



Mesoscale and  
Microscale Meteorology

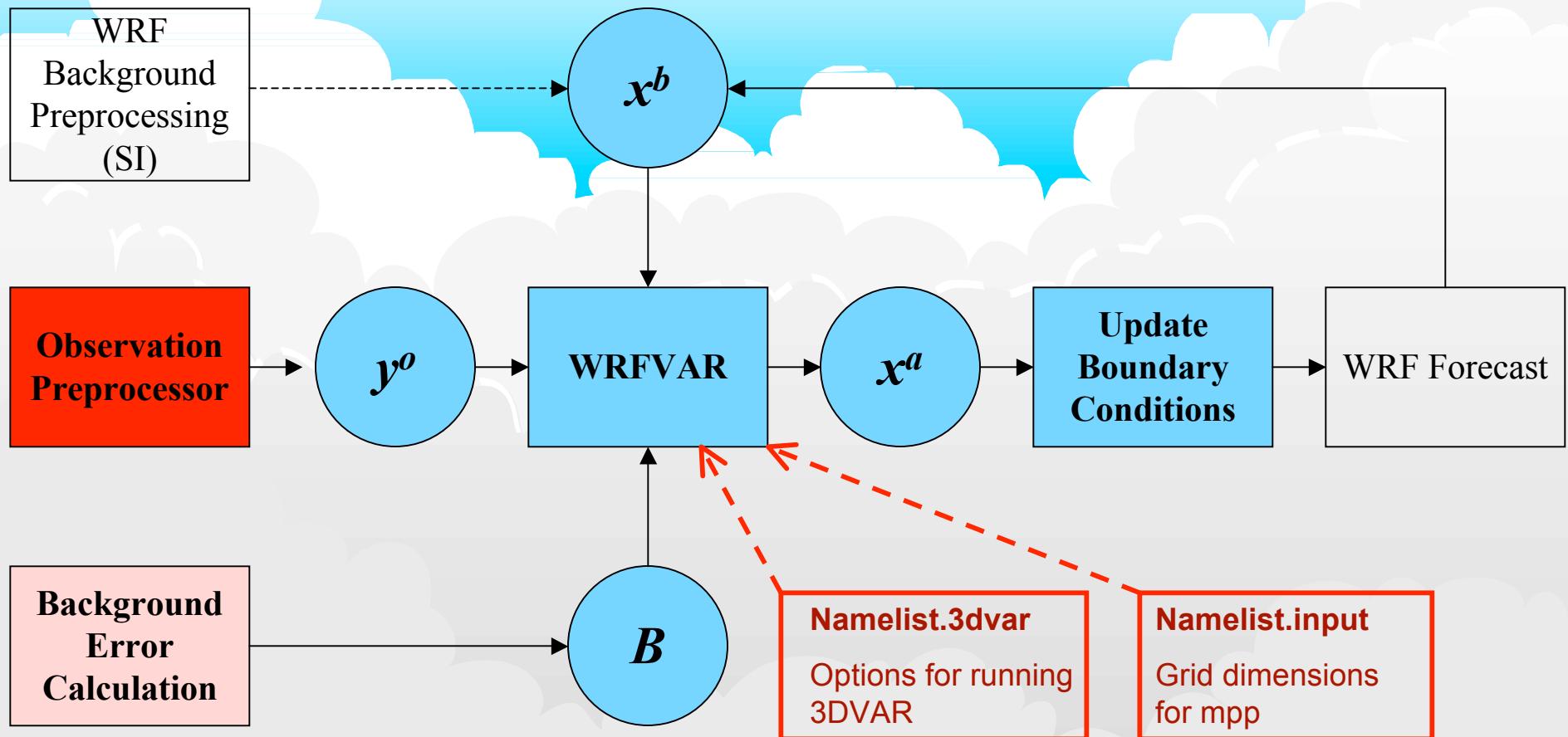
# Observation Preprocessor for WRF-Var

Yong-Run Guo

National Center for Atmospheric Research  
P.O.Box 3000, Boulder, CO 80307

Presented in WRF Tutorial, 26 July 2007

## WRF-Var in the WRF Modeling System



The observations are one of the important input files for WRF-Var, no observations will be no data assimilation.

# Why do we need the OBS preprocessor?

To prepare the OBS data file suitable for WRF-Var needs

- Screening of the conventional observations (time window, domain, duplication, etc.), and keep the necessary information for WRF-Var assimilation
  
- Save the OBS data processing time, **one OBS file** can be repeatedly used for multiple times of WRF-Var experiments.
  
- The output file from 3DVAR\_OBSPROC is still to be an ACSII format for easy manipulation (read, check, and edit), especially for research purpose.

## The data format of observations accepted by WRF-Var

### 1. BUFR and preBUFR --- format

Radiance data --- BUFR

Conventional data --- NCEP preBUFR (under testing)

ob\_format = 1 in WRF-Var

### 2. ASCII format --- for research and operational:

Current working format,

ob\_format = 2 in WRF-Var

fully supported

## Web sites for 3DVAR\_OBSPROC

The 3DVAR OBS preprocessor program can be obtained from:

[http://www.mmm.ucar.edu/wrf/users/download/get\\_source2.html](http://www.mmm.ucar.edu/wrf/users/download/get_source2.html)

By click “WRF-Var observation pre-processor”.

The files inputted to and Outputted from 3DVAR\_OBSPROC are ASCII files.

