

User's Manual

THE MESOSCALE MODEL
INTERFACE PROGRAM (MMIF)
Version 3.4.1, 2019-03-11

Prepared for:

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Air Quality Modeling Group
Mail Code C439-01
Research Triangle Park, NC 27711

Prepared by:

Bart Brashers
Chris Emery
Ramboll US Corporation

19020 33rd Ave W Ste 310
Lynnwood, WA 98036

773 San Marin Drive
Novato, CA 94998

11 March 2019

ACKNOWLEDGMENTS

The model authors would like to acknowledge EPA Contract No. EP-D-07-102 (Work assignments 2-06, 4-06, 5-08, and 10-1), EPA Contract No. EP-D-12-044 (Work assignments 4-07, the University of North Carolina, and the Electric Power Research Institute for providing support for the development of MMIF.

CONTENTS

	<i>Page</i>
1. Introduction	1
1.1 Background.....	2
2. Formulation.....	7
2.1 Output to Verify Domain and Point Selection	7
2.2 Output to Support CALPUFF	8
2.3 Output to Support AERMOD	9
2.3.1 ONSITE, UPPERAIR, and AERSURFACE files	10
2.3.2 AERCOARE	12
2.3.3 AERMET SFC and PFL	13
2.4 Output to Support SCICHEM.....	15
2.5 Grid Structure	16
2.5.1 Horizontal Grid Structure	17
2.5.2 Vertical Grid Structure	20
2.5.3 Vertical Aggregation	22
2.5.3.1 Aggregation and CALPUFF’s Lowest Layer	23
2.5.3.2 Aggregation and Monthly MMIF Runs for CALPUFF	26
2.5.4 Vertical Interpolation	26
2.5.4.1 “Pass-Through” Vertical Interpolation	30
2.6 Surface Characteristics	30
2.7 PBL Parameters.....	34
2.8 Cloud Cover.....	35
2.9 Diagnosis of Stability Class	36
3. Code Structure and Compilation	39
3.1 Compiling MMIF on Windows.....	40
3.2 Compiling MMIF on LINUX/UNIX	42
3.3 A Note on Binary Input/Output Files	43
4. How to Run MMIF	45
4.1 Control File Format	47
4.2 Control File Keyword Listing.....	51
4.3 A Note on Time Stamps	61
4.4 A Note on Precipitation Processing	62
5. Limitations	65
6. References.....	67

LIST OF ACRONYMS

EM	Equatorial Mercator projection
FSL	Upper-air data format available from the NCDC or NOAA/ESRL radiosonde database on the internet
LCC	Lambert Conformal Conic projection
LST	Local Standard Time
MCIP	Models-3 Meteorology-Chemistry Interface Processor
MMIF	Mesoscale Model Interface program
MM5	Penn State/NCAR Mesoscale Meteorological model version 5
MM5CAMx	MM5 pre-processor for the CAMx modeling system
NCAR	National Center for Atmospheric Research
NLCD	National Land Cover Database
NOOBS	No Observations, a setting in CALMET
PBL	Planetary Boundary Layer
PG	Pasquill-Gifford Stability Class
PS	Polar Stereographic projection
SCICHEM	Second-order Closure Integrated Puff Model with Chemistry
WRF	Weather Research and Forecasting model

1. INTRODUCTION

Skip straight to Section 4, "How to Run MMIF" (page 45) to learn about the control file format, command line options, and how to run MMIF.

The Mesoscale Model Interface Program (MMIF) converts prognostic meteorological model output fields to the parameters and formats required for direct input into dispersion models. MMIF v1.0 (Brashers and Emery, 2009) supported only CALPUFF, and was intended to be an alternative to CALMET in generating three-dimensional meteorological input fields for long-range transport assessments in support of regulatory air quality impact analyses. Processing data directly from the prognostic meteorological models into a CALPUFF-ready form brings a much higher level of consistency to the review process and will allow the use of specific or specialized meteorological simulations without additional modification or additional performance evaluation.

Support for AERMOD and SCICHEM dispersion models was added in MMIF v2.0. AERMOD is further supported in three modes. MMIF can act as a replacement for AERMET, the meteorological pre-processor from the AERMOD modeling system, mimicking its output directly. MMIF can also output data suitable as an input to AERMET, using the ONSITE and UPPERAIR pathways. MMIF can also output data suitable as an input to AERCOARE, an alternative to AERMET suitable for over-water dispersion. Support for SCICHEM is via the MEDOC file format, containing the required 2- and 3-dimensional fields. SCICHEM can also accept AERMET-style files (SFC and PFL files) for near-field (< 50 km) applications, which can be generated by MMIF.

In MMIF v2.1, support for vertical interpolation (as opposed to vertical aggregation) was added. This version also added support for WRF runs that used the NLCD 50-category land-use scheme, updated the surface-layer diagnostic routines to fix extremely stable conditions, and made a few more minor changes. MMIF v2.2 included only bug fixes and increased clarity for the time range when in CALPUFFv6 mode. MMIF v2.3 included enhancements and bug fixes associated with SCICHEM support. In MMIF v 3.2, new keywords were added to specify minimum values of mixing heights and turbulence parameters associated with AERMOD simulations. MMIF version 3.3 featured additional mixing height updates and enhanced keyword descriptions in the control file.

MMIF specifically processes geophysical and meteorological output fields from the Fifth Generation Mesoscale Model (MM5, version 3) and the Weather Research and Forecasting (WRF) model (Advanced Research WRF [ARW] core, versions 2 and 3). The program diagnoses certain required parameters that are not directly available from various versions of MM5 or WRF. It also offers the option to directly pass through planetary boundary layer (PBL) heights from the meteorological models, or to independently diagnose them from other variables. The program does not perform interpolation to different map projections or grid resolutions (all gridded fields remain on the same projection and resolution as the meteorological model). However, it can produce outputs for any sub-domain of the meteorological modeling grids. All program code is well organized and documented, and able to compile and run on both Windows and Unix/Linux platforms.

This user's manual documents MMIF v3.4.1 and includes descriptions of the algorithms, the program code, user input, and runtime instructions. The remainder of this section provides background information. The initial development of MMIF was sponsored by the U.S. Environmental Protection Agency (EPA), Office of Air Quality Planning and Standards (OAQPS), under Contract EP-D-7-102, Work Assignment 2-06. Further development was under Work Assignments 4-06 and 5-08. SCICHEM support between v2.2 and v2.3, and for v3.1, was provided by the Electric Power Research Institute (EPRI). Development continued under Contract EP-D-12-044, Work Assignment 4-07, and resulted in MMIF v3.3. MMIF v3.4 development was unsupported, the result of some minor changes and quick enhancements the model author found useful. Similarly, MMIF v3.4.1 was unsupported, with some minor bug fixes and a new option controlling the source of the cloud cover.

1.1 Background

The idea for a direct CALPUFF meteorological pre-processor in the style of EPA's Models-3 Meteorology-Chemistry Interface Processor (MCIP) was first discussed as a potential option for the CALPUFF system at EPA's 8th Conference on Air Quality Modeling in September 2005 (Anderson, 2008). CALMET has long supported the use of raw output data from MM5 and other prognostic meteorological models to define background or first-guess wind fields, and this capability has proven to be particularly useful in data-sparse areas. A direct interface processor, on the other hand, offers the advantage of minimizing the manipulation of prognostic meteorological data before input to CALPUFF. Prognostic meteorological model

output quality has advanced in recent years and many users find it preferable to use the data without further manipulation, for downstream air quality modeling applications.

The EPA/OAQPS, along with several EPA regional offices and Federal Land Management (FLM) agencies, have sponsored the development of MMIF for regulatory applications as an alternative to CALMET. Such a tool brings a much higher level of consistency to the review process, and it is recognized that many other organizations and projects could benefit from the use of a single-purpose direct converter, since many modelers often rely on CALMET to simply pass MM5 meteorological fields through to CALPUFF in the "NOOBS" mode.

The EPA developed a prototype meteorological re-formatting tool (Anderson, 2008) designed to process MM5 data directly to CALPUFF input formats. The prototype was based on a combination of ENVIRON's MM5CAMx converter and EPA's MCIP.

In 2009, ENVIRON was contracted by EPA/OAQPS to thoroughly review, update, test and debug the prototype to ensure sufficient quality for regulatory review and public distribution. The code update included an entire re-write in Fortran 90, employing dynamic memory allocation to maximize code portability, efficiency and ease of use. The update also expanded the capabilities of the processor to optionally process output fields from WRF-ARW versions 2 and 3. The code is organized within a highly modular structure, and in-code documentation was added to facilitate external review and future upgrades. The product of that work was MMIF version 1.0. Limited preliminary testing of MMIF v1.0 was conducted by ENVIRON using pre-existing MM5 and WRF datasets that were available in-house.

Interest in driving AERMOD directly from MM5/WRF output for remote regions has grown in recent years, as has EPA/FLM interest in SCICHEM. In 2010, ENVIRON was tasked with enhancing MMIF to support AERMOD (both directly, and via AERMET and AERCOARE) and SCICHEM. Interest in supporting over-water application of AERMOD (the so-called AERMOD-COARE method described in the EPA Model Clearinghouse memo of May 6, 2011 ⁽¹⁾) spurred both the development of a FORTRAN-based tool named AERCOARE (developed under a separate Work

⁽¹⁾http://www.epa.gov/ttn/scram/guidance/mch/new_mch/Model%20Clearinghouse%20Review%20of%20AERMOD-COARE.pdf

Assignment under the same Contract) and the enhancement of MMIF to support AERCOARE. MMIF was enhanced in several other ways as well, to improve flexibility and ease of use. The product of this work was MMIF version 2.0

Other Work Assignments under the same Contract applied MMIF in side-by-side comparisons between CALPUFF with CALMET, CALPUFF with MMIF, and SCICHEM (with MMIF). These tasks included plume study evaluations (short time periods) as well as typical AORV assessments (long time periods) and helped test and improve the final product.

In 2011, ENVIRON was tasked with several more enhancements to MMIF, resulting in MMIF v2.1. Support was added for vertical interpolation, and for processing of WRF files that used the NLCD 50-category land-use scheme. The treatment of the surface layer parameterization used to calculate required variables that were not found in the meteorological model output was fixed to give reasonable results even in very stable conditions. The user control of minimum and maximum PBL heights was removed, replaced by internally reasonable constraints. The second output file when run in SCICHEM mode was changed from the terrain file to a sampler locations (receptor) file containing the full 3-D grid.

The 2011 Work Assignment also included on-going support for MMIF, resulting in MMIF v2.2. In late 2012, the Electric Power Research Institute funded a fix for the vertical structure of the MEDOC file that MMIF writes. On-going support from EPA also provided for several bug fixes, and MMIF version 2.3 was released in the spring of 2013.

At the request of EPA, a major re-write of MMIF was undertaken in 2013. MMIF was re-programmed to support multiple outputs from a single run. The user can create multiple 3-D gridded outputs (CALPUFF, CALPUFFv6, and SCICHEM) at the same time, as well as any number of 1-D AERMOD, AERMET, and AERCOARE runs, limited only by the underlying operating systems maximum number of open files allowed. To better facilitate a varying number of output files, the control file format was drastically changed from a fixed-format to a keyword-driven format. The result was MMIF version 3.0, released in fall 2013. This version included bug fixes, and added support for most of the WRF landuse datasets, but did not include support for any new dispersion models or new major features.

In the spring of 2014, EPRI sponsored the development of a new version of SCICHEM. The file format for the MEDOC file headers produced by MMIF were changed accordingly, to include more information about the WRF/MM5 grid projection. This enhancement, along with better support for polar stereographic projections, resulted in MMIF v3.1.

A number of minor enhancements were added during the following year, including two new keywords to control minimum values of mixing height and the Monin-Obukhov length when running in AERMOD mode. MMIF v3.2 was released just prior to EPA's 11th Modeling Conference.

MMIF v3.3 was released in December 2016 and contains bug fixes, additional mixing height updates, improved SCICHEM support, and more descriptive comments in the control file. One change to the control file keywords that is *not* backwardly-compatible was required: MIXHT changed to CALSCI_MIXHT to avoid confusion with AER_MIXHT. Additionally, the "useful" file in AERMET and AERMOD modes are now identical; use the new keywords BAT or CSH (pick one) to establish the basename for output.

MMIF v3.4 was released in July of 2018, and contains a bug fix that affected CALPUFF and SCICHEM modes (CALMET-like output file headers were incorrect, and did not match the values in the useful info file). This bug was introduced in MMIF v3.3, and does not affect AERMET/AERMOD/AERCOARE modes. It also included enhancements to support EM (equatorial Mercator) projections, following user feedback. To support concurrent SCICHEM development, the units (e.g. m/s, kg/kg) of the parameters written to a MEDOC file are now written in that file's header. MMIF v3.4 now detects skipped/missing hours, and prints a warning. It no longer writes "run completed normally" when there were warnings. Following the 2018 MMIF Guidance update, aer_min_speed now defaults to 0. This also triggers setting the keyword THRESHOLD 0.0 in the MMIF-generated AERMET control file (*.IN1). The ADJ_U* setting is now turned on by default. Added features included a new keyword QAPLOT, similar to CALPUFF's IQAPLOT keyword. This causes MMIF to write a BLN or BNA file showing the full WRF/MM5 domain and the selected 3-D sub-domain, similar to CALPUFF's qagrid.bna file. It can also write a DAT file showing the requested single-point output locations, or a KML file containing both the domains and points. BLN, BNA, and DAT files can be plotted with the same plotting programs that many CALPUFF users are familiar with, and use the

projected coordinate system (LCC, PS, EM), the same as CALPUFF's IQAPLOT output files.

MMIF v3.4.1 was released in March of 2019, and contains a bug fix that affected CALPUFF and SCICHEM modes when reading MM5 data (the same bug was fixed for reading WRF data in MMIF v3.4). This release introduced a new keyword, CLOUDCOVER, that controls the source of cloud cover fraction written in CALPUFF, SCICHEM, and AERMOD modes. Cloud cover is not written to the ONSITE data file in AERMET mode. See Section 2.8 for more information.

2. FORMULATION

MMIF was developed from ENVIRON's² MM5CAMx and EPA's MCIP meteorological interface software. Key features include:

- Applicability on either Linux/Unix or Windows platforms;
- A simple text-based user interface "control" file;
- Two options to determine Pasquill-Gifford (PG) stability class (relevant only for CALPUFF);
- Options to re-diagnose or pass through PBL depth;
- An option to generate output on a sub-set of the meteorological modeling grid;
- An optional mass-weighted vertical aggregation of multiple MM5/WRF layers;
- An optional mass-weighted vertical interpolation from MM5/WRF sigma layers to height-above-ground layers.

MMIF supports three "downstream" dispersion models: CALPUFF, AERMOD, and SCICHEM.

2.1 Output to Verify Domain and Point Selection

If the output file format is set to QAPLOT, then MMIF will produce four files (BLN, BNA, DAT, and KML). The first three are intended to mimic CALPUFF's output triggered by setting IQAPLOT = 1. The full WRF/MM5 domain and the requested sub-domain are written in BLN/BNA format, and the individual extraction POINTs for AERMOD are written in the DAT file. These can be plotted in any program, including those very familiar to CALPUFF users.

The fourth output (KML) produces a file containing both the domains (full and sub-) and the individual points in a file suitable for viewing in the software program "Google Earth."

It is hoped that these QAPLOTs will speed the selection/revision of a requested sub-domain, and allow the users to tell very quickly if the requested POINT output locations are where they want.

² ENVIRON merged with Ramboll, used "Ramboll Environ" for a few years, and is now Ramboll.

2.2 Output to Support CALPUFF

If the output file format is set to CALPUFF, then MMIF will produce three output files (TXT, DAT, and GRD). The Useful Info File (TXT) will be a text file containing certain keywords and corresponding values, useful for inclusion in a CALPUFF control file. Keywords include the projection/datum information, grid origin and spacing, etc.

The First Output File, a Data Version 2.0 CALMET.DAT-formatted file, can be used by CALPUFF version 5.8.x directly. The Second Output File will be a Surfer⁽³⁾ Text Grid file (GRD) containing the MM5/WRF terrain elevations for the output sub-grid. This file can be useful for interpolating elevations of sources in a CALPUFF run.

If the output file format is set to CALUFFv6, then MMIF can produce four output files (TXT, DAT, GRD, and AUX). The Useful Info File (TXT) is as above.

The First Output File is a Data Version 2.1 CALMET.DAT-formatted file, which can both be used by CALPUFF v6.x directly. The Second Output File is a Surfer Text Grid file (GRD), as above. The Third Output File, a CALMET.AUX-formatted file (containing the 3-dimensional cloud liquid water mixing ratio) can also be used by CALPUFF version 6.x directly.

The following variables are written to a CALMET.DAT file format:

Time-invariant fields

- 2-D surface roughness length, m (Z0);
- 2-D landuse code, dimensionless (ILANDU);
- 2-D topographic elevation, m (ELEV);
- 2-D leaf area index, dimensionless (XLAI).

