos com mensagens/boletins meteorológicos que são transmitidos pelo GTS/INMET. Separa as mensagens criando arquivos distintos por tipo. O núcleo do sistema de pré-processamento é composto por programas de decodificação de cada tipo de dado e de controle de qualidade das observações: um nível sintático, onde as observações que não estiverem no formato pré-definido pela OMM são rejeitadas e outro, onde é verificada a confiabilidade das informações (Norris, 1990). A observação recebe uma marca (“*flag*”) de qualidade. Esta marca é informada em conjunto com a codificação das observações em formato binário BUFR (*Binary Universal Format for the Representation – FM 94*), vide publicação WMO no. 306 – Part B; este formato criado pela OMM garante um formato compacto para armazenamento dos dados e possível de ser lido por diversos centros meteorológicos. Estas observações em formato BUFR são armazenadas no Banco de Dados de observações do CPTEC (BDM). Os tipos de dados decodificados/codificados pelo sistema de pré-processamento atual são: SYNOP, SHIP, METAR, BUOY, TEMP, PILOT, AIREP, SATEM, SATOB, GRID, GRIB. Uma rotina separada foi feita para adquirir e pré-processar dados ATOVS.

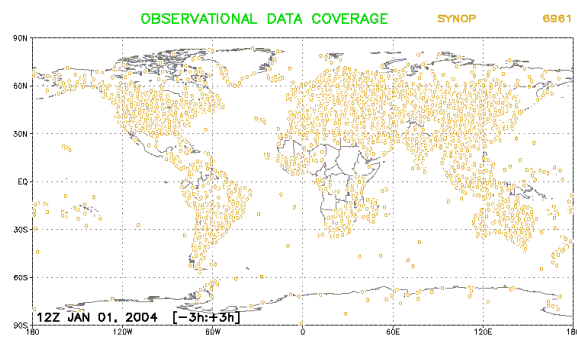
Para automatizar/operacionalizar a rotina de aquisição e pré-processamento de dados no centro foram desenvolvidos cerca de *quarenta scripts* divididos em quatro principais módulos de execução com a função de: (a) aquisição de dados, (b) execução de codificação /decodificação das mensagens, (c) preparação do arquivo de entrada para análise objetiva, (d) limpeza dos diretórios de execuções e arquivos de 48 horas anteriores. Ver documentação do sistema em Cintra e Borges (1999). Todos os dados recebidos no CPTEC são armazenados no seu formato original para possíveis re-análises ou re-processamento se necessários.

Mapas de cobertura global para cada um destes sistemas de observações descritas para um período de 6 horas centralizados em 12 GMT para 1 de janeiro de 2004 são mostrados nas figuras abaixo. Estas figuras ilustram a heterogeneidade do GOS bem como a cobertura típica de cada tipo de dado. Através do BDM do CPTEC e do sistema de Pré-processamento de dados, obtivemos a estatística de recebimento de dados do CPTEC (05.01.2004).

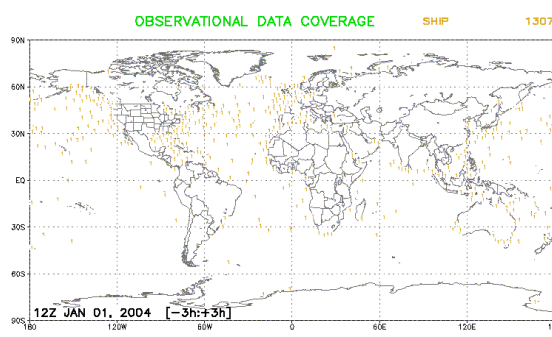
Tabela 1 – Quantidade de observações recebidas em média pelo CPTEC.

Tipo de Observação	Código GOS	Quantidade 12GMT	Quantidade diária
Estações de ar superior	TEMP	507	1114
Estações de ar superior	PILOT	280	781
Aeronaves	AIREP	4195	18574
Estações de superfície terrestre	SYNOP <sup>1</sup>	6995	28535
Estações de superfície terrestre	METAR	10412	43258
Estações de superfície marítima	SHIP <sup>1</sup>	1409	5724
Bóias a deriva e oceanográficas	BUOY	3348	11884
Satélite geoestacionário	SATOB	3598	12862

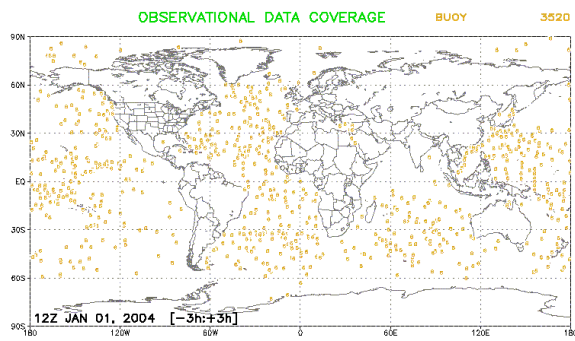
<sup>1</sup> Estas observações geram dois tipos diferentes de dados para armazenamento, portanto a quantidade recebida é aproximadamente 60 % da quantidade indicada.