## Advantages of ASCII format observations

- The *input file* to 3DVAR\_OBSPROC is the LITTLE\_R format (see below)

This is a report-based format, so all types of the observation data can be easily ‘cat’ted together to form a monolithic file.

It is easily to read, edit,.... with an ASCII file.

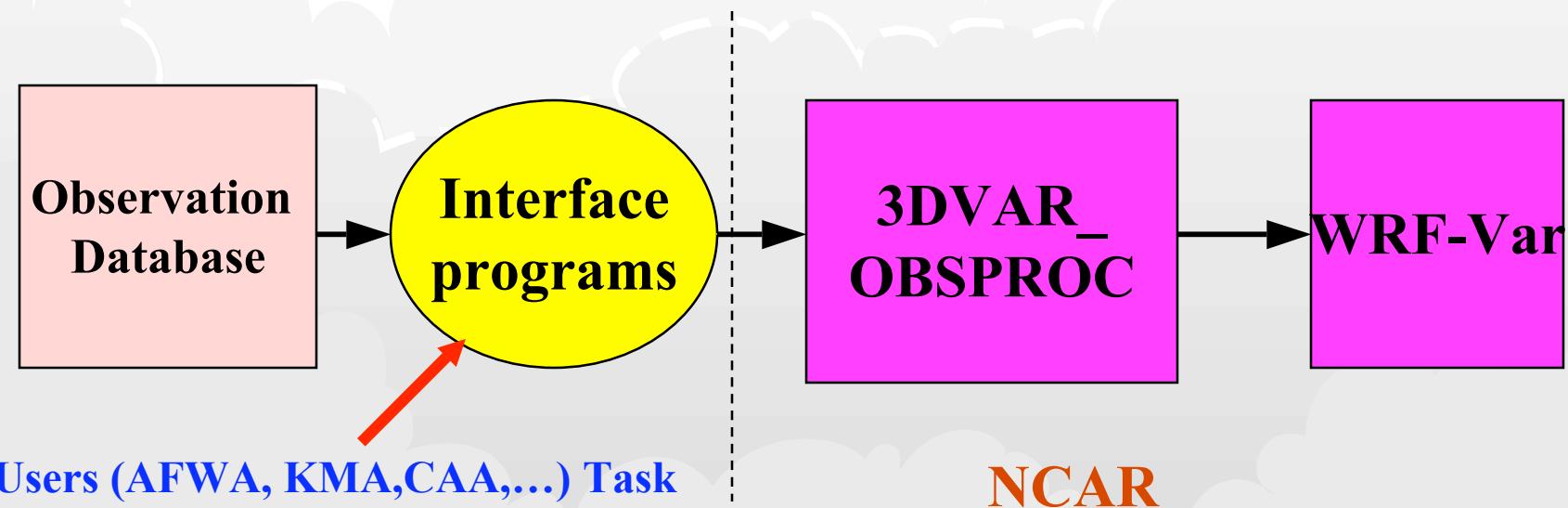
- The *output file* from observation preprocessor, 3DVAR\_OBSPROC program is still an ASCII file. It is easily manipulated for specific purposes of research, such as single point data tests, etc.
- Users’ duty is just to convert their own observations in any format to the LITTLE\_R format. Then that data can be processed by 3DVAR\_OBSPROC, and an observation file suitable for assimilation with WRF-Var is created.

## Disadvantages of ASCII format observations

The size of the observation file is larger in compared with the BUFR format file, not suitable for the remote sensing data, such as Radiance, etc.

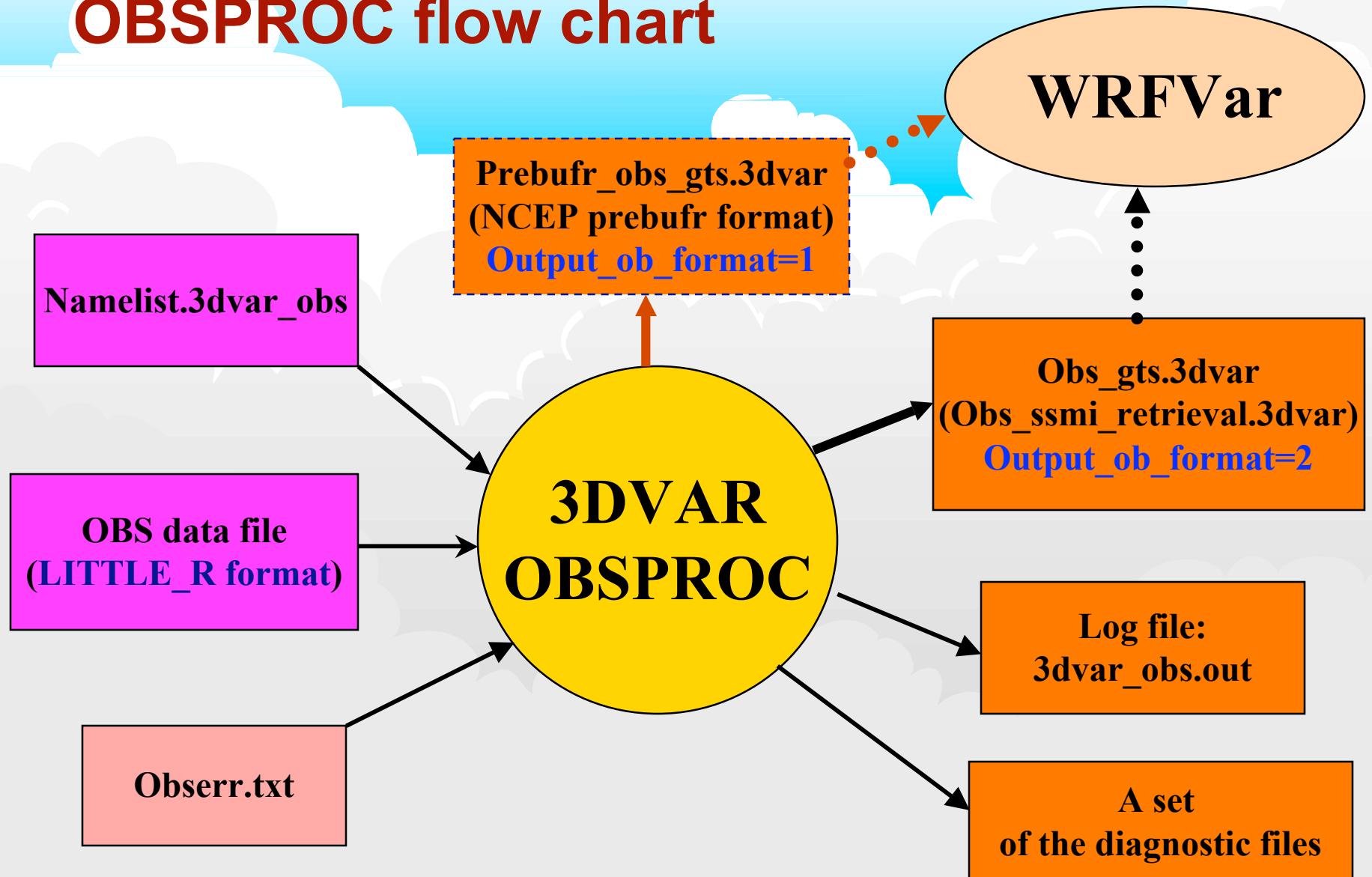
# Interface programs in AFWA, KMA, CAA,...