Time-variant fields

- 3-D U-component (west-east) scalar wind, m/s (U-LEV);
- 3-D V-component (south-north) scalar wind, m/s (V-LEV);
- 3-D W-component (vertical) scalar wind, m/s (WFACE);

⁽³⁾ Surfer is a commercial plotting program, made by Golden Software (<http://www.goldensoftware.com>). Many programs in the CALPUFF modeling system create files that Surfer can plot.

- 3-D temperature, K (T-LEV);
- 3-D cloud liquid water mixing ratio (CLDMR) to CALMET.AUX file;
- 2-D PG stability, dimensionless (IPGT);
- 2-D surface friction velocity scale, m/s (USTAR);
- 2-D PBL depth, m (ZI);
- 2-D Monin-Obukhov length, m (EL);
- 2-D convective velocity scale, m/s (WSTAR);
- 2-D rainfall rate, mm/hr (RMM);
- 2-D surface temperature, K (TEMPK);
- 2-D density, kg/m³ (RHO);
- 2-D surface solar flux, W/m² (QSW);
- 2-D relative humidity, percent (IRH);
- 2-D precipitation code, dimensionless (IPCODE).

MMIF writes CALPUFF-ready input files equivalent to the CALMET “NOOBS” option (number of surface, upper air, and over-water sites are all zero) and the CALGRID option (3-D gridded fields of wind components and temperature).

2.3 Output to Support AERMOD

The AERMOD modeling system includes the program AERMET, a meteorological pre-processor. AERMET contains an internal model for the growth of the daytime mixed layer, which (a) may not be appropriate for over-water situations where much of the incoming solar energy is converted to latent heat flux (evaporation) rather than sensible heat flux, and (b) is arguably a less advanced model than many versions of MM5/WRF PBL parameterizations.

MMIF supports AERMOD in three ways:

- create ONSITE, UPPERAIR, and mimic AERSURFACE data; then run AERMET
- create AERCOARE input files; then run AERCOARE (replaces AERMET)
- create AERMET-like output surface and profile files (replaces AERMET)

MMIF does not interpolate between MM5/WRF grid points to the exact point requested in the control file. It simply uses the center of the (single) grid cell containing the requested point.

2.3.1 ONSITE, UPPERAIR, and AERSURFACE files

If the output file format is set to AERMET, then MMIF can produce up to seven output files (BAT/CSH, IN1, IN2, IN3, AERSURFACE, DAT, and FSL).

The Useful Info File will be a text file with some lines suitable for inclusion in the 'ME' section of an AERMOD control file. Keywords include SURFFILE, PROFFILE, PROFBASE, SURFDATA, and UAIRDATA. It is identical to the useful file produced in AERMOD mode.

Using the keyword BAT (batch) or CSH, a file *.BAT for Windows or a *.CSH/*.SH for Linux platforms is created. In general, the user should pick either BAT or CSH, not both. The filename given following the BAT/CSH keyword serves as a base name, in addition to being a script that will run AERMET with each of the three stages of input files. MMIF will then write stage 1, stage 2, and stage 3 input files using the same base filename as the BAT/CSH file, but changing the extension to IN1, IN2, and IN3. Of course, these can be edited by the user to change any of the methods assumed by MMIF.

MMIF can write the run-length average roughness length, albedo, and Bowen ratio for the grid cell in the same format as AERSURFACE, the land-use pre-processor in the AERMOD modeling system. MMIF writes MONTHLY values, and outputs "0.99" for the months not covered by the MMIF run to prevent AERMET from stopping with out-of-range errors.

MMIF uses only one sector (0° to 360°) in its AERSURFACE-like output. Typical MM5/WRF model runs do not have sufficient resolution to calculate the 1 km geometric averages for each 30° sector (i.e. typical MM5/WRF runs have grid spacing greater than 1 km). If the requested location for a MMIF run is over land, it may be tempting to run AERSURFACE using traditional data sources, and use its output file in the AERMET run instead. However, this might produce a mismatch between the MM5/WRF winds/profiles and the AERSURFACE roughness lengths, Bowen ratio, and Albedo.

The ONSITE output file will be a text file containing the following values:

- Time in LST (XDATES keyword)
- Solar radiation, W/m² (INSO)
- Precipitation rate, mm/hr × 100 (PRCP)

- Surface pressure, mb × 10 (PRES)
- PBL height, m (MGHT)
- 2m height, m (HT01)
- 2m relative humidity, percent (RH01)
- 2m air temperature, °C (TT01)
- Air temperature difference between lowest MM5/WRF level and 2m, °C (DT01)

If the lowest MM5/WRF level is more than 10 meters above ground, MMIF writes the following:

- 10m height, m (HT02)
- 10m wind speed, m/s (WS02)
- 10m wind direction, degrees (WD02)
- Temperature, °C (TTnn)
- Relative humidity, percent (RHnn)

MMIF then writes the following fields for each requested output level to the ONSITE output File:

- Height of the measurement, m (HTnn)
- Wind speed, m/s (WSnn)
- Wind direction, m/s (WDnn)
- Temperature, °C (TTnn)
- Relative humidity, percent (RHnn)

The UPPERAIR output file will contain a sounding in simplified FSL⁽⁴⁾ format⁽⁵⁾.

Although MMIF attempts to mimic this format, it does not interpolate the MM5/WRF profile to mandatory reporting levels. All MM5/WRF levels are written using the code '5' (significant level) rather than '6' (wind level).

⁽⁴⁾Upper air data in FSL format can be download from NOAA/ESRL at <http://esrl.noaa.gov/raobs/>

⁽⁵⁾The FSL format is shown at http://esrl.noaa.gov/raobs/intl/fsl_format-new.cgi.

Note that if using WRF or MMIF re-calculated mixing heights (`AER_MIXHT` = WRF or MMIF) then no FSL file is required to run AERMET. The ONSITE file contains values for MGHT, and AERMET's model of the daytime growth of the mixed layer is not needed.

AERMET v11059 has an unintentional limit to the frequency of UPPERAIR files. In `mppbl.for`, around line 1110, the array `tmp` is used to find the "next" sounding within 2 days of the current time-stamp. It has 24 elements, but should have 48 elements to allow for up to 48 hourly soundings. As a workaround to this issue, and because users may find it generally useful, MMIF can output soundings in the FSL file every *N* hours. The default is to output every 12 hours. Using the keyword `FSL_INTERVAL` followed by an integer will cause MMIF to skip the hours not evenly divisible by this number. For example, including "`FSL_INTERVAL 3`" will generate output at 0Z, 3Z, 6Z, 9Z... 21Z. Supplying a 12 after the filename will generate data at 0Z and 12Z, the standard times for the global network of soundings.

The fields for station IDs (WBAN and WMO numbers) are filled with 9's.

2.3.2 AERCOARE

If the output file format is set to AERCOARE, then MMIF can produce two output files (Useful Info and CSV). The Useful Info file is a sample control file for AERCOARE, and can of course be edited by the user to change input values from MMIF's defaults if desired.

The DATA output file will be a comma separated values (CSV) data file, and includes the following values:

- Time (LST)
- Wind speed, m/s (`wspd`)
- Wind direction, degrees (`wdir`)
- Sea surface temperature, °C (`tsea`)
- 2m air temperature, °C (`tair`)
- 2m relative humidity, percent (`relh`)
- Surface pressure, mb (`pres`)
- Downward solar radiation at the surface, W/m² (`srad`)
- Downward longwave radiation at the surface, W/m² (`rdow`)
- Precipitation rate, mm/hr (`rain`)

- Total sky cover, tenths (tsky)
- Mixing height, m (mixh)
- Vertical potential temperature gradient above the PBL, K/m (VPTG)
- Heights of the wind, temperature, and relative humidity measurements, m (zwsp, ztem, zrel)
- Depth of the sea surface temperature measurement, m (zdep)

MMIF will output either the 10m winds, or the lowest model level winds, whichever is lower. MMIF also assumes the sea surface temperature values from MM5 and WRF are skin temperatures, so sets *zdep* to zero.

MMIF will issue a warning if the selected output point is over land, because AERCOARE is an over-water replacement for AERMET. MMIF will still write the file, because this file format can be useful for a quick examination of the WRF data at a point.

2.3.3 AERMET SFC and PFL

If the output file format is set to AERMOD, then MMIF can produce three output files (Useful Info, SFC, and PFL). The SFC and PFL files can be used by AERMOD directly.

The Useful Info File will be a text file with some lines suitable for inclusion in the 'ME' section of an AERMOD control file. Keywords include SURFFILE, PROFFILE, PROFBASE, SURFDATA, and UAIRDATA.

The First Output File will be a surface (SFC) file, and includes the following values:

- Time (LST)
- Sensible heat flux, W/m^2 (shflux)
- Monin-Obukhov wind speed scaling parameter, m/s (ustar)
- Convective velocity scale, m/s (wstar)
- Vertical potential temperature gradient above the PBL, K/m (VPTG)
- Convective mixing height (always PBL height), m (Zic)
- Mechanical mixing height (also PBL height), m (Zim)
- Monin-Obukhov length, m (L)
- Roughness length, m (z0)
- Bowen ratio, dimensionless (bowen)

- Albedo, fraction (albedo)
- 10m Wind speed, m/s (speed)
- 10m Wind direction, degrees (dir)
- Height of the wind measurement, m (Zwind)
- 2m air temperature, K (T)
- Height of the temperature measurement, m (Ztemp)
- Precipitation code: '11' if $T_{\text{air}} > 0^{\circ}\text{C}$, '22' otherwise ⁽⁶⁾ (ipcode)
- Precipitation rate, mm/hr (rain)
- 2m relative humidity, percent (RH)
- Surface pressure, mb (press)
- Cloud cover, tenths (cloud)
- Wind speed adjustment tag (always 'NAD-OS' for MMIF)

Note that the field width (number of characters) used to write the roughness length (z_0) has been increased from 7 characters to 9 characters (FORTRAN formats F7.4 to F9.6). AERMET uses format F7.4, but over water it is possible to valid z_0 values less than 0.0001 meters. Using AERMET's format would give some values over water with $z_0 = 0.0000$, which is incorrect.

The Second Output File will be a profile (PFL) file, and includes the following values:

- Time (LST)
- Height of measurement, m
- Flag for last level
- Wind direction, degrees
- Wind speed, m/s
- Air temperature, $^{\circ}\text{C}$
- Sigma theta, degrees, always '99' (missing) in MMIF output
- Sigma W, degrees, always '99' (missing) in MMIF output

⁽⁶⁾ See page C-13 of the AERMET Users Guide addendum v11059

2.4 Output to Support SCICHEM

If the output file format is set to SCICHEM, then MMIF can produce up to five output files (Useful Info, two versions of MCW, TER, and SMP). The Useful Info file will be a text file containing a few of the keywords used in the control file (SCI) and meteorological scenario (MSC) control files for SCICHEM. These parameters include the grid origin and size, the vertical grid structure, etc.

The primary output file will be a MEDOC-formatted file (which customarily uses the MCW extension). MMIF can write either a text (using the keyword ASCII) MEDOC file, or a binary MEDOC file (using the keyword BINARY). MMIF versions before v3.4 did not write the units of each output parameter in the MEDOC header, instead relying on SCICHEM to assume default units. Unfortunately, SCICHEM v3.1 assumes liquid water content in g/kg, which MMIF wrote it in the WRF-standard kg/kg. Starting with MMIF v3.4, the units are explicitly written to the header of the MEDOC file.

Using the keyword SAMPLER will produce a text file suitable for inclusion as the SMPFILE, the sampler locations file. This file will include the X, Y, and Z of all the points in the output sub-domain. The Z values are in meters, and are the height above ground of the transformed cell centers. See Section 2.5.2 for more information on the transformation the MEDOC file format requires for the cell centers. The heights of the transformed and un-transformed layers are equal only for grid points at sea level. Current best-practices for SCICHEM include *not* using a SMPFILE, but instead specifying extraction points after the SCICHEM run – during the sciDOSpost run.

Using the keyword TERRAIN will produce a text file with the terrain in a format readable by SCICHEM.

SCICHEM can also be run using AERMET surface (SFC) and profile (PFL) meteorological files, which can be produced by MMIF. Using AERMET-like SFC and PFL files in SCICHEM is recommended only for near-field (<50 km) applications.

The MEDOC files contain the following values:

- 3D easterly wind speed, m/s (U)
- 3D northerly wind speed, m/s (V)
- 3D vertical wind speed, m/s (W)

- 3D air temperature, K, (TA)
- 3D water vapor mixing ratio, kg/kg (H)
- 3D cloud liquid water mixing ratio, kg/kg (CLD)
- 2D terrain elevation, m (TOPO)
- 2D PBL height, m (ZI)
- 2D surface (sensible) heat flux, W/m^2 (HFLX)
- 2D precipitation rate, mm/hr (PRECIP)
- 2D estimate of cloud cover, fraction (CC)

2.5 Grid Structure

MMIF generates gridded meteorological fields on the same horizontal map projection and grid resolution as defined by the meteorological models. No interpolation to a different projection or resolution is performed (de-staggering of horizontal winds are the only minor exception, as described below). MMIF supports the three projections that MM5, WRF, and CALPUFF have in common: Lambert Conformal Conic (LCC), Polar Stereographic (PS), and Equatorial Mercator (EM). SCICHEM only requires that the grid be regular in some projection, and does not need to “know” the details of the actual projection. MMIF does allow the user to define a 3-dimensional sub-domain of the meteorological model grid on which to generate output fields. For CALPUFF and SCICHEM output, the user specifies the lower-left and upper-right corners of the desired sub-domain. For any of the AERMOD output formats, a single point (essentially a 1×1 grid) is specified, but it must be within the 3-D domain used by CALPUFF and SCICHEM.

Both MM5 and WRF using a pressure-based vertical coordinate, which all three of AERMOD, CALPUFF, and SCICHEM use a distance-based vertical coordinate. MMIF allows the user to either (a) aggregate multiple vertical meteorological model input layers to a subset of output layers, or include all input layers, or (b) interpolate to user-specified output layers. The user has the option to request the output layers be defined as the horizontal average (over the output sub-domain) of the input layers. It is not necessary to request a vertical layer structure that extends to the top of the meteorological model domain.

The aggregation routines induce a small error in the layer heights, in the sense that they use the average layer heights over the entire output domain. If the user desires a “pass-through” of the input meteorological model with as little change as

possible, it is suggested to use interpolation to the horizontally-averaged layer heights, rather than aggregation.

2.5.1 Horizontal Grid Structure

The horizontal grid structure of CALPUFF is defined such that all variables are valid at grid cell center ([Figure 2-1](#)). Both MM5 and WRF carry the horizontal wind components (U and V) in a staggered arrangement relative to the state (scalar) variables, referred to as “Arakawa B” and “Arakawa C” staggering, respectively. Therefore, the U and V wind components from the meteorological model are averaged to cell center within MMIF. For MM5 data, the four corner wind points are averaged with uniform weighting to cell center. For WRF data, the two U-face and two V-face winds are separately averaged with uniform weighting from their locations to cell center.

MMIF allows a sub-domain to be extracted from the full MM5 or WRF grid. [Figure 2-2](#) shows an example of the grid indexing convention used in both MM5 and WRF. The meteorological models index their grids according to the grid cell corners (MM5) or interfaces (WRF), which are noted by the external tick marks labeled 1 through 11 around the edge of the domain. These are the locations at which the wind components are carried by MM5 and WRF. The actual grid cells are indexed by the red values 1 through 10 (note there is always 1 less grid cell than the number of interface/corners). When selecting a sub-domain in MMIF, the output grid is defined by a range of grid cells (red values); in the example shown in [Figure 2-2](#), the output grid ranges from grid cell 4 through 7 in the x-direction, and from grid cell 3 through 7 in the y-direction.

In CALPUFF, the lower-left corner point coordinate used to geo-locate the grid is taken from the respective MM5/WRF *corner* point, which is at (4, 3) in [Figure 2-2](#). The values of XORIGKM, YORIGKM are written to the Useful Info File, along with the other parameters that define the grid.

In SCICHEM, the lower-left corner point coordinate is taken from the respective MM5/WRF grid cell *center* point, which is at (4.5, 3.5) in [Figure 2-2](#). This is because of the fundamental difference between the grid definition method in SCICHEM compared to CALPUFF, MM5, and WRF. All values in SCICHEM, including coordinates and corresponding latitude/longitude values, are defined at the *center* of each grid cell. The grid spacing is therefore defined not as the width of each grid cell, but the center-to-center distance between grid cells. Because the grids are

regular, these two definitions are functionally identical, except for the definition of the grid's lower-left and upper-right extents (grid cell *corners* for CALPUFF, grid cell *centers* for SCICHEM).

For the sake of simplicity and user-friendliness, MMIF uses the CALPUFF/MM5/WRF convention (grid corners) for all supported downstream models. It transparently writes the grid cell centers to the MEDOC file, when run in SCICHEM mode. This allows the user to simply change the Output Format in a control file, and retain the same grid for inter-model comparisons.

When run in any of the modes supporting AERMOD, MMIF uses a 1×1 output grid (a single point). Scalars such as temperature are averages over the grid cell, and winds have been de-staggered to represent the entire grid cell. Of course, each 1×1 grid still has cell corners (the corners of the single cell) which are printed to the screen during a MMIF run.

