

---

# **CCPP Technical Documentation**

*Release v3.0.0*

**J. Schramm, L. Bernardet, L. Carson,  
G. Firl, D. Heinzeller, L. Pan, and M. Zhang**

**Jun 17, 2019**

For referencing this document please use:

Schramm, J., L. Bernardet, L. Carson, G. Firl, D. Heinzeller, L. Pan, and M. Zhang, 2019. CCPP Technical Documentation Release v3.0.0. 91pp. Available at [https://dtcenter.org/GMTB/v3.0/ccpp\\_tech\\_guide.pdf](https://dtcenter.org/GMTB/v3.0/ccpp_tech_guide.pdf).

# CONTENTS

<b>1</b>	<b>CCPP Overview</b>	<b>1</b>
1.1	How to Use this Document . . . . .	4
<b>2</b>	<b>CCPP-Compliant Physics Parameterizations</b>	<b>7</b>
2.1	General Rules . . . . .	8
2.2	Input/output Variable (argument) Rules . . . . .	9
2.3	Coding Rules . . . . .	10
2.4	Parallel Programming Rules . . . . .	11
2.5	Scientific Documentation Rules . . . . .	12
2.5.1	Doxygen Comments and Commands . . . . .	12
2.5.2	Doxygen Documentation Style . . . . .	13
2.5.3	Doxygen Configuration . . . . .	18
2.5.4	Using Doxygen . . . . .	20
<b>3</b>	<b>CCPP Configuration and Build Options</b>	<b>23</b>
<b>4</b>	<b>Constructing Suites</b>	<b>27</b>
4.1	Suite Definition File . . . . .	27
4.1.1	Groups . . . . .	27
4.1.2	Subcycling . . . . .	27
4.1.3	Order of Schemes . . . . .	28
4.2	Interstitial Schemes . . . . .	28
4.3	SDF Examples . . . . .	29
4.3.1	Simplest Case: Single Group and no Subcycling . . . . .	29
4.3.2	Case with Multiple Groups . . . . .	29
4.3.3	Case with Subcycling . . . . .	30
4.3.4	Operational GFS v15 Suite . . . . .	30
<b>5</b>	<b>Autogenerated Physics Caps</b>	<b>35</b>
5.1	Dynamic Build Caps . . . . .	35
5.2	Static Build Caps . . . . .	38
<b>6</b>	<b>Host Side Coding</b>	<b>43</b>
6.1	Variable Requirements on the Host Model Side . . . . .	43
6.2	Metadata Variable Tables in the Host Model . . . . .	43
6.3	CCPP Variables in the SCM and UFS Atmosphere Host Models . . . . .	45
6.4	CCPP API . . . . .	46
6.4.1	Data Structure to Transfer Variables between Dynamics and Physics . . . . .	46
6.4.2	Adding and Retrieving Information from cdata (dynamic build option) . . . . .	47
6.4.3	Initializing and Finalizing the CCPP . . . . .	48
6.4.4	Running the physics . . . . .	49

6.4.5	Initializing and Finalizing the Physics . . . . .	49
6.5	Host Caps . . . . .	50
6.5.1	SCM Host Cap . . . . .	52
6.5.2	UFS Atmosphere Host Cap . . . . .	54
<b>7</b>	<b>CCPP Code Management</b>	<b>57</b>
7.1	Organization of the Code . . . . .	57
7.1.1	UFS Atmosphere . . . . .	57
7.1.2	SCM . . . . .	58
7.1.3	Directory Structure of ccpp/framework . . . . .	58
7.1.4	Directory Structure of ccpp/physics . . . . .	59
7.2	GitHub Workflow (setting up development repositories) . . . . .	59
7.2.1	Creating Forks . . . . .	59
7.2.2	Checking out the Code (NEMSFv3gfs) . . . . .	59
7.2.3	Checking out the Code (SCM) . . . . .	60
7.3	Committing Changes to your Fork . . . . .	62
7.4	Contributing Code, Code Review Process . . . . .	63
7.4.1	Creating a PR . . . . .	63
<b>8</b>	<b>Technical Aspects of the CCPP <i>Prebuild</i></b>	<b>65</b>
8.1	<i>Prebuild</i> Script Function . . . . .	65
8.2	Script Configuration . . . . .	65
8.3	Running ccpp_prebuild.py . . . . .	68
8.4	Troubleshooting . . . . .	69
<b>9</b>	<b>Building and Running Host Models</b>	<b>75</b>
9.1	SCM . . . . .	75
9.1.1	System Requirements, Libraries, and Tools . . . . .	75
9.1.2	Building and Running the SCM . . . . .	76
9.2	UFS Atmosphere . . . . .	80
9.2.1	System Requirements, Libraries, and Compilers . . . . .	80
9.2.2	Building the UFS Atmosphere . . . . .	81
9.2.3	Running the UFS Atmosphere Using the Regression Tests (RTs) . . . . .	82
<b>10</b>	<b>Tips for Adding a New Scheme</b>	<b>85</b>
<b>11</b>	<b>Acronyms</b>	<b>87</b>
<b>12</b>	<b>Glossary</b>	<b>89</b>
	<b>Index</b>	<b>91</b>

## CCPP OVERVIEW

**Warning:** The information in this document is up to date with the CCPP and GMTB SCM v3 public release as of June 17, 2019. If you are a developer looking for more current information, please obtain and build the up-to-date Technical Documentation from the master branch of the `ccpp-framework` code repository:

```
git clone https://github.com/NCAR/ccpp-framework
cd ccpp-framework/doc/CCPPtechnical
make html
make latexpdf
```

Ideas for this project originated within the Earth System Prediction Capability (ESPC) physics interoperability group, which has representatives from the US National Center for Atmospheric Research (NCAR), the Navy, National Oceanic and Atmospheric Administration (NOAA) Research Laboratories, NOAA National Weather Service, and other groups. Physics interoperability, or the ability to run a given physics *suite* in various host models, has been a goal of this multi-agency group for several years. An initial mechanism to run the physics of NOAA's Global Forecast System (GFS) model in other host models was developed by the NOAA Environmental Modeling Center (EMC) and later augmented by the NOAA Geophysical Fluid Dynamics Laboratory (GFDL). The *CCPP* expanded on that work by meeting additional requirements put forth by NOAA, and brought new functionalities to the physics-dynamics interface. Those include the ability to choose the order of parameterizations, to subcycle individual parameterizations by running them more frequently than other parameterizations, and to group arbitrary sets of parameterizations allowing other computations in between them (e.g., dynamics and coupling computations).

The architecture of the CCPP and its connection to a host model is shown in [Figure 1.1](#). There are two distinct parts to the CCPP: a library of physical parameterizations (*CCPP-Physics*) that conforms to selected standards and an infrastructure (*CCPP-Framework*) that enables connecting the physics to a host model.

The host model needs to have functional documentation for any variable that will be passed to or received from the physics. The *CCPP-Framework* is used to compare the variables requested by each physical *parameterization* against those provided by the host model<sup>1</sup>, and to check whether they are available, otherwise an error will be issued. This process serves to expose the variables passed between physics and dynamics, and to clarify how information is exchanged among parameterizations. During runtime, the CCPP-Framework is responsible for communicating the necessary variables between the host model and the parameterizations.

There are multiple options to build the CCPP (see more detail in [Chapter 3](#)). For example, with the CCPP dynamic build, all the CCPP-compliant parameterizations are compiled into a library which is linked to the host model at runtime. Conversely, with the CCPP static build, one or more physics suites are compiled into a library and linked to the host model when it is compiled. The dynamic build favors flexibility as users can select the parameterizations and their order at runtime, while the static build favors performance as it provides superior execution time and a smaller

<sup>1</sup> As of this writing, the CCPP has been validated with two host models: the Global Model Test Bed (GMTB) Single Column Model (SCM) and the atmospheric component of NOAA's Unified Forecast System (UFS) (hereafter the UFS Atmosphere) that utilizes the Finite-Volume Cubed Sphere (FV3) dycore. The CCPP can be utilized both with the global and standalone regional configurations of the UFS Atmosphere. Work is under way to connect and validate the use of the CCPP-Framework with NCAR and Navy models.

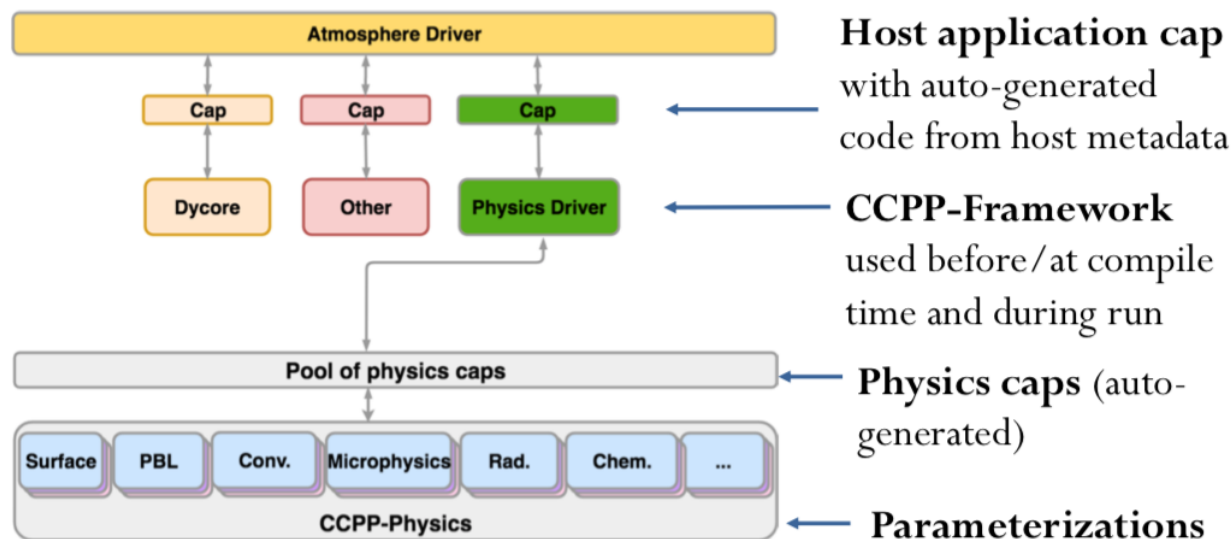


Fig. 1.1: Architecture of the CCPP and its connection to a host model, represented here as the driver for an atmospheric model (yellow box). The dynamical core (dycore), physics, and other aspects of the model (such as coupling) are connected to the driving host through the pool of physics caps. The CCPP-Physics is denoted by the gray box at the bottom of the physics, and encompasses the parameterizations, which are accompanied by physics caps.

