

APRESENTAÇÃO DO MÉTODO DE GERAÇÃO DE DADOS PARA A ASSIMILAÇÃO DO MODELO ETA/CPTEC

Rosangela Cintra
José Antonio Aravéquia
Centro de Previsão de Tempo e Estudos Climáticos-CPTEC/INPE
Rodovia Presidente Dutra, km 40
12630-000 -Cachoeira Paulista - SP
e-mail:rcintra@cptec.inpe.br,araveq@cptec.inpe.br

ABSTRACT

All observational data received at “Centro de Previsão de Tempo e Estudos Climáticos” (CPTEC) from the “Global Telecommunications Systems”(GTS) are archived with a consistent format BUFR (Binary Universal Form for the Representation of meteorological data). The BUFR data are pre-processed, selecting data types and performing a few simple quality check and corrections. Useful by-products of the assimilation system are the innovation files (Observation Data Stream – ODS) which contain the differences between the observations and the first guess of ETA model, interpolated to the observation location. This paper presents the structure of ODS file and how the GTS data are archived in it.

1 – Introdução

O “Global Telecommunications System” (GTS) distribui dados observacionais globais que são recebidos pelo “Centro de Previsão de Tempo e Estudos Climáticos” (CPTEC) e servirão para alimentar o modelo regional de Previsão Numérica de Tempo (PNT) ETA. Estes dados são preprocessados de forma a validá-los sintática e qualitativamente, com limites grosseiros, são armazenados em arquivos com formato BUFR (Binary Universal Form for the Representation–FM 94 -WMO), e então reunidos em janelas de tempo de 6 horas, tendo como hora central a hora sinóptica. Os dados observacionais serão então preparados para atender ao sistema de assimilação de dados utilizando o PSAS (Physical-space Statical Analysis System) da DAO (Data Assimilation Office) – NASA /EUA. Um novo formato de arquivo de dados observacionais chamado ODS (Observations Data Stream) é utilizado pelo PSAS Neste artigo será descrita a estrutura destes arquivos e como os dados são validados e preparados para serem arquivados nesse formato. Os arquivos ODS servirão de entrada primária para a matriz de covariância do modelo, e também para monitoramento e validação dos sistemas em desenvolvimento.

2- A estrutura da Entrada de Dados

Os arquivos ODS devem ser suficientemente autodescritivos contendo todo tipo de informação adicionais 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 associado a sua fonte 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).

O *tipo de dado*, denominado *kt*, é um inteiro que identifica a observação, conforme *Tabela 1*, que possui um índice associado a uma descrição e unidades. Por exemplo: $kt = 3$, apresenta um dado de pressão a nível do mar em hPa.

A *fonte da observação*, agora denominado *kx*, é um inteiro que identifica a origem da medição, conforme *Tabela 2*, que possui um índice associado a uma descrição. Por exemplo: $kx = 3$, apresenta um dado de superfície medido no navio (Ship).

Tabela 1 - Lista de tipos (*kt index*) de dados usados para assimilação de dados PSAS.

kt	variable	units	description
1*	us	m/sec	Surface (10m) zonal wind
2*	vs	m/sec	Surface (10m meridional wind
3*	slp	hpa	Sea level pressure
4*	u	m/sec	Upper-air zonal wind
5*	v	m/sec	Upper-air meridional wind
6*	h	m	Upper-air geopotential heigth
7*	w	g / kg	Upper-air water vapor mixing ratio
8*	T	Kelvin	Upper-air temperature
9*	Td	Kelvin	Upper-air dew-point temperature
10	rh	%	Upper-air relative humidity
11	q	g / kg	Upper-air specific humidity
12	Us	m/sec	Surface (10m) wind speed
13	Ts	Kelvin	Surface (10m) temperature
14	Tds	Kelvin	Surface (10m) dew-point temperature
15	rhs	%	Surface (10m) relative humidity
16	qs	g / kg	Surface (10m) specific humidity
17	P	mm/day	Precipitation rate
18	Q	mm	Total precipitable water
19	Qt	mm	Total cloud liquid water
20	c	percent	Fractional cloud liquid water
21	wx	none	Present weather code

*tipos utilizados na atual versão dos arquivos ODS.