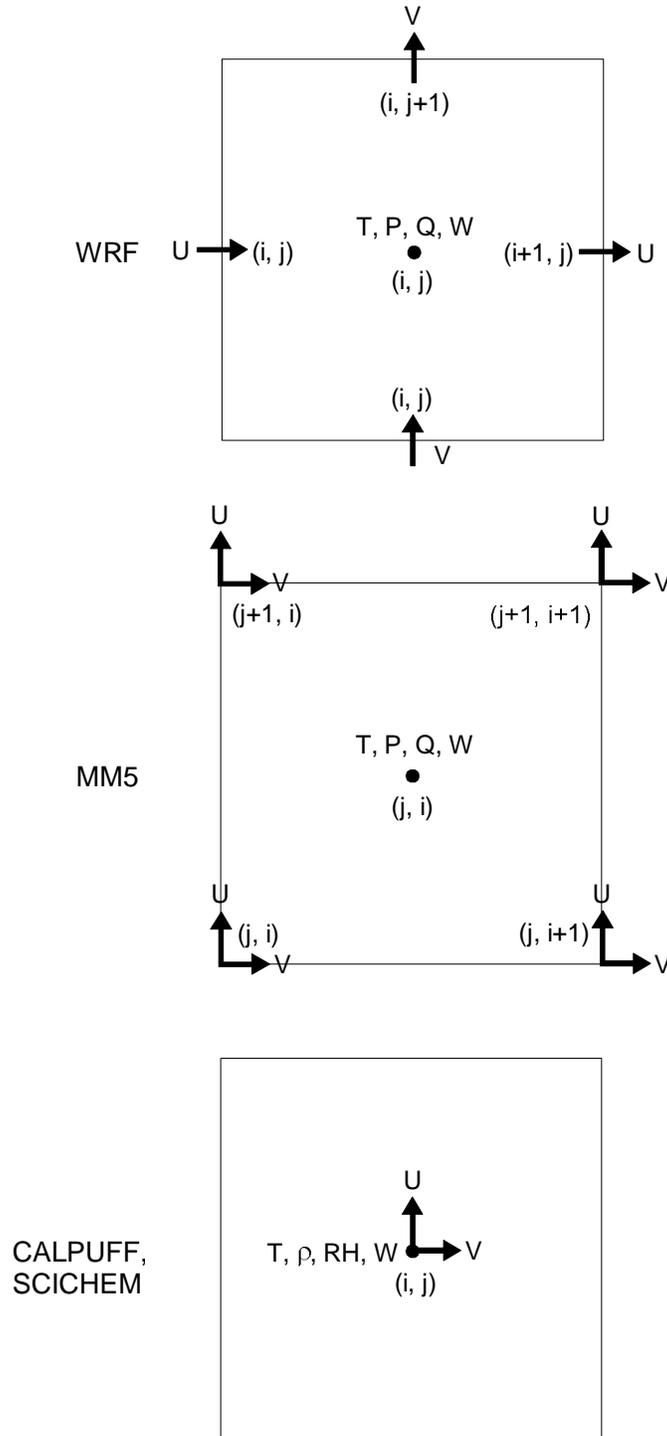


Figure 2-1. Horizontal grid cell arrangements for wind (U, V, W) and state (T, P, Q, ρ , RH) variables among three models addressed by MMI F. SCICHEM uses the same arrangement as CALPUFF.

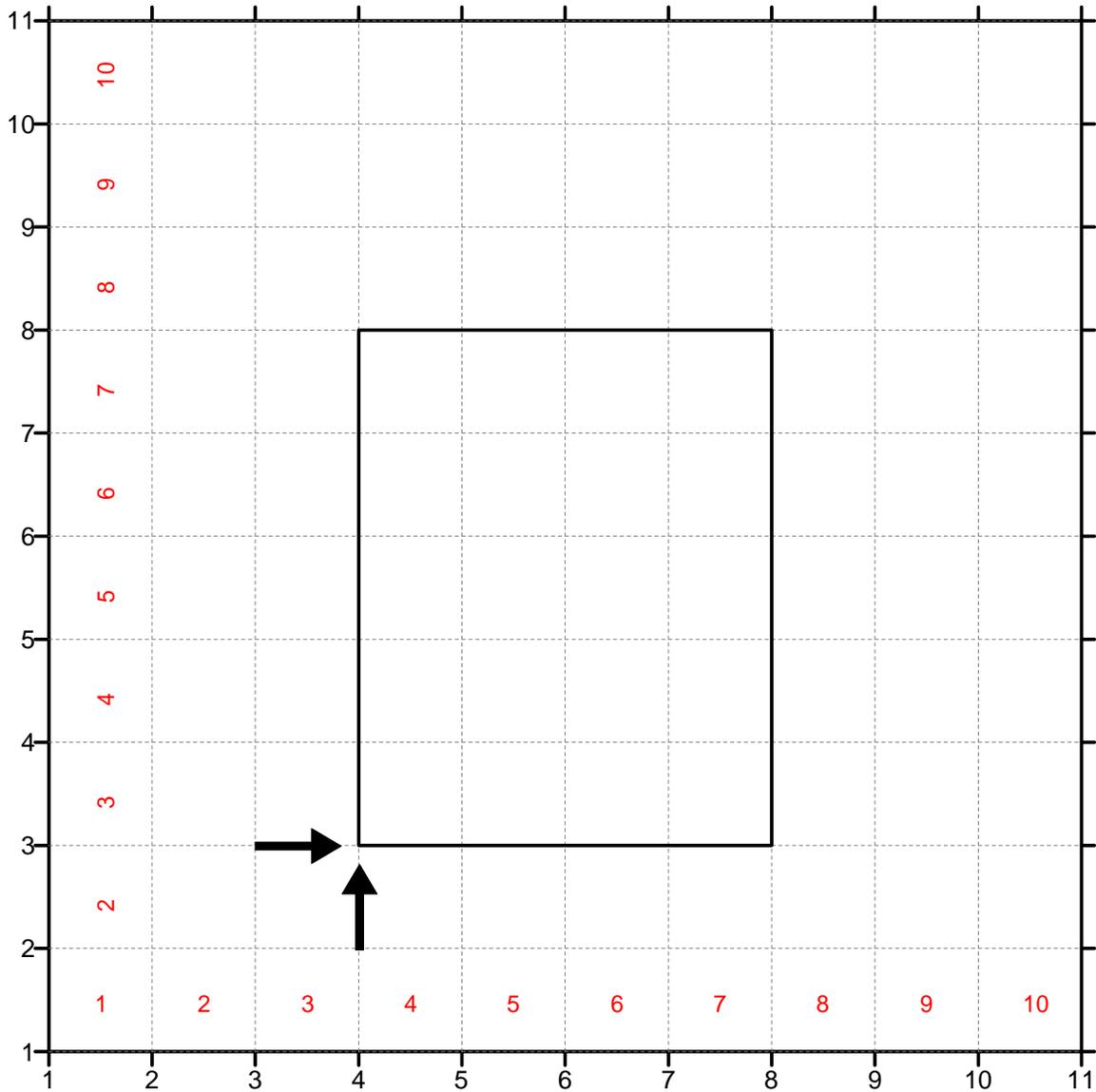


Figure 2-2. An example illustration of a full MM5/WRF grid (outer grid with 11x11 grid points and 10x10 grid cells) and a CALPUFF sub-domain (inner grid with 4x5 grid cells). The CALPUFF sub-domain is defined as ranging from grid cells 4 through 7 in the x-direction, and grid cells 3 through 7 in the y-direction. The arrows point to the lower-left corner point of the CALPUFF grid, which defines its reference coordinate location. SCICHEM carries all variables, including geo-locating variables, at grid cell centers.

2.5.2 Vertical Grid Structure

The MM5 and WRF meteorological models' vertical grid structure is based on a normalized pressure system, generally referred to as "sigma". While the sigma layer structure is constant, the corresponding layer heights vary in space (MM5 and

WRF) and time (WRF) according to underlying topographic elevation and surface pressure. Sigma layers expand and contract in height across valleys and ridges, respectively. In MM5, the thickness in meters of sigma layers is time-invariant; but in WRF the thickness can vary with time. Conversely, the layer structure in CALPUFF is defined to be a single static vertical height profile above terrain (referred to as “ZFACE” in CALPUFF).

When processing MM5 data, MMIF defines the input MM5 layer heights from a conversion of the sigma coordinate assuming a 1000 mb surface pressure, and using the standard temperature/pressure lapse rates provided in the MM5 output file. In the case of WRF data, the sigma coordinate from each output time-stamp is converted to a spatially-variant height coordinate for all grid columns across the input domain.

In the vertical, all CALPUFF variables are carried at layer centers (between the vertical levels defined by the ZFACE variable) except for vertical velocity, which is carried at the layer interfaces (i.e. at the ZFACE levels). AERMOD assumes all variables are point measurements, which corresponds with grid-cell average values. All SCICHEM variables, including vertical velocity, are grid-cell averages and assigned to the center of each grid cell. For SCICHEM output, the vertical velocity is interpolated from cell faces to cell centers at the time of output. The exception is the vertical velocity of the highest output layer, which is set to zero (a SCICHEM requirement) to avoid advecting puffs out of the top of the modeling domain.

CALPUFF’s vertical levels are specified relative the ground – it uses a “terrain-following coordinate system”, where each level is at a constant-height above ground. The top of the modeling domain is therefore not flat, but follows the undulations of the terrain below. AERMOD’s vertical coordinates are also specified relative to the ground.

SCICHEM has a fundamentally different vertical grid structure than CALPUFF and AERMOD. SCICHEM uses a transformed vertical coordinate system that is described in Section 10.3.2 of the SCICHEM Technical Documentation (Santos and Sykes, 2000) or Section 10.3.2 of the SCIPUFF Technical Documentation (Sykes et al., 2004). The top of the modeling domain D in SCICHEM is specified relative to mean sea level (not ground level), and is a constant over the entire domain (“flat”). The transformed vertical coordinate $SZ(k)$ can be written as

$$SZ(k) = \frac{Z_{AGL}(i,j,k)}{1 - elev(i,j)/D}$$

Where $elev(i,j)$ is the topographic elevation at a grid point, and $Z_{AGL}(i,j,k)$ is the height above ground level of the k^{th} grid cell center above that point. If the grid point has an elevation of zero (i.e. sea level) then the transformed and un-transformed vertical coordinates are identical. All grid points with non-zero elevation will have shallower layers than the transformed layers $SZ(k)$.

It is recommended by the SCICHEM authors that D be set to $SZ(k_{max})$, and be at least twice the maximum elevation within the modeling domain. MMIF will stop with an error if the user has selected output layers that do not meet this criteria, unless the "--force" option is given.

The vertical layers $SZ(k)$ are specified by the user in the MMIF control file. These are invariant, and are contained within the header of the MEDOC output file (described in Section 13.2.2 of the SCIPUFF Technical Documentation [Sykes et al., 2004], or Section 4.5.3 of the SCICHEM-2015 User's Guide [Chowdhury et al., 2015]). MMIF accomplishes the transformation by first inverting the above equation to find the un-transformed levels $Z_{AGL}(i,j,k)$ that will yield the correct transformed levels $SZ(k)$, then interpolating the MM5/WRF profile above each grid point (i,j) to the $Z_{AGL}(i,j,k)$ levels.

It is not possible to use aggregation to transform MM5/WRF levels to the $SZ(k)$ levels. MMIF issues an error and stops if the user requests both output for SCICHEM and aggregation.

2.5.3 Vertical Aggregation

When aggregating levels, the first time-stamp's spatially-varying heights are averaged over the output sub-domain to define a single ZFACE height profile. The horizontal wind, temperature, humidity, pressure, and cloud liquid water variables are aggregated to a subset of layers defined by the user. The aggregation is performed on a mass-weighted basis (using the MM5 and WRF pressure coordinates). Vertical velocity, however, is set to the value of the MM5/WRF vertical velocity at corresponding levels with no weighting.

[Figure 2-3](#) depicts aggregation schematically.

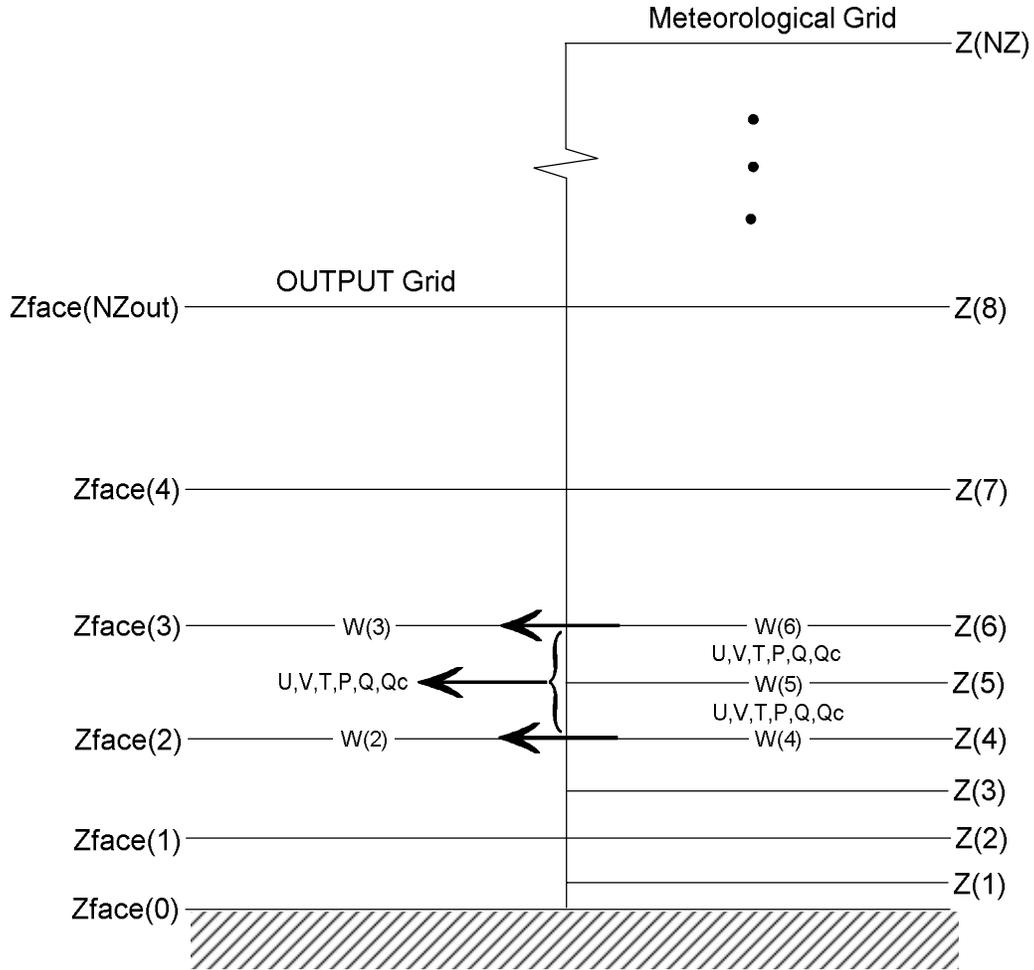


Figure 2-3. Vertical layer structure arrangements for the meteorological model (right) and aggregated output (left). OUTPUT vertical velocities (W) are located at layer interfaces and are assigned from the meteorological model; SCICHEM vertical velocities are interpolated to grid cell centers; output horizontal winds (U, V), temperature (T), humidity (Q), and cloud liquid water (Qc) are averaged from the meteorological layers (U, V, T, Q, Qc) in a mass-weighted manner using pressure (P).

2.5.3.1 Aggregation and CALPUFF's Lowest Layer

CALPUFF assumes that the first layer depth is 20 m, resulting in a 10 m layer midpoint that is consistent with the height at which most meteorological measurement probes are located. The meteorological model grid structure will usually not match this constraint.

If the user requests interpolation and specifies the top of the lowest layer be 20m, then no special treatment is necessary. MMIF will halt execution and print an error if the user requests CALPUFF output and does not request $ZFACE(1) = 20m$.

If the user requests aggregation, MMIF generates a 20 m first layer, using 10 m winds output by the meteorological model (if available, or by diagnosing them using surface similarity theory). This will usually lead to inconsistencies in the definition of the depth of layer 2 between CALPUFF and the meteorological models. Four different cases have been identified in which layer 2 winds and temperature and level 1 vertical velocity must be re-diagnosed using interpolation:

1. $ZFACE(1) < 19\text{ m}, ZFACE(2) < 39\text{ m}$
 Layer indices 3 through NZ_{out} are shifted down, thereby removing one layer;
 $ZFACE(1)$ is reset to 20 m;
 Layer 2 winds and temperature are interpolated from original layers 2 and 3;
 Level 1 vertical velocity is interpolated from original levels 1 and 2.
2. $ZFACE(1) < 19\text{ m}, ZFACE(2) \geq 39\text{ m}$
 Layer indices 2 through NZ_{out} remain the same;
 $ZFACE(1)$ is reset to 20 m;
 Layer 2 winds and temperature are interpolated from layers 2 and 3;
 Level 1 vertical velocity is interpolated from levels 1 and 2.
3. $19\text{ m} \leq ZFACE(1) < 39\text{ m}$
 Layer indices 2 through NZ_{out} remain the same;
 $ZFACE(1)$ is reset to 20 m;
 Layer 2 winds and temperature are interpolated from layers 1 and 2;
 Level 1 vertical velocity is interpolated between level 1 and the ground.
4. $ZFACE(1) \geq 39\text{ m}$
 Layer indices 1 through NZ_{out} are shifted up, thereby adding one layer;
 $ZFACE(1)$ is reset to 20 m;
 Layer 2 winds and temperature are interpolated from original layers 1 and 2;
 Level 1 vertical velocity is interpolated between original level 1 and the ground.

