8 Model Initialization and Preprocessing

8.1. Introduction

The ARPS model has a variety of initialization options, which can tailor the model to a specific application. The simplest initialization option involves a horizontally homogeneous base state and an initial thermal bubble with no terrain. Apart from a single sounding, this configuration requires no external data. For applications that require real terrain and surface data properties, preprocessors, ARPSTERN and ARPSSFC are available to generate the needed fields from databases. A program, EXT2ARPS, that interpolates external data sets to ARPS grid, and generates data sets for 3-D initialization of ARPS is also provided. This chapter describes each of these preprocessors in Sections 8.2, 8.3 and 8.5. In addition, Section 8.4 discusses the options for the horizontally homogeneous initial state and Section 8.6 describes the use of external boundary conditions.

8.2. Terrain Data Preprocessor

This section describes ARPS terrain data: preprocessor and the supported terrain data sets. ARPS terrain initialization process includes two steps: creating a terrain database in an efficient direct-access format, and generating an analyzed terrain field for use in ARPS.

Program ARPSTERN (in file *arpstern12.f*) performs the analysis of terrain data on ARPS model grid. The current version has provision for one analysis scheme, the Barnes multi-pass technique. This scheme allows for flexible user control over the analyzed terrain field (*e.g.* cutoff wavelength). In addition, the analysis scheme can be applied to an irregularly spaced database.

The following sections contain a description of the terrain database, the terrain processing programs and the Barnes analysis scheme. A more detailed description of the database can be found in Appendix C.

8.2.1. ARPS Terrain Data Preprocessors

At present, ARPS terrain preprocessor supports three data sets:

1) $1° \times 1°$ (~110×110 km^2) data covering the entire globe,

2) $5' \times 5'$ (~10×10 km^2) data covering North America and Europe, and

3) $30" \times 30"$ (~1×1 km^2) data covering the Continental United States. The original data files *elev.dat* and *dma_elev.dat* are stored in ASCII format and can be obtained through NCAR Data Services for a nominal fee. The contact is:

Data Support Section
Scientific Computing Division
National Center for Atmospheric Research
Boulder , CO 80307
Tel: (303)-497-1216
E-mail: datahelp@ncar.ucar.edu

The first step of ARPS terrain preprocessing converts the original ASCII files into unformatted binary direct-access files. These files are then used by the analysis program ARPSTERN to create the final smoothed terrain field. The direct-access files are platform-dependent (*e.g.* IBM RISC/6000 or Cray), and require the conversion and analysis programs to be run on the same platform. In general, the conversion has to be done only once for each computer system.

a) ASCII to binary data conversion

The following files are required to convert the ASCII data sets into direct access files, and should be found in the main ARPS directory:

dir1deg.f	Program DIR1DEG for 1° data conversion
dir5min.f	Program DIR5MIN for 5' data conversion
dir30sec.f	Program DIR30SEC for 30" data conversion
arpstern.input	Control input parameters for all three programs as well as ARPSTERN
load1deg.file	File containing special loader flags for use on CRAY
load5min.file	File containing special loader flags for use on CRAY
load30sec.file	File containing special loader flags for use on CRAY

makearps	Shell script for general make control for ARPS
make.arps40	Makefile for ARPS system
elev.dat	NCAR 5' and 1° ASCII data file
dma_elev.dat	NCAR 30" ASCII data file

The data files, *elev.dat* and *dma_elev.dat*, may reside in a separate directory. The name of this directory (*terndir*) is specified in input file *arpstern.input* (see subsection c).

To compile and link DIR1DEG, DIR5MIN and DIR30SEC, enter the commands:

> **makearps** *dir1deg*
> **makearps** *dir5min*
> **makearps** *dir30sec.*

Before executing these three programs, several parameters must be set in *arpstern.input*, which is a NAMELIST format input file that is also used by the analysis program ARPSTERN. These parameters include *comtype*, a flag for computer type, *terndir*, a string specifying the directory containing the ASCII data files, and *tdatadir*, the name of the directory into which the output data files will be written. After the parameters are set, programs DIR1DEG, DIR5MIN and DIR30SEC are executed by entering:

> **dir1deg** *< arpstern.input> dir1deg.out*
> **dir5min** *< arpstern.input> dir5min.out*
> **dir30sec** *< arpstern.input> dir30sec.out*

where **dir1deg**, **dir5min** and **dir30sec** are the executable commands produced in the compilation and link step. The standard output files (**.out*) contain the diagnostic information from the programs.

The following data files are produced (in directory *tdatadir*) by the conversion programs:

dir1deg.dat	1° direct-access data set
dir1deg.hdr	Header file containing the record numbers of *dir1deg.dat*
dir5min.dat	5' direct-access data set
dir5min.hdr	Header file containing the record numbers of *dir5min.dat*
dir30sec.dat	30" direct-access data set
dir30sec.hdr	Header file containing the record numbers of *dir30sec.dat*.

It is important to note that, depending on the location and the resolution of the model domain, one, two or all three of these data sets may be required by the analysis program ARPSTERN. Therefore all three data sets should be converted.

b) Analysis of terrain data to ARPS grid - ARPSTERN

The main program (ARPSTERN) for ARPS terrain analysis, *arpstern12.f,* and the input file, *arpstern.input*, are located in the main ARPS directory. To compile and run the program (after the direct access data sets have been created), enter the following commands:

> **makearps** *-ncarg arpstern*
> **arpstern** *< arpstern.input > arpstern.output*

Option *-ncarg* is provided in the makearps command so that the NCAR Graphics library is linked for ARPSTERN to plot the analyzed terrain fields. This option should be omitted when NCAR Graphics is not available.

Files *terrain.inc* and *arpstern.input* must be edited to set the control parameters properly before the execution of **arpstern**. The definitions of these parameters are explained in detail in the next subsection.

The terrain file generated by ARPSTERN for ARPS is called *arps40.trndata*.