Tabela 2 - Lista de fontes (*kx index*) de dados usados para assimilação de dados PSAS. (Contêm somente os códigos utilizados na atual versão dos arquivos ODS.)

kx	Description
1	Surface Land Obs –1
2	Surface Land Obs –2
3	Surface Ship Obs – 1
4	Surface Ship Obs – 2
5	Environment Buoy
6	Drifting Buoy
7	Rawinsonde
8	Pilot Wind
9	Ship Released Rawinsonde
10	Dropwinsonde
16	Aircraft Report
17	Aircraft Coded Report
19	Cld Trk Wind – Wisc
33	NESDIS SH Land – (120 km)
34	NESDIS SH Land - (500 km)
87	Pseudo-1000mb Heights

Tabela 3 – Lista dos atributos de dados que formam o ODS file.

Variable name	Description	Fortran Interface	Units	Valid Range	Internal Storage (bytes)
kt	Data type index	Integer		[1,255]	1
kx	Data source index	Integer		[1,65535]	2
ks	Sounding index	Integer		[1,65535]	2
km	Metadata index	Integer		[0,2 ³¹ -1]	4
lat	Latitude	Real	degrees north	[-90,+90]	2
lon	Longitude	Real	degress east	[-180,+180]	2
level	Level or channel	Real	hPa/none		4
julian	Julian day	Integer			1
time	Time stamp	Inteeger	minutes	[0,1439]	2
obs	Observation value	Real	Depends on <i>kt</i>		4
omf	Observation minus 6h forecast	Real			4
oma	Observation minus analysis	Real			4
Qc_flag	Quality control flag	Integer		[0,65534]	2
Mod_flag	Modification flag	integer		[0,255]	1

Cada observação tem um conjunto de atributos que devem ser assinalados, veja na *Tabela 3* acima, e que acompanham a codificação kx e kt. Estes atributos especificam a medição, por exemplo, o “índice da rádio-sondagem”, ou seja, identifica uma única observação a cada nível

Cada atributo do dado supra listado tem a identificação a seguir:

- *Sounding index(ks)*, deve conter a identificação dos perfis de uma mesma sondagem de um determinado horário sinótico. Por exemplo, a rádio-sonda reporta perfis de vento, temperatura e umidade, que são tipos de dados diferentes mas devem ter a identificação que pertencem a uma observação, são pedaços associados por um mesmo índice.
- *Metadada index(km)*, deve conter o ponteiro da informação que a completa.
- *Latitude(lat)*, deve conter a latitude da observação em graus negativos no hemisfério sul e positivo no hemisfério norte.
- *Longitude(lon)*, deve conter a longitude da observação em graus, a partir de 180W até 180E, longitudes a oeste serão negativas e a leste positivas.
- *Level(lev)*, deve conter o nível da observação, nos casos convencionais de radiosondagens se refere ao nível vertical, e com conceito estendido para os canais de medidas das radiancias, para os dados de superfície deve conter o nível de pressão se disponível.
- *Julian day (julian)*, deve conter a data juliana da medida, nesta convenção dia 23 de maio de 1968 corresponde ao dia juliano 2.440.000, a partir daí é calculado o dia juliano da observação.
- *Time stamp(time)*, o atributo contém o “elapsed time” em minutos desde 0 GMT, se refere ao tempo atual da medida e não o da hora sinótica.
- *Observation(Obs)*, deve conter o valor atual da observação, como é definido, u e v são componentes do vento e constituem 2 observações, o mesmo com perfis de temperatura para 18 níveis, são 18 observações separadas.
- *Inovation(omf)*, deve conter a diferença entre a observação e a previsão de 6 horas da análise estatística, interpolado no ponto da observação, usada ou não na análise. O relevante operador da observação será aplicado para o estado da previsão se necessário.
- *Observation minus analysis(oma)*, como o *omf*, deve conter a diferença entre a observação e a análise, interpolada para o ponto de observação. Este campo será preenchido após a execução da análise.
- *Quality control flag(qc_flag)*, contém o resultado de vários testes de controle de qualidade da observação, listados na *Tabela 4* abaixo.
- *Modification flag(mod_flag)*, acontecem testes simples “pass/fail”, ambos on- e off-line sistemas de controle de qualidade podem alterar o valor da observação para remover desvios, correções para erros de transmissão, ou combinar várias observações para uma simples “average” super observação. O *mod_flag* guarda o registro das modificações feitas durante o processo de assimilação, conforme a *Tabela 5* abaixo.