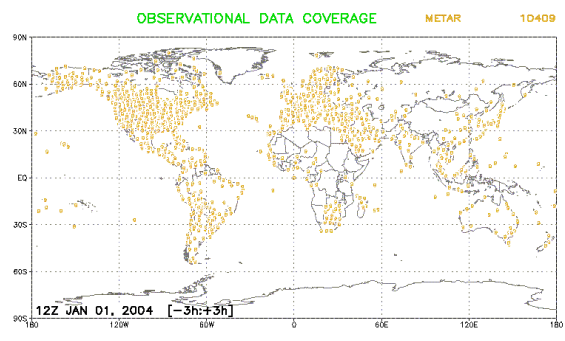
a)



b)

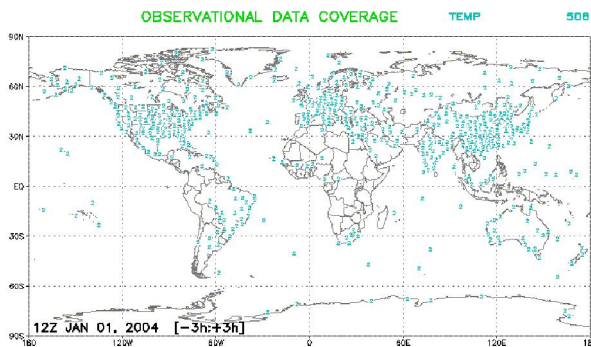


c)

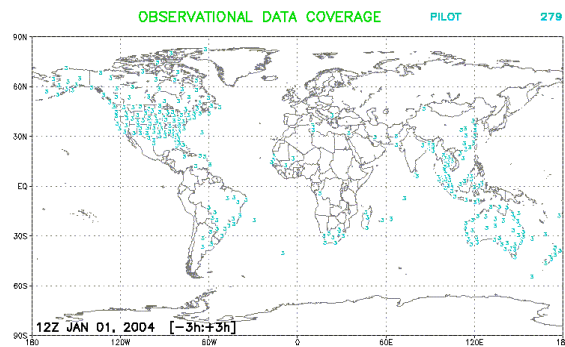


d)

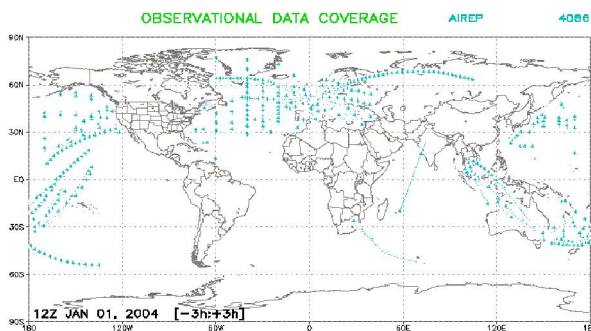
Figura 4 - Dados de superfície pré-processados a) SYNOP, b) SHIP, c) BUOY, d) METAR



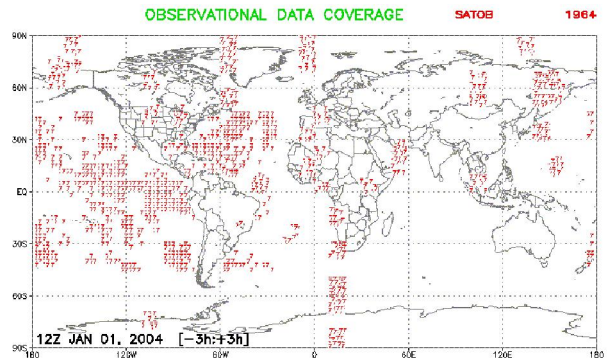
a)



b)



c)



d)

Fig. 5 – Dados de ar-superior: a) TEMP; b) Pilot; c) Pilot Dados de Satélite geoestacionário SATOB

## METODO DE PREPARAÇÃO DAS OBSERVAÇÕES PARA O SISTEMA DE ASSIMILAÇÃO DE DADOS PSAS

As observações chegam ao CPTEC preferencialmente pelo GTS (Sistema de Telecomunicações Global) e são armazenadas em um formato decodificado no BDM (Banco de dados Meteorológico), através do sistema de pré-processamento de dados. Antes da assimilação de dados, as observações



seriam extraídas do banco de dados. Estes dados já sofreram controle de qualidade rudimentar, por exemplo, uma aferição do formato e posição da observação, para limites climatológicos e hidrostáticos como também uma consistência interna e temporal, respectivamente. Então um arquivo de observação satisfatório para assimilação é criado em um módulo tipo “*pré-processamento*” da observação, através de programa interface que cria a entrada do sistema de assimilação. Isto requer conversões de variáveis da observação, é verificada a confiabilidade da observação constante do código BUFR. O arquivo resultante contém toda a informação de observações da janela de dados (atualmente seis horas) e é uma contribuição inicial para um sistema de previsão integrado, ou ciclo de assimilação. A observação é examinada e pré-seleciona observações únicas.

Uma quantia muito grande de dados observados está disponível para uso da assimilação de dados e sistemas de previsão. Em um período de 6-horas há um total de mais de 500.000 dados para serem usados. Todo dado disponível é considerado para uso na análise. Exceções são observações de satélite (dados localmente densos) onde é aplicada alguma seleção. Para alguns tipos de dados, há necessidade de uma transformação de variáveis físicas necessárias para melhorar o procedimento de análise. Os vários tipos de dados têm características muito diferentes em termos de cobertura geográfica, estrutura vertical, distribuição temporal e habilidade corrigir o estado modelo. A necessidade desta prática de *pré-seleção* é devido ao fato de que cada vez mais temos observações da atmosfera e que os algoritmos de assimilação não estão preparados numericamente ou computacionalmente para todas elas, então é necessário compatibilizar algoritmos e tamanho do vetor de observações. Todos estes fatores suportam um outro “*pré-processamento*” de observações que justificam o desenvolvimento do software que gera o arquivo de entrada para sistema de assimilação de dados.