- Interface programs to convert the data format from AFWA, KMA, CAA,... observation database to **LITTLE\_R** format to provide the OBS file to WRF-Var OBS preprocessor



- In NCAR, MM5 Utility **fetch.csh** will obtain the OBS data file in LITTLE\_R format from NCAR archive.

# OBSPROC flow chart



# Input OBS (LITTLE\_R) file to preprocessor

- *OBS decoded file* in LITTLE\_R format containing *Reports*
- *Report* containing *Records* (*header*, *data*, ..., and *ending*) and 3 *tail integers* (3I7)
- *Record* containing *fields*
  - The *fields* in the *header record* (Fortran format in parenthesis)
  - The *fields* in the *data record* (Fortran format in parenthesis)
  - The *fields* in the *ending record*

| No | Field                 | No | Filed                    | No | Field                     |
|----|-----------------------|----|--------------------------|----|---------------------------|
| 1  | Latitude (f20.5)      | 2  | Longitude (f20.5)        | 3  | ID (a40)                  |
| 4  | Name (a40)            | 5  | Platform (a40)           | 6  | Source (a40)              |
| 7  | Elevation (f20.5)     | 8  | Num_vld_fld (i10)        | 9  | Num_error (i10)           |
| 10 | Num_warning (i10)     | 11 | Seq_num (i10)            | 12 | Num_dupd (i10)            |
| 13 | Is_sound (L10)        | 14 | Bogus (L10)              | 15 | Discard (L10)             |
| 16 | Valid_time%sut (i10)  | 17 | Valid_time%julian (i10)  | 18 | Valid_time%date_char(a20) |
| 19 | Slp%data (f13.5)      | 20 | Slp%qc (i7)              | 21 | Ref_pres%data (f13.5)     |
| 22 | Ref_pres%qc (i7)      | 23 | Ground_t%data (f13.5)    | 24 | Ground_t%qc (i7)          |
| 25 | SST%data (f13.5)      | 26 | SST%qc (i7)              | 27 | Psfc%data (f13.5)         |
| 28 | Psfc%qc (i7)          | 29 | Precip%data (f13.5)      | 30 | Precip%qc (i7)            |
| 31 | T_max%data (f13.5)    | 32 | T_max%qc (i7)            | 33 | T_min%data (f13.5)        |
| 34 | T_min%qc (i7)         | 35 | T_min_night%data (f13.5) | 36 | T_min_night%qc (i7)       |
| 37 | P_tend03%data (f13.5) | 38 | P_tend03%qc (i7)         | 39 | P_tend24%data (f13.5)     |
| 40 | P_tend24%qc (i7)      | 41 | Cloud_cvr%data (f13.5)   | 42 | Cloud_cvr%qc (i7)         |
| 43 | Celling%data (f13.5)  | 44 | Celling%qc (i7)          | 45 | Pw%data (f13.5)           |
| 46 | Pw%qc (i7)            | 47 | Tb19v%data (f13.5)       | 48 | Tb19v%qc (i7)             |
| 49 | Tb19h%data (f13.5)    | 50 | Tb19h%qc (i7)            | 51 | Tb22v%data (f13.5)        |
| 52 | Tb22v%qc (i7)         | 53 | Tb37v%data (f13.5)       | 54 | Tb37v%qc (i7)             |
| 55 | Tb37h%data (f13.5)    | 56 | Tb37h%qc (i7)            | 57 | Tb85v%data (f13.5)        |
| 58 | Tb85v%qc (i7)         | 59 | Tb85h%data (f13.5)       | 60 | Tb85h%qc                  |