[Figure 2-4](#) depicts these cases schematically.

Note that the case where *two* MM5/WRF layers are within the first CALMET layer (i.e. $ZFACE(2) < 19\text{ m}$) has not been accounted for in the above scheme. Nor has the possibility that more than one MM5/WRF layer exist between 20 m and 39 m. For these cases it is suggested to request interpolation rather than aggregation.

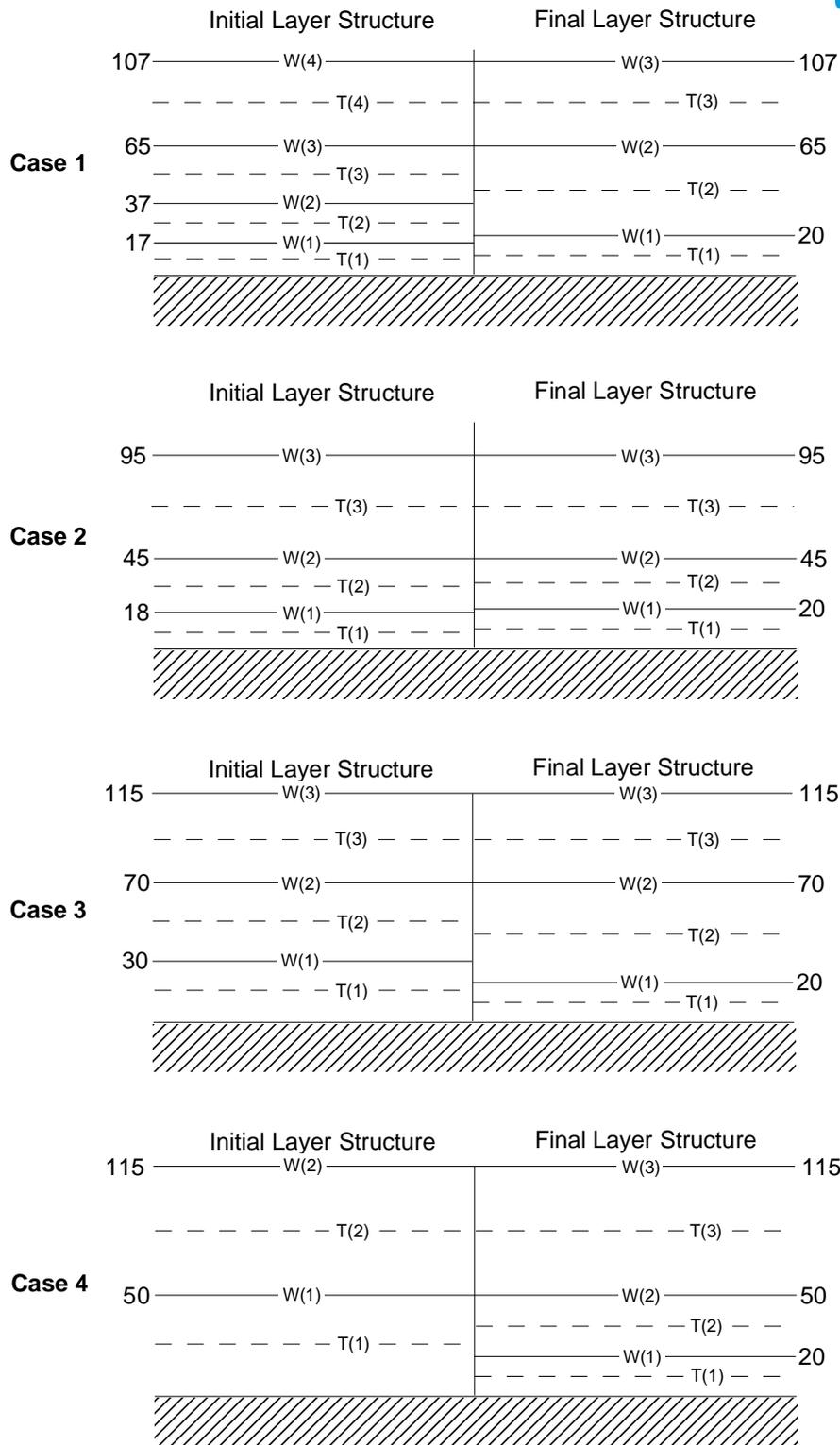


Figure 2-4. Schematic examples of the approach to introduce a 20-m deep first layer for CALPUFF. The four cases are described in the text: [1] $ZFACE(1) < 19$ m, $ZFACE(2) < 39$ m; [2] $ZFACE(1) < 19$ m, $ZFACE(2) \geq 39$ m; [3] $19 \text{ m} \leq ZFACE(1) < 39$ m; and [4] $ZFACE(1) \geq 39$ m. Locations of vertical velocity (W) and temperature (T) are shown in each case.

2.5.3.2 Aggregation and Monthly MMIF Runs for CALPUFF

It is convenient to run MMIF in monthly segments, to avoid creating very large CALMET.DAT files. A CALPUFF run can use NMETDAT = 13 to perform an annual run (12 months, plus a few days of spin-up to populate the grid with puffs). Annual CALPUFF runs are very convenient for calculating annual averages, or multi-year averages of the annual 98th percentile of some parameter.

Because the ZFACE values assigned by MMIF when aggregating are output-domain-averaged using the first output time-stamp, they may vary from one month-long run to the next. This is most true when processing WRF output, but even MM5 output may vary its base state with time. CALPUFF assumes that all ZFACE values remain constant over a series of CALMET.DAT files.

The user is cautioned to either run monthly CALPUFF runs, or switch from aggregation to interpolation to avoid this problem.

2.5.4 Vertical Interpolation

It is perhaps common to think about data such as profiles of temperature or wind speed in terms of a continuous curve, like the thicker line in [Figure 2-5](#). When considering MM5/WRF output, these profiles are quantized (averages over a layer, defined at the grid cell's midpoints) and we often think of them as straight line segments like the thinner line in [Figure 2-5](#).

To convert pressure levels to height-above-ground (AGL) levels, most WRF post-processors (UPP, ARWpost, RIP4, MCIP, CDO) use linear interpolation with pressure (or a pressure derivative) as the vertical coordinate. However, all of these surveyed post-processors only consider the input layer midpoints immediately above and below the new layer's midpoint, and ignore any intervening layers.

If the output vertical grid levels are much sparser than the input levels, the situation shown by the red line in [Figure 2-5](#) below could occur. This obviously under-estimates the local maximum and would over-estimate a local minimum, resulting in smoother vertical profiles that do not accurately reproduce the character of the profile.

It is common to use far fewer layers in a CALPUFF simulation than in the underlying MM5/WRF simulation. Typical values might be ~35 vertical layers from the surface to 100 millibar in MM5/WRF, and 10 vertical layers from the surface to 4000 meters in CALPUFF. If this down-scaling of the vertical resolution was performed using

typical “nearest neighbor” linear interpolation, then local maxima and minimum in e.g. wind speed would not be preserved.

Because the MM5/WRF data are averages over a layer (defined using a vertical coordinate $P(z)/P_0$, often called sigma) perhaps an easier way to think of the situation is represented by the boxes in [Figure 2-6](#).

MMIF’s interpolation scheme conserves the total area of the boxes. Here the X-axis is the value in question (e.g. temperature, humidity, wind speed, etc.) and the Y-axis is sigma. Given output layers shown by the red lines in [Figure 2-6](#), each output (red) layer contains the area of the portions of the input (blue) boxes it spans. The area of the red box is the same as the red-striped portion of the blue boxes.

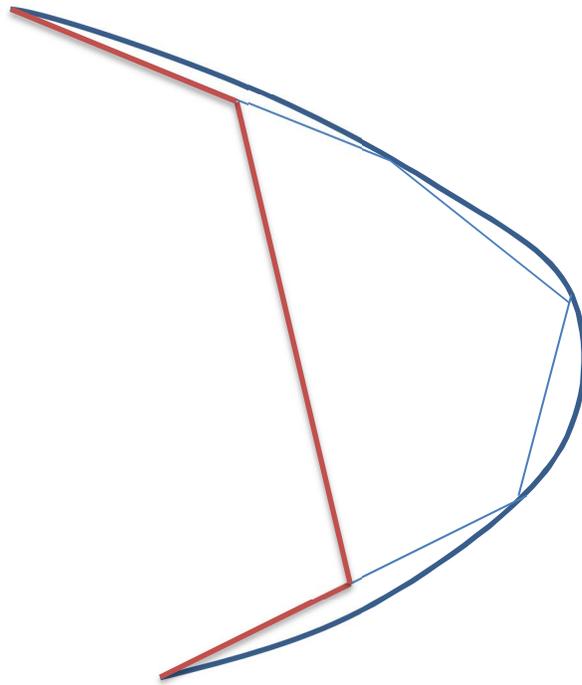


Figure 2-5. Cartoon vertical profile of some parameter such a temperature or wind speed. The Y-axis is height, and the X-axis is the value of the parameter. The thick blue line represents the continuous profile, the thin blue line represents the quantized MM5/WRF values, and the red line represents a profile resulting from traditional “nearest neighbor” linear interpolation with sparse output layers.

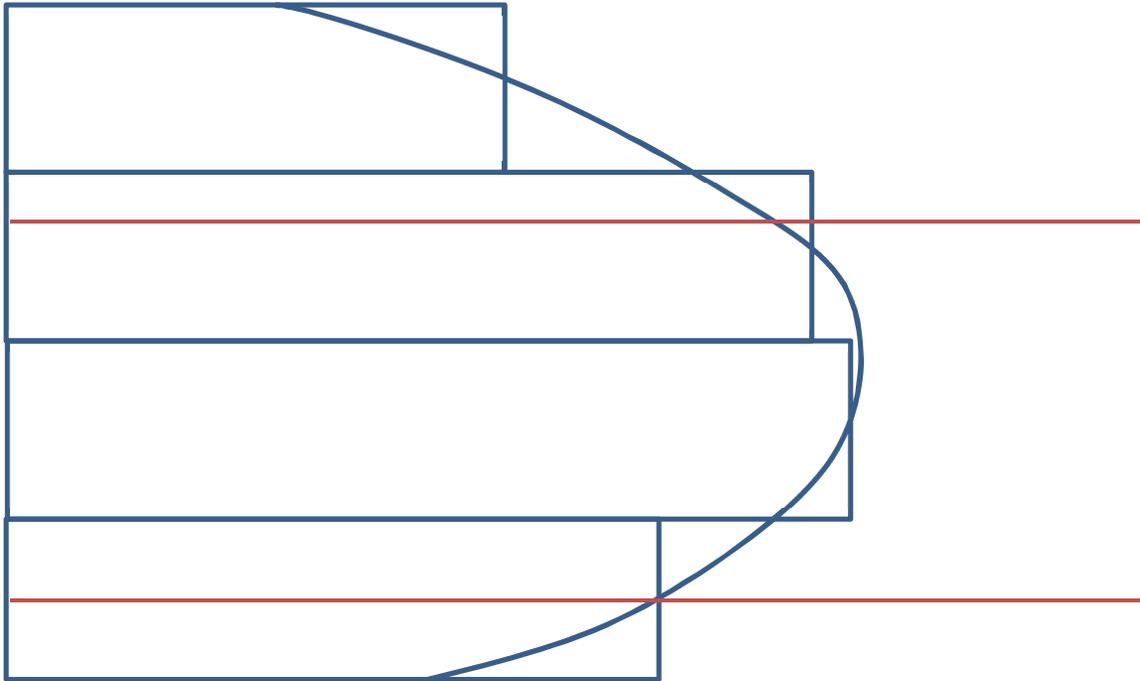


Figure 2-6. Cartoon vertical profile of some parameter such a temperature or wind speed. The Y-axis is sigma (P/P_0), and the X-axis is the value of the parameter. The thick blue line represents the continuous profile, blue boxes represent the quantized MM5/WRF values, and the red lines represent an output layer.

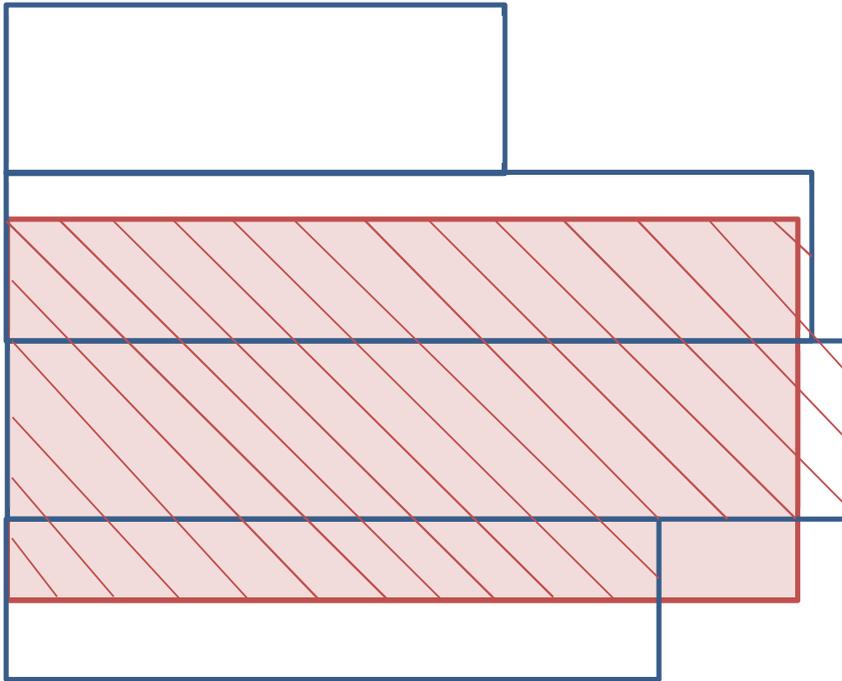


Figure 2-6. Cartoon vertical profile of some parameter such a temperature or wind speed. The Y-axis is sigma (P/P_0), and the X-axis is the value of the parameter. The blue boxes represent the quantized MM5/WRF values, and the red box represents an output layer. The area of the red (output) box is equal to the area of the stippled portion of the blue (input) boxes.

Because the vertical coordinate is sigma ($\sigma = P/P_0$), this is effectively a pressure (mass) weighted averaging scheme. The height above ground for any sigma level is known (e.g. [PH+PHB]/g in WRF).

The interpolation scheme steps are as follows:

- Convert the desired output level heights $Z_{out}(i)$ (top and bottom, i.e. ZFACE levels) to sigma values $\sigma_{out}(j)$ using the current time-step's values of PH.
- Loop over output levels $\sigma_{out}(i)$, where i ranges from 1 (surface) to NZ (top):
 - Find the closest input level $\sigma_{in}(j)$ above $\sigma_{out}(i)$
 - Add $[\sigma_{out}(i) - \sigma_{in}(j)] * S_{in}(j)$ to the sum, where S is the scalar values in question
 - Step up in j until $\sigma_{out}(i+1) > \sigma_{in}(j+k)$, adding $\Delta\sigma_{in} * S_{in}$ for each full level to the sum
 - Add $[\sigma_{in}(j+k+1) - \sigma_{out}(i+1)] * S_{in}(j+k)$ to the sum

$$o \quad S_{out}(i) = \text{Sum} / [\sigma_{out}(i) - \sigma_{out}(i+1)]$$

If an output layer is fully contained within an MM5/WRF layer, then $S_{out} = S_{in}$.

If an output layer is fully contained within the lowest MM5/WRF layer, then the Monin-Obukhov similarity profiles (the “log layer” profiles) are evaluated at the output layer’s mid-point. To be absolutely correct, each output value should be averaged over the layer rather than evaluated at its middle. However, performing a numerical integral of the profile over a layer will likely produce as large an error as simply evaluating the profile at the mid-point, and will add to program runtimes.

Vertical wind speed is defined in MM5, WRF and CALPUFF at the levels (cell interfaces); and in SCICHEM at the layer centers. The interpolation subroutine performs simple (pressure-weighted) vertical interpolation between adjacent levels to support all the downstream dispersion models, rather than the layer-weighted output described above. For SCICHEM, the vertical wind speed is interpolated during output to grid cell centers.

2.5.4.1 “Pass-Through” Vertical Interpolation

If the user desires a “pass-through” of the MM5/WRF data that minimizes the changes that interpolation can sometimes incur, yet possibly also wishes to reduce the total number of layers for the downstream dispersion model, the following methodology is suggested:

1. Perform a short (1 timestamp of output) MMIF run using aggregation (use “K”), including all MM5/WRF layers.
2. Inspect the messages written to the screen during the run, noting the input and output ZFACE values.
3. Possibly repeat the short MMIF run using aggregation, reducing the total number of layers by eliminating middle layers.
4. When a compromise between a reduced number of output layers and reasonable vertical spacing is achieved, copy the resulting output ZFACE values printed to the screen to the control file, and switch to interpolation (use “TOP”).
5. Perform the production MMIF runs using interpolation.