ARPSTERN includes subroutines that set up the analysis and terrain grids, obtain the required terrain data, perform the desired analysis on the initial terrain field, calculate and plot the response function, and produce contour plots of the terrain fields. The calling tree of ARPSTERN is:

```
ARPSTERN-+-STRLNTH
         +-{OPNGKS}
         +-SETGRID-+-SETMAPR
         |         +-LLTOXY
         |         +-XYTOLL
         |         +-RESPONSE-+-{AGSETC}
         |                    +-{AGSETI}
         |                    +-{EZXY}
         +-GETTER--READTER
         +-BARNES
         +-PLOT (1)-+-{SET}
         |          +-{WTSTR}
         |          +-{SUPMAP}
         |          +-{CONREC}
         |          +-{FRAME}
         +-RMSDIF--PLOT see 1
         +-{CLSGKS}
```

In the program, subroutine SETGRID sets up the terrain and analysis grids. Subroutine GETTER retrieves, by calling READTER, the terrain data from the data base, and if necessary interpolates data from coarser sets to finer resolution data points for use in the analysis program. Subroutine BARNES performs the actual data analysis. The definitions of the other ARPSTERN related subroutines can be found in the subroutine glossary section of Chapter 5.

The terrain data read by GETTER are a function of the geographical location of the analysis grid and of the type of terrain resolution desired by the user. If the header file of the desired resolution contains the needed 1°×1° latitude / longitude block, the record number and file name are passed to READTER and the data are read and added to the initial terrain array. If no match is found the next coarsest resolution data set is searched until the desired data block is found. In the case that no match occurs (this is possible) the program will print a stop message and terminate. Note that the program is designed to search in the direction of increasing terrain data spacing.

c) Control parameters description for terrain data preprocessors

This section provides a guide for setting the control parameters in, unless stated otherwise, input file *arpstern.input. arpstern.input* is used by both the ASCII-to-binary conversion programs and the analysis program ARPSTERN. The default values are given for the example in Section 8.2.3.

Control Parameters for Terrain Preprocessors
(set in *terrain.inc* and *arpstern.input*)

Parameter	Definition/Purpose	Options/Suggested Values
nx *(set in terrain.inc)*	Number (non-dimensional, ND) of ARPS grid points in the ξ (or *x*) direction.	Same as in *dims.inc* set for ARPS. Default = 61.
ny *(set in terrain.inc)*	Number (ND) of ARPS grid points (ND) in the η (or *y*) direction.	Same as in *dims.inc* set for ARPS. Default = 61.
ndx *(set in terrain.inc)*	Number (ND) of data points per 1°×1° data block in the longitudinal direction.	= 2 if *itertype* = 3. = 12 if *itertype* = 2. = 120 if *itertype* = 1. Default = 120.

ndy *(set in* *terrain.inc)*	Number (ND) of data points per 1°×1° data block in the latitudinal direction.	= 2 if *itertype* = 3. = 12 if *itertype* = 2. = 120 if *itertype* = 1. Default = 120.
nptx *(set in* *terrain.inc)*	The minimum number (ND) of terrain data points needed to complete the analysis in the longitudinal direction.	Variable depending on the type of terrain data to be used, the grid spacing and value of k_0 (see Eq. 8.2.7) Default = 300.
npty *(set in* *terrain.inc)*	The minimum number (ND) of terrain data points needed to complete the analysis in the latitudinal direction.	Variable depending on the type of terrain data to be used, the grid spacing and value of k_0 (see Eq. 8.2.8) Default = 300.
dx	ARPS grid spacing (m) in *x*-direction.	Same as in *arps40.input* as is set for ARPS (see Chapter 4). Default = 1000.0.
dy	ARPS grid spacing (m) in *y*-direction.	Same as in *arps40.input* as is set for ARPS (see Chapter 4). Default = 1000.0.
ctrlat	The latitude location (degrees north) of the center of model domain.	Same as in *arps40.input* as is set for ARPS (see Chapter 4). Default = 40.7.
ctrlon	The longitude (degrees east) location of the center of model domain.	Same as in *arps40.input* as is set for ARPS (see Chapter 4). Default = 248.0
analtype	Option (ND) for the terrain data analysis schemes.	1 = Barnes analysis, 2 = Bilinear interpolation with smoothing. Default = 1.
mapproj	Map projection option (ND) used by ARPS model.	Same as in *arps40.input* as is set for ARPS (see Chapter 4). Default = 1.

trulat1 **trulat2**	1st true latitude (degrees north) of map projection. 2nd true latitude (degrees north) of map projection.	Same as in *arps40.input* as is set for ARPS. Default = 40.7.
trulon	True longitude (degrees east) of map projection.	Same as in *arps40.input* as is set for ARPS. Default = 248.0.
sclfct	Scaling factor (ND) for map projection.	Same as in *arps40.input* as is set for ARPS. Default = 1.0.
itertype	Type (ND) of maximum resolution terrain data desired by the user.	1 = 30" data (on the order of 1 km). 2 = 5' data (on the order of 10 km). 3 = 1° data (on the order of 110 km). Ideally, the resolution of the selected data is slightly higher than the model grid resolution. Default = 1.
rmsopt	Option (ND) to calculate the *r.m.s.* differences for a one or multi-pass scheme. Note: *r.m.s.* differences will be calculated automatically for multi-pass analyses.	0 = *r.m.s.* calculated. 1 = no *r.m.s.* calculation. Default = 0.
comtype	Flag for the type of computer used to run the terrain data preprocessors and analysis.	1 = IBM and other computers supporting integer *2 arrays. 4 = CRAY and other computers supporting a minimum of 8 byte integers. Default = 1.
tdatadir	Path name of the directory that contains the direct access data and header files. Character *tdatadir*80.	Character string in quotes. Not to exceed 80 characters in length. Default = '*arpstern.data*'.
terndir	Path name of the directory containing the original ASCII data files *elev.dat* and *dma_elev.dat*. Character *terndir*80.	Character string in quotes. Not to exceed 80 characters in length. Default = '*arpstern.data*'.

knot	Shape factor (ND) used in denominator of the Barnes weight function (k_0).	0 .0 = determined by program, > 0.0, any positive value. Recommended value = average grid-to-closest-data point spacing squared. Default = 0.0.
gamma	Convergence factor (ND) used in the multiple pass analysis (γ).	Positive values between $0.2 < \gamma < 1.0$. Value 1.0 recommended for retaining long wavelength structures and minimizing the return of wavelength features. Default = 1.0.
ipass	Number (ND) of iterations through the Barnes analysis scheme.	1 = single pass. >1 = multiple pass. Multiple pass analyses produce steeper response functions at the cutoff wavelength. Default = 4.
tol	Lower limit (ND) of the weight function/ response considered to be significant. Values below this level are neglected. Used to determine *knot*.	Small number on the order of 10^{-7} (smaller values force larger number of data points to be included in the computational box) Default = 10^{-7}.
wdn	Wavelength in terms of the analysis (ARPS) grid spacing used with the desired initial response *rdnot.*	Any positive number. Value 2 recommended for the removal of small scale features. Default = 2.0.
rdnot	Desired response (ND) of waves of length *wdn* by the first pass of the Barnes analysis.	$0 < rdnot < 1$. $rdnot < 0.05$ recommended. Default = 0.006.