Tabela 4 – Convenções para o “flag” de controle de qualidade (qc_flag).

Bit Position	Test Description	Possible Values
1	Off-line gross limit check	0 or 1
2	Off-line climatological check	0 or 1
3	Off-line hydrostatic check	0 or 1
4	Off-line sea level pressure check	0 or 1
5	Off-line integrity test	0 or 1
6	Off-line black-list mark	0 or 1
7-8	Off-line Complex QC test	00 or 01 or 11
9-10	Off-line subjective test	00 or 01 or 11
11	On-line gross limit check	0 or 1
12	On-line buddy check	0 or 1
13	On line final QC decision	0 or 1
14	On-line passive data type mark	0 or 1
15-16	Bits currently not used	00

Tabela 5 – Convenções para o “flag” de modificação.

Bit Position	Description	Possible Values
1	Observation is a super-ob	0 or 1
2	Observation modified by Complex QC system	0 or 1
3	Bias correction applied	0 or 1
4	Derived data type	0 or 1
5-8	Bits not used	0000

Para cada arquivo ODS pré- ou pós- análise é necessário 35 bytes para cada observação, Para uma média de 4.000 observações para cada horários sinóptico, 4 vezes ao dia , requer mais ou menos 5 Mbytes de armazenamento por dia.

3. Método de Validação e preparação dos dados.

Nos arquivos em formato BUFR, os arquivos globais divididos em tipos/ subtipos, conforme *Tabela 6*, são reunidos em janelas de tempo de 6 horas, tendo a hora sinóptica como hora central. Note-se que neste formato os dados já foram preprocessados, ou seja, passaram por um tratamento de validação de dados, baseada em limites grosseiros, atribuindo um grau de confiabilidade ao valor reportado na observação, com controle de posicionamento das medições, etc., de acordo com Norris(1990).

Destes dados já consistidos, são solicitados somente os da área de atuação da América do Sul (60S a 20N, 90W a 20W). Os *dados de superfície* (kx 1, 3, 6) geram *us, vs, slp e h 1000mb*. O dado de *h 1000mb*(kt6 e kx 87) é calculado a partir da pressão reduzida a nível do mar e da temperatura do primeiro nível sigma do first-guess do ETA, utiliza da equação hipsométrica $Z = H \ln (p1/p2)$, onde $H=R\cdot T/g$ e $R = 287$, $g = 9.81$, $p1= slp$ e $p2 = 1000$. Dos *dados de ar-superior* (kx 7, 8, 9, 10, 16 e 17) geram *u, v, h, w* nos diversos níveis. Dos *dados de satélite NESDIS* (kx 33, 34, 35) geram *h* a partir da temperatura média virtual da camada, utiliza a equação hipsométrica supra, onde $H=R\cdot T_v/g$, $p1 =$ primeiro nível e $p2 =$ segundo nível, e serão incrementados às próximas camadas. Dos *dados de satélite Geoestacionário* (kx 19) são gerados *u e v* para cada nível observado.