2.6 Surface Characteristics

MMIF attempts to read two-dimensional surface fields of 10m winds, roughness length, albedo, Monin-Obukhov length, and leaf area index (LAI) from the MM5/WRF data, and generates these fields if they are not found in the files.

MMIF requires that MM5 was run using the 24-category U.S. Geological Survey (USGS) landuse/landcover dataset. For WRF, MMIF support the following landuse datasets:

- USGS 24- and 33-category datasets
- NLCD 40- and 50-category datasets
- MODIS 20- and 33-category dataset
- MODIFIED_IGBP_MODIS_NOAH dataset

If any other dataset is diagnosed from the meteorological output files, MMIF will stop with an error⁽⁷⁾.

Values of roughness length, Monin-Obukhov length, LAI, etc. are included by default in some, but not all, of the versions of MM5 and WRF that MMIF supports. Starting with WRF version 3.2, the user can request that certain fields be included in its output files via the `iofields_filename` parameter in the `&time_control` namelist. This specifies the filename (one for each domain) of a text file that must be present in the run-time directory during a WRF run. If the file contains the line

```
+ :h : 0 : ZNT , RMOL , LANDUSEF
```

then the time-varying roughness length (ZNT), the inverse of the Monin-Obukhov length (RMOL) and the land-use fraction by category (LANDUSEF) will be included in WRF's output file. MM5 can also be forced to include similar fields.

If MMIF does not find any of the above fields, it calculates them using the techniques found in this section and the following section.

The roughness length, albedo, and LAI can be diagnosed from the landuse codes according to the methodology developed by U.S. Forest Service (USFS) for use in the BlueSky/RAINS modeling system (MM52GEO program). This methodology assigns seasonally dependent (winter/summer) albedo, Bowen ratio, soil heat flux, anthropogenic heat flux, LAI and surface roughness to each of the landuse categories. As an example, the 24-category USGS values are shown in [Table 2-1](#). The summer season is defined to run from Julian date 105 to Julian date 287 (April

⁽⁷⁾ See the '--force' option to prevent stopping with this error.

15 to October 14, respectively, for non-leap years). MMIF only uses the albedo, LAI, and surface roughness, and always extracts or calculates the soil heat flux and Bowen ratio.

The roughness length and LAI can be diagnosed from the other supported landuse codes according to a methodology provided by the FWS, using the look-up tables found in MCIP version 4.1 and WRF version 3.5 (albedo and roughness length are given in the file LANDUSE.TBL). All versions of WRF that are recent enough to use the NCLD landuse scheme are also recent enough that albedo is written, by default, to the WRF output files.

MMIF always calculates the Bowen ratio, because both MM5 and WRF always include sensible and latent heat fluxes in their output. The Bowen ratio is used only by AERMOD, and limits have been placed (following MM5AERMOD) on its values.

Table 2-1. Surface characteristics assigned to USGS 24-category landuse/landcover in the USFS BlueSky/RAINS MM52GEO program.

Landuse	Albedo (%)		Bowen Ratio		Soil Heat Flux (W/m ²)		LAI		Surface Roughness (cm)	
	Sum.	Win.	Sum.	Win.	Sum.	Win.	Sum.	Win.	Sum.	Win.
Urban	18	18	1.5	1.5	0.25	0.25	1.5	1	50	50
Dryland crop/pasture	17	23	1.0	1.0	0.15	0.15	0	0	15	5
Irrigated crop/pasture	18	23	1.0	1.0	0.15	0.15	1	0	15	5
Mixed crop/pasture	18	23	1.0	1.0	0.15	0.15	0.5	0	15	5
Crop/grass mosaic	18	23	1.0	1.0	0.15	0.15	0	0	14	5
Crop/woodland mosaic	16	20	1.0	1.0	0.15	0.15	3	0.5	20	20
Grassland	19	23	1.0	1.0	0.15	0.15	0	0	12	10
Shrubland	22	25	1.0	1.0	0.15	0.15	2	0.5	10	10
Mix shrub/grass	20	24	1.0	1.0	0.15	0.15	1	0.25	11	10
Savanna	20	20	1.0	1.0	0.15	0.15	2	0.5	15	15
Deciduous broadleaf	16	17	1.0	1.0	0.15	0.15	6	0	50	50
Deciduous needle	14	15	1.0	1.0	0.15	0.15	5	0	50	50
Evergreen broadleaf	12	12	1.0	1.0	0.15	0.15	5	5	50	50
Evergreen needle	12	12	1.0	1.0	0.15	0.15	8	7	50	50
Mixed forest	13	14	0.5	0.5	0.25	0.25	4	2	50	50
Water bodies	8	8	0	0	1.00	1.00	0	0	1	1
Herbaceous wetland	14	14	0.5	0.5	0.25	0.25	2	1	20	20
Wooden wetland	14	14	0.5	0.5	0.25	0.25	5	3	40	40

Barren sparsely vegetated	25	25	1.0	1.0	0.15	0.15	0	0	10	10
Herbaceous tundra	15	60	0.5	0.5	0.15	0.15	1	0	10	10
Wooded tundra	15	50	0.5	0.5	0.15	0.15	1	0	30	30
Mixed tundra	15	55	0.5	0.5	0.15	0.15	1	0	15	15
Barren tundra	25	70	0.5	0.5	0.15	0.15	0	0	10	5
Snow/ice	55	70	0.5	0.5	0.15	0.15	0	0	5	5

2.7 PBL Parameters

MMIF reads certain PBL and surface similarity parameters from the meteorological model output files, including PBL depth, friction velocity (U^*), ground temperature, and 10-meter wind components (if available). Using these data as well as surface roughness (z_0) and lowest layer gridded winds, temperature, pressure, and humidity, MMIF calculates two-dimensional surface fields of air density, 10-meter temperature, Monin-Obukhov length (L) and convective velocity scale (W^*). MMIF will also diagnose 10-meter wind components if not available from the meteorological model files. Finally, the program will optionally re-diagnose PBL depths if specified by a user input flag. In MMIF v3.3 and newer, setting the CALSCI_MIXHT keyword to 'WRF' indicates that PBL heights from WRF will be used in CALPUFF and SCICHEM modes, and setting it to 'MMIF' requests that MMIF-recalculated PBL heights be written.

The calculation of surface similarity variables L and W^* (if not found in the MM5/WRF file) are determined from the Richardson-number based methodology of Louis (1979). These values together with U^* and z_0 are then applied in a standard surface-layer scaling algorithm to derive 10-meter wind and temperature from input layer 1 winds and temperature in each grid cell. Note that the 10-meter winds are output for AERMOD and CALPUFF; but the 10-meter temperatures are output only for CALPUFF. SCICHEM does not use any of the 2-meter or 10-meter parameters.

The optional re-diagnosis of PBL depth is determined from the bulk Richardson approach of Vogelezang and Holtslag (1996) using the Louis (1979) surface

parameters determined above. The re-diagnosis also incorporates the methodology of Gryning and Batchvarova (2003), which varies a critical Richardson number for over-water PBL heights. Experiments have shown that a critical Richardson of 0.05 for over-water PBL heights yielded better agreement when using the Vogelezang and Holtslag bulk Richardson method.

In MM5/WRF output files, the PBL depth values have no finer resolution than the vertical grid structure – the PBL depth is “rounded” up or down to the nearest vertical grid cell center. Adjacent grid points can therefore have PBL depths that vary abruptly, non-smoothly. MMIF’s method of re-diagnosis of PBL depth refines the depth to within 1/20th of the local MM5/WRF cell height, or within 10m, whichever is less.

In non-AERMOD modes, the PBL depth (whether passed-through or re-diagnosed) is limited to be between the heights of the lowest and highest MMIF output vertical grid cell centers. This prevents MMIF from writing a set of inconsistent files, where the maximum PBL depth exceeds the modeling domain. It is also consistent with the PBL depths found in MM5/WRF output files, which are discretized to the grid’s vertical mid-points. For AERMOD mode, the PBL depth is limited to 4000m, following AERMET (see `MP2.INC` in the AERMET source code package). When running MMIF v3.3 and newer in AERMET mode, the source of mixing height values can be specified (WRF, MMIF, AERMET) through the `AER_MIXHT` keyword.

2.8 Cloud Cover

Cloud cover is required in AERMET’s surface (SFC) file format, and in SCICHEM’s MEDOC file format. However, it is generally not available in MM5 or WRF files, and must be estimated or derived.

MMIF v2.2 used the same method as EPA’s MM5AERMOD program. It used a method following Randall (1994) and Hong et al. (1998) where both grid-scale and convective (sub-grid-scale) cloud amounts are estimated and the maximum over all vertical model layers taken as the 2-D value. If the sum of the MM5/WRF-reported cloud liquid water mixing ratio and ice mixing ratio are less than 1×10^{-12} , then the cloud fraction for that level is taken to be zero. If the relative humidity of that level is 100%, then the cloud fraction for that level is taken to be 100%. Values between those extremes use a modified version of Randall (1994) and Zhao (1995) to estimate the cloud cover.

Starting with MMIF v2.3, the cloud cover was calculated following Angevine et al. (2012), which is essentially the same as the method used by COAMPS. The fractional cloud fraction is given by

$$F_{cloud}(k) = \frac{RH(k) - RH_{crit}}{100 - RH_{crit}}$$

Where RH is the relative humidity in percent, and RH_{crit} is 80% over water and 70% over land. The total fractional cloud cover is taken as the maximum at any level k .

The use of these types of parameterizations was because before WRF 3.6, the CLDFRA variable in WRF output was either 0 or 1, with no fractional values. In WRF 3.6, many improvements were made to this output, including a complete accounting of sub-grid-scale clouds from any of the Cumulus parameterizations.

Starting with MMIF v3.4.1, the user has a choice for the source of the cloud cover: WRF, ANGEVINE, or RANDALL. Specifying WRF (or the synonym CLDFRA) will use WRF's output, specifying ANGEVINE (or the synonym COAMPS) will retain the behavior of MMIF v2.2 to 3.4, and specifying RANDALL (or the synonym MM5AERMOD) will revert to the behavior of MMIF before version 2.2.

2.9 Diagnosis of Stability Class

The Phase II guidance (EPA, 1998) from the Interagency Workgroup on Air Quality Modeling (IWAQM) recommends the use of PG stability class in CALPUFF long range transport analyses. The EPA Model Clearinghouse reaffirmed this approach in early 2006 (EPA, 2006). Hence, the FLMS typically will not accept turbulence-based dispersion modeling for long range transport analyses at Class I receptor locations. Furthermore, a number of algorithms inside CALPUFF (such as puff splitting and chemistry) are dependent upon PG stability class even if turbulence-based dispersion is selected. It is therefore necessary that MMIF constructs gridded PG fields from available MM5/WRF data fields. Gridded PG fields are not required to support AERMOD or SCICHEM.

CALMET calculates PG stability class based on the Turner (1970) method. CALMET supports two different methods for calculating PG class:

- Observation-based (ceiling height and cloud cover from standard surface data);

- “NOOBS” method (ceiling height and cloud cover derived from MM5 hydrometeors)

Note that the “NOOBS” method introduced by Robe (Earth Tech, 2001) constructs PG from MM5-estimated ceiling heights and diagnostic cloud cover.

Two options are available in MMIF for calculating PG stability class; they both differ from the CALMET approach in that they do not rely on the diagnosis of cloud cover and ceiling height:

- SRDT method
 - PG is based upon the wind speed, solar radiation, and the “Delta-T” (SRDT) method published in Supplement C to the Guideline on Air Quality Models (EPA, 1993).
 - Daytime stability is derived from the Turner method using 10-meter wind speed and solar radiation to estimate an insolation class. If the 10-meter wind speed is not found in the MM5/WRF file, then the lowest level’s winds speed is used. Nighttime stability is derived from the sign of the temperature difference between 10 and 2 meter temperatures (or lowest MM5/WRF level and 2 m temperatures).
 - The code was implemented directly from the Meteorological Processor for Regulatory Models (MPRM).
- Golder (1972) method
 - PG is based upon relationships among Monin-Obukhov lengths and surface roughness.
 - The code was implemented from the AERMOD LTOPG subroutine.

[Table 2-2](#) displays PG stability in the SRDT method according to wind speed, insolation class, and temperature difference. Insolation is taken from the gridded downward solar radiation output by the meteorological models. The advantage of this approach is that the modeled solar radiation includes attenuation from modeled cloud cover, eliminating the need to indirectly derive cloud cover from gridded

hydrometeor data. The nighttime stability (Delta-T method) is activated when insolation equals zero.

Note that the SRDT method will not produce a stable PG category during the day, or an unstable PG category during the night, regardless of the actual vertical potential temperature gradient or the sign of the surface heat fluxes.

Table 2-2. PG stability class according to surface wind speed (S , m/s), daytime solar radiation (R , W/m^2), and nighttime vertical temperature gradient (DeltaT) in the SRDT algorithm. Stability ranges from very un-stable (1) to very stable (6).

		Wind Speed Category					
		$0 \geq S > 2$	$2 \geq S > 2.5$	$2.5 \geq S > 3$	$3 \geq S > 5$	$5 \geq S > 6$	$S \geq 6$
Day	$R \geq 925$	1	1	1	2	3	3
	$675 \geq R > 925$	1	2	2	2	3	4
	$175 \geq R > 675$	2	3	3	3	4	4
	$0 \geq R > 175$	4	4	4	4	4	4
Night	$\Delta T < 0$	5	4	4	4	4	4
	$\Delta T \geq 0$	6	5	4	4	4	4

3. CODE STRUCTURE AND COMPI LATION

MMIF is written in Fortran90 (F90) and consists of a main driving program and several subroutines and F90 modules that are (mostly) separated into individual files. The program is highly modular to allow for easy addition or substitution of alternate routines in the future. All code is well documented to include such information about the tasks performed and a history of modifications. All of the program's global data structures are dynamically allocated during program startup according to the dimensions of the meteorological models' grid definition. This alleviates the need to customize and re-compile the program for a particular application.

The program code is arranged in the following file structure:

<code>mmif.f90</code>	Main driving program
<code>aggregate.f90</code>	Vertical aggregation routines
<code>avg_zface.f90</code>	Calculate domain-average ZFACE values of MM5/WRF levels
<code>cloud_cover.f90</code>	Derive cloud cover
<code>functions.f90</code>	F90 module of useful functions, mostly thermodynamic
<code>interpolate.f90</code>	Vertical interpolation routines
<code>landuse.f90</code>	Surface characteristic look-up tables by landuse type
<code>met_fields.f90</code>	F90 module defining and allocating global variable arrays
<code>module_llxy.f90</code>	F90 module for calculating projections (from WRF v3.8.1)
<code>output_aercoare.f90</code>	Output routines to write AERCOARE inputs
<code>output_aermet.f90</code>	Output routines to mimic AERMET outputs
<code>output_calmet.f90</code>	Output routines to mimic CALMET outputs
<code>output_onsite.f90</code>	Output routines to write AERMET inputs
<code>output_scichem.f90</code>	Output routines to write MEDOC files for SCICHEM
<code>parse_control.f90</code>	F90 module to read the command line and parse the control file
<code>pasquill_gifford.f90</code>	SRDT and GOLDER Pasquill-Gifford stability class calculations
<code>pbl_height.f90</code>	PBL height re-diagnostic routine
<code>read_mm5.f90</code>	Reading routine for MM5 output data
<code>read_wrf.f90</code>	Reading routine for WRF output data
<code>sfc_layer.f90</code>	Calculate Monin-Obukhov length, and 10m winds & temperature
<code>timesubs.f90</code>	Group of date/time manipulation routines
<code>wrf_netcdf.f90</code>	F90 module containing NetCDF I/O routines

It is possible to allow a single code to support multiple platforms using conditional compilation and "-D" macros, as MMIF v1.0 did. The MMIF v1.0 file `command_line.f90` required passing through the FORTRAN pre-processor, but all three supported compilers (Portland Group, Intel, and GNU FORTRAN) all use exactly the same format, and the conditional compiling was removed from MMIF 2.0 for the sake of simplicity.

3.1 Compiling MMIF on Windows

The full distribution of MMIF includes a pre-compiled version of `MMIF.EXE`. This section describes how to re-compile MMIF, if necessary.

Commercial compilers such as Portland Group (pgf90) or the Intel FORTRAN compiler (ifort) typically produce faster code (shorter run-times) compared to free compilers – but only when calculations (CPU usage) rather than storage throughput (disk I/O) is the limiting factor. In the case of MMIF, relatively few calculations are performed, and disk I/O is most often the limiting factor. MMIF is limited by the speed of the underlying disk systems, so using a commercial compiler will most likely not yield significantly shorter run-times.

Instead, MMIF was compiled for Windows using open-source software: GNU FORTRAN (gfortran) supplied by the Minimalist GNU for Windows (MinGW) compiler suite and the Minimal System (MSYS) development environment. MinGW was used by UNC to compile the Windows version of CALMET2NETCDF, developed as a separate Work Assignment under the same Contract as MMIF v2.0 development.