8.2.2. Barnes Analysis Scheme Description

The primary goal of a terrain preprocessor is to generate a numerically smoothed yet realistic terrain data set suitable for use by ARPS. The current terrain analysis program for ARPS utilizes the Barnes (1964) method. To avoid introducing small scale noise near the analysis (model) grid boundary, a

buffered data area surrounding the analysis grid is employed. This buffer zone supplies data outside the analysis area for use in the Barnes analysis.

a) Barnes analysis scheme

The Barnes analysis scheme iteratively applies a Gaussian shaped weight function, *w*, to the input data. This analysis technique has the form:

$$g^n(i,j) = g^{n-1}(i,j) + \frac{\sum_{m=1}^{M}\left[w_m(f_o(x,y) - g^{n-1}(x,y))\right]}{\sum_{m=1}^{M} w_m} \tag{8.2.1}$$

where,

$$g^n(i,j) \qquad = \text{the analysis value at the analysis point } (i,j),$$
$$g^{n-1}(i,j) \qquad = \text{a background field or value of the previous pass at the analysis point,}$$
$$f_o(x,y) \qquad = \text{the data value at a data point } (x,y),$$
$$g^{n-1}(x,y) \qquad = \text{the previous pass analysis at a data point.}$$

The weight function for the Barnes scheme is:

$$w_m = \exp-(\frac{r_m^2}{\gamma k_1}) \tag{8.2.2}$$

where,

$$r_m \qquad = \text{distance between the analysis point and the data point,}$$
$$k_o \qquad = \text{shape or influence factor,}$$
$$k_1 \qquad = \gamma \, k_o,$$
$$\gamma \qquad = \text{convergence factor with } \gamma = 1 \text{ for an one pass scheme and } 0.2 \leq \gamma \leq 1.0 \text{ for a multiple pass scheme.}$$

w is the weight of each terrain data point a distance *r* from the analysis point for a predetermined shape factor k_o. For a multiple pass application, a small γ on the order of 0.2 will revive smaller scale structures removed by the initial pass. The shape factor k_o can be specified by the user or determined by the program. This method allows the user to define a low response at the Nyquist wavelength and for other small-scale structures in the analyzed field. The response of the first pass in spectral space is:

$$D^o = \exp(-k_o \, (\frac{\pi}{\lambda})^2) \tag{8.2.3}$$

where,

$$D^o \quad = \text{initial response of wavelength } (\lambda)$$

λ = wavelength in terms of *wdn* times the maximum grid spacing in x and y directions,

k_o = shape factor.

Solving (8.2.3) for k_o, we have

$$k_o = -\ln(D^o)\,(\frac{\lambda}{\pi})^2 \qquad (8.2.4)$$

The use of maximum data spacing in the x and y directions ensures the removal of smaller scale features due to data spacing variation. The Barnes analysis incorporates a tolerance or cut-off radius to improve the program efficiency. The tolerance parameter (*tol*) and shape factor k_o define the size of the computational box surrounding each analysis point. The computational box is described in the next subsection.

b) Initial terrain data area requirements

When the user selects the multiple pass option (*ipass>1*), the analysis program requires a larger initial terrain area than would be required for a single pass analysis. This is due to the fact, referencing equation (8.2.3), that the second pass adds to the background field (the first pass analysis at the analysis points) the Barnes analysis applied to the difference field between the initial terrain data and the first Barnes analysis at the terrain points. Additional points are required outside the analysis grid area in order to perform an analysis that converges and is free of edge effects.

For this particular application, the analysis at the analysis / model grid points are performed at the same locations for each pass. The analysis at the terrain points is conducted over a smaller and smaller area with each pass.

Before the terrain domain size requirements are defined, a number of intermediate variables need to be determined. The number of data points required for the Barnes scheme to converge to within a pre-set tolerance limit is calculated by the following relations (x and y directions, respectively):

$$nbufx = (\frac{\ln(tol)*k_o}{\text{min. terrain spacing x dir.}})^{\frac{1}{2}} + 1, \qquad (8.2.5)$$

$$nbufy = (\frac{\ln(tol)*k_o}{\text{min. terrain spacing y dir.}})^{\frac{1}{2}} + 1. \qquad (8.2.6)$$

The computational box size is defined by $2*nbuf(x,y) + 1$ data points. Figure 8.1 depicts the area around each analysis point required to obtain a convergent analysis.
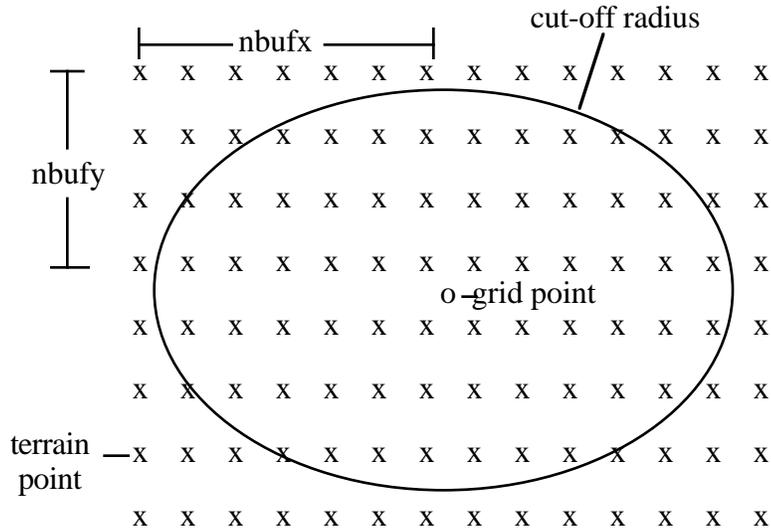
Figure 8.1 Computational box indicating grid points, terrain/data points and the cut-off radius (area of influence) according to a prescribed tolerance level.