Todos os dados são interpolados com o “first-guess”, previsão de 6 horas, gerado pelo modelo ETA, hora central a hora sinóptica, a mesma janela de tempo do arquivo BUFR, e geram a diferença entre a observação e o first-guess. Estas informações e seus atributos são colocados no arquivo *ODS* que será analisado pelo *PSAS*. Após a análise o arquivo *ODS pós-análise*, é acrescido de informações por ela gerada como a diferença da análise com o “first-guess”(oma) e o estado da observação (qc_flag), ou seja com os “flags” de controle de qualidade .

Tabela 6 – Tipos/Subtipos de dados em formato BUFR.

Category	description	Bufr type	Bufr Sub-type	Category	description	Bufr type	Bufr Subtype
Surface data-land	Synop-land	0	1	Single level	Satob sec 2	5	82
	Synop rec2	0	2	upper air	Satob sec 3	5	83
	Synop (auto)	0	3	satelite-geos	Satob sec 4	5	84
	Synop (auto) 2	0	4		Satob sec 5	5	85
Surface data-sea	Synop ship abbr	1	9	Vertical	Pilot	2	91
	Synop ship	1	11	Soundings	Pilot ship	2	92
	Synop ship rec2	1	12	(other than	Temp land	2	101
	Synop ship (a)	1	13	Satellite)	Temp ship	2	102
	Synop ship (a) 2	1	14		Temp drop	2	103
	Synop ship red	1	19	Oceanographic	Dribu/buoy	31	131
	Buoy/ drifter	1	21	data	Bathy	31	132
	Bathy	1	22		Tesac	31	133
	Tesac	1	23	Single level	Codar	4	141
	Tovs-120 km	3	51	Upper air	Airep	4	142
Vertical Soundings (satellite)	Satem-500 km ll	3	61	(other than	Colba	4	143
	Satem-500 km w	3	62	Satellite)	Amdar	4	144
	Satem-500 km hl	3	63		Acar	4	145

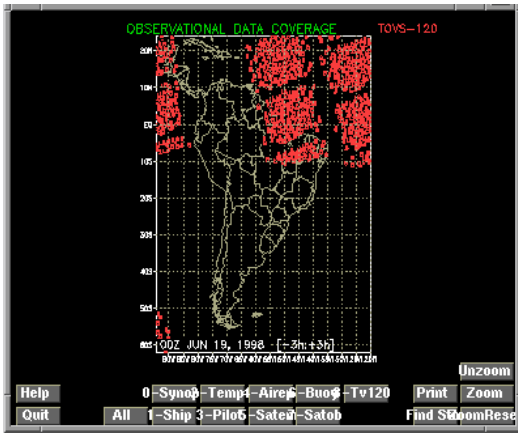


Fig1. Dados TOVS-120 Km – BUFR do dia 19/06/1998,00Z

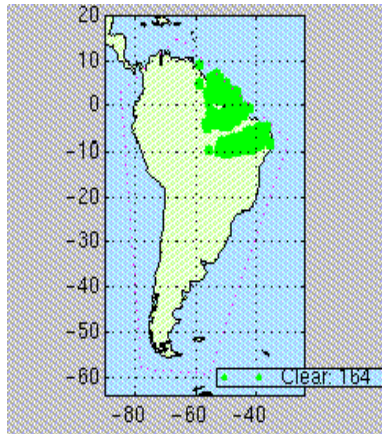


Fig2. ODS h (kt 6) pré-análise 19/06/1998,00Z.