memory footprint. The type of build defines several differences in the creation and use of the auto-generated code, many of which are not exposed to the user. The differences pertain to the interfaces between CCPP-Framework and the physics (physics caps) and the host model (host model cap), as well as in the procedures for calling the physics. In addition, the building option varies with choice of the host model. When the CCPP is connected to the SCM, the dynamic build option should be used. When the CCPP is used with the *UFS Atmosphere*, users have multiple building options and should choose the one that best matches their needs.

The *CCPP-Physics* contains the parameterizations and suites that are used operationally in the UFS Atmosphere, as well as parameterizations that are under development for possible transition to operations in the future. The CCPP aims to support the broad community while benefiting from the community. In such a CCPP ecosystem (Figure 1.2), the CCPP can be used not only by the operational centers to produce operational forecasts, but also by the research community to conduct investigation and development. Innovations created and effectively tested by the research community can be funneled back to the operational centers for further improvement of the operational forecasts.

Both the CCPP-Framework and the CCPP-Physics are developed as open source code, follow industry-standard code management practices, and are freely distributed through GitHub (<https://github.com/NCAR/ccpp-physics> and <https://github.com/NCAR/ccpp-framework>).

The first public release of the CCPP took place in April 2018 and included all the parameterizations of the operational GFS v14, along with the ability to connect to the SCM. The second public release of the CCPP took place in August 2018 and additionally included the physics suite tested for the implementation of GFS v15. Since then, the GFS v14 suite has been retired and additional parameterizations have been made CCPP-compliant in order to encompass the physics suites that are under consideration for GFS v16. The suites that are currently supported in the CCPP can be found in Table 1.1. In addition to the schemes listed, more schemes are under consideration for inclusion into the CCPP in the future.

The CCPP is governed by the groups that contribute to its development. The governance of the CCPP-Physics is currently led by NOAA, and the GMTB works with EMC and the Next Generation Global Prediction System (NGGPS) Program Office to determine which schemes and suites to be included and supported. The governance of the CCPP-Framework is jointly undertaken by NOAA and NCAR (see more information at <https://github.com/NCAR/ccpp-framework/wiki> and <https://dtcenter.org/gmtb/users/ccpp>). Please direct all inquiries to [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu).

### Common Community Physics Package (CCPP) Ecosystem

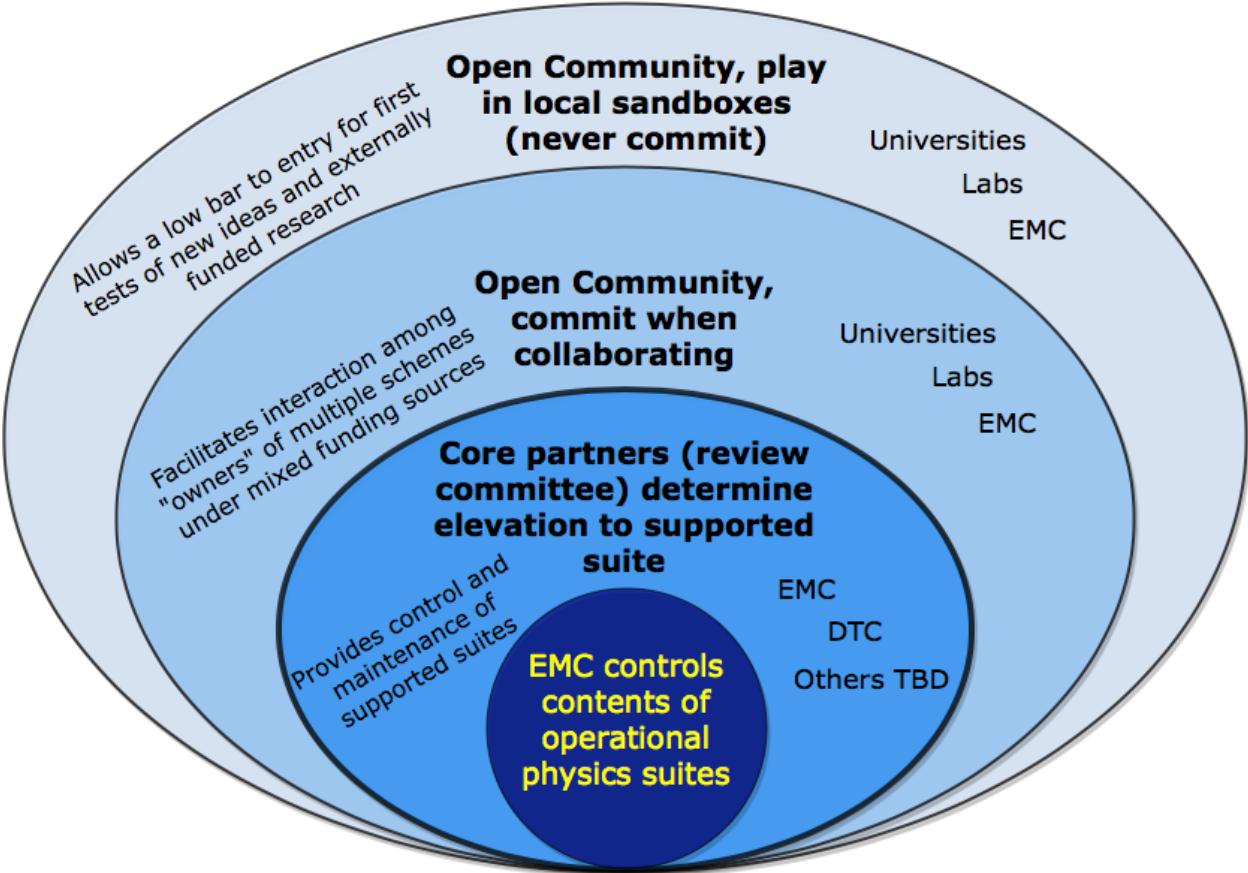


Fig. 1.2: CCPP ecosystem.

Table 1.1: Suites supported in the CCPP

	<b>GFS v15</b>	<b>GFS_v15plus</b>	<b>CPT_v0</b>	<b>GSD_v0</b>
Microphysics	GFDL	GFDL	aaMG3	aaThompson
PBL	EDMF	TKE EDMF	EDMF	MYNN
Deep convection	saSAS	saSAS	CSAW	GF
Shallow convection	saSAS	saSAS	saSAS	MYNN and GF
Radiation	RRTMG	RRTMG	RRTMG	RRTMG
Surface layer	GFS	GFS	GFS	GFS
Land surface	Noah	Noah	Noah	RUC
Ozone	NRL 2015	NRL 2015	NRL 2015	NRL 2015
H <sub>2</sub> O	NRL 2015	NRL 2015	NRL 2015	NRL 2015

The suites that are currently supported in the CCPP are listed in the first row. The types of parameterization are denoted in the first column, where H<sub>2</sub>O represent the stratospheric water vapor parameterization. The GFS\_v15 suite includes the GFDL microphysics, a Eddy-Diffusivity Mass Flux (EDMF) planetary boundary layer (PBL) scheme, scale-aware (sa) Simplified Arakawa-Schubert (SAS) convection, Rapid Radiation Transfer Model for General Circulation Models (RRTMG) radiation, the GFS surface layer, the Noah Land Surface Model (LSM), and the 2015 Navy Research Laboratory (NRL) ozone and stratospheric water vapor schemes. The other three suites are candidates for future operational implementations. The GFS\_v15plus suite is the same as the GFS\_v15 suite except using the Turbulent Kinetic Energy (TKE)-based EDMF PBL scheme. The Climate Process Team (CPT) v0 suite (CPT\_v0) uses the aerosol-aware (aa) Morrison-Gettleman 3 (MG3) microphysics scheme and Chikira-Sugiyama convection scheme with Arakawa-Wu extension (CSAW). The NOAA Global Systems Division (GSD) v0 suite (GSD\_v0) includes aaThompson microphysics, Mellor-Yamada-Nakanishi-Niino (MYNN) PBL and shallow convection, Grell-Freitas (GF) deep convection schemes, and the Rapid Update Cycle (RUC) LSM.

## 1.1 How to Use this Document

This document contains documentation for the Common Community Physics Package (CCPP). It describes the:

- physics schemes and interstitials
- suite definition files
- CCPP-compliant parameterizations
- adding a new scheme/suite
- host-side coding
- fundamentals of obtaining, pre-building, building and running the CCPP with NEMSFv3gfs
- CCPP code management and governance

For the latest version of the released code, please visit the [GMTB Website](#)

Please send questions and comments to the help desk: [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu)

This table describes the type changes and symbols used in this guide.

