CSU – RAMS

Standard Input Formats for Pressure Coordinate and Observational Data (Data Prep Stage)

This document contains detailed notes on how gridded pressure level datasets (such as NCEP-reanalysis, GFS analysis, NARR reanalysis, etc.), rawindsonde observations, and surface observations need to be formatted for being ingested into RAMS Isentropic analysis for creating variable initialization files (Var-files). "Dprep" or "DeGrib" software is made available for performing this data preparation.

Updated by:

Stephen Saleeby
Department of Atmospheric Science
Colorado State University

Last updated: 31 October, 2019

IMPORTANT – READ THIS FIRST

Note first that in order to create variable initialization files “varfiles” for RAMS case study simulations (when doing a RAMS execution with RUNTYPE = ‘MAKEVFILE’) you need to first create the dataprep “dp” files as shown here in the section on “Gridded Data File”. This takes gridded pressure level data (such as the NCEP/NCAR Reanalysis or ERA-Interim) and converts that GRIB data into the RALPH-2 format shown below in the “Gridded Data File” section. The code to do the creation of dp files from GRIB pressure level data is compiled and executed in the main RAMS distribution in directories bin.dp.grib1 and bin.dp.grib2, depending on whether your data is in GRIB-1 or GRIB-2 format. Commands for processing Grib-1 vs Grib-2 are slightly different. You can run the executable with no command line arguments to get an output list of data that can be processed with each Grib type. Also can see the dprep/dgrib source code to examine differences.

Note that only gridded pressure level, snow, and soil data is needed for making dp files and the subsequent varfiles. Pressure level grids are mandatory, but snow and soil grid are optional. Defaults will be used if snow and soil grids are excluded.

The sections below on “Upper air observations file” and “Surface observations file” are for ingesting observed rawinsonde data and surface data as purely optional choices. Use of the rawindsonde and surface observations is ingested in RAMS via the section on “ODA” seen in the RAMSIN namelist. Typically, only the pressure level gridded data is used.

Note that RAMS can handle gridded pressure level data in projections of latitude-longitude, Lambert-conformal, and polar stereographic. Others are not supported at this time.

To run the compiled dataprep / dgribber code, run something like this:

```
dgrib-executable [-t #] [-d YYYYMMDDHH] [-h fcsthr] [-f gribfile]

-t option: (FOR GRIB-1 format only; different for Grib-2)
 # = 1: NcepReanalysis 1 or 2 Global Lat/lon at 2.5-deg
        NO soil/snow; Using R / RELH
        Upper air names: UGRD,VGRD,TMP,HGT,RH
        SWE,SnowDepth names: NOTHING,NOTHING
        Soil Moisture,Temp names: NOTHING,NOTHING
 # = 2: GDAS-FNL Global Lat/Lon at 1.0deg UNTIL Jan 14, 2015 00Z
        YES soil/snow; Using R / RELH
        Upper air names: UGRD,VGRD,TMP,HGT,RH
        SWE,SnowDepth names: WEASD,WEASD
        Soil Moisture,Temp names: SOILW,TMP
 # = 3: GDAS-FNL Global Lat/Lon at 1.0deg AFTER Jan 14, 2015 00Z (or)
        GDAS-FNL Global Lat/Lon at 0.25deg after Jul 8, 2015 (or)
        Forecast grids for GFS and HRRR Lat/Lon Grid (or)
        RAP Rapid Refresh Analysis 32km Awips Grid 221
        YES soil/snow; Using R / RELH
        Upper air names: UGRD,VGRD,TMP,HGT,RH
```

```
        SWE,SnowDepth names: WEASD,SNOD
        Soil Moisture,Temp names: SOILW,TSOIL
 # = 4: NARR 32km Lambert Conformal Grid
        YES soil/snow; Using Q / SPFH
        Upper air names: UGRD,VGRD,TMP,HGT,SPFH
        SWE,SnowDepth names: WEASD,SNOD
        Soil Moisture,Temp names: SOILW,TSOIL
 # = 5: ERA-Interim Global Lat/Lon at various resolutions (or)
        ERA5 Global Lat/Lon or area subset at 0.25deg
        YES soil/snow; Using R / RELH
        Upper air names: U,V,T,Z,R
        SWE,SnowDepth names: SD,SD
        Soil Moisture,Temp names: SWVL2,SWVL1,STL2,STL1
 # = 6: ERA5 Lat/Lon Global or area subset at 0.25deg
        YES soil/snow; Using Q / SPFH
        Upper air names: U,V,T,Z,Q
        SWE,SnowDepth names: SD,SD
        Soil Moisture,Temp names: SWVL2,SWVL1,STL2,STL1

-d option: YYYYMMDDHH is the time
-h option: forecast hour (should be 0 for reanalysis)
-f option: gribfile is the name of the file to extract
```

Note the following:

1. ERA5 relative humidity is with respect to water or ice depending on temperature. But RAMS does initialization assuming RH with respect to water, so estimates can be high at colder temperatures. Better to use ERA5 specific humidity. Yes soil/snow.