### **Arquivo ODS - *Observation Data Stream***

Para o sistema de assimilação de dados que está sendo implantado no CPTEC, o GPSAS, é necessário um conjunto de observações que disponibilizem variáveis de estado que possam compor o vetor observação para solução do problema de análise. Este conjunto de observações, que retrata a atmosfera medida num determinado momento, deve estar de acordo com os requisitos de entrada do sistema de assimilação de dados, para isto foram desenvolvidos programas-interface que preparam estas variáveis. Um novo formato de arquivo de dados de observação chamado ODS (*Observations Data Stream*) é utilizado especificamente pelo PSAS, deve ser suficientemente autodescritivo contendo todo tipo de informação adicional para caracterizar a observação. A ferramenta da autodescrição da observação é a lista de tipos de dados e fontes usadas no horário da sua criação, ou

seja, cada observação possui o tipo de dado associada a sua fonte (instrumento) com os atributos que lhe acompanham. A implementação do conceito ODS foi feita em FORTRAN-90, usando a interface NetCDF para o “*Hierarchical Data Format*” (HDF), desenvolvido no “*National Center for Supercomputer Applications*” (NCSA. Como o software do GPSAS foi adotado do Global Modelling Assimilation Office (GMAO)/NASA, também adotamos o formato ODS/HDF. O formato assegura portabilidade do software e arquivos de dados entre diversos conjuntos de plataformas de hardware e fazer arquivos ODS suficientemente autodescritivos e prover vários pedaços adicionais de informação sobre a observação como identificadores de sondagens e de tempo. O formato do arquivo ODS varia entre *pré-análise* e *pós-análise*. Um arquivo de pré-análise deveria ser produzido somente pela descompactação dos dados do sistema de pré-processamento e não conter os atributos que só podem ser providos pelo sistema de assimilação (por exemplo, a previsão de 6 horas, o valor de análise interpolado). Um arquivo pós-análise deve ser produzido pelo sistema de assimilação, tendo a mesma informação de arquivos ODS de pré-análise complementada pela informação específica de assimilação. Não obstante, um arquivo ODS pós-análise também deve ser uma contribuição válida para sistemas de assimilação de dados, que provê a mesma informação do arquivo ODS pré-análise adicionado das marcas de controle de qualidade da assimilação original. Esta capacidade é extremamente útil para isolar os efeitos do subsistema de controle de qualidade ao testar o impacto de modificações para o sistema (da Silva, 1995).

### Estrutura do Arquivo

Todo arquivo ODS contém informação de *cabeçalho* sobre a observação informada e do próprio arquivo. Muito dos nomes de informação dos atributos da observação, conjuntos de intervalos válidos e unidades.

- Atributos globais são parâmetros do arquivo ODS e não mostram informações sobre as observações (tabela 2).
- Listas contém informação sobre índices válidos para os atributos da observação (tabela 3).

Tabela 2 - Atributos globais

Nome do Atributo	Descrição	Atual colocação
source	Fonte do arquivo ODS	Data Assimilation Office, NASA/GSFC
title	Título do arquivo ODS	
type	Tipo de arquivo ODS	
version	Versão do arquivo ODS	2.10
data_info	Ponto de contato	data@dao.gsfc.nasa.gov
history	Histórico dos dados	

Tabela 3 - Conjunto de listas que contêm informação dos índices para os atributos de observação.

Nome	tamanho	Descrição	Veja
kt names	nkt	Nomes dos tipos de dados	Tabela 5.5
kt units	nkt	Unidades dos tipos de dados	Tabela 5.5
kx names	nkx	Nomes de fontes de dados	Tabela 5.6
Kx meta	nkx	kx específicos para informação de meta-dado	Tabela 5.6
qcx names	nqcx	Significado para cada possível valor de qcexcl	Tabela 5.7

Um arquivo ODS consiste de observações que foram relatadas e organizadas em *segmentos*, que contêm toda a informação da observação para uma determinada data e hora sinótica. Cada segmento é definido via ponteiro que contem a localização e o número da observação do segmento, o valor da observação vem acompanhado de uma série de atributos que descrevem a medida. A lista destes atributos consta da Tabela 4.

Tabela 4 – Lista dos atributos de dados que formam o arquivo ODS.

Nome do Atributo	Descrição	Fortran Interface	Unidades	Intervalo válido	Número de “bytes”
lat	Latitude	Real	degrees north	[-90,+90]	2
lon	Longitude	Real	degress east	[-180,+180]	2
level	Nível ou canal	Real	hPa/none		4
Julian	Dia juliano	Integer			1
time	Minutos do dia juliano	Integer	minutes	[180, 65354]	2
kt	Índice do tipo de dado	Integer		[1,255]	1
kx	Índice da fonte de dado	Integer		[1,65535]	2
ks	Índice da sondagem	Integer		[1,65535]	2
xm	Índice do metadado	Real			4
qcexcl	Flag de exclusão do controle de qualidade	Integer		[0, 255]	1
Qchist	Marca da hitória do controle de qualidade	Integer		[0, 65534]	2
obs	Valor da observação	Real	Depende do kt		4
omf	Observation menos previsão 6h	Real	Depende do kt		4
oma	Observação menos análise	Real	Depende do kt		4

Esta é a lista de atributos de observação. A interface de FORTRAN se refere a este atributo de escrita/leitura para arquivar usando o software desenvolvido pela GMAO/NASA.