After the Barnes shape factor k_o, tolerance, *nbufx* and *nbufy* are determined, an estimate of the number of terrain data points (*nptx, npty*) required to complete the analysis can be made using the following relations: (Note: the user is required to estimate and specify *nptx* and *npty* in *terrain.inc*)

$$nptx = 1.3*((ipass+1)*2*nbufx+\text{analysis area (degrees lon.)}*ndx \quad (8.2.7)$$
$$npty = 1.3*((ipass+1)*2*nbufy+ \text{analysis area (degrees lat.)}*ndy \quad (8.2.8)$$

Parameters *nptx* and *npty* are set in the *terrain.inc* and define the minimum number of terrain data points required in the *x* and *y* directions respectively. Parameters *ndx* and *ndy* prescribe the number of data points per 1 degree block and are set in *terrain.inc* (see Section 8.2.2c).

c) Barnes response function

The response function for a specific n-pass analysis is calculated and plotted by ARPSTERN. For a single pass scheme D^o is obtained through the use of Eq. (8.2.3):

$$D^o = \exp(-k_o*(\frac{\pi}{\lambda})^2). \quad (8.2.9)$$

For a two-pass scheme (*ipass* = 2):

$$D^{1*} = D^o + (1 - D^o)*D^1 \quad (8.2.10)$$

where, $D^1 = (D^o)^\gamma$.

For multiple pass schemes (*ipass > 2*):

$$D^n{}^* = D^o + (1 - D^o)(D^1 + \varphi^n) \qquad (8.2.11)$$

where,

$$\varphi^n = \sum_{i=1}^{i=n} D^i{}^* \prod_{j=2}^{i} (1 - D^{j-1}), \quad \text{for } 2 < n < \infty \text{ and } n = 2 \text{ is for the third}$$

pass (Koch *et al.*, 1985).

8.2.3. Example of a Multi-pass Analysis

An example is given for a 4-pass analysis over Northern Utah. This area is highly variable in terrain elevation and illustrates the characteristics of the multiple pass Barnes scheme. This example used the 30" Continental US data set which, at the latitude of Northern Utah, has an approximate resolution of 700 by 900 meters in the longitudinal and latitudinal directions, respectively. The parameters used by this example are:

In file *terrain.inc*, *nx*=61, *ny*=61, *ndx*=120, *ndy*=120, *nptx*=300, *npty*=300.

In *arpstern.input*, *dx*=1000.0 m, *dy*=1000.0 m, *mapproj*=1, *turlat1*=40.7, *trulon*=248.0, *scfct*=1.0, *itertype*=1, *rmsopt*=1, *comtype*=1, *tdatadir*='arpstern.data', *terndir*='arpstern.data', *knot*=0.0, *gamma*=1.0, *ipass*=4, *tol*=0.0000001, *wdn*=2.00, *rdnot*=0.006.

The final terrain field of the analysis is shown in Figure 8.2. The example area includes a flat basin represented by the Great Salt Lake, gentle valleys, and steep mountainous terrain.

The computational requirements of the terrain preprocessor vary with the number of passes, grid sizes and desired response function characteristics. On an IBM RISC 590 workstation, approximately 30 CPU seconds was required to complete this 4-pass example. When determining which data sets to apply, the user should be aware of the analysis grid spacing to terrain data resolution ratio. Ratios of order 2 will provide a useful analysis field without excess computations. Ratios on the order of 10, for example, require significant amounts of CPU time and memory, without much improvement in the final product.

Figure 8.2. Analyzed terrain contours (contour interval = 100 m) on a 60×60 km grid.

8.3. Surface Characteristics Preprocessor ————————————

8.3.1. Introduction

The preprocessor program, ARPSSFC, prepares the surface characteristics data set for use in ARPS. The data are retrieved from a surface characteristics database and mapped to ARPS grid. The output parameters for ARPS include soil type, vegetation type, leaf area index and surface roughness.

8.3.2. Surface Characteristics Data Base

The original soil and vegetation data were obtained from the Global Ecosystems Database (GED) Version 1.0 on CD-ROM. The original data sets are in IBM PC format and were converted into ASCII integer files for use by ARPSSFC. The ASCII files were created using the following format:

```
integer colmn, rows      ! Columns and Rows
```

```
      integer proj              ! Map projection type??
      real resl                 ! Resolution  (degree??)
      integer data(colmn,rows)  ! Data array

      WRITE(20,'(a20,i8,i8,e20.10,i8)')
&              title,colmn,rows,resl,proj
      WRITE(20,'(20i4)')
&              ((data(i,j),i=1,colmn),j=1,rows)
```

The original data sets can be ordered from:

> *Global Ecosystems Database Project*
> *NOAA/NGDC*
> *Mail Code E/GC1*
> *325 Broadway*
> *Boulder, CO 80303-3328*
> *Phone:(303) 497-6125*
> *email: info@mail.ngdc.noaa.gov*
> *http://www.ngdc.noaa.gov/ngdc.html*

The surface data in ASCII format include the following files:

whsoil.data - the soil type data set is based on Wilson and Henderson-Sellers (1985) soil type classification. It has a spatial resolution of $1° \times 1°$ with global coverage.

owe14d.data - the vegetation type data set is based on Olson *et al.* (1985) world ecosystems classification. It also has a spatial resolution of $1° \times 1°$ with global coverage.

ndvinnnn.data - Kidwell (1990) monthly Normalized Difference Vegetation Index (NDVI) data. *nnnn* in the file name indicates the year and month of the data, *e.g.*, *ndvi9006.data* is for June, 1990. The data covers a period from Jan., 1985 to Dec., 1990. The spatial resolution is $10' \times 10'$ and is global in coverage.