2. Note that soil/snow data MUST be on the same grid type (e.g.lat-lon, lambert-conformal) and resolution (e.g. 1.0-deg, 0.25-deg) as the atmospheric pressure level data.

3. ERA5 data found at the following on the Copernicus data server:

Pressure level reanalysis:
"ERA5 hourly data on pressure levels from 1979 to present":
https://cds.climate.copernicus.eu/cdsapp#!/dataset
/reanalysis-era5-pressure-levels?tab=form

Snow, soil moisture, soil temperature reanalysis:
"ERA5 hourly data on single levels from 1979 to present":
https://cds.climate.copernicus.eu/cdsapp#!/dataset
/reanalysis-era5-single-levels?tab=form

Example for ERA5 data in Grib-1 format:

```
dgrib-6.2.11 -t 6 -d 2016071618 -h 0 -f era5_0.25deg_2016071618.grb1
```

Gridded Data File (ie. DataPrep or "DP" files)

Following is a listing of the first few lines of a sample gridded pressure level data file.

```
2007    2   11  1200    0   26  360   181
1      1.00      1.00   -90.0  -180.0     90.0    179.0      0.000      0.000
  1000      975      950      925      900      850      800     750     700     650      600      550      500
        450      400      350      300      250      200      150     100       70       50       30       20       10
     0.90000     0.90000     0.80000     0.70000     0.60000     0.50000     0.50000     0.40000
     0.30000     0.20000     0.10000     0.00000     0.00000    -0.10000    -0.20000    -0.30000
….
```

Record	Field	Name	Description
1	1	Year	Integer – 4 digit year
	2	Month	Integer – 2 digit month
	3	Date	Integer – 2 digit date
	4	Time	Integer – 4 digit UTC data time
	5	Valid time increment	Integer – time past data time when data is valid (see below)
	6	# of pressure levels	Integer – number of vertical pressure level in dataset
	7	x-points	Integer – number of points in x-direction
	8	y-points	Integer – number of points in y-direction
2	1	Projection	Integer – projection flag 1 – latitude/longitude (LL) 2 – Lamber-conformal (LC) 3 – True polar-stereographic (PS) (assumes valid at 60N)
	2	x-spacing	Real – grid spacing in X-direction. Units depend on projection. LL – degrees, LC – meters, PS – meters (at 60N)
	3	y-spacing	Real – grid spacing in Y-direction. Units depend on projection LL – degrees, LC – meters, PS – meters (at 60N)
	4	Projection parameter	Real – projection dependent parameter LL, LC, PS – southwest latitude of grid (degrees)
	5	Projection parameter	Real – projection dependent parameter LL, LC, PS – southwest longitude of grid (degrees)
	6	Projection parameter	Real – projection dependent parameter LL, LC, PS – northeast latitude of grid (degrees)
	7	Projection parameter	Real – projection dependent parameter LL, LC, PS – northeast longitude of grid (degrees)
	8	Projection parameter	Real – projection dependent parameter LL – unused, LC – intersection/tangency latitude of projection (degrees), PS – tangency latitude of projection (degrees) (Set to 90.; only North hemisphere currently)
	9	Projection parameter	Real – projection dependent parameter LL – unused, LC – center longitude (degrees), PS – center longitude (degrees)
3	2+	Levels	Integer – level coordinates. Units dependent on vertical coordinate type flag. Must have as many values as Record 2 : field 6.

Notes:

• All data is space-delimited so that FORTRAN can do a free format, list-directed read. C can also read it without a specific format specification.

• The time field is the data time of the gridded data or when a particular forecast was started. The valid time increment is the amount of time past the data time when a forecast data is applicable. The increment is formatted in an hhhmm (h-hours,mminutes) format. As an example, consider the case of a 36 hour RAMS forecast starting at 1200 UTC. The initial field will have a data time of 1200 and the increment will be 0. The 12-hour forecast will have a data time of 1200 and an increment of 1200. The 36-hour forecast will have a data time of 1200 and an increment of 3600. All dates refer to the data time.

After the header information, the data is written as follows. Each variable at each level is written at a time. In FORTRAN, it is:

write(unit, format) ((data(i,j),i=1,num_x),j=1,num_y)

Each record then starts in the southwest corner (lower left) and proceeds row-wise.