O índice de *tipo de dado*, denominado *kt*, é um inteiro que identifica a observação, possui um índice associado a uma descrição e às unidades. Esta é a identidade de índice observável. Por exemplo, uma observação de pressão ao nível do mar tem  $kt = 3$ . Uma lista dos tipos de dados poderá identificar cada tipo de dados e suas unidades. A lista parcial de tipos de dados usada em DAS (índice *kt*) segue na tabela 5. A versão atual e completa desta lista deve ser incluída em cada arquivo ODS.

Tabela 5 - Lista de alguns tipos (*kt index*) de dados usados na assimilação de dados

<b>kt</b>	<b>variable</b>	<b>units</b>	<b>descrição</b>
1*	us	m/sec	vento zonal de superfície (10m)
2*	vs	m/sec	vento meridional de superfície (10m)
3*	slp	hpa	Pressão ao nível do mar
4*	u	m/sec	Vento zonal de ar superior
5*	v	m/sec	Vento meridional de ar superior
6*	h	m	Altura geopotencial de ar superior
7*	w	g / kg	Razão de mistura do vapor d'água de ar superior
8	T	Kelvin	Temperatura de ar superior
9	Td	Kelvin	Temperatura no ponto de orvalho de ar superior
23*	thck	m	Thickness

\*tipos utilizados na atual versão dos arquivos ODS.

O índice da *fonte da observação*, denominado *kx*, é um inteiro que identifica a origem da medição, possui um índice associado a uma descrição. Por exemplo, uma medida com  $kx = 7$  foi feita por uma radiossonda. A lista parcial de fontes de dados usada em DAS (índice *kx*), veja tabela 6. A versão atual e completa desta lista deve ser incluída em cada arquivo ODS.

Tabela 6 - Lista de algumas fontes (*kx index*) de dados usados para assimilação de dados PSAS.

<b>kx</b>	<b>Description</b>	<b>Data</b>
1	Estação de superfície terrestre - 1	Synop
3	Estação de superfície marítima - 1	Ship
6	Bóia a deriva	Buoy
7	Radiossonda	Temp
8	Balão piloto de vento	Pilot
9	Radiossonda de navio	Temp-ship
16	Informação de aeronave	Airep
19	Vento por localização de nuvens	Satob
42	Retrievals de temperatura do NESDIS SH marítimo	Atovs
90	Estação de superfície terrestre de aeroporto	Metar

*Sounding index* (ks), este índice é necessário permitir a junção de perfis que vem da mesma sondagem de uma determinada hora sinóptica. Por exemplo, uma radiossonda que informa perfis de ventos, temperatura e umidade, deveria ter o mesmo índice de sondagem associado com cada pedaço do dado.

Inovação (*omf*), é a observação menos previsão usada na análise estatística interpolada ao local de observação. Estes serão calculados para determinar se a observação foi usada na análise ou não. E a observação menos análise (*oma*), a observação menos a análise interpolada ao local de observação. Como a inovação, o operador de observação pertinente será aplicado ao estado de análise.

## **MÉTODO PREPARAÇÃO DAS OBSERVAÇÕES E PRÉ-SELEÇÃO DOS DADOS**

Atualmente, estão implantados dois softwares distintos que geram análises, um para o Modelo de PNT Regional Eta e outro para a análise do Modelo de PNT Global do CPTEC, temos dois sistemas de assimilação e um software de interface para cada software, a diferença reside no fato de que o arquivo ODS para o RPSAS é um ODS de pós-análise, ou seja, um arquivo consta o atributo *omf*, portanto faz a interpolação dos pontos de grade do modelo para o ponto de observação, a documentação desta interface esta descrita em relatório técnico interno da documentação da interface e em Cintra e Aravéquia (1998).

### **Arquivos de entrada para o GPSAS**

Os arquivos de dados globais são reunidos em janelas de tempo de 6 horas, tendo a hora sinóptica como hora central, o corte da recepção dos dados (*cutoff*) é feito com 2:30 horas depois do ultimo horário da janela de tempo, por exemplo, para reunir observações para gerar o arquivo da análise da hora sinóptica 12Z, com janela de - 3 horas e + 3 horas, seria das 9:00 às 14:59 hs., o *cutoff* é feito às 17:30 horas, e todas as observações daquela janela de tempo que forem recebidas após este arquivamento, não entram para a análise. Existem dois programas distintos que criam arquivos ODS pré-análise, ambos partem de arquivos em formato BUFR:

### **Arquivo ODS de dados convencionais e dados de vento de satélite**

Foi desenvolvido o programa *bufr\_to\_ods.x*, que faz leitura do arquivo de observações preparado pelo sistema de pré-processamento, verifica-se o “flag” do controle de qualidade das observações, onde foi estabelecido o primeiro critério de seleção: as informações com

confiabilidade menor que 40% não são utilizadas. As observações de boa qualidade são selecionadas de acordo com cada tipo/subtipo do dado informado pelo BUFR:

- Para os dados de superfície – observações SHIP e BUOY informam ventos (direção e velocidade) sobre os oceanos, e geram componentes de ventos. Observações SYNOP e METAR informam pressão reduzida ao nível do mar para as observações de superfície terrestre gerando também pressão reduzida ao nível do mar para o ODS. Os níveis destas observações são assinalados com 2000 hPa (convenção do ODS).
- Para os dados de ar-superior - observações TEMP informam direção e velocidade do vento, geopotencial, temperatura e temperatura do ponto de orvalho, geram para o ODS vento zonal e meridional ( $u, v$ ), altura geopotencial ( $h$ ) e razão de mistura do vapor d'água ( $w$ ) para todos os níveis medidos. As observações de aeronaves (AIREP e AMDAR) informam ventos e geopotencial, geram também componentes do vento ( $u, v$ ) e altura geopotencial ( $h$ ) nos níveis informados.
- Para os dados de satélite Geoestacionário – observações SATOB informam vento (direção e velocidade) e geram componentes de ventos ( $u$  e  $v$ ) para cada nível observado.
- Para os dados de satélite NESDIS - observações SATEM (produto de resolução 500 km) informam temperatura ( $t$ ) e temperatura do ponto de orvalho( $td$ ) para gerar altura geopotencial ( $h$ ) a partir das sondagens de temperatura para diversos níveis.

Foi necessário fazer *conversões* de variáveis para obter as observações requisitadas para o processo de assimilação de dados PSAS:

- Direção e velocidade do vento são medidas por instrumentos meteorológicos de superfície e de altitude. Direções são medidas do Norte. O vetor deve ser decomposto em duas componentes de fluxo ortogonais,  $u$  e  $v$ .  $u$  é conhecido como vento zonal e é positivo em direção a Leste.  $v$  é o vento meridional e é positivo de Norte. E devem ser computados como:

$$u = - S \cos(d) \quad (5.1)$$

$$v = - S \sin(d) \quad (5.2)$$

sendo  $S$  a velocidade do vento em  $m s^{-1}$  e  $d$  a direção em graus.

- A altitude da radiossonda é reportada em unidades de geopotencial sobre o nível médio do mar. A elevação da estação deve ser a posição geométrica inspecionada, e será convertida a altura de geopotencial. Geopotencial, designado  $\phi$ , é a energia potencial de uma parcela do ar de massa unitária, sendo que o seu valor referencial de zero encontra-se, em geral, no nível médio do mar. Geopotencial é:

$$\phi = \int_0^z g dz \quad (5.3)$$

- As unidades são  $m^2 s^{-2}$ . A unidade mgp (metro geopotencial) é definida como uma re-escala geopotencial, e é dado por:

$$h = \frac{1}{9.80665} \int g(z) dz \quad (5.4)$$

$$1 \text{ mgp} = 981 \text{ cm s}^{-2} \times 100 \text{ cm} = 9,81 \times 10^4 \text{ cm}^2 \text{s}^{-2} = 9,81 \text{ m}^2 \text{s}^{-2}$$

A altura geopotencial é obtida dividindo o geopotencial pela aceleração de gravidade referencial do planeta ( $g$ ), pois este varia pouco com a altitude na camada atmosférica, então é considerado como a constante (9,81) ao longo da vertical em toda camada.

- As informações reportadas pela radiossonda são *temperatura* e *temperatura do ponto de orvalho* que serão convertidas em *razão de mistura do vapor d'água*. A razão de mistura ( $r$ ) é definida como a razão entre a massa de vapor d'água e a massa de ar seco. É a medida da quantidade de vapor d'água que uma parcela de ar contém, calculada através da seguinte equação:

$$r = 0.622 * e / (p - e) \quad (5.5)$$

onde “ $p$ ” é a pressão atmosférica (hPa), “ $e$ ” a pressão de vapor (hPa), determinado por:

$$e = T_d * es / 100 \quad (5.6)$$

sendo “ $es$ ” a pressão de saturação do vapor em (hPa), definida como

$$es = 6.11 * \exp(10 * (7.5 * T) / (237.3 + T)) \quad (5.7)$$

onde  $T_d$  é a temperatura em ponto de orvalho e  $T$  é a temperatura do ar em graus K. Estas equações foram colocadas em uma função “*mixratio*”, onde os argumentos são as variáveis  $T$  e  $T_d$  das radiossondas (Vianello, 1991).

### **Arquivo ODS de dados convencionais e dados de vento de satélite**

O segundo programa desenvolvido *atovs\_to\_ods.x*, utiliza arquivo da rotina paralela de dados ATOVS adquiridos através do NESDIS, também agrupado em janela de 6 horas. Este arquivo, tem informações BUFR *tipo 3, subtipo 55* da tabela 5.8. Para atribuir o (*kx 33 a 56*) verifica-se o perfil de acordo com o hemisfério, horário, continente, se é observação de céu claro ou com nuvens. Para cada perfil deve-se adicionar o atributo “*ks*” para identificar os pedaços de observação como sendo de determinado perfil. Estes perfis são dados reconstituídos de radiâncias de satélite, recebidos como temperatura ( $T$ ), em graus Kelvin e razão de mistura ( $r$ ), em gramas de vapor por gramas de ar úmido, para um determinado nível de pressão. Estas variáveis não são assimiladas no GPSAS então foram codificadas as necessárias conversões. As variáveis  $T$  e  $r$  são transformadas em

espessura e em seguida em altura geopotencial, utilizando a equação hipsométrica, essa equação é uma aplicação importante da equação do balanço hidrostático. A equação hipsométrica relaciona a espessura entre duas superfícies isobáricas com a temperatura, verticalmente mediada entre esses dois níveis.