The above three data sets are used by ARPSSFC to derive *soiltyp*, the soil type, *vegtyp,* the vegetation type, *lai*, the leaf area index and *roufns*, the surface roughness length. ARPSSFC first converts the soil and vegetation data into the categories consistent with ARPS. The soil and vegetation types and NDVI from the latitude / longitude grid are mapped onto ARPS grid, by selecting values from the nearest data point (because the type classification is discrete). The leaf area index and surface roughness are derived from the vegetation type and NDVI on ARPS grid then smoothed by a 25-point filter. In the following tables, the conversions from the original classification to the categories used in ARPS are given.

Table 8.1. Conversion of Soil Types from Wilson and
Henderson-Sellers (1985) Soil Classes

USDA Soil Type Classes used by ARPS	Definition	W&H Soil cases
01	Sand	
02	Loamy sand	11 17 23
03	Sandy loam	14 20 26 27
04	Silt loam	
05	Loam	12 18 24
06	Sandy clay loam	15 21 28
07	Silty clay loam	
08	Clay loam	13
09	Sandy clay	19 25
10	Silty clay	16 22
11	Clay	29 30 31
12	Ice	34
13	Water	00

Table 8.2. Conversion to ARPS Vegetation Type from Olson
World Ecosystem Classes

Vegetation Type	Definition	Olson World Ecosystem Classes	Grass/tree	Roughness
01	Desert	01 08 50 69 71	grass	0.002
02	Tundra	42 53	grass	0.020
03	Grassland	16 30 37 40 52	grass	0.010
04	Grassland with shrub cover	02 41 43 49	tree	0.100
05	Grassland with tree cover	58	tree	0.150
06	Deciduous forest	24 25 26 27 29 56 61	tree	0.750
07	Evergreen forest	06 20 21 22 23 46 47 48 57 60 62	tree	1.000
08	Rain forest	32 33 54	tree	2.000
09	Ice	17 70	grass	0.005
10	Cultivation	28 31 36 38 39 55	grass	0.020
11	Bog or marsh	44 45	grass	0.020
12	Dwarf shrub	59 63 64	grass	0.060
13	Semidesert	51	grass	0.040
14	Water	00 65 66 67 68 72 73	grass	0.001

The Leaf Area Index, *lai*, is calculated from NDVI for herbaceous vegetation and trees, respectively. The relations between the NDVI and LAI for herbaceous vegetation (Asrar *et al*., 1984) are:

lai = - ln[1-NDVI / 0.915) / 0.83] / 0.96

and for trees (Nemani and Running, 1989):

lai = 1.625 exp(NDVI / 0.34)

8.3.3. Description of ARPSSFC Program

The driver program for ARPSSFC is contained in file *arpssfc11.f*. ARPSSFC shares with ARPS40 a number of utility subroutines and the compilation and linking are controlled by the same shell script *makearps* and the makefile *make.arps40*. The calling tree of ARPSSFC is:

```
ARPSSFC  +- INITPARA -+- GTLFNKEY
         |            -+- JULDAY
         |            -+- STRLNTH
         |            -+- PRTLOG -+- STRLNTH
         |                       -+- RETUNIT
         |
         +- SETGRD -+- SETMAPR
         |          -+- LLTOXY
         |          -+- SETORIG -+- XYTOLL
         |                       -+- LLTOXY
         |
         +- XYTOLL
         +- GTSFCDT -+- STRLNTH
         |           -+- GETUNIT
         |           -+- RETUNIT
         |
         +- GTSOILTYP
         +- GTSFCDT
         +- GTVEGTYP
         +- GTSFCDT
         +- GTNDVI
         +- GTLAI
         +- SMOOTH25P
         +- GTRFNS
         +- SMOOTH25P
         +- WRTSFCP -+- STRLNTH
         |           -+- GETUNIT
         |           -+- RETUNIT
         |
         +- {OPNGKS}
         +- PLOT -+- {SET}
         |        -+- {WTSTR}
         |        -+- {SUPMAP}
         |        -+- {CONREC}
         |        -+- {FRAME}
         |
         +- {CLSGKS}
```

Subroutines INITPARA and SETGRD are used to read in the input parameters from *arps40.input* and set up ARPS computational grid.

To compile and run program ARPSSFC, enter the following:

makearps *-ncarg arpssfc*
arpssfc *< arps40.input > arpssfc.output*

where **makearps** is the general ARPS make shell script. Script **makearps** creates executable file **arpssfc**.

A number of NCAR Graphics routines are called (those in { } in the calling tree) to generate graphic plotting of the surface property fields. When NCAR Graphics is not available, *-ncarg* option should be omitted in the makearps command. In this case, no graphic output will be generated.

Before the compilation and execution, a set of parameters must be set properly. They are described in the following subsection.

8.3.4. Control Parameters for ARPSSFC

The control parameters used by ARPSSFC are set in *dims.inc*, *arpssfc.inc,* and *arps40.input*. Table 8.3 describes the setting of these parameters.

A more detailed description of the parameters set in *dims.inc* and *arps40.input* can be found in Chapter 4, ARPS Run Time Configuration Guide.

Table 8.3. Control Parameters for the Surface Property Data
Preprocessor - ARPSSFC

Parameter	Definition	Options/Default setting	Set in file
nx	Number of ARPS grid points in x direction	As for ARPS	dims.inc
ny	Number of ARPS grid points in y direction	As for ARPS	dims.inc
colst	Columns for soil type data (e-w)	360	arpssfc.inc
rowst	Rows for soil type data (s-n)	180	arpssfc.inc
colvt	Columns for vegetation type data (e-w)	2160	arpssfc.inc
rowvt	Rows for vegetation type data (s-n)	1080	arpssfc.inc
colnv	Columns for NDVI data (e-w)	2160	arpssfc.inc
rownv	Rows for NDVI data (s-n)	1080	arpssfc.inc
dx	Grid spacing (m) in ξ-direction.	As for ARPS	arps40.input
dy	Grid spacing (m) in η-direction.	As for ARPS	arps40.input
ctrlat	The latitude (degrees north) of the center of model domain.	As for ARPS	arps40.input
ctrlon	The longitude (degrees east) of the center of model domain.	As for ARPS	arps40.input
mapproj	Map projection option	As for ARPS	arps40.input
trulat1 trulat2	True latitudes of map projection	As for ARPS	arps40.input
trulon	True longitude of map projection.	As for ARPS	arps40.input
sclfct	Scaling factor for map projection	As for ARPS	arps40.input
stypout	Output flag for soil type data	1 = yes, 0=no	arps40.input
vtypout	Output flag for vegetation type data	1 = yes, 0=no	arps40.input
laiout	Output flag for leaf area index	1 = yes, 0=no	arps40.input
rfnout	Output flag for surface roughness	1 = yes, 0=no	arps40.input
stypfl	File name of vegetation class data	'whsoil.data'	arps40.input
vtypfl	File name of vegetation class data	'owe14d.data'	arps40.input
ndvifl	File name of NDVI data	'ndvi.data'	arps40.input
sfcdtfl	File name of output surface property data for ARPS	'arpssfc.data'	arps40.input