For a global gridded dataset such as GDAS, ERA5, NCEP reanalysis, this means the “dp” file starts with data at -90S, -180W then writes from west to east around the earth, then increments one grid point northward, then writes from west to east around the earth, and again increments northward. This repeats until reaching 90N. So, the first entry for the first variable at the top of the “dp” file is the southwest corner of the data (-90S, -180W) and the last entry for the same variable written further down in the “dp” file is the northeast corner (e.g. 90N,179E). The southwest data point has to start at -90S, -180W but the final point to the northeast could be something like 90N, 179E depending on the resolution of the data. The 1-deg GDAS would end at 90N, 179E, while the 2.5-deg NCEP reanalysis would end at 90N, 177.5E. RAMS will add the “extra” N/S column at 180E in the varfile creation.

Concerning the order of the data, first the lowest coordinate level is written with each variable following in the following order and units:

• east-west (u) velocity component (m/s)
• north-south (v) velocity component (m/s)
• temperature (K)
• geopotential height (m)
• relative humidity (fraction)

This set of variables is then written for each coordinate level up to the top of the data.

After the pressure level data is written, 6 surface fields follow written in the same way as the upper level fields:
• second soil level down, volumetric soil moisture (m3/m3) (typical 0-10cm depth from datasets)
• top soil level, volumetric soil moisture (m3/m3) (varies between datasets)
• second soil level down, soil temperature (K) (typically 0-10cm depth from datasets)
• top soil level, soil temperature (K) (varies between datasets)
• snow water equivalent (kg/m2)
• snow depth (m) (Some datasets give this and others do not. If not, we assume
snow depth = snow water equivalent x 10)

We apply dataset top soil level data to RAMS top soil level. Next we apply dataset next level down soil data to all lower soil levels in RAMS.

RAMS/ISAN currently assumes that the wind components are earth-relative for the lat-lon grid and grid-relative for the Lambert-Conformal and true polar-stereographic projections.

RAMS/ISAN will read and use the soil and snow data if they are in the data prep files. Otherwise the model will fill the variable initialization file with missing data values. In this case, user needs to make sure not to try and use soildata (ISOILDAT=1 in RAMSIN namelist) if its filled with missing data values. At runtime, RAMS will tell you if soil data is bad and the model will stop.

• A separate file must exist for each analysis time.

Upper air observations file

Following is a sample of the upper air observations file:

```
999999 2
1993  7  4   0000   01001       0024  0009    70.93000    -8.67000     9.00000
   97600.000  000       212.409  000       7.60  000    0.7883  000
   92500.000  000       653.065  000       4.80  000    0.8249  000
   85000.000  000      1338.865  000       1.00  000    0.8200  000
   70000.000  000      2871.519  000      -9.50  000    0.8591  000
   67900.000  000      3106.343  000    -11.30  000    0.9223  000
   63800.000  000      3583.331  000    -12.90  000    0.7672  000
   56100.000  000      4553.193  000    -19.10  000    0.5360  000
   50000.000  000      5400.000  000    -25.30  000    0.6762  000
   45700.000  000      6048.044  000    -29.30  000    0.7641  000
   40000.000  000      6986.854  000    -36.10  000    0.5949  000
   37300.000  000      7468.442  000    -39.90  000    0.4660  000
   35300.000  000      7843.197  000    -42.30  000    0.6018  000
   30300.000  000      8856.608  000    -51.10  000    0.5448  000
   30100.000  000      8899.660  000    -51.50  000    0.5433  000
   30000.000  000      8921.284  000    -51.30  000    0.4798  000
   27400.000  000      9519.279  000    -44.90  000    0.0555  000
   27000.000  000      9617.594  000    -45.10  000    0.0264  000
   25000.000  000     10134.874  000    -42.50  000    0.0085  000
   24100.000  000     10383.210  000    -41.50  000    0.0090  000
   20000.000  000     11650.147  000    -41.30  000    0.0090  000
   16500.000  000     12958.866  000    -40.90  000    0.0079  000
   15000.000  000     13605.299  000    -42.50  000    0.0085  000
   11900.000  000     15164.091  000    -44.50  000    0.0076  000
   10000.000  000     16332.646  000    -43.50  000    0.0080  000
     1334.000  000       7.72  000      65.00  000
     2867.000  000       6.17  000      45.00  000
     5400.000  000       5.14  000      45.00  000
     6980.000  000      12.35  000      55.00  000
     8910.000  000      12.35  000      75.00  000
   10120.000  000       7.20  000      75.00  000
   11640.000  000       5.66  000      90.00  000
   13590.000  000       1.54  000     145.00  000
   16320.000  000       1.54  000     260.00  000
1993  7  4   0000   01028       0038  0014    74.52000    19.01999    18.00000
  100800.000  000        18.783  000      1.60  000    0.9715  000
  100000.000  000        83.118  000      1.80  000    0.9437  000
    98800.000  000       180.666  000      2.00  000    0.9576  000
    95900.000  000       420.757  000      0.40  000    0.9570  000
    89900.000  000       946.058  000      6.60  000    0.7868  000
    87700.000  000      1149.788  000      7.00  000    0.3581  000
…
…
```