<b>Typeface or Symbol</b>	<b>Meaning</b>	<b>Example</b>
AaBbCc123	The names of commands, files, and directories; on-screen computer output	Edit your <code>.bashrc</code> Use <code>ls -a</code> to list all files. host\$ You have mail!



Following these typefaces and conventions, shell commands, code examples, namelist variables, etc. will be presented in this style:

```
mkdir ${TOP_DIR}
```



## CCPP-COMPLIANT PHYSICS PARAMETERIZATIONS

The rules for a scheme to be considered CCPP-compliant are summarized in this section. It should be noted that making a scheme CCPP-compliant is a necessary but not guaranteed step for the acceptance of the scheme in the pool of supported CCPP-Physics. Acceptance is dependent on scientific innovation, demonstrated value, and compliance with the rules described below. The criteria for acceptance of a scheme into the CCPP is under development.

It is recommended that parameterizations be comprised of the smallest units that will be used. For example, if a given set of deep and shallow convection schemes will always be called together and in a pre-established order, it is acceptable to group them within a single scheme. However, if one envisions that the deep and shallow convection schemes may someday operate independently, it is recommended to code two separate schemes to allow more flexibility.

Some schemes in the CCPP have been implemented using a driver as an entry point. In this context, a driver is defined as a wrapper that sits on top of the actual scheme and provides the CCPP entry points. In order to minimize the layers of code in the CCPP, the implementation of a driver is discouraged, that is, it is preferable that the CCPP be composed of atomic parameterizations. One example is the implementation of the MG microphysics, in which a simple entrypoint leads to two versions of the scheme, MG2 and MG3. A cleaner implementation would be to retire MG2 in favor of MG3, to put MG2 and MG3 as separate schemes, or to create a single scheme that can behave as MG2 and MG3 depending on namelist options.

The implementation of a driver is reasonable under the following circumstances:

- To preserve schemes that are also distributed outside of the CCPP. For example, the Thompson microphysics scheme is distributed both with the Weather Research and Forecasting (WRF) model and with the CCPP. Having a driver with CCPP directives allows the Thompson scheme to remain intact so that it can be synchronized between the WRF model and the CCPP distributions. See more in `mp_thompson_hrrr.F90` in the `ccpp-physics/physics` directory.
- To deal with optional arguments. A driver can check whether optional arguments have been provided by the host model to either write out a message and return an error code or call a subroutine with or without optional arguments. For example, see `mp_thompson_hrrr.F90`, `radsw_main.f`, or `radlw_main.f` in the `ccpp-physics/physics` directory.
- To perform unit conversions or array transformations, such as flipping the vertical direction and rearranging the index order, for example, `cu_gf_driver.F90` in the `ccpp-physics/physics` directory.

Schemes in the CCPP are classified into two categories: primary schemes and interstitial schemes. Primary schemes are the major parameterizations, such as PBL, microphysics, convection, radiation, surface layer parameterizations, etc. Interstitial schemes are modularized pieces of code that perform data preparation, diagnostics, or other “glue” functions and allow primary schemes to work together as a suite. They can be categorized as “scheme-specific” or “suite-level”. Scheme-specific interstitial schemes augment a specific primary scheme (to provide additional functionality). Suite-level interstitial schemes provide additional functionality on top of a class of primary schemes, connect two or more schemes together, or provide code for conversions, initializing sums, or applying tendencies, for example. The rules and guidelines provided in the following sections apply both to primary and interstitial schemes.

## 2.1 General Rules

A CCPP-compliant scheme is in the form of Fortran modules. *Listing 2.1* contains the template for a CCPP-compliant scheme (ccpp/framework/doc/DevelopersGuide/scheme\_template.F90), which includes four essential components: argument metadata tables, the `_init`, `_run`, and `_finalize` subroutines.

```

module scheme_template

  contains

  subroutine scheme_template_init ()
end subroutine scheme_template_init

  subroutine scheme_template_finalize()
end subroutine scheme_template_finalize

!> \section arg_table_scheme_template_run Argument Table
!! | local_name | standard_name | long_name |
↪units | rank | type | kind | intent | optional |
!! |-----|-----|-----|-----|-----|-----|
↪---|-----|-----|-----|-----|-----|
!! | errmsg | ccpp_error_message | error message for error handling in CCPP |
↪none | 0 | character | len=* | out | F |
!! | errflg | ccpp_error_flag | error flag for error handling in CCPP |
↪flag | 0 | integer | | out | F |
!!

  subroutine scheme_template_run (errmsg, errflg)

    implicit none

    !--- arguments
    ! add your arguments here
    character(len=*), intent(out) :: errmsg
    integer, intent(out) :: errflg

    !--- local variables
    ! add your local variables here

    continue

    !--- initialize CCPP error handling variables
    errmsg = ''
    errflg = 0

    !--- initialize intent(out) variables
    ! initialize all intent(out) variables here

    !--- actual code
    ! add your code here

    ! in case of errors, set errflg to a value != 0,
    ! assign a meaningful message to errmsg and return

    return

  end subroutine scheme_template_run

```

(continues on next page)

(continued from previous page)

```
end module scheme_template
```

Listing 2.1: Fortran template for a CCPP-compliant scheme showing an argument table and the `_init`, `_run`, and `_finalize` subroutines.

More details are found below:

- Each scheme must be in its own module and must include three (`_init`, `_run`, and `_finalize`) subroutines (entry points). The module name and the subroutine names must be consistent with the scheme name. The `_init` and `_finalize` subroutines are run automatically when the CCPP-Physics are initialized and finalized, respectively. These two subroutines may be called more than once, depending on the host model's parallelization strategy, and as such must be idempotent (the answer must be the same when the subroutine is called multiple times). The `_run` subroutine contains the code to execute the scheme.
- Each non-empty CCPP entrypoint subroutine requires a commented argument table (Listing 2.1). Empty subroutines do not require an argument table (e.g., `scheme_template_init` in Listing 2.1), since no variables need to be passed.
- The argument table contains the metadata of the variables required by the scheme. The table must precede the entry point subroutine (`_init`, `_run`, and `_finalize`) and must start with `!> \section arg_table_subroutine_name Argument Table` and end with a line containing only `!!`
- The current metadata attributes of a variable include `local_name`, `standard_name`, `long_name`, `units`, `rank`, `type`, `kind`, `intent`, and `optional` (see more in section *Doxygen Modules* of this chapter).
- If the width of an argument table exceeds 250 characters, the table should be wrapped in C preprocessor directives:

```
#if 0
!> \section arg_table_scheme_template_run Argument Table
!> ...
!!
#endif
```

- For better readability, the columns in the argument table are aligned.
- All external information required by the scheme must be passed in via the argument list. Statements such as `'use EXTERNAL_MODULE'` should not be used for passing in data and all physical constants should go through the argument list.
- Note that standard names, variable names, module names, scheme names and subroutine names are all case sensitive.
- Interstitial modules (`scheme_pre` and `scheme_post`) can be included if any part of the physics scheme must be executed before (`_pre`) or after (`_post`) the module `scheme` defined above. These situations are described in more detail in Section 5.1.

## 2.2 Input/output Variable (argument) Rules

- Variables available for CCPP physics schemes are identified by their unique `standard_name`. While an effort is made to comply with existing `standard_name` definitions of the Climate and Forecast (CF) conventions (<http://cfconventions.org>), additional names are used in the CCPP (see below for further information).
- A list of available standard names and an example of naming conventions can be found in `ccpp/framework/doc/DevelopersGuide/CCPP_VARIABLES_${HOST}.pdf`, where `${HOST}` is the name of the host model. Running the CCPP *prebuild* script (described in Chapter 8) will generate a LaTeX source file that can

be compiled to produce a PDF file with all variables defined by the host model and requested by the physics schemes.

- A `standard_name` cannot be assigned to more than one local variable (`local_name`). The `local_name` of a variable can be chosen freely and does not have to match the `local_name` in the host model.
- All variable information (units, rank, index ordering) must match the specifications on the host model side, but sub-slices can be used/added in the host model. For example, when using the UFS Atmosphere as the host model, tendencies are split in `GFS_typedefs.F90` so they can be used in the necessary physics scheme:

```
!! | IPD_Data(nb)%Intdiag%dt3dt(:, :, 1) | cumulative_change_in_temperature_due_to_
↪longwave_radiation
!! | IPD_Data(nb)%Intdiag%dt3dt(:, :, 2) | cumulative_change_in_temperature_due_to_
↪shortwave_radiation_and_orographic_gravity_wave_drag
!! | IPD_Data(nb)%Intdiag%dt3dt(:, :, 3) | cumulative_change_in_temperature_due_to_PBL
```

- The two mandatory variables that any scheme-related subroutine must accept as `intent(out)` arguments are `errmsg` and `errflg` (see also coding rules in [Section 2.3](#)).
- At present, only two types of variable definitions are supported by the CCPP-framework:
  - Standard Intrinsic Fortran variables are preferred (`character`, `integer`, `logical`, `real`). For character variables, the length should be specified as `*` in order to allow the host model to specify the corresponding variable with a length of its own choice. All others can have a `kind` attribute of a `kind` type defined by the host model.
  - Derived data types (DDTs). While the use of DDTs is discouraged, some use cases may justify their application (e.g. DDTs for chemistry that contain tracer arrays or information on whether tracers are advected). It should be understood that use of DDTs within schemes forces their use in host models and potentially limits a scheme’s portability. Where possible, DDTs should be broken into components that could be usable for another scheme of the same type.
- It is preferable to have separate variables for physically-distinct quantities. For example, an array containing various cloud properties should be split into its individual physically-distinct components to facilitate generality. An exception to this rule is if there is a need to perform the same operation on an array of otherwise physically-distinct variables. For example, tracers that undergo vertical diffusion can be combined into one array where necessary. This tactic should be avoided wherever possible, and is not acceptable merely as a convenience.
- If a scheme is to make use of CCPP’s subcycling capability, the loop counter can be obtained from CCPP as an `intent(in)` variable (see [Listing 6.2](#) for a mandatory list of variables that are provided by the CCPP-Framework and/or the host model for this and other purposes).