Note that as of fall 2013, neither the MinGW nor the MinGW-64 (64-bit) versions of the gfortran compiler properly define `off_t` as a “long long” integer. This causes the large file support in NetCDF to fail to compile, which means the Windows version of MMIF cannot currently read WRF files larger than 2 gigabytes (GB). The Windows version of the commercial Portland Group compilers relies on Cygwin for its operational shell, and has a similar problem (64-bit compiler but 32-bit linker). Attempts to use NetCDF 4.3 and NetCDF-Fortran 4.2 to support large WRF files failed, for the same reason.

Compiling the source code in Windows requires three steps:

1. Download and install MinGW/MSYS
2. Download and compile the NetCDF libraries
3. Compile MMIF

The MinGW installer can be downloaded from <http://www.mingw.org>, or <http://sourceforge.net/projects/mingw/files/>. MMIF v3.4.1 was compiled using the 20110802 version. Simply download the installer (e.g. `mingw-get-inst-20110802.exe`) and run it. After installation, use the Windows Start Button to find the MinGW Shell, and run it. A command prompt will appear, running the MSYS

shell (a bash shell). The prompt is a \$, as opposed to the Windows/DOS cmd shell prompt of C:\>. Drive letters (including mapped drives) are accessed via the directories /L where L is a drive letter (C, D, etc.). A few useful commands are:

```
pwd  Print working directory, show the directory you are in
cd   Change directory
ls   List files, like the DOS command dir
mv   Move (or rename) files, like the DOS command move
rm   Remove (delete) files, like the DOS command del
```

Use the `cd` command to navigate to the directory containing the MMIF source code. For example, if you extracted the MMIF source code into a mapped drive X:\software\mmif-3.4.1, type

```
cd /x/software/mmif-3.4.1
```

Compiling the MMIF source code requires access to the NetCDF libraries, regardless of whether MM5 or WRF output data are to be processed. Downloading and compiling NetCDF version 4.1.1 has been automated via the script `compile_netcdf_windows.sh`, which is included in the MMIF distribution. Run it by typing

```
compile_netcdf_windows.sh
```

It installs the `wget` utility if it's not already in the MSYS system, and uses that to download the NetCDF source package `netcdf-4.1.1.tar.gz` from the Unidata website. Other versions of NetCDF did not compile cleanly without further user changes under MinGW 20110802, including versions 3.6.x, 4.1.2, and 4.1.3. The script then runs the typical Linux series of `configure; make; make install` to compile the NetCDF libraries and programs. This results in a sub-directory named `netcdf-4.1.1-mingw`. Useful programs for viewing NetCDF files can be found in the `netcdf-4.1.1-mingw\bin` directory. The FORTRAN include files and libraries are in `netcdf-4.1.1-mingw\include` and `netcdf-4.1.1-mingw\lib`, respectively. These paths have already been entered into the Windows version of the makefile, `makefile.windows`.

To compile MMIF, type

```
compile_mmif_windows.sh
```

This script simply calls `make` using the file `makefile.windows`. The result is the file `mmif.exe`. MMIF > v2.x uses static libraries, alleviating the need for a DLL to be placed somewhere where the system can find it (as was required by MMIF v1.0). The executable `mmif.exe` must still be placed somewhere the system can find it, either in the run-time directory (where MMIF is being run) or in a directory in the user's `%PATH%`.

MMIF v1.0 included a batch file named `compile.bat`, which could be used to compile the program on Windows using Intel FORTRAN (`ifort`). MMIF v1.0 was tested with `ifort` version 11.073 on Windows XP. The file `compile.bat` has been updated for MMIF v2.x, but has not been tested due to the way the Intel FORTRAN compiler trial software has changed since MMIF v1.0 was released in 2009.

Users of Intel FORTRAN will need a copy of `netcdf.dll` that is compatible with their version of `ifort`. They can be downloaded from <ftp://ftp.unidata.ucar.edu/pub/netcdf/contrib/win32/>. If no version of `netcdf.dll` can be found that matches the user's version of `ifort`, it is best to compile it by hand.

If the Intel FORTRAN compiler `ifort` is installed on the target computer, MMIF may be compiled by either double-clicking `compile.bat` in Windows Explorer; or by opening a Command Prompt (DOS box), changing to the appropriate directory, and typing `compile.bat` at the prompt. This was tested for MMIF v1.0, but remains untested for newer versions of MMIF.

3.2 Compiling MMIF on LINUX/UNIX

The MMIF source package includes a `makefile` to facilitate compilation of the program on Linux/Unix platforms. The `makefile` can be used to compile the MMIF source code into an executable program, and currently supports the Portland Group F90 compiler (`pgf90`), the Intel FORTRAN compiler (`ifort`), and the GNU FORTRAN compiler (`gfortran`).

The user may edit the `makefile`, un-commenting the lines of variables for the desired compiler, and commenting out the blocks for the other compilers.

Compiling the source code requires access to NetCDF libraries, regardless of whether MM5 or WRF output data are to be processed. The NetCDF libraries should be obtained from <http://www.unidata.ucar.edu/software/netcdf/> and installed on

the computer on which MMIF is compiled. It is important to use the same compiler to compile NetCDF as will be used to compile MMIF. The user must alter the makefile variable `$NETCDF` for the specific path on their computer (often, but not always, the same as the environment variable `$NETCDFHOME`). MMIF has been tested with NetCDF versions 3.6.1, 3.6.2, and 4.1.1.

MMIF is compiled by issuing the command `make` at a shell prompt within the main source directory. It will generate an executable program called `mmif` that will reside in the source code directory.

3.3 A Note on Binary Input/Output Files

MM5/WRF output and CALPUFF (and optionally SCICHEM) input files are written as FORTRAN “unformatted” (binary) files. This means that the data are written directly to the output unit as represented in memory, without translation from binary to the ASCII character set. This reduces file volume and improves read/write speed. However, there are two ways to represent machine-level formats for storing binary information in memory: IEEE “big-endian” and “little-endian”. The difference between these is essentially the order of the bits in a word, and which order is used depends on the computer platform and its operating system. The native format for many Unix workstations is big-endian, and this includes Sun, SGI, HP, and IBM. Exceptions are DEC and Linux/Windows PCs, which use little-endian by default.

WRF output files are written as NetCDF files, which are platform-independent. This means that the usual method of supplying the “`-convert big_endian`” (ifort) or “`-byteswapio`” (pgf90) compiler flags will not work, as it applies globally and only MM5 files (not WRF files) require this conversion. Fortunately, most modern compilers (pgf90, ifort, gfortran) now support the “`convert=big_endian`” option in the FORTRAN `open()` statement.

In general, MMIF can be run on machines that use either the big- or little-endian binary formats, as long as MM5 and AERMOD/CALPUFF/SCICHEM are run on the same type of platform. If any component of the modeling system is run on a different platform using the opposite binary representation, I/O files will not be properly read and will likely lead to a program crash. A typical run-time error message from using the wrong binary format is “input record too long,” so if you

get this error message, check for consistency between your binary files and compiler options.

The binary compatibility between Windows (ifort) and Linux (pgf90) has been verified. For example, PRTMET-compiled-on-Linux can read MMIF-compiled-on-Windows binary output, and vice-versa.

4. HOW TO RUN MMIF

When executing, the MMIF program will (if no command-line parameters are given) open and read a control file named `mmif.inp` that must exist in the current directory. If a filename is given on the command line, that file is read as the control file. The control file contains all of the user configuration options, flags, and pathnames to the meteorological output data files and the output files.

Run MMIF by typing its filename (with a possible path to it) at the command prompt, and optionally supplying a control file filename (with a possible path). Some examples from the DOS prompt are shown below. The first example reads `mmif.inp` in the `code` directory. The second example reads the control file `test_mm5.inp` in the current directory, and saves the screen output to a file `test_mm5.out`. The third and fourth examples use the same control file, but will create the output files in different directories (unless absolute pathnames are specified in the control file).

```
C:\Projects\MMIF\code>MMIF
```

```
C:\Projects\MMIF\code>MMIF test_mm5.inp > test_mm5.out
```

```
C:\Projects\MMIF\code>MMIF ..\test_mm5\test_mm5.inp
```

```
C:\Projects\MMIF\test_mm5>MMIF ..\code\MMIF test_mm5.inp
```

The user can also double-click the MMIF executable in the Windows Explorer, although the utility of this method is reduced when the Command Prompt automatically closes after the run has finished, before its contents can be viewed by the user. Use the greater-than sign ">" in both Windows and Linux to re-direct the screen output to a file.

Useful information is printed to the screen during execution, including information on the MM5/WRF projection, horizontal and vertical extents of both the input (MM5/WRF) and output grids (including the 1×1 single-point grids). MMIF prints whether it found the 10m winds, roughness length, albedo, Monin-Obukhov length, and LAI in the input MM5/WRF file. Messages are printed as MM5/WRF files are opened and closed, and as time-stamps are written to the output file.

It is common to initially run MMIF for a single hour or day of output, view the information printed to the screen, re-load the KML showing the domain and points, edit the control file, and re-run MMIF, until the desired horizontal and vertical grid structure is achieved. This allows user to check the extent of the sub-grid (CALPUFF or SCICHEM) or output point (AERMET, AERCOARE, or AERMOD).

The coordinates are printed in three ways: I,J coordinates, Latitude and Longitude, and kilometers from the grid origin in the MM5/WRF projection. The latter may be LCC (Lambert Conformal) or PS (Polar Stereographic) or EM (Equatorial Mercator) projections.

When aggregating, inspecting the MM5/WRF input sigma (and height) layers and the MMIF output vertical grid structure messages that are printed to the screen can be very helpful in picking which layers to include.

When interpolating, if the user wishes to “pass through” the MM5/WRF data as unchanged as possible, yet also wishes to reduce the number of output layers from the number of layers found in the MM5/WRF file, it can be instructive to first do a short run using aggregation to arrive at ZFACE values. Then enter those ZFACE values into the control file and switch to interpolation.

MMIF responds to a few optional command-line parameters:

```
C:\> mmif --help
Usage: mmif [-h | --help] [--force] [--sample] [filename]
Where
  --force    don't stop execution after non-fatal errors
  --recalc   re-calculate u10,T2,q2 from lowest MET layer
  --sample   write a a sample control file to the screen
  --version  print the version and exit
  -h         show this help message
  --help    show this help message
  filename   control filename, default is 'mmif.inp'
```

Tip: use 'mmif mmif.inp > mmif.out' to re-direct the screen output to a file.

Tip: use 'mmif --sample > mmif.inp' to create a MMIF control file, for the user to edit. The sample control file contains more outputs than a typical user would want, to show the range of possible outputs.

Currently, the only non-fatal errors that `--force` can over-ride involve:

- Continue execution even if the MM5/WRF file did not use a supported land-use type. For example, TCEQ (Texas) commissioned some additional categories to be added, producing MM5 files with a non-standard 27 categories of land-use codes. The first 24 were identical to the USGS 24 categories, so the look-up tables were still mostly correct.
- For SCICHEM output, continue execution even if the mid-point of the highest layer to be output is not at least twice the maximum terrain elevation within the output domain. This requirement is due to the vertical coordinate transformations used internally in SCICHEM.

It is hoped that the `--sample` command-line argument will prove useful to new users of MMIF.

4.1 Control File Format

Starting with MMIF version 3.0, the control file is keyword-driven. Many of the keywords were used in the older fixed-format control file, so their meaning will not be entirely new to previous users of MMIF. Some important features of the new control file are:

- Blank lines and any portion of a line after a comment character (`# ; !`) are ignored.
- Keywords and values must be either space-delimited or comma-delimited, or a mixture.
- Any number of spaces can be used between words (elements) in a line.
- All words in the control file are case insensitive, except for filenames which are used verbatim.
- Filenames may contain spaces and commas, if enclosed by either single or double quotes.
- Many of the keywords have acceptable synonyms.
- Keywords may be repeated: the values of the last one encountered will be used.
- The order of the lines containing keywords does not matter, except for the POINT and AER_LAYERS keywords (see below). The location specified by POINT is re-used for all following 1-D (point) OUTPUT, until the next POINT keyword is encountered – and similarly for AER_LAYERS.
- The INPUT MM5/WRF files may appear at the top of a control file, or even interspersed between the other lines. The INPUT files must,

however, be listed in chronological order (MMIF does not attempt to sort the list of input files).

- Primary keywords must be the first word on a line in the control file, but leading spaces are allowed.
- Most keywords are optional: MMIF assumes reasonable/useful default values.

A minimal control file using only the four required keywords is given in the following example:

```
Start    2008 01 06 01
Stop     2008 01 10 00
Output   CALPUFF calmet calmet_2008_jan.dat
Input    "X:\epa 12km conus\wrfout_d01_2008-01-05_12_00_00"
```

The above would create a CALPUFF v5.8 ready CALMET.DAT formatted file, using all but the outer five (5) points along the edge of the WRF grid, assuming a time zone of zero (GMT). Technically, the keywords OUTPUT and INPUT are not required, but running MMIF without them would be pointless.

A minimal control file used to create inputs for AERMOD is given in the following example:

```
Start    2008 01 06 01
Stop     2008 01 10 00
POINT    LATLON  35.892  -78.782  -5
Output   AERMOD  SFC    KRDU.SFC
Output   AERMOD  PFL    KRDU.PFL
Input    'X:\epa 12km conus\wrfout_d01_2008-01-05_12_00_00'
```

This would create AERMOD-ready SFC and PFL files, using the WRF data from the cell containing the Raleigh-Durham airport.

Typing "mmif --sample" at the command prompt gives the following sample control file, which illustrates all of the keywords:

```
; This file can be space-delimited or comma-delimited, or a mixture.
; Comment characters are #, ;, and !. Blank lines are ignored.
; Omitting optional keywords is the same as giving their default values.
; START, STOP, and TimeZone are the only required keywords, the rest are optional.
; Keywords are case in-sensitive, filenames are not (depends on your OS).
; Filenames may contain spaces, if enclosed in quotes.

# START and STOP can be either of the forms below, or YYYY-MM-DD_HH:mm:ss.
```

```

start      2008 07 04 01 ; start time in LST for TimeZone, hour-ending format
stop       2008070600   ; end   time in LST for TimeZone, hour-ending format

# TimeZone is relative to GMT, i.e. -5 (GMT-05) is the US East Coast

TimeZone   -10   ! default is zero, i.e. GMT-00

# MMIFv3.x auto-detects if INPUT files are MM5 or WRF files, so METFORM
# needs to be included only if MMIF guesses wrong, and you need to over-ride.

# MetForm WRF

# ORIGIN (optional) can be used to OVER-RIDE the origin of X,Y projected
# coordinate system, which is normally set from the parameters of the MM5/WRF
# file. This keyword is REQUIRED for Mercator projections.

# origin 40.0 -97.0 ! RPO Projection

# GRID has three options: IJ, LL (or latlon), or KM (or PROJ,LCC,PS,EM),
# followed by two lower-left coordinates, and two upper-right coordinates.
# Default is to output the whole grid, after trimming 5 points off each edge.

grid       IJ -5,-5 -5,-5   ! default

# LAYERS has three options: TOP, MID, or K; followed by the values to be used.
# TOP and MID are in meters. Default is from the EPA/FLM 2009 Guidance.
# TOP is preferred: MMIF interpolates between MID points to get TOPs.

layers top 20 40 80 160 320 640 1200 2000 3000 4000 ! FLM CALMET Guidance

# PG STABILITY class calculation method is either SRDT or GOLDER (default)
# PG stability is used only for CALPUFF output.

stability  GOLDER   ! default

# CLOUDCOVER (or CC) source is one of WRF, ANGEVINE, or RANDALL
#   WRF      use WRF's internal CLDFRA variable
#   ANGEVINE use Angevine et al. (2012) RH function (default in MMIF >= 2.2)
#   RANDALL  use Randall (1994)/Zhao (1995) method (default in MMIF < 2.2)

CLOUDCOVER ANGEVINE ! default

# CALSCI_MIXHT is either WRF (default) or MMIF, to pass-through or re-
# calculate the WRF mixing height for CALPUFF and SCICHEM outputs.
# Use AER_MIXHT (below) for AERMET and AERMOD modes.

CALSCI_MIXHT WRF      ! default

# AER_MIXHT (WRF, MMIF, or AERMET) controls the source of mixing height
#   values you want to use in AERMET mode.
# AER_MIN_MIXHT is the lower bound on both Convective and Mechanical
#   Mixing Heights in AERMOD mode.
# AER_MIN_OBUK  is the lower bound on Monin-Obukhov length, such that
#   ABS(L) > AER_min_Obuk, in AERMOD mode.
# AER_MIN_SPEED is the lower bound on windspeed in AERMOD mode,
#   passed through to THRESHOLD in AERMET mode.