As transformações foram programadas para obter os perfis de altura geopotencial. Inicialmente obtendo os perfis de temperatura ( $T$ ) e umidade ( $r$ ), converte-se para *Temperatura Virtual* ( $T_V$ ) iniciando do topo (1.0 hPa) da atmosfera até 1000 hPa. A  $T_V$  é a temperatura do ar seco à mesma densidade do ar úmido, com a mesma pressão. O ar úmido é menos denso que o ar seco; portanto, a temperatura virtual é sempre maior que a temperatura real. Contudo, mesmo para ar seja muito quente e úmido, a temperatura virtual excede a temperatura real apenas por poucos graus. Calcula-se  $T_V$  utilizando a seguinte fórmula:

$$T_V = T(1 + 0.608 * r) \quad (5.8)$$

onde  $T$  e  $r$  são as variáveis informadas pelos perfis do ATOVS para um determinado nível de pressão, com  $T$  e  $T_V$  em graus K absolutos e  $r$  em (g/g). Não temos a informação de  $T_V$  para o nível de 925 hPa e este nível é requerido pelo GPSAS, então é feita a interpolação do nível 920 hPa e 950 hPa, para obter a  $T_V$  do nível requerido. Deseja-se obter o perfil de *altura geopotencial* com o perfil de temperatura virtual calculado; a partir da superfície até ao topo da camada (ordem inversa), utiliza-se a equação hipsométrica (5.9), que descreve a relação entre a pressão e a altitude na atmosfera, e possibilita calcular a espessura de camadas atmosféricas que possuam perfil de temperatura virtual linear:

$$\ln \frac{p_{i+1}}{p_i} = - \frac{g \Delta h}{R * T_{VM}} \quad (5.9)$$

onde  $p$  é a pressão em hPa;  $g$  é a aceleração da gravidade (padrão), igual a 9,80665, de acordo com as normas da OMM;  $R^*$  é a constante dos gases para o ar seco, igual a 287,04 m<sup>2</sup> s<sup>-2</sup>K<sup>-1</sup>;  $T_{VM}$  é a temperatura virtual média da camada, igual a  $[T_V(i + 1) + T_V(i)] / 2$  em graus K; e  $\Delta h$  é a espessura da camada em metros; e  $i$  refere-se ao índice dos níveis.

A espessura é designada para cada nível superior da camada em metros geopotenciais e pode ser somada para se obter a altura geopotencial  $h_n$  para cada nível de pressão. Então,

$$h_n = \sum_i^n \Delta h_i . \quad (5.10)$$

Assim é calculada a altura geopotencial requerida pelo sistema de assimilação, faz-se o somatório do  $\Delta h_i$ , partindo do primeiro nível 950 hPa informado, até o topo do perfil 0,4 hPa, totalizando 40 níveis. Estes perfis recebem o  $kt = 23$ .



Com critérios de pré-seleção obtém-se aproximadamente de 300.000 observações de ATOVS para a assimilação. Calcula-se o dia Juliano, atribuem-se as colunas *Julian e time* e se faz a gravação do arquivo para entrada do GPSAS. Com este método foram obtidos dados consolidados em arquivos ODS pré-análise.

### **Arquivo ODS de dados convencionais e dados de vento de satélite**

Para incorporação de informações de radiossondas, coletadas no experimento de campo do projeto SALLJEX de janeiro de 2003 na assimilação de dados utilizando o GPSAS e obter re-análises do modelo global, foram desenvolvidos dois programas: *lê\_rs.x*, programa que prepara as informações colocando-as em formato intermediário, a partir da leitura do arquivo tipo planilha com dados coletados; e o segundo programa *rs\_to\_gods.x*, que a partir do arquivo em formato intermediário criado pelo programa anterior, obtém-se um arquivo ODS pré-análise, possibilitando a inclusão deste arquivo com dados adicionais em conjunto com os outros arquivos ODS de dados convencionais e dados de satélite.

Para incluir observações de experimentos como radiossondas do LBA (1999) foi desenvolvido um programa *rs\_to\_ods.f*, onde são lidos arquivos tipo planilha em caractere com as informações de nível coletado (que ultrapassam os níveis mandatários e significativos das radiossondas operacionais). Foi criado um arquivo intermediário em formato próprio para ser lido no programa de interface do RPSAS (Cintra & Aravéquia, 2000).