## 2.3 Coding Rules

- Code must comply to modern Fortran standards (Fortran 90/95/2003).
- Labeled `end` statements should be used for modules, subroutines and functions, for example, `module scheme_template → end module scheme_template`.
- Implicit variable declarations are not allowed. The `implicit none` statement is mandatory and is preferable at the module-level so that it applies to all the subroutines in the module.
- All `intent(out)` variables must be set inside the subroutine, including the mandatory variables `errflg` and `errmsg`.
- Decomposition-dependent host model data inside the module cannot be permanent, i.e. variables that contain domain-dependent data cannot be kept using the `save` attribute.
- `goto` statements are not allowed.

- common blocks are not allowed.
- Errors are handled by the host model using the two mandatory arguments `errmsg` and `errflg`. In the event of an error, a meaningful error message should be assigned to `errmsg` and set `errflg` to a value other than 0, for example:

```
write (errmsg, '(*(a))') 'Logic error in scheme xyz: ...'
errflg = 1
return
```

- Schemes are not allowed to abort/stop the program.
- Schemes are not allowed to perform I/O operations except for reading lookup tables or other information needed to initialize the scheme, including `stdout` and `stderr`. Diagnostic messages are tolerated, but should be minimal.
- Line lengths of up to 120 characters are suggested for better readability (exception: CCPP metadata argument tables).

Additional coding rules are listed under the *Coding Standards* section of the NOAA NGGPS Overarching System team document on Code, Data, and Documentation Management for NOAA Environmental Modeling System (NEMS) Modeling Applications and Suites (available at [https://docs.google.com/document/u/1/d/1bjnyJpJ7T3XeW3zCnhRLTL5a3m4\\_3XIAUeThUPWD9Tg/edit#heading=h.97v79689onyd](https://docs.google.com/document/u/1/d/1bjnyJpJ7T3XeW3zCnhRLTL5a3m4_3XIAUeThUPWD9Tg/edit#heading=h.97v79689onyd)).

## 2.4 Parallel Programming Rules

Most often shared memory (OpenMP: Open Multi-Processing) and MPI (Message Passing Interface) communication are done outside the physics in which case the physics looping and arrays already take into account the sizes of the threaded tasks through their input indices and array dimensions. The following rules should be observed when including OpenMP or MPI communication in a physics scheme:

- Shared-memory (OpenMP) parallelization inside a scheme is allowed with the restriction that the number of OpenMP threads to use is obtained from the host model as an `intent (in)` argument in the argument list (*Listing 6.2*).
- MPI communication is allowed in the `_init` and `_finalize` phase for the purpose of computing, reading or writing scheme-specific data that is independent of the host model's data decomposition. An example is the initial read of a lookup table of aerosol properties by one or more MPI processes, and its subsequent broadcast to all processes. Several restrictions apply:
  - The implementation of reading and writing of data must be scalable to perform efficiently from a few to millions of tasks.
  - The MPI communicator must be provided by the host model as an `intent (in)` argument in the argument list (*Listing 6.2*).
  - The use of `MPI_COMM_WORLD` is not allowed.
- Calls to MPI and OpenMP functions, and the import of the MPI and OpenMP libraries, must be guarded by C preprocessor directives as illustrated in the following listing. OpenMP pragmas can be inserted without C preprocessor guards, since they are ignored by the compiler if the OpenMP compiler flag is omitted.

```
#ifdef MPI
  use mpi
#endif
#ifdef OPENMP
  use omp_lib
#endif
...
```

(continues on next page)

(continued from previous page)

```

#ifdef MPI
  call MPI_BARRIER(mpicomm, ierr)
#endif

#ifdef OPENMP
  me = OMP_GET_THREAD_NUM()
#else
  me = 0
#endif

```

- For Fortran coarrays, consult with the GMTB helpdesk ([gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu)).

## 2.5 Scientific Documentation Rules

Technically, scientific documentation is not needed for a parameterization to work with the CCPP. However, scientific and technical documentation is important for code maintenance and for fostering understanding among stakeholders. As such, it is required of physics schemes in order to be included in the CCPP. This section describes the process used for documenting parameterizations in the CCPP. Doxygen was chosen as a tool for generating human-readable output due to its built-in functionality with Fortran, its high level of configurability, and its ability to parse inline comments within the source code. Keeping documentation with the source itself increases the likelihood that the documentation will be updated along with the underlying code. Additionally, inline documentation is amenable to version control.

The purpose of this section is to provide an understanding of how to properly document a physics scheme using doxygen inline comments. It covers what kind of information should be in the documentation, how to mark up the inline comments so that doxygen will parse it correctly, where to put various comments within the code, and how to configure and run doxygen to generate HTML (or other format) output. Part of this documentation, namely subroutine argument tables, have functional significance as part of the CCPP infrastructure. These tables must be in a particular format to be parsed by Python scripts that “automatically” generate a software cap for a given physics scheme. Although the procedure outlined herein is not unique, following it will provide a level of continuity with previous documented schemes.

Reviewing the documentation for CCPP parameterizations is a good way of getting started in writing documentation for a new scheme. The CCPP Scientific Documentation can be converted to html format (see [https://dtcenter.org/GMTB/v3.0/sci\\_doc/](https://dtcenter.org/GMTB/v3.0/sci_doc/)).

### 2.5.1 Doxygen Comments and Commands

All doxygen commands start with a backslash (“\”) or an at-sign (“@”). The doxygen inline comment blocks begin with “!>”, and subsequent lines begin with “!!”, which means that regular Fortran comments using “!” are not parsed by doxygen.

In the first line of each Fortran file, a brief one-sentence overview of the file purpose is present using the doxygen command “\file”:

```

!> \file gwdps.f
!! This file is the parameterization of orographic gravity wave
!! drag and mountain blocking.

```

A parameter definition begins with “!<”, where the sign ‘<’ just tells Doxygen that documentation follows. Example:



```
integer, parameter, public :: NF_VGAS = 10    !< number of gas species
integer, parameter       :: IMXCO2  = 24    !< input CO2 data longitude points
integer, parameter       :: JMXCO2  = 12    !< input CO2 data latitude points
integer, parameter       :: MINYEAR = 1957 !< earliest year 2D CO2 data available
```

## 2.5.2 Doxygen Documentation Style

To document a physics suite, a broad array of information should be included in order to serve both software engineering and scientific purposes. The documentation style could be divided into four categories:

- Doxygen Files
- Doxygen Pages (overview page and scheme pages)
- Doxygen Modules
- Bibliography

### Doxygen files

Doxygen provides the “\file” tag as a way to provide documentation on the Fortran source code file level. That is, in the generated documentation, one may navigate by source code filenames (if desired) rather than through a “functional” navigation. The most important documentation organization is through the “module” concept mentioned below, because the division of a scheme into multiple source files is often functionally irrelevant. Nevertheless, using a “\file” tag provides an alternate path to navigate the documentation and it should be included in every source file. Therefore, it is prudent to include a small documentation block to describe what code is in each file using the “\file” tag, e.g.:

```
!> \file gwdps.f
!! This file is the parameterization of orographic gravity wave
!! drag and mountain blocking.
```

The brief description for each file is displayed next to the source filename on the doxygen-generated “File List” page:

### Doxygen Overview Page

Pages in Doxygen can be used for documentation that is not directly attached to a source code entity such as file or module. They are external text files that generate pages with a high-level scientific overview and typically contain a longer description of a project or suite. You can refer to any source code entity from within the page.

The GMTB maintains a main page, created by the Doxygen command “\mainpage”, containing an overall description and background of the CCPP. Physics developers do not have to edit the file with the mainpage, which has a user-visible title, but not label:












```
/**
\mainpage Introduction
...
*/
```

All other pages listed under the main page are created using the Doxygen tag “\page” described in the next section. In any Doxygen page, you can refer to any entity of source code by using Doxygen tag “\ref” or “@ref”. Example in GFSv15\_suite.txt:

The GFS v15 physics suite uses the parameterizations in the following order, as defined in

## File List

Here is a list of all files with brief descriptions:

 <a href="#">calprecipitype.F90</a>	This file contains the subroutines that calculates dominant precipitation type
 <a href="#">funcphys.F90</a>	This file includes API for basic thermodynamic physics
 <a href="#">gfdl_cloud_microphys.F90</a>	This file contains the CCPP entry point for the column GFDL cloud microphysics ( Chen and Lin (2013) [13] )
 <a href="#">gfdl_fv_sat_adj.F90</a>	This file contains the fast saturation adjustment in the GFDL cloud microphysics. and it is an "intermediate physics" implemented in the remapping Lagrangian to Eulerian loop of FV3 solver
 <a href="#">GFS_MP_generic.F90</a>	This file contains the subroutines that calculate diagnostics variables before/after calling any microphysics scheme:
 <a href="#">gscond.f</a>	This file contains the subroutine that calculates grid-scale condensation and evaporation for use in Zhao and Carr (1997) [106] scheme
 <a href="#">gwdc.f</a>	This file is the original code for parameterization of stationary convection forced gravity wave drag based on Chun and Baik (1998) [18]
 <a href="#">gwdps.f</a>	This file is the parameterization of orographic gravity wave drag and mountain blocking
 <a href="#">h2ophys.f</a>	This file include NRL H2O physics for stratosphere and mesosphere
 <a href="#">mfpl.f</a>	This file contains the subroutine that calculates the updraft properties and mass flux for use in the Hybrid EDMF PBL scheme
 <a href="#">module_bfmicrophysics.f</a>	This file contains some subroutines used in microphysics

```
\c FV3_GFS_v15 :
- \ref fast_sat_adj
- \ref GFS_RRTMG
- \ref GFS_SFCLYR
- \ref GFS_NSST
- \ref GFS_NOAH
- \ref GFS_SFCSICE
- \ref GFS_HEDMF
- \ref GFS_GWDPS
- \ref GFS_RAYLEIGH
- \ref GFS_OZPHYS
- \ref GFS_H2OPHYS
- \ref GFS_SAMFdeep
- \ref GFS_GWDC
- \ref GFS_SAMFshal
- \ref GFDL_cloud
- \ref GFS_CALPRECIPTYPE
- \ref STOCHY_PHYS
```

The HTML result is [here](#). You can see that the “-” signs before “@ref” generate a list with bullets. Doxygen command “\c” displays its argument using a typewriter font.