```

```

aer_mixht      WRF      ! default
aer_min_mixht  1.0      ! default (same as AERMET)
aer_min_obuk   1.0      ! default (same as AERMET)
aer_min_speed  0.0      ! default (following Apr 2018 MMIF Guidance)

# See the Users Guide for the OUTPUT keyword details

Output qaplot   BLN      domain.bln
Output qaplot   BNA      domain.bna
Output qaplot   DAT      points.dat
Output qaplot   KML      qaplot.kml

Output calpuff  useful   calmet.info.txt
Output calpuff  calmet   calmet.met
Output calpuff  terrain  terrain.grd

Output calpuffv6 useful   calmetv6.info.txt
Output calpuffv6 calmet   calmetv6.met
Output calpuffv6 aux      calmetv6.aux
Output calpuffv6 terrain  terrainv6.grd

Output scichem  useful   scichem.info.txt
Output scichem  binary   scichem.bin.mcw
Output scichem  ascii    scichem.asc.mcw
Output scichem  sampler  scichem.smp
Output scichem  terrain  scichem.ter

point latlon    21.203   -157.925

Output aercoare useful   aercoare.near.Honolulu.info.inp
Output aercoare data     aercoare.near.Honolulu.csv

POINT LL        21.324   -157.929  -9   ! in GMT-9 timezone
AER_layers     1         4   ! write 2m, 10m, and the 4 lowest WRF layers
Output aermnet BAT      PHNL.BAT ! basename for SFC/PFL files
Output aermnet CSH     PHNL.csh ! basename for SFC/PFL files
Output aermnet useful   PHNL.useful.txt
Output aermnet onsite   PHNL.dat
Output aermnet upperair PHNL.fsl
Output aermnet aersfc   PHNL.aersfc.dat

FSL_INTERVAL   6         ! output every 6 hours, not 12 (the default)
POINT IJ       73        32
Output aermnet FSL      'Upper air at PHTO.FSL'
POINT KM       60.0     -12.0
Output aermnet FSL      "Upper air at PHOG.FSL"

POINT latlon    20.963   -156.675  -9   ! in GMT-9 timezone
AER_layers     0         0   ! write only 2m and 10m data
Output aermnet useful   PJHJ.info.txt
Output aermnet sfc      PJHJ.sfc
Output aermnet PFL      PJHJ.pfl
# INPUT gives filenames of either MM5 or WRF files

INPUT test_problems\wrf\wrfout_d02_2008-07-04_00_00_00
INPUT test_problems\wrf\wrfout_d02_2008-07-04_12_00_00

```

```
INPUT test_problems\wrf\wrfout_d02_2008-07-05_00_00_00
INPUT test_problems\wrf\wrfout_d02_2008-07-05_12_00_00
INPUT test_problems\wrf\wrfout_d02_2008-07-06_00_00_00
INPUT test_problems\wrf\wrfout_d02_2008-07-06_12_00_00
```

4.2 Control File Keyword Listing

The following describes each keyword of the control file. Several examples are given for many of the keywords, but the user should, for most keywords, include only one of them in a control file. Italics are used to indicate number values.

Default values are indicated.

```
Start YYYY MM DD HH
Stop YYYY MM DD HH
```

Date/hour to start processing (year, month, day, and hour). See the note on time-stamps in Section 4.3.

Start and *Stop* can also be in *YYYYMMDDHH* format, i.e. no spaces. They can also be given in MM5/WRF format (*YYYY-MM-DD_HH:mm:ss*), though only the year, month, day, and hour are used – the minutes and seconds are not used.

```
TimeZone HH
```

Time zone shift in hours from UTC (0=GMT, -5=EST, -6=CST, -7=MST, -8=PST) for model output. CALPUFF, and AERMOD are run in local time, but MM5 and WRF use UTC. SCICHEM version 3 and newer assume the MEDOC-formatted file is in UTC. Note that CALPUFF v5.8 and AERMOD always assume they are run in the Western hemisphere and use a positive time zone shift. MMIF and CALPUFF v6+ use the actual time zone shift (negative for the Western hemisphere) to maximize global flexibility. Because CALPUFF requires the user to pick a single time zone even if a CALPUFF domain spans several time zones, MMIF assigns a single time zone to the CALPUFF output. However, MMIF can shift each individual *POINT* output (AERMOD modes) to its own time zone – see the *POINT* keyword description below.

Tip: when setting the time zone for each output point using the keyword *POINT* (described below), it is important to remember that the beginning of the extraction period is specified by the *Start* and *TimeZone* keywords above. To avoid truncating the beginning of one of your *POINT* output files, set the global *TimeZone* to (at least) the eastern-most time zone for any of your requested points. For

output in the western hemisphere, `TimeZone` can be set to 0 (zero) if no gridded output is desired during the same MMIF run.

If MMIF encounters a MM5/WRF file containing time-stamps that have already been processed, those time-stamps are skipped (except for precipitation processing, see Section 4.3). This allows consecutive MM5/WRF initializations to include some spin-up hours that overlap, a common practice. MMIF also skips any sub-hourly time-stamps found in the WRF file.

```
METFORM    MM5
METFORM    WRF
```

MMIF automatically detects whether an `INPUT` file is from MM5 or from WRF. If for some reason the auto-detection fails, this keyword can be included to skip the auto-detection. In general, it need not be included in a MMIF input file (the line can be deleted).

```
ORIGIN    LAT LON
```

Including this keyword causes MMIF to over-ride the values for the origin of the X,Y grid found in the input MM5/WRF file (the default behavior). It may not be *useful* to specify e.g. an LCC grid origin LON that is different from `STAND_LON`, but it is technically *possible*, and MMIF allows it.

The Lambert Conformal Conic (LCC) projection is fully defined by only three parameters: the Standard Longitude, and the two True Latitudes. This essentially defines a plane, and the user is free to define the origin of a coordinate system (as well as the units, e.g. km vs. meters). The origin of the coordinate system is the point defined by the Origin Latitude and Origin Longitude.

It is convenient to define the origin of the LCC coordinate system at the center of the coarse MM5/WRF grid. In MM5, it was required that the Origin Latitude/Longitude equal the Central Latitude/Longitude. WRF allows the user to specify `REF_X` and `REF_Y` in the `&geogrid` section of the WPS namelist. Together with the values of `REF_LAT` and `REF_LON`, this defines the origin of the LCC grid. The point (`REF_LON`, `REF_LAT`) in LAT-LON space is the same point as (`REF_X`, `REF_Y`) in I-J space. [Note the difference in terminology: WRF uses (X, Y) to refer to grid cell numbers, which is perhaps more commonly referred to with (I, J). (X, Y) is more commonly used to refer to distance (e.g. in kilometers) on the grid.]

If the WRF user does not specify (REF_X, REF_Y) then they default to the center of the coarse grid (known as the Mother Of All Domains, or MOAD). In this case, the WRF parameter MOAD_CEN_LAT has the same value as the Origin Latitude. The WRF parameter CEN_LON will also coincide with the STAND_LON (Standard Longitude). The origin of the LCC grid is therefore (MOAD_CEN_LAT, STAND_LON). Both of these parameters can be found in the global attributes section of WRFOUT files. This is probably the most common way to run WRF.

However, if the WRF user specified a (REF_X, REF_Y) combination, then the origin of the LCC grid may not coincide with either the Central Longitude or the Central Latitude. Unfortunately, WRFOUT files do not contain the STAND_LAT parameter. This is perhaps reasonable, since only STAND_LON, TRUELAT1 and TRUELAT2 are required to fully define the projection, and the downstream user is free to define the latitude of the origin of the grid. As long as the user is self-consistent, any point, even the Central (LAT, LON), can be used.

The ORIGIN keyword can be used to specify the origin (X, Y) = (0, 0) latitude and longitude for any of the supported projections. For the Mercator projection, no suitable longitude (other than the central longitude of each domain) is included in the WRFOUT file's global attributes section. Therefore, the ORIGIN keyword is *required* when using WRF data run on a Mercator (PROJ=3) projection.

```

GRID IJ iLL  jLL  iUR  jUR
GRID LL LatLL LonLL LatUR LonUR
GRID KM xLL  yLL  xUR  yUR
GRID IJ -5    -5    -5    -5    (default)

```

Specify the requested output sub-grid's lower-left (LL) and upper-right (UR) corners. The 2nd keyword after GRID must be one of the following values:

- IJ Use MM5/WRF grid point index (I, J). I is the West-to-East coordinate, and J is the South-to-North coordinate, contrary to MM5 nomenclature.
- LL Use latitude (positive in the Northern hemisphere) and longitude (positive in the Eastern hemisphere). LATLON is a synonym for LL.
- KM Use the MM5/WRF projected grid coordinate system, e.g. Lambert conformal conic (LCC), polar stereographic (PS), or Equatorial Mercator (EM). Each of these can be used as synonyms for KM.

Tip: When using the IJ method, entering 0 0 for both corners will automatically select the entire MM5/WRF grid. Entering a negative value (e.g. -5 -5) will “trim” the number points from the edges of the MM5/WRF grid (e.g. all but the first 5 points around the perimeter of the MM5/WRF domain). This is useful because in nested grids, the 5 perimeter points are smoothed for numerical reasons, and contain unphysical data.

```
POINT IJ I J [TimeZone]
POINT LL Lat Lon [TimeZone]
POINT KM X Y [TimeZone]
```

Specify the requested output point, for AERCOARE, AERMET, and AERMOD outputs (meaningless for gridded outputs). The 2nd keyword after POINT is similar to that of GRID above, but only one point is required (instead of two points that define a “box”).

The optional 5th keyword is the local point’s time-zone shift (in hours from GMT, like TIMEZONE). If it is not given, this POINT output will be assigned the global time-zone shift. Use -5 for the US East coast, -8 for the US West coast, 0 for GMT, etc.

Tip: see the note above about setting the global TimeZone to the eastern-most of the POINT TimeZones.

Unlike GRID, it is very useful to repeat the keyword POINT in a MMIF control file. As MMIF parses the control file, it re-uses the values of POINT found until another line containing POINT is encountered. This would allow the user to output both AERMET and AERMOD files for the same point without having to repeat the line starting with POINT. Note, however, that MMIF attempts to guess the output filenames associated with that point, and will write the filenames of the *next* encountered outputs at any particular I,J location to the Useful Info file (see below). This means an OUTPUT USEFUL line should come *before* the rest of the outputs for this point.

```
LAYERS K 1 2 3 5 7 10 15
LAYERS TOP 20 40 80 160 320 640 1200 2000 3000 4000 (default)
LAYERS MID 10 30 60 120 240 480 920 1600 2500 3500
```

Specify the requested output vertical layer structure. The 2nd keyword after LAYERS must be one of the following keywords, followed by a series of values:

K Aggregate in the vertical, using the vertical index (as in "I,J,K") to specify which layers to output. The values that follow must be a series of integers, where output layer interfaces match MM5/WRF layer interfaces.

Example: 1,2,3,4,5,6,7,9,12,15,18,21 means output layers 1-7 exactly match WRF layers 1-7, output layer 8 interface matches WRF layer interface 9 (so WRF layers 8 and 9 are aggregated/collapsed to output layer 8), output layer 9 interface matches WRF layer interface 12 (so WRF layers 10, 11, and 12 are aggregated/collapsed to output layer 9), etc. See Section 2.4 for details.

For SCICHEM output, aggregation is not possible, and MMIF will stop after printing an error message.

TOP Interpolate in the vertical, using layer-top interface heights above the surface (i.e. ZFACE values) to specify the output layers. The values that follow must be a series of real values, specifying the top interface of each layer. MMIF assumes the first layer bottom is the surface, so do not start the list with a zero.

Example: 20,40,80,160,320,640,1200,2000,3000,4000 means interpolate to match the EPA-FLM recommended heights as of August 20, 2009⁽⁸⁾. These are the default values.

MID Interpolate in the vertical, using layer mid-point heights above the surface to specify the output layers. The values that follow must be a series of values specifying the grid cell centers of each layer.

Example: 10,30,60,120,240,480,920,1600,2500,3500 means interpolate to get similar, but not identical, layer structure as in the TOP example above. The top of the highest layer is extrapolated.

MMIF will count the number of values given, dynamically allocate arrays, and convert any commas to spaces before reading the values.

STABILITY GOLDER (default)
STABILITY SRDT

Specify the Pasquill-Gifford stability class calculation method, see section 2.9 for details. The PG stability class values are written only to CALMET files.

CLOUDCOVER ANGEVINE ! default
CLOUDCOVER WRF

⁽⁸⁾ See <http://www.nature.nps.gov/air/permits/flag/index.cfm>

CLOUDCOVER RANDALL

Specify the source of the Cloud Cover values to be written by MMIF. The 2nd keyword after CLOUDCOVER (or its synonym CC) must be one of:

WRF MMIF's 2-D cloud cover output will contain the maximum value (of the column) of the cloud fraction from WRF's 3-D CLDFRA field.

ANGEVINE Use the Angevine (2012) COAMPS method.

RANDALL Use the Randall (1994)/Hong et al. (1998) method.

See Section 2.8. RANDALL was the default before MMIF 2.2. ANGEVINE was (and is) the default starting with MMIF 2.2. This cloud cover is used by CALPUFF and SCICHEM, and in the SFC file written in AERMOD (direct) mode. It is *not* written to the ONSITE datafile in AERMET mode.

CALSCI_MIXH WRF (default)
CALSCI_MIXHT MMIF

Setting CALSCI_MIXHT to WRF causes MMIF to pass through PBL depth from WRF with no changes. WRF values of PBL height are quantized to the nearest layer midpoint, and each PBL physics choice subroutine does its own calculation of PBL height.

Setting CALSCI_MIXHT to MMIF causes MMIF to re-diagnose PBL depth using a bulk Richardson method with 20 times the vertical resolution of the MM5/WRF data. See section 2.7 for details.

AER_MIXHT WRF (default)
AER_MIXHT MMIF
AER_MIXHT AERMET

The AER_MIXHT keyword specifies the source of mixing height values when MMIF is run in AERMET mode. The default source is WRF, but mixing heights from MMIF and AERMET can also be used.

AER_MIN_MIXHT VALUE (default = 1.0 m)

Specify the minimum allowed mixing height (m) for AERMOD SFC output (see below). The default value of 1.0 m was taken from the AERMET source code.

AER_MIN_OBUK VALUE (default = 1.0 m)

Specify the minimum allowed absolute value of the Monin-Obukhov length (m) for AERMOD SFC output (see below). The default value of 1.0 m was taken from the AERMET source code: $|L| > 1$.

AER_MIN_SPEED VALUE (default = 0.0 m/s)

Specify the minimum allowed wind speed (m/s) for AERMOD SFC output (see below). The default value of 0.0 m/s comes from recent guidance⁹ issued by the US EPA. MM5 and WRF can output very small, yet non-zero (non-calm) values for the wind speed (e.g. 0.001 m/s) in cases where a physical anemometer would register a calm wind. However, AERMOD has a minimum wind speed of 0.2828 m/s.

FSL_INTERVAL VALUE (default = 12 hours)

Specify the interval, in hours, between successive writes to the AERMET FSL (UPPERAIR) output (see below). The default value of 12 hours will produce the customary 0Z and 12Z soundings in the output file. Using a value of 6 would produce soundings at 0Z, 6Z, 12Z, and 18Z for each day of the MMIF run. Note: the default in MMIF 3.1 and earlier was 1 hour.

AER_LAYERS VALUE (default = 1, Number_of_Layers)

Specify the lowest and highest layer indexes (two integers) to write to AERMOD PFL and AERMET ONSITE outputs (does not affect AERMET FSL output). All layers whose index is between the first and 2nd values, inclusive, will be written.

Increasing the 1st (lower) value is useful to suppress output of an MM5/WRF layer that is too close to the 10 m speed/direction extracted from the MM5/WRF variables U10, V10. If the user, after running MMIF in AERMET mode, receives the following error when running AERMET:

```
****      ERROR MESSAGES      ****
JOB          E27 OSTEST   : Duplicate ONSITE heights specified at   10.00 meters!
```

then increase the 1st value of AER_LAYERS to suppress output of the lowest MM5/WRF layer.

⁹ See https://www3.epa.gov/ttn/scram/models/relat/mmif/MMIF_Guidance.pdf

Reducing the 2nd (upper) value given is useful if the user did not give the keyword `LAYER`, and does not want the *.PFL file to emulate a 3500 m tall meteorological tower with instruments at 10 levels.

Giving two zeroes (0 0) will cause MMIF to write only the 2 m (temperature) and 10 m (winds, temperature) levels and skip all the values extracted from the MM5/WRF layers. This simulates what AERMET outputs when run with traditional NWS airport data.

`OUTPUT` `MODEL` `FORMAT` `FILENAME`

Specifies the outputs from MMIF. Any number of lines containing `OUTPUT` can be included in the control file. Both 3-D `GRID` and 1-D `POINT` data can be written during the same MMIF run, but all the requested points must be inside the requested output sub-grid specified by the `GRID` keyword. When requesting only 1-D `POINT` data, it is suggested to omit the `GRID` keyword. Multiple 3-D `GRID` outputs are allowed (i.e. MMIF can write `CALMET`, `CALMETv6`, and `SCICHEM` files in a single run) but they must use the same grid. The number of files created is limited only by the computer system's maximum number of open files, and available memory.

The filename (4th word of the line) can include spaces if it is enclosed in either single or double quotes.

The 2nd keyword specifies the model for which the MMIF output is intended (think "which model do I want to run next?"). Valid values include `CALPUFF`, `CALPUFFv6`, `SCICHEM`, `AERMET`, `AERCOARE`, `AERMOD`, and `QAPLOT` (an output type that isn't really a model). Each `MODEL` has several file `FORMATS` it supports, specified by the 3rd keyword of an `OUTPUT` line.