### **Ferramenta de visualização ODSVIEW**

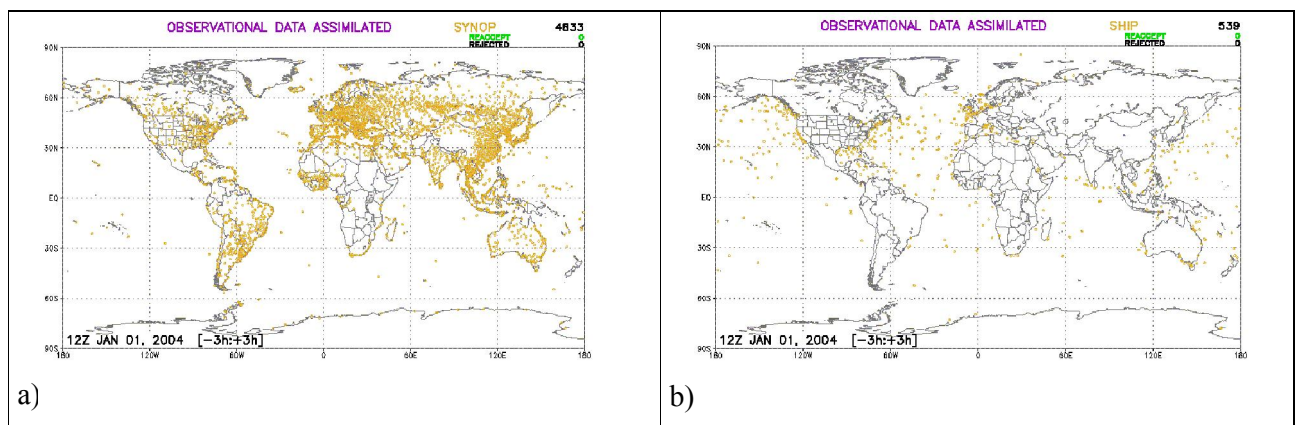
Para visualização das observações dos arquivos binários HDF em formato ODS, foi desenvolvida uma ferramenta chamada *odsview.x* (separadamente para os arquivos regionais e globais) que lê as informações e classifica-as por tipo, gerando arquivos (*station data*) para serem visualizados no GrADS (Grid Analysis and Display System), uma ferramenta interativa de fácil manipulação, e visualização de dados da ciência da terra (<http://grads.iges.org/grads/grads.html>). O programa *odsview.x* foi desenvolvido para leitura das observações; as informações foram ordenadas e divididas por instrumento, por perfil ou ponto da grade desestruturada das observações; para cada observação, foram gravados o valor da observação *obs* (no caso do ODS pré-análise) e os valores do *omf* e *oma* (no caso do ODS pós-análise). Para visualização dos arquivos foi desenvolvido um programa (*odsview.gs*) interpretado pelo GrADS, que disponibiliza a imagem de sua distribuição espacial. Indica também a quantidade de pontos (para os tipos de dados) que entrarão ou entraram na assimilação de dados. Para os arquivos pós-análise que possuem as marcas de controle de

qualidade, são contabilizadas e apresentadas espacialmente, as observações rejeitadas, e as observações re-aceitas.

## RESULTADOS E CONCLUSÃO

A Assimilação de Dados atmosféricos é um campo de estudo das ciências da terra que gera uma estimativa do estado da atmosfera num determinado tempo, através da combinação de uma previsão de tempo curta, fornecida por um modelo e de observações, ou seja, dados meteorológicos de diferentes fontes, no mesmo instante de tempo, o resultado é atualmente, o principal meio para fornecer condições iniciais para modelos de PNT.

Com estes método e desenvolvimento para preparação dos dados observacionais apresentados, obtivemos arquivos consistentes e adequados para o inserir no ciclo de assimilação de dados que está sendo implantado no CPTEC. Estes arquivos são o componente observacional que propiciará ao CPTEC obter de análise própria para o centro, independência de condição inicial para o modelo e independência de dados para a de geração de suas análises. Não há como realizar assimilação de dados sem os dados, e deve ser capaz de manipular a entrada de dados para pesquisas em assimilação. O método utilizado atentou para necessidade de obtenção das variáveis de estado para o sistema de análise e para atender as necessidades do sistema estatístico de análise objetiva em espaço físico PSAS, utilizando a estrutura de arquivos requerida por ele. As figuras abaixo apresentam a cobertura de dados resultante dos arquivos ODS pré-análise gerados para a análise do dia 01.01.2004 às 12 horas.