## Physics Scheme Pages

Each major scheme in CCPP should have its own scheme page containing an overview of the parameterization. These pages are not tied to the Fortran code directly; instead, they are created with a separate text file that starts with the command “\page”. Scheme pages are stored in the `ccpp-physics/physics/docs/pdftxt` directory. Each page has a label (e.g., “GFS\_SAMFdeep” in the following example) and a user-visible title (“GFS Scale-Aware

Simplified Arakawa-Schubert (sa-SAS) Deep Convection Scheme” in the following example). It is noted that labels must be unique across the entire doxygen project so that the “\ref” command can be used to create an unambiguous link to the structuring element. It therefore makes sense to choose label names that refer to their context.

```
/**
\page GFS_SAMFdeep GFS Scale-Aware Simplified Arakawa-Schubert (sa-SAS) Deep
↳Convection Scheme
\section des_deep Description
The scale-aware mass-flux (SAMF) deep convection scheme is an
updated version of the previous Simplified Arakawa-Schubert (SAS) scheme
with scale and aerosol awareness and parameterizes the effect of deep
convection on the environment (represented by the model state variables)
in the following way ...

\section intra_deep Intraphysics Communication
\ref arg_table_samfdeepcnv_run

\section gen_al_deep General Algorithm
\ref general_samfdeep

*/
```

The physics scheme page will often describe the following:

1. Description section (“\section”), which usually includes:

- Scientific origin and scheme history (“\cite”)
- Key features and differentiating points
- A picture is worth a thousand words (“\image”)

To insert images into doxygen documentation, you’ll need to have your images ready in a graphical format, such as Portable Network Graphic (png), depending on which type of doxygen output you are planning to generate. For example, for LaTeX output, the images must be provided in Encapsulated PostScript (.eps), while for HTML output the images can be provided in the png format. Images are stored in `ccpp-physics/physics/docs/img` directory. Example of including an image for HTML output:

```
\image html gfdl_cloud_mp_diagram.png "Figure 1: GFDL MP at a glance (Courtesy of S.
↳J. Lin at GFDL)" width=10cm
```

2. Intraphysics Communication Section (“\section”)

The argument table for CCPP entry point subroutine `{scheme}_run` will be in this section. It is created by inserting a reference link (“\ref”) to the table in the Fortran code for the scheme.

3. General Algorithm Section (“\section”)

The general description of the algorithm will be in this section. It is created by inserting a reference link (“\ref”) in the Fortran code for the scheme.

The symbols “/\*\\*” and “\*/” need to be the first and last entries of the page. See an example of GFS Scale-Aware Simplified Arakawa-Schubert (sa-SAS) Deep Convection Scheme page in the previous page.

Note that separate pages can also be created to document something that is not a scheme. For example, a page could be created to describe a suite, or how a set of schemes work together. Doxygen automatically generates an index of all pages that is visible at the top-level of the documentation, thus allowing the user to quickly find, and navigate between, the available pages.

## Doxygen Modules

The CCPP documentation is based on doxygen modules (note this is not the same as Fortran modules). Each doxygen module pertains to a particular parameterization and is used to aggregate all code related to that scheme, even when it is in separate files. Since doxygen cannot know which files or subroutines belong to each physics scheme, each relevant subroutine must be tagged with the module name. This allows doxygen to understand your modularized design and generate the documentation accordingly. [Here](#) is a list of module list defined in CCPP.

A module is defined using:

```
!>\defgroup group_name group_title
```

Where `group_name` is the identifier and the `group_title` is what the group is referred to in the output. In the example below, we're defining a parent module "GFS radsw Main":

```
!> \defgroup module_radsw_main GFS radsw Main
!! This module includes NCEP's modifications of the RRTMG-SW radiation
!! code from AER.
!! ...
!!\author Eli J. Mlawer, emlawer@aer.com
!!\author Jennifer S. Delamere, jdelamer@aer.com
!!\author Michael J. Iacono, miacono@aer.com
!!\author Shepard A. Clough
!!\version NCEP SW v5.1 Nov 2012 -RRTMG-SW v3.8
!!
```

One or more contact persons should be listed with author. If you make significant modifications or additions to a file, consider adding an author and a version line for yourself. The above example generates the Author, Version sections on the page. All email addresses are converted to mailto hypertext links automatically:

**Author** Eli J. Mlawer, [emlawer@aer.com](mailto:emlawer@aer.com)

Jennifer S. Delamere, [jdelamer@aer.com](mailto:jdelamer@aer.com)

Michael J. Iacono, [miacono@aer.com](mailto:miacono@aer.com)

Shepard A. Clough

**Version** NCEP SW v5.1 Nov 2012 -RRTMG-SW v3.8

In order to include other pieces of code in the same module, the following tag must be used at the beginning of a comment block:

```
\ingroup group_name
```

For example:

```
!>\ingroup module_radsw_main
!> The subroutine computes the optical depth in band 16: 2600-3250
!! cm-1 (low - h2o, ch4; high - ch4)
!-----
!      subroutine taumol16
!.....
```

In the same comment block where a group is defined for a physics scheme, there should be some additional documentation. First, using the "`\\brief`" command, a brief one or two sentence description of the scheme should be included. After a blank doxygen comment line, begin the scheme origin and history using "`\\version`", "`\\author`" and "`\\date`".

Each subroutine that is a CCPP entry point to a parameterization, should be further documented with a documentation block immediately preceding its definition in the source. The documentation block should include at least the following components:

- A brief one- or two-sentence description with the "`\\brief`" tag
- A more detailed one or two paragraph description of the function of the subroutine
- An argument table that includes entries for each subroutine argument
  - The argument table content should be immediately preceded by the following line:

```
!!\section arg_table_SUBROUTINE_NAME
```

This line is also functional documentation used during the CCPP *prebuild* step. The first line of the table should contain the following “header” names

- a. `local_name`: contains the local subroutine variable name
- b. `standard_name`: CF-compliant standard name
- c. `long_name`: a short description
- d. `units`: format follows “unit exponent”, i.e.  $m^2 s^{-2}$  for `m:sup:2/s:sup:2`
- e. `rank`: 0 for scalar, 1 for 1-D array, 2 for 2-D array, etc.
- f. `type`: integer, real, logical, character, DDT, etc.
- g. `kind`: the specified floating point precision kind (at present, to be extended to different integer kinds in the future)
- h. `intent`: in, out, inout
- i. `optional`: T/F

The argument table should be immediately followed by a blank doxygen line “!!”, which is needed to denote the end of an argument table. Here is an example :

```
!! \section arg_table_scheme_X_run Argument Table
!! | local_name | standard_name | long_name
↪ | | | | | | | | |
!! |-----|-----|-----|-----|-----|-----|
↪-----|-----|-----|-----|-----|-----|
!! | im | horizontal_loop_extent | horizontal loop extent
↪ | | | 0 | integer | | in | F |
!! | levs | vertical_dimension | vertical layer dimension
↪ | | | 0 | integer | | in | F |
```

The order of arguments in the table does not have to match the order of actual arguments in the subroutine.

- A section called “General Algorithm” with a bullet or numbered list of the tasks completed in the subroutine algorithm
- At the end of initial subroutine documentation block, a “Detailed algorithm” section is started and the entirety of the code is encompassed with the “`!> @{}`” and “`!> @}`” delimiters. This way, any comments explaining detailed aspects of the code are automatically included in the “Detailed Algorithm” section.

For subroutines that are not a CCPP entry point to a scheme, no argument table is required. But it is suggested that following “`\\ingroup`” and “`\\brief`”, use “`\\param`” to define each argument with local name, a short description and unit, i.e.,

```

!> \ingroup HEDMF
!! \brief This subroutine is used for calculating the mass flux and updraft_
↪properties.
!! ...
!!
!! \param[in] im      integer, number of used points
!! \param[in] ix      integer, horizontal dimension
!! \param[in] km      integer, vertical layer dimension
!! \param[in] ntrac   integer, number of tracers
!! \param[in] delt    real, physics time step
!! ...
!! \section general_mfpbl mfpbl General Algorithm
!! -# Determine an updraft parcel's entrainment rate, buoyancy, and vertical_
↪velocity.
!! -# Recalculate the PBL height ...
!! -# Calculate the mass flux profile and updraft properties.
!! \section detailed_mfpbl mfpbl Detailed Algorithm
!> @{
    subroutine mfpbl(im,ix,km,ntrac,delt,cnvflg,          &
        &  z1,zm,thvx,q1,t1,u1,v1,hpbl,kpbl,          &
        &  sflx,ustar,wstar,xf,tcko,qcko,ucko,vcko)
        ...
    end subroutine mfpbl
!> @}

```

## Bibliography

Doxygen can handle in-line paper citations and link to an automatically created bibliography page. The bibliographic data for any papers that are cited need to be put in BibTeX format and saved in a .bib file. The bib file for CCPP is included in the repository, and the doxygen configuration option `cite_bib_files` points to the included file.