8.4. Horizontally Homogeneous Model Initial State ——————

ARPS can be initialized with a horizontally homogeneous initial state or a three-dimensionally varying analysis. In the horizontally homogeneous case, five options are available. They are: initialization using an external sounding, specification of an isentropic atmosphere, an isothermal atmosphere, an atmosphere with constant static stability, and the Weisman and Klemp (1982) analytic moist thermodynamic sounding. The horizontally homogeneous options are described in detail in the following subsections and the corresponding code can be found in subroutine INIBASE. The three dimensional initialization procedure is presented in Section 8.5.

8.4.1. Initialization using external sounding(*inibasopt=1*)

A single column sounding file for ARPS contains the values of thermodynamic quantities, moisture and wind at either height or pressure levels. The data do not have to be equally spaced or defined at the model grid levels. When the data are not at the model grid levels, they are linearly (in height) interpolated to the model grid levels. It is important to note that the sounding data must cover the entire extent of the physical model domain. Otherwise, extrapolation will be performed that might yield spurious results. It is recommended that the user check the sounding data after the interpolation. The printout of these data can be found in the standard output file or a separate sounding file (named *runname.sound*). The latter can be fed into your favorite sounding plotting program (*e.g.* Skew-T plot or Tephigram). At CAPS, a program called ARPSKEWT is available as one of the ARPStools.

a) Example sounding

A sample input sounding file (provided in the ARPS release as *may20.snd*) is shown below (those inside [ ] are annotations). A sounding file must have exactly the same format as the example.

```
Sounding Input for ARPS              [header line, not used by model]
supercell storm                      [description of sounding]
15:00:00 CST                         [time of sounding]
20 May, 1977                         [date of sounding]
FT. SILL, OK                         [location of sounding]
'height' 'pot. temp.' 'rel. humidity'  'uv'
                                     [data types (see below)]
0.0    0.965e5                       [height (m) and pressure (pa) at ground level]
33                                   [number of data levels in sounding: lvlsnd]
Z, THETA, RH, U,  V                  [data column labels not used by ARPS]
```

16250.000	406.000	0.00000	14.000	-5.100
15750.000	398.000	0.00000	15.000	-5.400
15250.000	390.000	0.00000	16.000	-5.700
14750.000	382.000	0.00000	17.000	-6.000
14250.000	374.000	0.00000	18.000	-6.300
13750.000	366.000	0.00000	19.000	-6.800
13250.000	358.000	0.00000	20.000	-7.600
12750.000	351.000	0.00000	21.000	-8.800
12250.000	346.000	0.00000	21.800	-10.200
11750.000	343.600	0.00000	22.600	-11.200
11250.000	341.000	0.00000	23.400	-11.700
10750.000	338.500	0.00000	24.000	-11.800
10250.000	336.000	0.34960	24.400	-11.000
9750.000	333.500	0.53010	24.300	-10.000
9250.000	331.000	0.58590	23.700	-8.500
8750.000	329.000	0.57420	22.400	-6.000
8250.000	327.500	0.56300	20.400	-3.000
7750.000	326.000	0.63030	18.400	1.000
7250.000	324.600	0.63160	16.000	6.600
6750.000	323.200	0.60140	13.400	11.400
6250.000	321.800	0.64950	11.000	13.500
5750.000	320.500	0.71920	8.600	14.100
5250.000	319.000	0.77210	6.100	13.700
4750.000	317.500	0.78080	3.600	12.500
4250.000	315.500	0.76890	1.500	10.800
3750.000	313.500	0.74610	-0.500	9.600
3250.000	311.200	0.82920	-3.500	8.800
2750.000	309.000	0.87740	-4.400	8.000
2250.000	307.000	0.89300	-6.200	7.000
1750.000	305.000	0.86640	-7.800	6.000
1250.000	304.000	0.85490	-9.200	4.700
750.000	303.000	0.80820	-10.400	2.000
250.000	303.000	0.71020	-12.000	-3.000
0.000	303.000	0.71020	-12.000	-3.000

In the above example, the data are specified at 500 m increments. Their formats are described next. Please note that the data have to be given in the order of decreasing height or increasing pressure.

b) Sounding data types (see also subroutine SOUNDG):

In the sounding data file, the temperature and moisture variables can be specified at either height or pressure levels and in several types. The types are designated by character strings in the sixth record of the file. Only the <u>first character</u> of each string is examined (case insensitive), and for this reason <u>the first character should not be left blank</u>. The data types for height/pressure, temperature, moisture and winds, respectively, are:

> <u>H</u>eight AGL (meters, default)
> <u>P</u>ressure (Pa)
>
> <u>P</u>otential Temperature (K, default)
> <u>T</u>emperature (K)
>
> <u>S</u>pecific Humidity (mixing ratio) (kg/kg, default)
> <u>R</u>elative Humidity (nondimensional, ranging from 0 - 1).
> <u>D</u>ew Point temperature (K)
>
> <u>X</u>-y components (m/s)
> <u>D</u>irection-speed (degree, m/s)
> <u>K</u>not-direction-speed(degree, knot).

In the example given, the base state is specified by potential temperature and relative humidity at height levels. The horizontal winds are given in their *x* and *y* velocity components. When the winds are given as direction and speed, the direction is the angle between the north and the direction from which the wind comes from (*e.g.*, a $(u, v) = (10, 0)$ m/s wind will be given as direction = $270°$, speed = 10 m/s). Pressure at height levels is diagnosed from the hydrostatic relation.