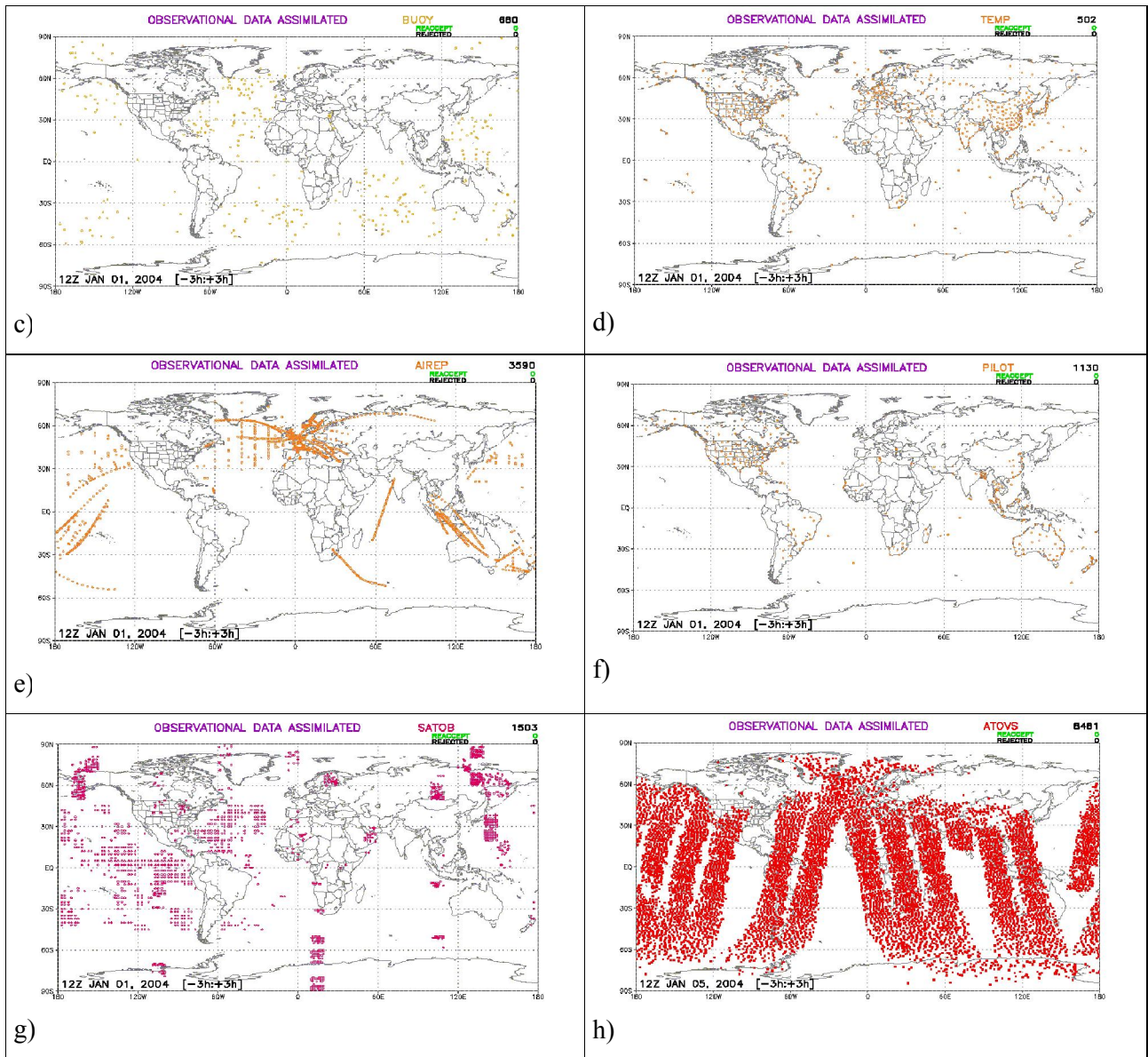


Fig. 6 – Distribuição de dados globais em arquivos ODS pré-análise: a), b), c) superfície; d), e), f) ar-superior; g) e h) satélite.

A recepção de dados através do NESDIS, como foi descrito, é uma alternativa para resolver o problema da falta destes sobre a América do Sul, existente nos dados recebidos do GTS/INMET. Preferencialmente adquirimos os dados da primeira fonte, mas se não são adquiridos por algum motivo, temos preparado na rotina de aquisição de dados, decodificar os dados do INMET com os mesmos “retrievals” preparados pelo DAS, onde obteremos um arquivo com cobertura quase completa.

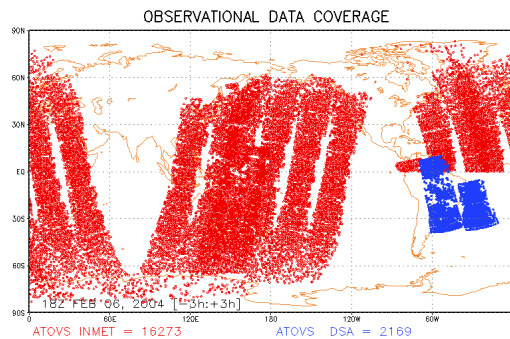


Fig. 7 - Distribuição de dados globais em arquivos ODS pré-análise: dados ATOVS, provenientes do INMET e D.S.A

## REFERÊNCIAS BIBLIOGRÁFICAS

Atlas, R.: Atmospheric Observations and Experiments to Assess Their Usefulness in Data Assimilation. Journal of the Meteorological Society of Japan, 1997, v.75, no. 1B, pp.111-130.

Cintra, R. e Aravéquia, J. A.: *Apresentação do Método de Geração de Dados para Assimilação do Modelo ETA/CPTEC*. Anais do X Congresso Brasileiro de Meteorologia. Brasília – DF, 1998.

Cintra, R., D. L. Herdies, J. N. Tamaoki. *Otimização do Sistema de Assimilação de Dados Atmosféricos PSAS para o Modelo Global do CPTEC*. Anais do XII Congresso Brasileiro de Meteorologia, 2002 Foz do Iguaçu - PR.

Daley, R. *Atmospheric Data Analysis*, New York: Cambridge University Press, 1991.

Espinoza E.S., R. Cintra, D. L. Herdies, J.A. Aravéquia, J. P. Bonatti, J. Bustamante, J.L. Gomes, S. H. S. Ferreira e W. F. A. Lima. *Sistema de Assimilação ETA/RPSAS*. Anais do XII Congresso Brasileiro de Meteorologia, 2002. Foz do Iguaçu– PR.

Guo J. ; Da Silva A.<sup>(2)</sup> *Computational Aspects of Goddard's Physical-Space Statistical Analysis System. (PSAS)*. Greenbelt, Maryland, 1995. 45 p.(DAO office Note 96-03).

Herdies, D.; Ferreira, S.H.; Bonatti, J. P.; Cintra, R.; da Silva, Arlindo. *O Sistema de Assimilação de Dados Atmosféricos Global do CPTEC/INPE*. Anais do XII Congresso Brasileiro de Meteorologia, 2002 Foz do Iguaçu– PR.