Citations are invoked with the following tag:

```
\cite bibtex_key_to_paper
```

## Equations

See [link](#) for information about including equations. For the best rendering, the following option should be set in the Doxygen configuration file:

```

USE_MATHJAX           = YES
MATHJAX_RELPATH       = https://cdnjs.cloudflare.com/ajax/libs/mathjax/2.7.2

```

There are many great online resources to use the LaTeX math typesetting used in doxygen.

## 2.5.3 Doxygen Configuration

### Configuration File

The CCPP contains a doxygen configuration file `./ccpp/physics/physics/docs/ccpplatex_dox`, such that you don't need to create an additional one.

If starting from scratch, you can generate a default configuration file using the command:

```
doxygen -g <config_file>
```

Then you can edit the default configuration file to serve your needs. The default file includes plenty of comments to explain all the options. Some of the important things you need to pay attention to are:

- The name of your project:

```
PROJECT_NAME = 'your project name'
```

- The input files (relative to the directory where you run doxygen):

```
INPUT =
```

The following lines should be listed here: the doxygen mainpage text file, the scheme pages, and the source codes to be contained in the output. The order in which schemes are listed determines the order in the html result.

- The directory where to put the documentation (if you leave it empty, then the documentation will be created in the directory where you run doxygen):

```
OUTPUT_DIRECTORY = doc
```

- The type of documentation you want to generate (HTML, LaTeX and/or something else):

```
GENERATE_HTML = YES
```

If HTML is chosen, the following tells doxygen where to put the html documentation relative OUTPUT\_DIRECTORY:

```
HTML_OUTPUT = html
```

and

```
HTML_FILE_EXTENSION = .html
```

tells what the extension of the html files should be.

- Other important settings for a Fortran code project are:

```
OPTIMIZE_FOR_FORTRAN      = YES
EXTENSION_MAPPING         = .f=FortranFree      \
                          .F90=FortranFree     \
                          .f90=FortranFree
LAYOUT_FILE               = ccpp_dox_layout.xml
CITE_BIB_FILES            = library.bib
FILE_PATTERN              = *.f              \
                          *.F90            \
                          *.f90            \
                          *.txt
GENERATE_TREEVIEW         = yes
```

Doxygen files for layout (`ccpp_dox_layout.xml`), a html style (`ccpp_dox_extra_style.css`), and bibliography (`library.bib`) are provided with the CCPP. Additionally, a configuration file is supplied, with the following variables modified from the default:

## Diagrams

On its own, Doxygen is capable of creating simple text-based class diagrams. With the help of the additional software GraphViz, Doxygen can generate additional graphics-based diagrams, optionally in Unified Modeling Language

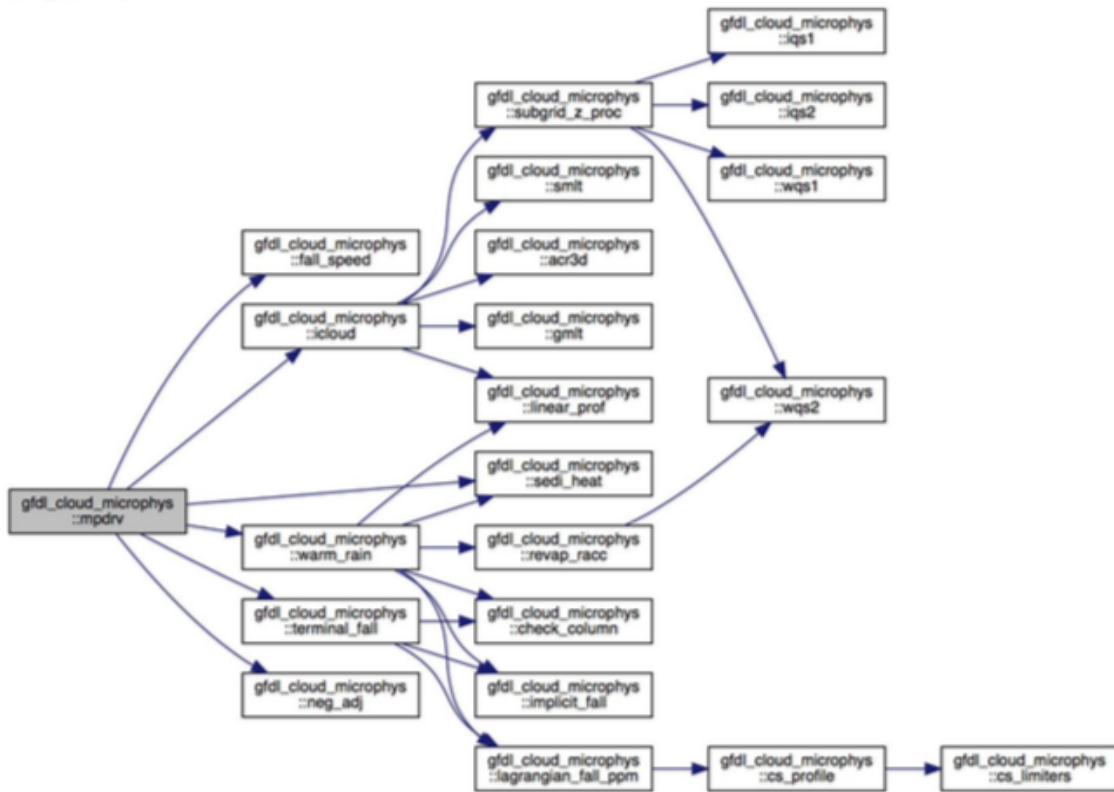
(UML) style. To enable GraphViz support, the configure file parameter "HAVE\_DOT" must be set to "YES".

You can use doxygen to create call graphs of all the physics schemes in CCPP. In order to create the call graphs you will need to set the following options in your doxygen config file:

```
HAVE_DOT           = YES
EXTRACT_ALL       = YES
EXTRACT_PRIVATE   = YES
EXTRACT_STATIC    = YES
CALL_GRAPH        = YES
```

Note that will need the DOT (graph description language) utility to be installed when starting doxygen. Doxygen will call it to generate the graphs. On most distributions the DOT utility can be found in the GraphViz package. Here is the call graph for subroutine *mpdrv* in GFDL cloud microphysics generated by doxygen:

Here is the call graph for this function:



### 2.5.4 Using Doxygen

In order to generate doxygen-based documentation, one needs to follow four steps:

1. Have the doxygen executable installed on your computer. For example, for the NOAA machine Theia, add the following line into `.cshrc` file in your home directory:

```
alias doxygen/scratch4/BMC/gmtb/doxygen-1.8.10/bin/doxygen
```

Source your `.cshrc` file.

2. Document your code, including doxygen main page, scheme pages and inline comments within source code as described above.



3. Prepare a Bibliography file in BibTeX format for papers referred to in the physics suites.
4. Create or edit a doxygen configuration file to control what doxygen pages, source files and bibliography file get parsed, how the source files get parsed, and to customize the output.
5. Run the doxygen command from the command line with the doxygen configuration file given as an argument:

```
$doxygen $PATH_TO_CONFIG_FILE/<config_file>
```

Running this command may generate warnings or errors that need to be fixed in order to produce proper output. The location and type of output (HTML, LaTeX, etc.) are specified in the configuration file. The generated HTML documentation can be viewed by pointing an HTML browser to the `index.html` file in the `./docs/doc/html/` directory.

For precise instructions on creating the scientific documentation, contact the GMTB helpdesk at [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu).



## CCPP CONFIGURATION AND BUILD OPTIONS

While the *CCPP-Framework* code can be compiled independently, the *CCPP-Physics* code can only be used within a host modeling system that provides the variables and the kind, type, and DDT definitions. As such, it is advisable to integrate the CCPP configuration and build process with the host model's. Part of the build process, known as the *prebuild* step since it precedes compilation, involves running a Python script that performs multiple functions. These functions include configuring the *CCPP-Physics* for use with the host model and autogenerating FORTRAN code to communicate variables between the physics and the dynamical core. The *prebuild* step will be discussed in detail in [Chapter 8](#).

There are some differences between building and running the SCM and the UFS Atmosphere. In the case of the UFS Atmosphere as the host model, there are several build options ([Figure 3.1](#)). The choice can be specified through command-line options supplied to the `compile.sh` script for manual compilation or through a regression test (RT) configuration file. Detailed instructions for building the code are discussed in [Chapter 9](#).