Valid `OUTPUT` `CALPUFF` formats include:

<code>USEFUL</code>	Produces a text file giving values required in a <code>CALPUFF</code> control file (e.g. <code>PMAP</code> , <code>RLATO</code> , <code>XLAT1</code> , <code>DATUM</code> , <code>XORIGKM</code> , <code>NX</code> , <code>NY</code> , <code>NZ</code> , <code>ZFACE</code> , etc.).
<code>CALMET</code>	Produces a <code>CALMET.DAT</code> formatted file, for use with <code>CALPUFF</code> v5.8.
<code>TERRAIN</code>	Produces a Golden Software Surfer ASCII *.GRD file of the MM5/WRF terrain, similar to output from other programs in the <code>CALPUFF</code> system.

Valid OUTPUT CALPUFFv6 formats include:

USEFUL	Same as for CALPUFF above
CALMET	Produces a CALMET.DAT formatted file, for use with CALPUFF v6.x.
AUX	Produces a CALMET.AUX formatted file containing the cloud liquid water, for use with CALPUFF v6.x.
TERRAIN	Same as CALPUFF above.

Valid OUTPUT SCICHEM formats include:

USEFUL	Produces a text file containing values useful in the various control files (e.g. XMIN, XREF, LONO, LATO, NZB, etc.).
ASCII	Produces an ASCII-style MEDOC file (MCW) for use with SCICHEM.
BINARY	Produces a BINARY-style MEDOC file (MCW) for use with SCICHEM.
TERRAIN	Produces a text file containing the MM5/WRF terrain, in a file format useful for SCICHEM. Note that terrain is also included in the MEDOC file format, and this file cannot be used together with a MEDOC file in the same SCICHEM run.

Note that SCICHEM v3.0 and later can also use AERMET-style SFC and PFL files for near-field (< 50 km) simulations.

Valid OUTPUT AERCOARE formats include:

USEFUL	Produces a control file for the AERCOARE program. Values of the input and output filenames, lat, lon, and some defaults, have been set.
DATA	Produces a comma-separated values (CSV) text file that contains near-surface data, for use with AERCOARE. See section 2.3.2. This output format is the most generic, and may be useful for users wishing to simply extract MM5/WRF timeseries data at a point.

Valid OUTPUT AERMET formats include:

BAT or CSH	Produces a text file containing either a DOS-style BAT file or a Linux-style CSH script, as determined by the keyword. This batch file runs the three complete AERMET input files (IN1, IN2, and IN3) that are also created, using the same base filename and changing the extension - <i>this filename sets the resulting SFC/PFL base filenames</i> . The values of AERMET keywords LOCATION, XDATES, ONSITE, UPPERAIR, AERSURF, etc. have been set appropriately. All filenames
------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(but not the path portion, if any, of the filename) will be converted to uppercase, an AERMET requirement.

- USEFUL** Produces a text file containing lines appropriate for inclusion in an AERMOD input file, specifying the “ME” pathway. It includes the MM5/WRF terrain elevation of the output point, as PROFBASE.
- ONSITE** Produces a text file formatted for use in the ONSITE pathway of AERMET. It contains surface parameters, 2m and 10m data, and upper-level data. Use the keywords `MIN_LAYER` and `MAX_LAYER`, or `LAYERS`, to control how many upper layers are written. See the IN1 file created by including `OUTPUT AERMET BAT` for the file format.
- FSL** Produces a text file formatted to mimic NOAA/ESRL FSL-formatted upper-air data. See section 2.3.1. The keyword `UPPERAIR` is a synonym.
- AERSFC** Produces a text file that mimics AERSURFACE output. See section 2.3.1.

Valid `OUTPUT AERMOD` formats include:

- USEFUL** Produces a text file containing lines appropriate for inclusion in an AERMOD input file, specifying the “ME” pathway. It includes the MM5/WRF terrain elevation of the output point, as PROFBASE.
- SFC** Produces a text file formatted for use with AERMOD as a `SURFFILE`.
- PFL** Produces a text file formatted for use with AERMOD as a `PROFFILE`.

Valid `OUTPUT QAPLOT` formats include:

- BLN** Produces a BLN file, a text file containing a header line giving the number *N* of vertices (lines) to follow followed by the coordinates of the vertices of a polygon. This block can be repeated. MMIF writes one block for the full WRF/MM5 domain, and one block for the requested sub-grid output (specified by the `GRID` keyword). This file format is commonly used with plotting software that many CALPUFF users are familiar with. The coordinates are given in the projected coordinate system (LCC, PS, EM) just like CALPUFF’s `IQAPLOT=1` output.
- BNA** Produces a BNA file, a text file very similar to a BLN file, but supports labels. CALPUFF produces a very similar QA file if the `IQAPLOT=1` setting is selected.
- DAT** Produces a text file containing the coordinates in the projected coordinate system (LCC, PS, EM) and in latitude/longitude of the 1-D

`POINT` output locations. These coordinates represent the center of the grid cell containing the requested `POINT`, so may differ from the coordinates in a MMIF input file.

`KML` Produces a text file in KML format. KML is a file format used to display geographic data in an Earth browser such as Google Earth. Both the `GRIDS` (full WRF/MM5 grid and requested sub-grid) and `POINT` data are included.

`INPUT` *FILENAME*

The lines starting with `INPUT` give the filenames of the MM5/WRF files to process. Spaces are allowed in MM5/WRF filenames, if enclosed in quotes. Case is preserved.

Any time-stamps found in MM5/WRF files that are before the requested `START` time-stamp are skipped (except for precipitation processing, see Section 4.4).

4.3 A Note on Time Stamps

CALPUFF v5.8 labels the first hour of a day as hour 1 (the hour between midnight and 1:00 AM). Similarly, the last hour of a day is 0 (the hour between 11:00 PM and midnight). The starting hour `HH` should be set to "1" if it is desired that 24-hour average concentrations should start and end at midnight.

CALPUFF v6.x changed from the "label" concept to a "beginning and ending instant in time" concept. In MMIF v2.1 this change caused some confusion. Starting with MMIF v2.2, the user should specify the beginning and ending time-stamps the same way, for both CALPUFF and CALPUFFv6 output. This allows the user to only change the output format line in the MMIF control file and have the CALPUFF and CALPUFFv6 output span the same time period.

AERMOD labels the first hour of the day as hour 1 (the hour between midnight and 1:00 AM) but labels the last hour of the day as hour 24 (the hour between 11:00 PM and midnight). The starting hour `HH` should be set to "1", just as in CALPUFF.

MMIF will properly handle both 2011 05 31 24 and 2011 06 01 00, and treat them as the same hour.

Neither MM5 nor WRF write hourly-averaged values, both write only instantaneous values at the time step closest to the requested output time (usually the top of the hour). Because the fields at this instant in time are the result of the "history" of the

preceding time period, these instantaneous values are treated as “hour ending” time labels. The WRF output at 03:00:00 is labeled hour 3 in CALPUFF and AERMOD output, and as 02:00 to 03:00 in CALPUFFv6 output.

MMIF currently skips any sub-hourly WRF fields (i.e. time-stamps with the minutes value not close to zero), even though WRF is capable of writing more than one output per hour, and CALPUFFv6 is capable of handling sub-hourly timesteps.

4.4 A Note on Precipitation Processing

Both MM5 and WRF save a field representing the accumulated precipitation (mm) since the beginning of the simulation. At each grid point, this value only grows with each time step. However, CALPUFF, AERMOD and SCICHEM all require the hourly precipitation rate (mm/hr) at each grid point. The MMIF program converts accumulated precipitation to hourly precipitation rate automatically, by subtracting the last hour’s field from the current hour’s field, point by point. This requires the user to pay special attention to the *first* hour of each MM5/WRF run that is to be processed.

If the first hour to be processed by MMIF is also the first hour of an MM5 or WRF simulation, then the accumulated precipitation and the hourly precipitation rate are the same, and the program writes these values to the output file. However, it is common practice to allow for some “spin-up” time when running MM5 or WRF. These first hours of the simulation are often written to separate output files, and ignored or discarded. If, for example, the first MM5/WRF file given in the MMIF.INP file starts at the 12th hour of a run, and that hour corresponds to the “start Output (LST)” time-stamp, then MMIF cannot subtract the precipitation field from the 11th hour. In these cases, MMIF sets the initial precipitation field to zero everywhere and prints a warning to the screen.

For “spin-up” cases, it is therefore important to supply MMIF with at least one hour of MM5/WRF data from *before* the “start Output (LST)” time-stamp, so that MMIF can subtract that hour’s precipitation field from the first requested output hour’s precipitation field. This is true both for the very first hour to be processed, as well as for the first hour of each subsequent MM5/WRF initialization to be processed. The former case is reasonably obvious, but the latter case is perhaps not as obvious. Consider the following example:

A user has an annual MM5 simulation for 2005 that was run as a series of 5-day runs with 12 hours of overlapping spin-up each, as is common practice. The first MM5 initialization starts at 2004-12-31_12 (YYYY-MM-DD_HH), the second run at 2005-01-04_12, and so on. Each hour's model output is written to a separate file. For simplicity's sake, let's assume the filenames have already been converted to LST during post-processing, so we can ignore the UTC-LST time shift. Because the user intends to ignore the first 12 hours of each run to allow for model spin-up, he enters the following filenames into the MMIF control file:

```
Start 2005 01 01 01
[...]
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_01
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_02
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_03
[...]
Input C:\mm5\2004-12-31\mmout_d1.2005-01-04_00
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_01
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_02
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_03
[...]
```

This file list will cause MMIF to set the entire precipitation field to zero for the first output hour (2005-01-01_01), because it does not have access to the previous hour's precipitation field. The solution is to include the previous file, using the line

```
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_00
```

as the first MM5 file listed in the control file.

A second, more subtle problem occurs at hour 1:00 on 2005-01-04, when we switch from one MM5 initialization to the next. MMIF will subtract the precipitation field found in 2004-12-31\mmout_d1.2005-01-04_00 from the precipitation field in 2005-01-04\mmout_d1.2005-01-04_01.

Because MM5 is imperfect, it is unlikely that frontal systems and rain bands will be perfectly aligned from one initialization to the next. This may result in negative precipitation rates at some grid points, when MMIF subtracts the one from the other. The solution is to include one hour's data from *before* the end of the spin-up period of *each* run.

The correct set of files to include in the MMIF control file for this example is:

```
Start 2005 01 01 01
[...]
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_00
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_01
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_02
Input C:\mm5\2004-12-31\mmout_d1.2005-01-01_03
[...]
Input C:\mm5\2004-12-31\mmout_d1.2005-01-04_00
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_00
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_01
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_02
Input C:\mm5\2005-01-04\mmout_d1.2005-01-04_03
[...]
```

Note that `mmout_d1.2005-01-04_00` is included *twice*, once from each initialization. Even though MMIF reads the MM5 data for time-stamp 2005-01-04_00 twice, it only *writes* the data from the first MM5 file. The data from the second MM5 file is read (but not written) and its precipitation field is stored, and subsequently subtracted when the next MM5 file containing a new time-stamp is read by the program.

If the user instead has MM5/WRF data files that contain more than one time-stamp of data, e.g. 24 hours per file, the same care must be taken to assure that the first time-stamp that MMIF reads for each initialization is not the first time-stamp to be written. This is true for both MM5 files (which are read “linearly”) and WRF files (which may be read “direct access”). Although WRF files are NetCDF files, for which it is possible to read only the desired data subset (i.e. read just one field from the *N*th time-stamp from the “middle” of a file), MMIF parses all the time-stamps in WRF files to assure the correct processing of precipitation.

Starting with MMIF v2.2, the initialization of the precipitation field (reading the hour prior to the first hour from each initialization that is to be written) is explicitly written to the screen, using both LST and UTC time-stamps.

MMIF does not currently use the `prec_acc_dt` feature introduced in WRFv3.4.

5. LIMITATIONS

The case where *two* MM5/WRF layers are within the first CALMET layer (i.e. $ZFACE(2) < 19$ m) has not been accounted for in the aggregation scheme. Nor has the possibility that more than one MM5/WRF layer exists between 20 m and 39 m. For these cases it is suggested to use interpolation rather than aggregation.

An AERMET SFC file contains both a convective and mechanical mixing height. Because MM5/WRF only produce a single value for the height of the PBL (which can be re-calculated by MMIF), MMIF writes the same value for both mixing heights in the SFC file when the Monin-Obukhov length is negative (indicating unstable conditions). The convective mixing height, W^* , and the vertical potential temperature gradient above the mixed layer are set to the missing value flag for stable conditions, to minimize AERMOD warnings.

In AERMET mode, the UPPERAIR (FSL) file does not contain mandatory reporting levels (unless the MM5/WRF layer happens to be centered on one). Certain downstream software, e.g. CALMET, which require mandatory levels, will likely not work with MMIF FSL files.

MMIF does not utilize the `prec_acc_dt` feature introduced in WRFv3.4.

MMIF supports only some of the land-use datasets available in WRF. It does not support the "OLD", "SiB", or "SSiB" land-use dataset available in WRFv3.5.1 and later, due to a lack of leaf area index (LAI) values for the look-up tables. Albedo and roughness length can be obtained from WRF's LANDUSE.TBL file, but WRF's treatment of LAI is more complex (see VEGPARAM.TBL, which gives maximum and minimum LAI values). Previous values for the LAI tables were obtained from MCIP, which has not been extended to support SiB or SSiB.

MMIF does not check that subsequent MM5/WRF files (after the first file is read) have the same grid parameters (projection, number of points, etc.).

When MMIF mimics AERSURFACE output, it uses only the values of surface roughness, Bowen ratio, and albedo for the grid cell containing the requested output point. If the input WRF grid spacing is less than 1km, this might contradict EPA's 2008 guidance on determining the roughness length (inverse-distance weighted geometric mean out to 1km). Similarly, MMIF does not perform a

10×10km unweighted geometric mean of the Bowen ratio, nor a 10×10km unweighted arithmetic mean of the albedo.

The Windows executable cannot read WRF files larger than 2 GB, due to a compiler issue described in Section 3.1.

6. REFERENCES

- Anderson, B. 2008. "The MM5-CALPUFF Project." Presented at the EPA/Regional/State/ Local Modeler's Workshop, Denver, CO (June 11, 2008).
- Angevine, Wayne M., Lee Eddington, Kevin Durkee, Chris Fairall, Laura Bianco, Jerome Brioude, 2012. "Meteorological Model Evaluation for CalNex 2010". *Mon. Wea. Rev.*, 140, 3885-3906.
- Chowdhury, B. et al. 2015. "SCICHEM-2015 User's Guide". Sage-Xator Management, Princeton NJ, 216 pp.
- EPA. 1998. "Interagency Workgroup on Air Quality Modeling (IWAQM): Phase 2 Summary Report and Recommendations for Modeling Long Range Transport Impacts." Technical Report prepared by the U.S. Environmental Protection Agency, Research Triangle Park, NC, 160 pp (EPA-454/R-98-009).
- EPA. 1993. "An Evaluation of a Solar Radiation/ Delta-T (SRDT) Method for Estimating Pasquill-Gifford (P-G) Stability Categories." Technical Report prepared by the U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC (EPA-454/R-93-055).
- EPA. 2005. "Guideline on Air Quality Models, 40 CFR Part 51, Appendix W." Published in the Federal Register, Vol. 70, No. 216, November 9, 2005.
- EPA. 2017. "Guideline on Air Quality Models, 40 CFR Part 51, Appendix W." Published in the Federal Register, Vol. 82, No. 10, January 17, 2017.
- Golder, D. 1972. "Relations among stability parameters in the surface layer." *Boundary Layer Meteor.*, 3, 47-58.
- Gryning, S.E. and Batchvarova. 2003. "Marine Boundary-Layer Height Estimated from NWP Model Output." *Int. J. Environ. Pollut.*, 20, 147-153.
- Louis, J.F. 1979. "A Parametric Model of Vertical Eddy Fluxes in the Atmosphere." *J. Atmos. Sci.*, 35, 187-202.
- Robe, F.R., Z-X. Wu and J.S. Scire. 2002. "Real-time SO₂ Forecasting System with Combined ETA Analysis and CALPUFF Modeling." Proceedings of the 8th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 14-17 October 2002, Sofia, Bulgaria.
- Turner, D.B. 1970. "Workbook of Atmospheric Dispersion Estimates." Technical Report prepared by the U.S. Environmental Protection Agency, Research Triangle Park, NC, 84 pp.

- Vogelezang D. and A. Holtslag. 1996. "Evaluation and model impacts of alternative boundary-layer height formulations. *Boundary Layer Meteor.*, 81, 245-269.
- Santos, L.P. and R.I. Sykes, 2000. "SCICHEM Version 1.2 Technical Documentation", prepared by Titan Corporation, Princeton, NJ, ARAP Report No. 722, prepared for EPRI. 245 pp.
- Sykes, R.I., S.F. Parker, and D.S. Henn, 2004. "Draft SCIPUFF Version PD2.2 Technical Documentation", prepared by Titan Corporation, Princeton NJ, 249 pp.