## The fields in the data record (Fortran format in parenthesis)

| No | Field                    | No | Field               |
|----|--------------------------|----|---------------------|
| 1  | Pressure%data (f13.5)    | 2  | Pressure%qc (i7)    |
| 3  | Height%data 9f13.5)      | 4  | Height%qc (i7)      |
| 5  | Temperature%data (f13.5) | 6  | Temperature%qc (i7) |
| 7  | Dew_point%data (f13.5)   | 8  | Dew_point%qc (i7)   |
| 9  | Speed%data (f13.5)       | 10 | Speed%qc (i7)       |
| 11 | Direction%data (f13.5)   | 12 | Direction%qc (i7)   |
| 13 | U%data (f13.5)           | 14 | U%qc (i7)           |
| 15 | V%data (f13.5)           | 16 | V%qc (i7)           |
| 17 | RH%data (f13.5)          | 18 | RH%qc (i7)          |
| 19 | Thickness%data (f13.5)   | 20 | Thickness%qc (i7)   |

## The fields in the ending record

| No | field         | No | field | No | field         | No | field |
|----|---------------|----|-------|----|---------------|----|-------|
| 1  | -777777.00000 | 2  | 0     | 3  | -777777.00000 | 4  | 0     |
| 5  | -888888.00000 | 6  | 0     | 7  | -888888.00000 | 8  | 0     |
| 9  | -888888.00000 | 10 | 0     | 11 | -888888.00000 | 12 | 0     |
| 13 | -888888.00000 | 14 | 0     | 15 | -888888.00000 | 16 | 0     |
| 17 | -888888.00000 | 18 | 0     | 19 | -888888.00000 | 20 | 0     |

## Types of observations to be processed

→18 types (**SYNOP, SHIPS, METAR, TEMP , AIREP, PILOT , AMDAR, PROFL, SATOB, SATEM, SSMT1, SSMT2, SSMI, GPSPW, GPSZD, GPSRF, QSCAT, BOGUS**)

Each type of the observations is identified by its WMO code in WRF-Var. If the standard WMO code is not available to a **new** data type, user should assign a 3-digit code to that data type.

## WMO code for each type of observations

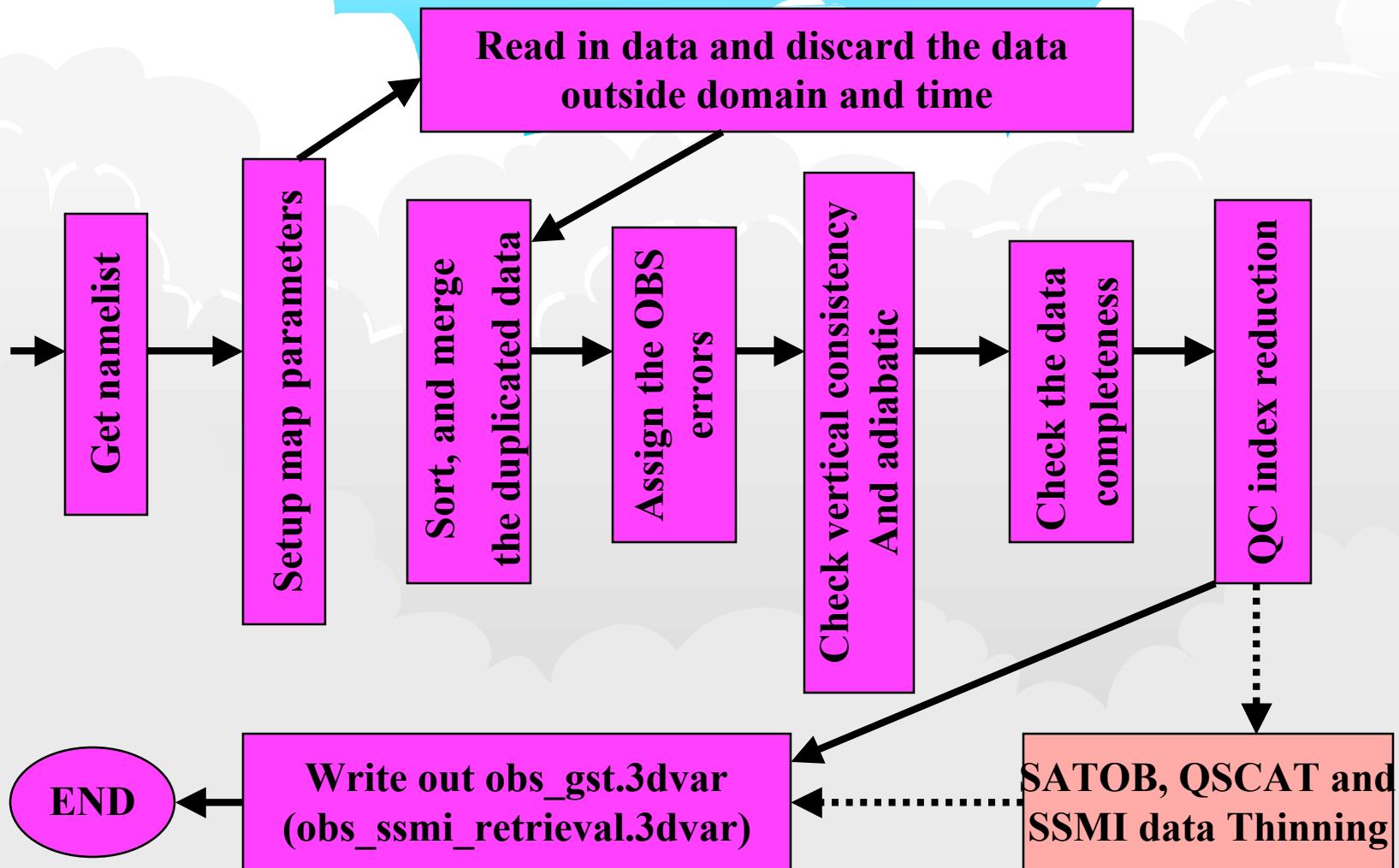
| Name  | WMO code       | WMO code name                               |
|-------|----------------|---|
| SYNOP | 12, 14         | SYNOP, SYNOP MOBIL                          |
| SHIP  | 13             | SHIP  |
| METAR | 15, 16         | METAR, SPECI                                |
| PILOT | 32, 33, 34     | PILOT, PILOT SHIP, PILOT MOBIL              |
| SOUND | 35, 36, 37, 38 | TEMP, TEMP SHIP, TEMP DROP, TEMP MOBIL      |
| AMDAR | 42             | AMDAR                                       |
| SATEM | 86             | SATEM                                       |
| SATOB | 88             | SATOB                                       |
| AIREP | 96, 97         | AIREP                                       |
| GPSPW | 111            | GPSPW (Ground-based GPS precipitable water) |
| GPSZD | 114            | GPSZD (Ground-based GPS Zenith Total Delay) |
| GPSRF | 116            | GPSRF (Space-based GPS Refractivity)        |
| SSMT1 | 121            | SSMT1                                       |
| SSMT2 | 122            | SSMT2                                       |
| SSMI  | 125            | SSMI  |
| PROFL | 132            | WIND PROFILER                               |
| BOGUS | 135            | TCBOU (Typhoon bogus), BOGUS (other bogus)  |
| QSCAT | 281            | Quik SCAT level-2B SeaWind                  |
| OTHER |                | UNKNOWN                                     |