The relevant options for building CCPP with the UFS Atmosphere can be described as follows:

- **Without CCPP (CCPP=N):** The code is compiled without CCPP and runs using the original UFS Atmosphere physics drivers, such as `GFS_physics_driver.F90`. This option entirely bypasses all CCPP functionality and is only used for RT against the unmodified UFS Atmosphere codebase.
- **With CCPP (CCPP=Y):** The code is compiled with CCPP enabled and restricted to CCPP-compliant physics. That is, any parameterization to be called as part of a suite must be available in CCPP. Physics scheme selection and order is determined at runtime by an external suite definition file (SDF; see [Chapter 4](#) for further details on the SDF). The existing physics-calling code `GFS_physics_driver.F90` and `GFS_radiation_driver.F90` are bypassed altogether in this mode and any additional code needed to connect parameterizations within a suite previously contained therein is executed from the so-called CCPP-compliant “interstitial schemes”. One further option determines how the CCPP-compliant physics are called within the host model:
  - **Dynamic CCPP (STATIC=N):** This option is recommended for research and development users, since it allows choosing any physics schemes within the CCPP library at runtime by making adjustments to the CCPP SDF and the model namelist. This option carries computational overhead associated with the higher level of flexibility. Note that the *CCPP-Framework* and *CCPP-physics* are dynamically linked to the executable.
  - **Static CCPP (STATIC=Y):** The code is compiled with CCPP enabled but restricted to CCPP-compliant physics defined by one or more SDFs used at compile time. This option is recommended for users interested in production-mode and operational applications, since it limits flexibility in favor of runtime performance and memory footprint. Note that the *CCPP-Framework* and *CCPP-physics* are statically linked to the executable.

Conversely, for the SCM, only the dynamic CCPP build is supported.

For all options that activate the CCPP, the `ccpp_prebuild.py` Python script must be run. This may be done manually or as part of a host model build-time script. In the case of the SCM, `ccpp_prebuild.py` must be run manually,

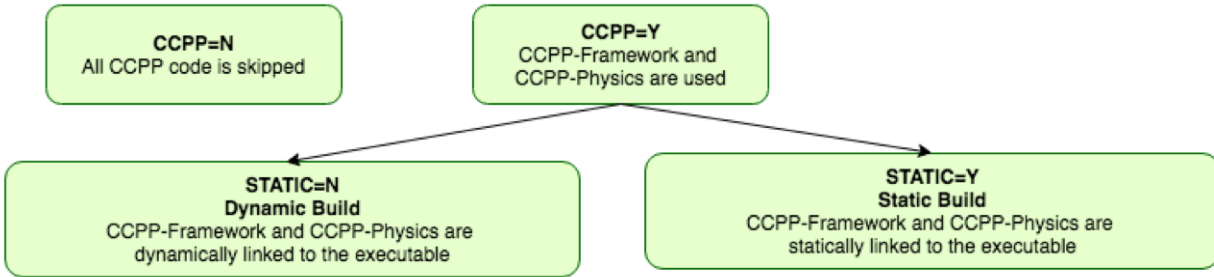


Fig. 3.1: Options for building the CCPP with the UFS Atmosphere. Runs where computational performance is valued over physics flexibility should use `STATIC=Y`.

as it is not incorporated in that model's build system. In the case of the UFS Atmosphere, `ccpp_prebuild.py` is run automatically as a step in the build system, although it can be run manually for debugging purposes.

The path to a host-model specific configuration file is the only required argument to `ccpp_prebuild.py`. Such files are included with the SCM and NEMSFv3gfs repositories, and must be included with the code of any host model to use the CCPP. Figure 3.2 depicts the main functions of the `ccpp_prebuild.py` script for the dynamic build. Using information included in the configuration file, the script reads metadata associated with all physics variables provided by the host model and metadata associated with all physics variables requested by all physics schemes in the CCPP (at least those whose entry/exit point source files are listed in the configuration file). The script matches the variables and autogenerates software caps for each physics scheme entry point so that they are available to be called by the host model, depending on the contents of a SDF provided at runtime.

In case of a static build, the basic function of the script is the same, but the output (namely, autogenerated software caps) take a different, more computationally-efficient approach. For a general idea of how the static option is different, compare Figure 3.3 for the static case with Figure 3.2 for the dynamic case. As mentioned, during the build step, one or more SDFs are used as an additional input. The script parses the SDF(s) and only matches provided/requested variables that are used within the particular physics suite(s). Rather than autogenerate software caps for each physics scheme (regardless of request), in static mode, the script autogenerates software caps for the physics suite(s) as a whole and for each physics group as defined in the SDF(s). At runtime, a single SDF is used to select the suite that will be executed in the run. This arrangement allows for more efficient variable recall (which is done once for all physics schemes within each group of a suite), leads to a reduced memory footprint of the CCPP compared to the dynamic option and speeds up execution to be on par with physics called using the existing physics-calling code.

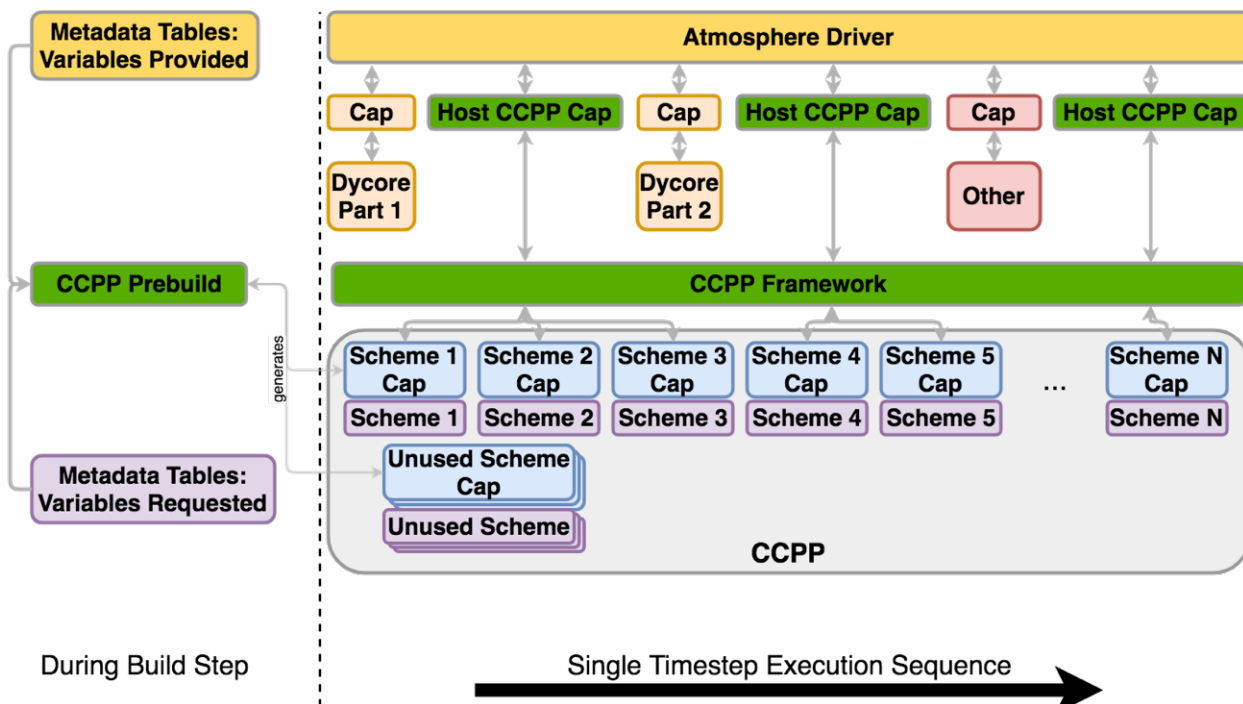


Fig. 3.2: This figure depicts one generalized example of the interaction between an atmospheric model and CCPP-Physics for one timestep using the **dynamic** build, with execution progressing toward the right. The “Atmosphere Driver” box represents model superstructure code, perhaps responsible for I/O, time-stepping, and other model component interactions. In this example, physics code must be executed between portions of the dycore code and another model component. The CCPP host cap allows multiple calls to groups of physics schemes defined in a SDF. The CCPP-Framework executes each physics group consisting of a collection of schemes and their associated software caps. The left side of the figure depicts the function of the `ccpp_prebuild.py` script, matching variables from the host model with CCPP-compliant schemes and autogenerating software caps for each scheme (regardless of its use within a particular suite).