The file starts again with:

Record	Field	Name	Description
1	1	999999	marker to start a new file or new section of a file
	2	Version	Integer – dataset version number (this is version 2)

After this, there is a section for each observation:

Section	Field	Name	Description
Header	1	year	Integer – 4 digit year
	2	month	Integer – 2 digit month
	3	date	Integer – 2 digit date
	4	time	Integer – 4 digit UTC data time
	5	Station ID	Character – up to 8 characters
	6	# P levels	Integer – number of pressure levels where thermodynamic information is reported
	7	# Z levels	Integer – number of height levels where wind data is reported
	8	Station latitude	Real – station latitude (degrees)
	9	Station longitude	Real – station longtitude (degrees)
	10	Station elevation	Real – station elevation (meters)
Presssure data	1	Pressure	Real – pressure (Pa)
	2	Pressure flags	Integer – 3 digit "quality" flag (see below)
	3	Geopotential height	Real – geopotential height (m)
	4	Geopotential height flags	Integer – 3 digit "quality" flag (see below)
	5	Temperature	Real – temperature (Celcius)
	6	Temperature flags	Integer – 3 digit "quality" flag (see below)
	7	Relative humidity	Real – relative humidity (fraction)
	8	Relative humidity flags	Integer – 3 digit "quality" flag (see below)
Height data	1	Height	Real – height above sea level (m)
	2	Height flags	Integer – 3 digit "quality" flag (see below)
	3	Wind speed	Real – wind speed (m/s)
	4	Wind speed flags	Integer – 3 digit "quality" flag (see below)
	5	Wind direction	Real – wind direction (degrees)
	6	Wind direction flags	Integer – 3 digit "quality" flag (see below)

Notes:
• Quality flags are a 3 digit code denoting what stages of quality control have been performed during the 3 stages of QC.
Some of the values are:
0 - not checked
9 - missing
5 - passed, good data
1 – checked, flagged as bad data
• Rawindsondes will generally have both the pressure and height level data reported. Wind profilers will only have the height levels, so these will have a 0 for the number of pressure levels.

• A separate file must exist for each analysis time.

Surface observations file

Following is a sample of the surface observations file:

```
999999 2
9
WINDSPEED    m/s
WIND_DIRECTION    deg
TEMPERATURE    C
DEWPOINT    C
STN_PRES    Pa
SLP    Pa
6-HR_PCP    mm
24-HR_PCP    mm
CLOUD_COVER    fraction
1993    09    01    0000    71066    58.620    -117.170    338.    2.57    000    350.   000 \
  16.8    000    11.5    000    96710.0    000    100770.0    000    -999.0   000    0.0   000   .80    000
1993    09    01    0000    71068    56.620    -115.170    340.    2.57    000    350.   000 \
  16.8    000    11.5    000    96710.0    000    100770.0    000    -999.0    000   0.0   000   .80    000
…..
…..
```

**** Note that each observation line is actually all one record and is just broken up here for documentation purposes. The file starts with header information:

The file starts with header information:

Record	Field	Name	Description
1	1	999999	marker to start a new file or new section of a file
	2	Version	Integer – dataset version number (this is version 2)

After this, there is a line for each observation:

Section	Field	Name	Description
Surface	1	year	Integer – 4 digit year
	2	month	Integer – 2 digit month
	3	date	Integer – 2 digit date
	4	time	Integer – 4 digit UTC data time
	5	Station ID	Character – up to 8 characters
	6	Station latitude	Real – station latitude (degrees)
	7	Station longitude	Real – station longtitude (degrees)
	8	Station elevation	Real – station elevation (meters)
	9+	Value/flag pairs	Real/integer – pairs of values and their quality flags. See below.

Notes:

• There must be as many value/flag pairs as specified in the header. Values are in the order specified in the header and are of the specified units in the header. ***This is a future capability; for now there only must be the first 5 variables in the units stated in the example above. All additional variables will currently be ignored.***

• A separate file must exist for each analysis time.