## ■ Input and output files for OBS preprocessor

### 3 *Input files*

- OBS decoded file (**Reports**) in **little\_r** format  
A report (F90 pointer linking structure)
  - ❖ header record (fields)
  - ❖ Level1 data record (fields)
  - .....
  - ❖ Leveln data record (fields)
  - ❖ Ending record (fields)
  - 3 Integers in format(3i7)
- Namelist file (**namelist.3dvar\_obs**) (See: **README.namelist**)
  - Record1: input file names
  - Record2: analysis times
  - Record3: Maximum number of observations allowed
  - Record4: quality control switches
  - Record5: print switches
  - Record6: define the reference state: ptop, etc.
  - Record7: Geographic parameters
  - Record8: Domain settings
  - Record9: Output format: prebufr, ascii, or both
- AFWA OBS errors file: **obserr.txt** (provided by 3DVAR system)

## Flow chart of OBS preprocessor



## Tasks of the OBS preprocessor: 3DVAR\_OBSPROC

- 1, To perform a time-windowed and, in case of regional application (domain\_check\_h = .TRUE.), geographically-filtered dump of the ingested observations

Currently, there is *no time-check for observation data in WRF-Var assimilation code*, so to select the observation data within a suitable time-window must be performed in 3DVAR\_OBSPROC.

For the regional application with the IPROJ = 1 (Lambert conformal), 2 (Polar Stereographic), or 3 (Mercator), there is a geographic-filtered performed based on the model domain settings. For the global application of WRF-Var, it should set IPROJ = 0 and no geographic-filtered is performed.

## Tasks of the OBS preprocessor: 3DVAR\_OBSPROC (cont.)

- 2, To retrieve the pressure or height based on the observed information with the hydrostatic assumption
- 3, To remove the duplicate reports of observations
- 4, To re-order (from bottom to top) and merge the data reports with the same platform, time, and location based on the pressure.
- 5, To discard the data above the model top ( $p < p_{top}$ ) in the upper-air observations (remove\_above\_lid = .TRUE.)

## Tasks of the OBS preprocessor: 3DVAR\_OBSPROC (cont.)

6, To assign the observation errors to the different types of observations

Observations errors

- NCEP OBS error (Parrish and Derber 1992)
- US Air Force (AFWA) OBS error file
- Directly from the observation reports

7, To perform the vertical consistency check and super adiabatic check for the multi-level observations

## Tasks of the OBS preprocessor: 3DVAR\_OBSPROC (cont.)

8, To complete thinning with the SATOB, SSMI, and QSCAT data

The data points nearest to the model grid-points will be picked up for assimilation.

9, To write out the OBS files in ASCII format as the WRF-Var input

→ GTS data (*obs\_gts.3dvar*): pressure, Wind, height, temperature, dewpoint, RH, etc.