8.4.2. Isentropic Atmosphere (*inibasopt = 2*)

For an isentropic atmosphere with constant potential temperature θ_0, other base state variables can be expressed as functions of height using the hydrostatic relation:

$$C_p \theta_o \frac{\partial \bar{\pi}}{\partial z} = -g$$

$$\bar{p} = p_o \bar{\pi}^{C_p/R} \qquad .$$

$$\bar{\rho} = \frac{\bar{p}}{R_d \, \theta_o \, \bar{\pi}}$$

Here the overbar indicates the base state variables. For this option, $\bar{q}_v$ is set to zero and $\bar{u}$ and $\bar{v}$ are set according to the value of option *viniopt*. $\theta_0 = 300$ K is hardwired in the code (subroutine INIBASE).

8.4.3. Isothermal Atmosphere (*inibasopt = 3*)

For an isothermal atmosphere with constant temperature T_0, we have the formulations for the other state variables as follows:

$$\bar{p} = p_0 \exp\left(- \frac{gz}{R_d T_0}\right),$$

$$\bar{\pi} = \left(\frac{\bar{p}}{p_0}\right)^{\kappa},$$

$$\bar{\theta} = T_0 / \bar{\pi},$$

$$\bar{\rho} = \bar{p} / (R_d T_0),$$

where $p_0 = 10^5$ Pa. In the program, $\bar{q}_v = 0$ and $\bar{u}$ and $\bar{v}$ are set according to the value of option *viniopt*. $T_0 = 250$ K is hardwired in subroutine INIBASE.

8.4.4. Constant Static Stability Atmosphere (*inibasopt = 4*)

The state variables of a constant static stability ($N = $ constant) atmosphere are given by

$$\bar{\theta} = \theta_0 \, exp\left(\frac{N^2 z}{g}\right),$$

$$\frac{\partial \bar{\pi}}{\partial z} = - g / (C_p \, \bar{\theta}),$$

$$\bar{p} = p_0 \, \bar{\pi}^{1/\kappa},$$

$$\bar{\rho} = \bar{p} / (R_d \theta_0 \, \bar{\pi}),$$

where $N = 0.01 \ s^{-1}$ and $\theta_0 = 300$ K.

8.4.5. Analytic Moist Thermodynamic Sounding (*inibasopt = 5*)

An analytic moist thermodynamic sounding after Weisman and Klemp (1982) is supported by ARPS. The potential temperature, and relative humidity are given in the following.

Define z_{tr} as the tropopause height, T_{tr} and θ_{tr} as the temperature and potential temperature at the tropopause level, respectively, and θ_0 as the surface potential temperature. Then, the vertical base state profiles of potential temperature $\bar{\theta}(z)$ and relative humidity $\overline{RH}(z)$ are given by:

$$\bar{\theta}(z) = \theta_o + (\theta_{tr} - \theta_o)\left(\frac{z}{z_{tr}}\right)^{5/4} \qquad\qquad \text{if } z \leq z_{tr}.$$

$$\bar{\theta}(z) = \theta_{tr} \exp\left[\frac{g}{C_p T_{tr}} (z - z_{tr})\right] \qquad\qquad \text{if } z > z_{tr}.$$

and

$$\overline{RH}(z) = 1 - \frac{3}{4}\left(\frac{z}{z_{tr}}\right)^{5/4} \qquad\qquad \text{if } z \leq z_{tr}.$$

$$\overline{RH}(z) = 0.25 \qquad\qquad \text{if } z > z_{tr}.$$

where $z_{tr} = 1.2 \times 10^4$ m, $T_{tr} = 213$ K, $\theta_{tr} = 343$ K, $\theta_0 = 300$ K are set inside subroutine INIBASE. A user can reset these values in subroutine SOUNDG for a particular application. Finally, the water vapor mixing ratio below 1200 m (assumed mixed layer top) is reset to 0.015 kg/kg. Again, the wind is set according to *viniopt*.

8.5. Initialization Using 3-D Analysis ——————————————————

8.5.1. Generalized data interpolator

Program EXT2ARPS interpolates data from a user-supplied 3-D analysis grid to ARPS grid. The program calls subroutine RDEXTFIL, which supplies the fields on the user's grid. These fields are typically read in from a file. Subroutine RDEXTFIL must be written by the user and must provide the ARPS variables (described below) and arrays of latitude and longitude coordinates of the user-supplied grid. The user-supplied data must cover a volume (horizontal and vertical extent) at least as large as the ARPS model. The interpolator will use linear (in height) interpolation to properly place the data on the staggered ARPS grid and initialize the base state and perturbation fields. A smoother is applied to the ARPS fields to eliminate any noise due to the interpolation. A higher order interpolation scheme may be provided in future ARPS releases.

After the external data are interpolated to ARPS grid, the wind fields are adjusted to ensure that the anelastic mass continuity equation is satisfied. This has been found to be important in reducing the initial oscillations in the mass (pressure) field.

The steps required to process an external file are outlined below.

a) Prepare extdims.inc

The file *extdims.inc* has dimensions of the external fields that will be read and processed. It is used in *ext2arps.f* and *rdextfil.f*.

b) Prepare rdextfil.f

A sample RDEXTFIL is provided that generates fake data on a 2.5° by 2.5° latitude/longitude grid. The file should be edited to read from the desired database and fill the output arrays with the read-in data. The data must be returned on an unstaggered grid, so some interpolation of the original data may be necessary. The units of your input data may need to be changed to match the ARPS units, which are specified below:

p_ext	Pressure (Pascal)
hgt_ext	Height (m)
t_ext	Temperature (K)
qv_ext	Mixing ratio (kg kg^{-1})
u_ext	Eastward wind velocity (ms^{-1})
v_ext	Northward wind component (ms^{-1})
qc_ext	Cloud water mixing ratio (kg kg^{-1})
qr_ext	Rain water mixing ratio (kg kg^{-1})
qi_ext	Ice mixing ratio (kg kg^{-1})
qs_ext	Snow mixing ratio (kg kg^{-1})
qh_ext	Hail mixing ratio (kg kg^{-1}).