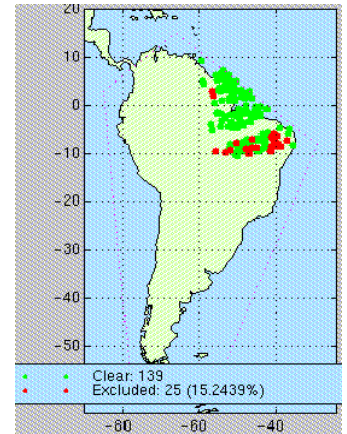


Fig3. ODS h(kt6) pós-análise 19/06/1998

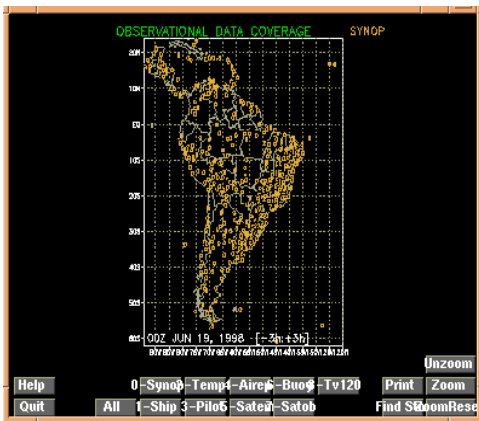


Fig4. Dados SYNOP – BUFR do dia 19.06.1998,00Z

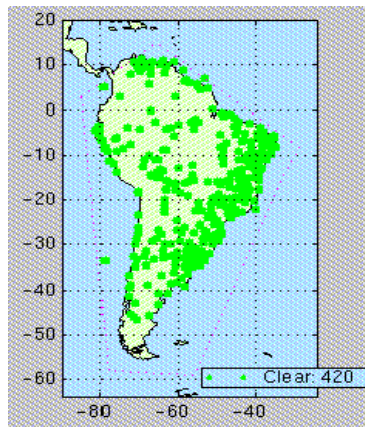


Fig5. ODS slp (kt 3) pré-análise 19.06.1998,00Z

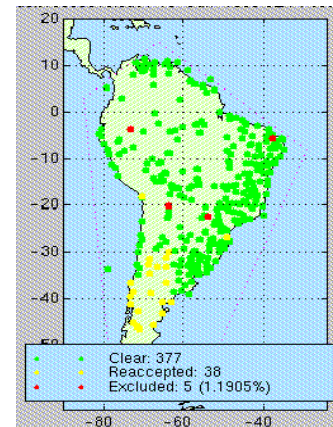


Fig6. ODS slp pós-análise 19.06.1998

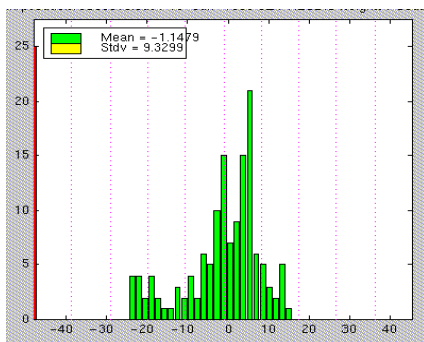


Fig7. ODS(omf) pseudo-heights de 19.06.98, 00Z.

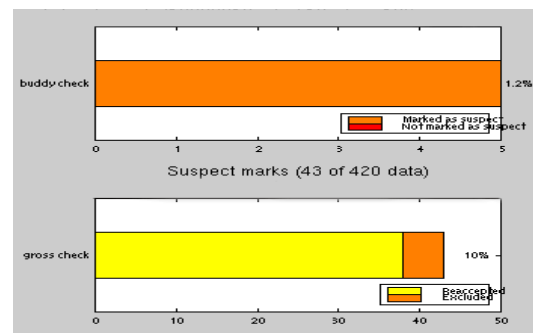


Fig8. ODS (qc-flag) – pseudo heights de 19.06.98,00Z

4. REFERÊNCIAS

DA SILVA, ARLINDO e REDDER, CHRISTOPHER, 1995 – Documentation of the GEOS/DAS Observation Data Stream (ODS)- Version 1.01

NORRIS, BRIAN, 1990 – ECMWF Meteorological Bulletin, Data checking and validation .

DRAGOSAVAC, MILAN, 1994 – ECMWF Meteorological Bulletin – BUFR – User’s Guide and Reference Manual.

HOLTON, J. R., 1992 – “An Introduction to Dynamic Meteorology”- 3rd. Ed. – Academic Press – 511 pags.