→ SSMI data (*obs\_ssmi\_retrieval.3dvar*): PW and surface wind speed

10, To plot the distribution for each type of observations

→MAP\_plot

## Output files

### 1, *Obs\_gts.3dvar* and *obs\_ssmi\_retrieval.3dvar*

**Header:** the information for this OBS file and data format

**Data** : *header record* and *data records for each of levels*

- These are the OBS input file to WRF-Var program
- *obs\_ssmi\_retrieval.3dvar* needed only when SSMI retrieval data available
- These files can be used as input to MAP\_plot to obtain the gmeta plot file with NCAR GRAPHICS

### 2, *3dvar\_obs.out* ---- a program execution log file

The printing out from the program execution. It can be used to monitor the execution and to identify the troubles if any

### 3, Diagnostic files depended on the print switches in namelist



## File: obs\_gts.3dvar

TOTAL = 8169, MISS. =-888888.,  
SYNOP = 1432, METAR = 164, SHIP = 86, BUOY = 0, TEMP = 179, AMDAR = 0,  
AIREP = 265, PILOT = 0, SATEM = 0, SATOB = 6043, GPSPW = 0, SSMT1 = 0,  
SSMT2 = 0, TOVS = 0, QSCAT = 0, PROFL = 0, OTHER = 0,  
PHIC = 28.50, XLONGC = 116.00, TRUE1 = 10.00, TRUE2 = 45.00, XIM11 = 1.00, XJM11 = 1.00,  
TS0 = 275.00, TLP = 50.00, PTOP = 7000., PS0 = 100000.,  
IXC = 67, JXC = 81, IPROJ = 1, IDD = 1, MAXNES= 10,  
NESTIX= 67, 67, 67, 67, 67, 67, 67, 67, 67, 67,  
NESTJX= 81, 81, 81, 81, 81, 81, 81, 81, 81, 81,  
NUMC = 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,  
DIS = 135.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00,  
NESTI = 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,  
NESTJ = 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,  
INFO = PLATFORM, DATE, NAME, LEVELS, LATITUDE, LONGITUDE, ELEVATION, ID.  
SRFC = SLP, PW (DATA,QC,ERROR).  
EACH = PRES, SPEED, DIR, HEIGHT, TEMP, DEW PT, HUMID (DATA,QC,ERROR)\*LEVELS.  
INFO\_FMT = (A12.1X,A19.1X,A40.1X,I6,3(F12.3,11X),6X,A5)  
SRFC\_FMT = (F12.3,I4,F7.2,F12.3,I4,F7.2)  
EACH\_FMT = (3(F12.3,I4,F7.2),11X,3(F12.3,I4,F7.2),11X,1(F12.3,I4,F7.2)))

| #           | FM-35 TEMP          | 1999-08-19_11:00:00        | AHMADABAD / INDIA | 17          | 23.070   | 72.630   | 55.000  | 42647       |          |             |          |                       |
|-------------|---------------------|----------------------------|-------------------|-------------|----------|----------|---------|-------------|----------|-------------|----------|-----------------------|
| -888888.000 | -88 200.00          | -888888.000                | -88 0.20          |             |          |          |         |             |          |             |          |                       |
| 100000.000  | 0 100.00            | 3.000                      | 0 1.10            | 225.000     | 0 5.00   | 53.000   | 0 7.00  | 305.550     | 0 1.00   | 298.550     | 0 1.00   | 65.984 0 15.00        |
| 92500.000   | 0 100.00            | 3.000                      | 0 1.10            | 220.000     | 0 5.00   | 745.000  | 0 7.48  | 300.550     | 0 1.00   | 295.550     | 0 1.00   | 73.714 0 12.60        |
| 91600.000   | 0 100.00            | -888888.000                | -88 1.10          | -888888.000 | -88 5.00 | 830.000  | 0 7.54  | 300.150     | 0 1.00   | 295.150     | 0 1.00   | 73.653 0 12.30        |
| 87300.000   | 0 100.00            | -888888.000                | -88 1.10          | -888888.000 | -88 5.00 | 1251.000 | 0 7.84  | 297.350     | 0 1.00   | 291.350     | 0 1.00   | 68.712 0 10.82        |
| 85000.000   | 0 100.00            | 6.000                      | 0 1.10            | 250.000     | 0 5.00   | 1498.000 | 0 8.00  | 295.350     | 0 1.00   | 290.750     | 0 1.00   | 74.809 0 10.00        |
| 79200.000   | 0 100.00            | -888888.000                | -88 1.13          | -888888.000 | -88 5.00 | 2104.000 | 0 8.22  | 290.750     | 0 1.00   | 288.550     | 0 1.00   | 86.755 0 10.00        |
| 78000.000   | 0 100.00            | -888888.000                | -88 1.18          | -888888.000 | -88 5.00 | 2238.000 | 0 8.27  | 292.950     | 0 1.00   | 281.950     | 0 1.00   | 48.599 0 10.00        |
| 70700.000   | 0 100.00            | -888888.000                | -88 1.39          | -888888.000 | -88 5.00 | 3076.000 | 0 8.57  | 289.950     | 0 1.00   | 268.950     | 0 1.00   | 23.245 0 10.00        |
| 70000.000   | 0 100.00            | 2.000                      | 0 1.40            | 30.000      | 0 5.00   | 3168.000 | 0 8.60  | -888888.000 | -11 1.00 | -888888.000 | -11 1.00 | -888888.000 -11 10.00 |
| 61100.000   | 0 100.00            | -888888.000                | -88 1.76          | -888888.000 | -88 5.00 | 4312.000 | 3 10.01 | 282.350     | 0 1.00   | 271.350     | 0 1.00   | 45.937 0 10.00        |
| 60000.000   | 0 100.00            | -888888.000                | -88 1.80          | -888888.000 | -88 5.00 | 4463.000 | 3 10.20 | 281.350     | 0 1.00   | 272.350     | 0 1.00   | 52.920 0 10.00        |
| 57400.000   | 0 100.00            | -888888.000                | -88 1.90          | -888888.000 | -88 5.00 | 4828.000 | 3 10.66 | 279.150     | 0 1.00   | 274.150     | 0 1.00   | 70.188 0 10.00        |
| 52700.000   | 0 100.00            | -888888.000                | -88 2.14          | -888888.000 | -88 5.00 | 5526.000 | 3 11.55 | 275.550     | 0 1.00   | 272.650     | 0 1.00   | 81.128 0 10.00        |
| 50300.000   | 0 100.00            | -888888.000                | -88 2.28          | -888888.000 | -88 5.00 | 5902.000 | 3 12.04 | 273.350     | 0 1.00   | 269.650     | 0 1.00   | 76.187 0 10.00        |
| 50200.000   | 0 100.00            | -888888.000                | -88 2.29          | -888888.000 | -88 5.00 | 5918.000 | 3 12.06 | 273.150     | 0 1.00   | 269.450     | 0 1.00   | 76.157 0 10.00        |
| 50000.000   | 0 100.00            | 3.000                      | 0 2.30            | 95.000      | 0 5.00   | 5950.000 | 0 12.10 | 272.850     | 0 1.00   | 269.250     | 0 1.00   | 76.682 0 10.00        |
| 46900.000   | 0 100.00            | 2.000                      | 0 2.42            | 100.000     | 0 5.00   | 6461.000 | 3 12.90 | 270.450     | 0 1.00   | 266.250     | 0 1.00   | 72.904 0 10.00        |
| FM-35 TEMP  | 1999-08-19_11:00:00 | BOMBAY / SANTACRUZ / INDIA | 25                | 19.120      | 72.850   | 14.000   | 430     |             |          |             |          |                       |