NOTES:

1) The winds should be earth relative (that is eastward and northward, which may not be the same as *u* and *v* relative to the x and y coordinates of the source grid).

2) If the cloud, rain, ice, snow or hail mixing ratios are not available they should be set to -999., and EXT2ARPS will set them to zero in the ARPS file. Filling the arrays with zero in RDEXTFIL will achieve the same effect, but less efficiently.

Map projection information for the external grid is also provided by RDEXTFIL as described by the following variables:

iproj_ext	Projection number of external grid
trlon_ext	True longitude of external map projection (degrees east) (longitude line parallel to the -*y* coordinate)
latnot_ext(2)	Latitude(s) where projection is true, that is where the projection surface intersects the sphere, (*latnot_ext(2)* is only used when *iproj_ext* = 2 or -2)
x0_ext	*x* coordinate of origin for external grid (m)
y0_ext	*y* coordinate of origin for external grid (m)

c) Modify make.arps40

You may need to add references to your database libraries or other routines needed for successful linking. Add any subroutines to the compilation list (EXTARPSSRCS) and the link list (EXTARPSOBJS), and add any new libraries to the library list (EXTLIBS).

*d) Create executable **ext2arps***

Enter the command: **makearps** *ext2arps*

e) Modify ARPS input file

Before running **ext2arps**, ARPS input file must be edited for the ARPS run you are preparing. The grid set up, map projection parameters, and control parameters for history data dump of the model run are used by **ext2arps** as well as **arps40**.

If you want to use **ext2arps** to generate boundary conditions for ARPS, in addition to the history files that can be used to initialize the model, set the variable *exbcdmp* to 1 in the NAMELIST block *output*. Before running **arps40** you will likely want to reset *exbcdmp* to 0.

The NAMELIST block *extdfile* contains variables that describe the directory of the input file, the external file's names and the valid time of the data in the files. These parameters can be used by your external data reader and will be passed to RDEXTFIL. Their definitions can be found in Chapter 4.

*f) Run **ext2arps***

Finally, issue command

ext2arps *< arps40.input >! ext2arps.out*

to execute program EXT2ARPS.

The program will create one or a set of history files (depending on how many external data sets you specify in *arps40.input* - the value of *nextdfil*). These files can be used to initialize the model (for initialization option *initopt* = 3) at the desired time. The format and directory of the history file are set in *arps40.input*. If *exbcdmp = 1*, external boundary condition file(s) are created at the same time.

8.5.2. Using External Data Sets in ARPS

CAPS is currently developing an analysis program specifically tailored for ARPS that uses a wide variety of observations as input. For recent field programs, OLAPS, Oklahoma Local Analysis and Prediction System (Brewster *et al.*, 1994) based on the LAPS software originally written by NOAA's Forecast Systems Lab (McGinley, 1995) was used as the real-time data ingest and background objective analysis system for ARPS. OLAPS grids at 10 km resolution have been archived as part of the VORTEX (Rasmussen *et al.*, 1994) data collection project in the U.S. Southern Plains. Also, we have had success creating initial and boundary conditions from the Rapid Update Cycle (RUC) model from U.S. National Meteorological Center.

RDEXTFIL subroutines are available to ingest RUC data in GEMPAK format and OLAPS analysis fields. The former requires the GEMPAK libraries (available from UNIDATA), and the latter the LAPS database routines available from the Forecast Systems Lab. Users interested in contacting those organizations for software should e-mail CAPS (*arpsuser@uoknor.edu*). for more details.

8.6. Time-Dependent Boundary Conditions

ARPS has the capability of forcing the lateral boundary solutions using an external data set. The external data may come from another model or from a coarser resolution run of ARPS. The latter practice can be referred to as one-way interactive nesting, which will be discussed more in next subsection.

A relaxation boundary condition similar to that discussed by Davies (1983) is employed in ARPS. With this scheme an extra forcing term of the form

$$\frac{\partial \bar{\rho}\phi}{\partial t} = -K_b \bar{\rho}(\phi - \phi_e) \qquad (8.6.1)$$

is added to the right hand side of u, v, θ, p and water/ice variable equations. In Eq. (8.6.1), ϕ represents a prognostic variable and ϕ_e is the corresponding external data value. In ARPS, ϕ_e is a value linearly interpolated between two times when external boundary condition data are available.

The relaxation coefficient K_b is defined as,

$$K_b = \frac{K_{b0}}{1 + [2(i - i_b)/(n_b - 1)]^2} \quad \text{for } |\ i - i_b\ | \le n_b\ \text{-}\ 1$$

or $\qquad K_b = 0.0 \qquad\qquad\qquad \text{for } |\ i - i_b\ | > n_b\ \text{-}\ 1 \qquad (8.6.2)$

where K_{b0} is the maximum relaxation coefficient at the boundary, n_b the width of the relaxation zone in terms of the number of grid points, i the grid index and i_b the index of boundary grid point. Apparently K_b is non-zero only within a zone near the lateral boundaries, and decreases from its maximum value at the boundary to 1/5 of this value at the inner edge of this zone. This relaxation procedure forces the model boundary solution towards the external value.

To control noise that might be produced by the inconsistency between the interior model solution and the externally forced value, enhanced computational mixing can be added in the boundary zone. This mixing has the same formation as the second order horizontal computational mixing as described in Section 6.4, and is given by:

$$\frac{\partial \bar{\rho}\phi}{\partial t} = K_c \left[\frac{\partial^2 \bar{\rho}(\phi - \phi_e)}{\partial \xi^2} + \frac{\partial^2 \bar{\rho}(\phi - \phi_e)}{\partial \eta^2} \right] \tag{8.6.3}$$

The coefficient K_c has the same form as K_b given in (8.6.2). Chapter 4 provides the suggested values for K_{b0}, K_{c0} and n_b. It also describes the way to specify external boundary condition files. The external boundary condition files can be produced at the same time history data are written by DTADUMP. A data conversion program, EXT2ARPS, that prepares the external boundary condition files for ARPS from external data sets is described in Section 8.5.

Given that ARPS has the capability of using time-dependent external boundary conditions, as discussed in Section 8.6, it is straightforward to run ARPS in a self-nested mode. This is discussed in Section 7.5.