# of observations

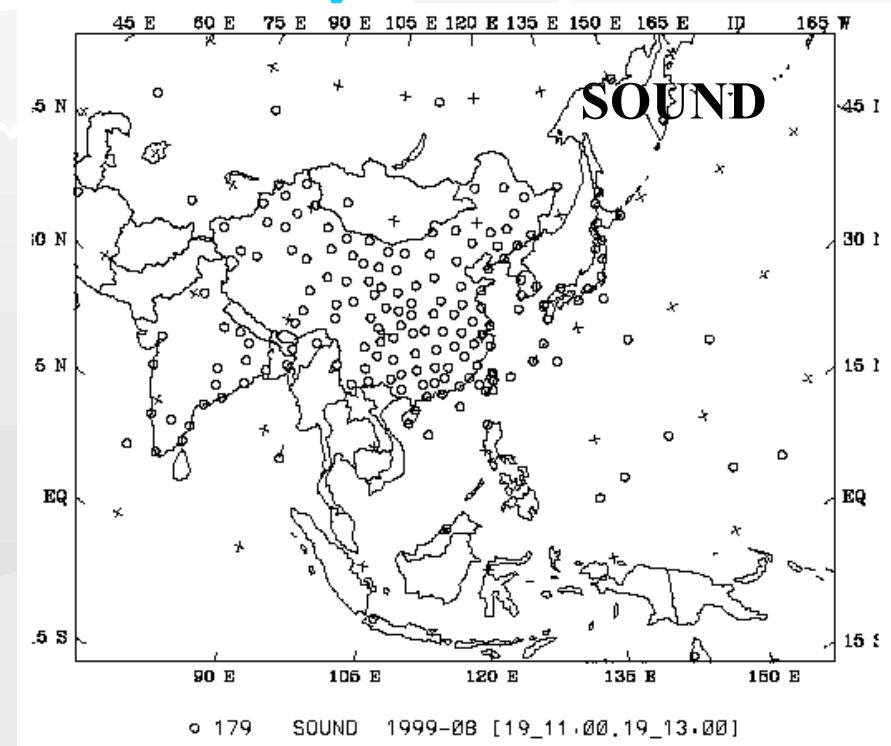
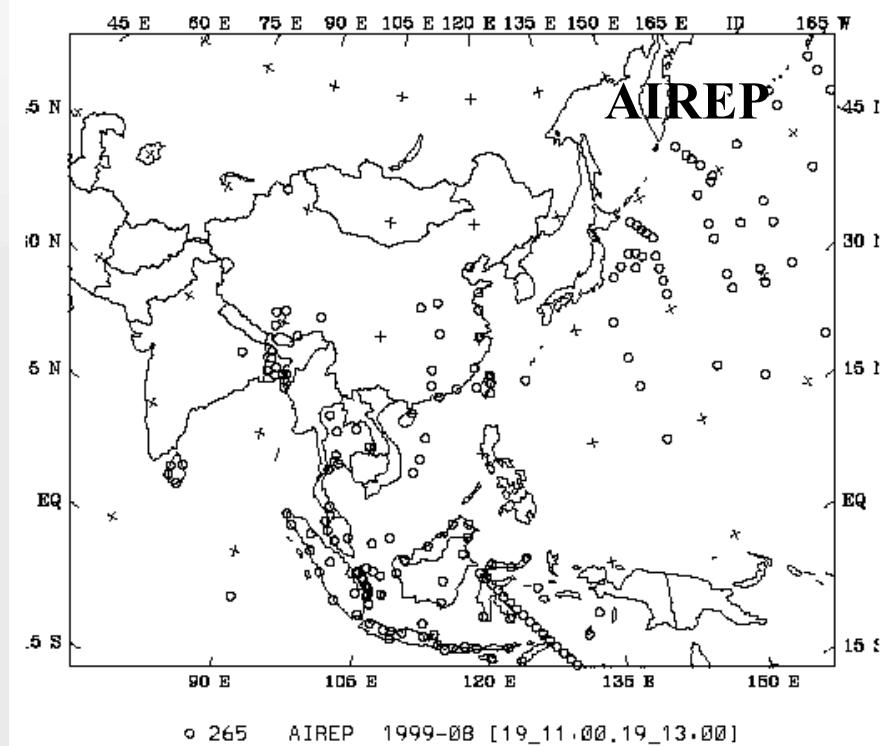
Model domain information

Data format

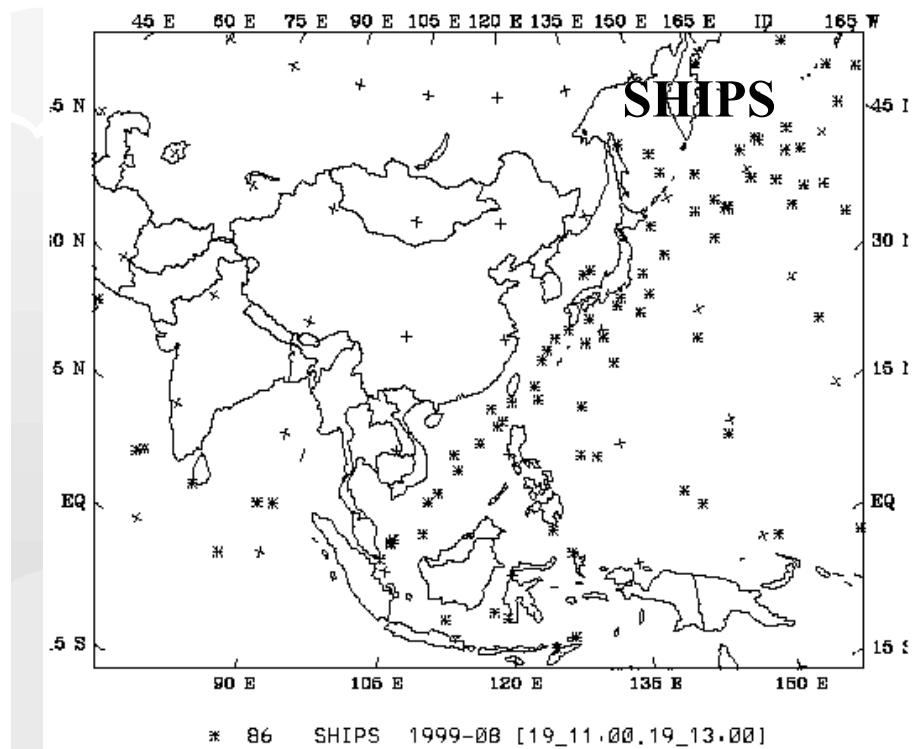
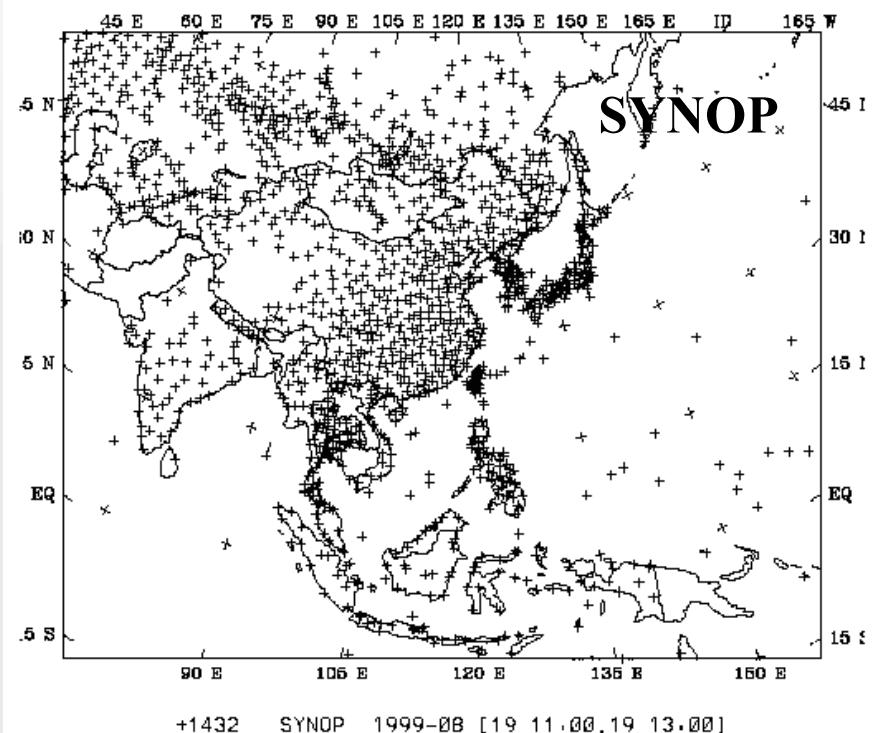
## How to plot the OBS distribution?

- Go to the directory ***3DVAR\_OBSPROC/MAP\_plot***
- Modify the shell script ***Map.csh***
  - » To fill in ***TIME\_ANALYSIS***, etc., and ***OBSDATA*** file name
- Run shell script ***Map.csh***
  - » You will have a gmeta file: ***gmeta.\${TIME\_ANALYSIS}*** to show the the distribution of observations contained in ***OBSDATA*** file.

## Distribution for each type of observations



## Distribution for each type of observations



## Distribution for each type of observations