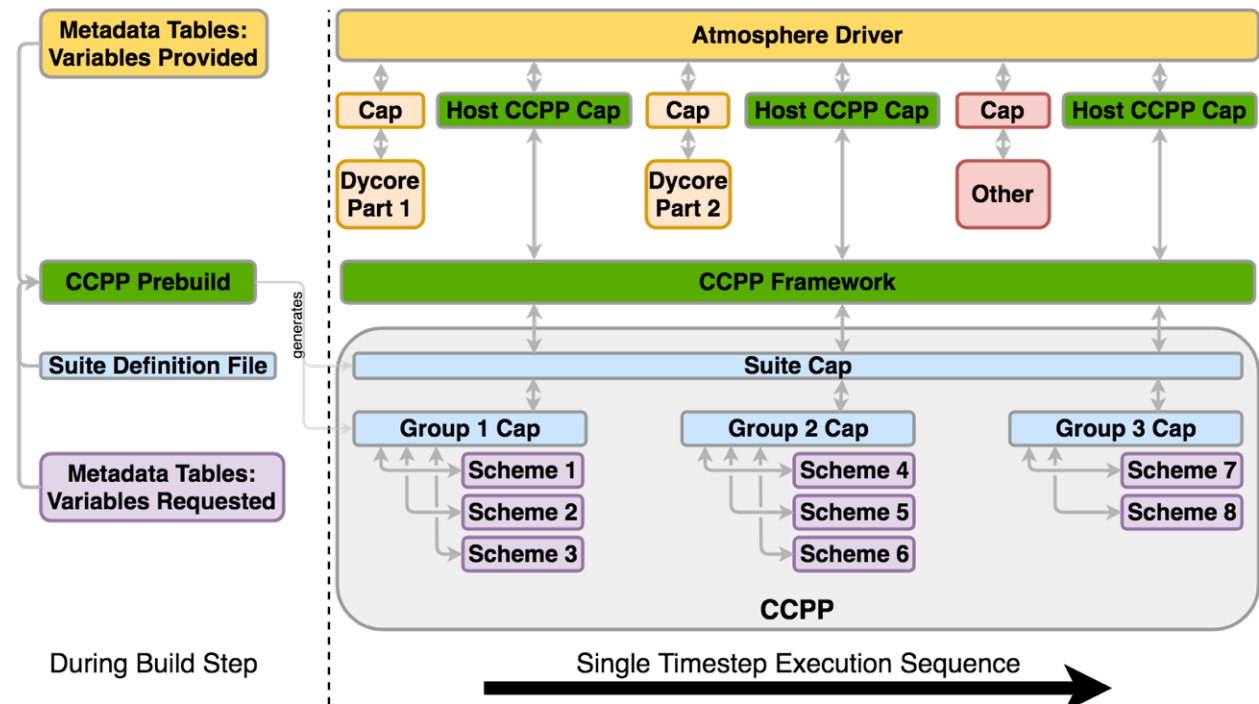


Fig. 3.3: As in Figure 3.2 , but for the **static** build using a single SDF. In this case, software caps are autogenerated for the suite and physics groups (defined in the SDF provided to the `ccpp_prebuild.py` script) rather than for individual schemes. The suite must be defined via the SDF at prebuild time. When multiple SDFs are provided during the build step, multiple suite caps and associated group caps are produced, but only one is used at runtime.

## CONSTRUCTING SUITES

### 4.1 Suite Definition File

The *SDF* is a file in XML format used to specify the name of the suite, the physics schemes to run, groups of physics that run together, the order in which to run the physics, and whether subcycling will be used to run any of the parameterizations with shorter timesteps. The *SDF* files are located in `ccpp/suites/suite_*.xml`.

In addition to the primary parameterization categories (such as radiation, boundary layer, deep convection, resolved moist physics, etc.), the *SDF* can have an arbitrary number of interstitial schemes in between the parameterizations to preprocess or postprocess data. In many models, this interstitial code is not obvious to the model user but, with the *SDF*, both the primary parameterizations and the interstitial schemes are listed explicitly.

The name of the suite is listed at the top of the *SDF* and must be consistent with the name of the *SDF*: file `suite_ABC.xml` contains `suite name='ABC'`, as in the example below. The suite name is followed by the `time_vary` step, which is run only once when the model is first initialized.

```
<suite name="ABC" lib="ccppphys" ver="3.0.0">
  <!-- <init></init> -->
  <group name="time_vary">
    <subcycle loop="1">
      <scheme>GFS_time_vary_pre</scheme>
      <scheme>GFS_rrtmg_setup</scheme>
      <scheme>GFS_rad_time_vary</scheme>
      <scheme>GFS_phys_time_vary</scheme>
    </subcycle>
  </group>
```

#### 4.1.1 Groups

The concept of grouping physics in the *SDF* (reflected in the `<group name="XYZ">` elements) enables “groups” of parameterizations to be called with other computation (such as related to the dycore, I/O, etc.) in between. One can edit the groups to suit the needs of the host application. For example, if a subset of physics schemes needs to be more tightly connected with the dynamics and called more frequently, one could create a group consisting of that subset and place a `ccpp_run` call in the appropriate place in the host application. The remainder of the parameterization groups could be called using `ccpp_run` calls in a different part of the host application code.

#### 4.1.2 Subcycling

The *SDF* allows subcycling of schemes, or calling a subset of schemes at a smaller time step than others. The `<subcycle loop = n>` element in the *SDF* controls this function. All schemes within such an element are called

n times during one `ccpp_run` call. An example of this is found in the `FV3_GFS_v15.xml SDF`, where the surface schemes are executed twice for each timestep (implementing a predictor/corrector paradigm):

```
<!-- Surface iteration loop -->
<subcycle loop="2">
  <scheme>sfc_diff</scheme>
  <scheme>GFS_surface_loop_control_part1</scheme>
  <scheme>sfc_nst_pre</scheme>
  <scheme>sfc_nst</scheme>
  <scheme>sfc_nst_post</scheme>
  <scheme>lsm_noah</scheme>
  <scheme>sfc_sice</scheme>
  <scheme>GFS_surface_loop_control_part2</scheme>
</subcycle>
```

Note that currently no time step information is included in the *SDF* and that the subcycling of schemes resembles more an iteration over schemes with the loop counter being available as integer variable with standard name `ccpp_loop_counter`. If subcycling is used for a set of parameterizations, the smaller time step must be an input argument for those schemes.

### 4.1.3 Order of Schemes

Schemes may be interdependent and the order in which the schemes are run may make a difference in the model output. For the static build, reading the *SDF*(s) and defining the order of schemes for each suite happens at compile time. For the dynamic build, no *SDF* is used at compile time (Figure 4.1). Instead, at runtime the *SDF* is read and the order of schemes is determined.

Some schemes require additional interstitial code that must be run before or after the scheme and cannot be part of the scheme itself. This can be due to dependencies on other schemes and/or the order of the schemes as determined in the *SDF*.

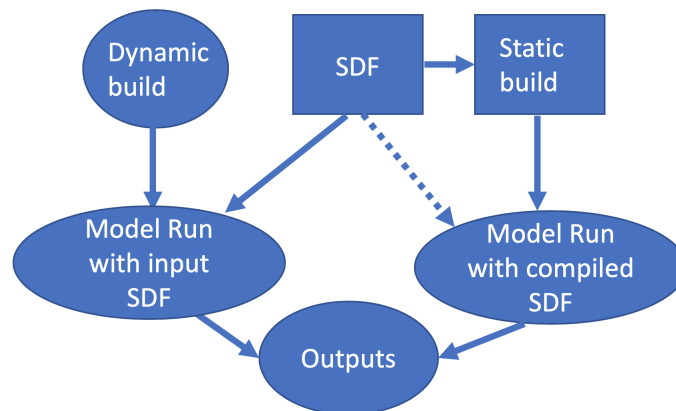


Fig. 4.1: Schematic of the use of the *SDF* in dynamic and static builds. Note that, for the static build, more than one *SDF* can be supplied at compile time, but only one can be used at runtime.

## 4.2 Interstitial Schemes

The *SDF* can have an arbitrary number of additional interstitial schemes in between the primary parameterizations to preprocess or postprocess data. There are two main types of interstitial schemes, scheme-specific and suite-level. The



scheme-specific interstitial scheme is needed for one specific scheme and the suite-level interstitial scheme processes data that are relevant for various schemes within a suite.

## 4.3 SDF Examples

### 4.3.1 Simplest Case: Single Group and no Subcycling

Consider the simplest case, in which all physics schemes are to be called together in a single group with no subcycling (i.e. `subcycle loop="1"`). The *SDF* `suite_Suite_A.xml` could contain the following:

```
<?xml version="1.0" encoding="UTF-8"?>
<suite name="Suite_A" lib="ccppphys" ver="3.0.0">
  ...
  <group name="physics">
    <subcycle loop="1">
      <scheme>Suite_A_interstitial_1</scheme>
      <scheme>scheme_1_pre</scheme>
      <scheme>scheme_1</scheme>
      <scheme>scheme_1_post</scheme>
      <scheme>scheme_2_generic_pre</scheme>
      <scheme>scheme_2</scheme>
      <scheme>scheme_2_generic_post</scheme>
      <scheme>Suite_A_interstitial_2</scheme>
      <scheme>scheme_3</scheme>
      ...
      <scheme_n</scheme>
    </subcycle>
  </group>
</suite>
```

Note the syntax of the *SDF* file. The root (the first element to appear in the xml file) is the `suite` with the name of the suite given as an attribute. In this example, the suite name is `Suite_A`. Within each suite are groups, which specify a physics group to call (i.e. `physics`, `fast_physics`, `time_vary`, `radiation`, `stochastics`). Each group has an option to subcycle. The value given for `loop` determines the number of times all of the schemes within the `subcycle` element are called. Finally, the `scheme` elements are children of the `subcycle` elements and are listed in the order they will be executed. In this example, `scheme_1_pre` and `scheme_1_post` are scheme-specific preprocessing and postprocessing interstitial schemes, respectively. The suite-level preprocessing and postprocessing interstitial schemes `scheme_2_generic_pre` and `scheme_2_generic_post` are also called in this example. `Suite_A_interstitial_2` is a scheme for `suite_A` and connects various schemes within this suite.

### 4.3.2 Case with Multiple Groups

Some models require that the physics be called in groups, with non-physics computations in-between the groups.

```
<?xml version="1.0" encoding="UTF-8"?>
<suite name="Suite_B" lib="ccppphys" ver="3.0.0">
  <group name="g1">
    <subcycle loop="1">
      <scheme>SchemeX</scheme>
      <scheme>SchemeY</scheme>
      <scheme>SchemeZ</scheme>
    </subcycle>
```

(continues on next page)

(continued from previous page)

```

</group>
<group name="g2">
  <subcycle loop="1">
    <scheme>SchemeA</scheme>
    <scheme>SchemeB</scheme>
    <scheme>SchemeC</scheme>
  </subcycle>
</group>
</suite>

```

### 4.3.3 Case with Subcycling

Consider the case where a model requires that some subset of physics be called on a smaller time step than the rest of the physics, e.g. for computational stability. In this case, one would make use of the subcycle element as follows:

```

<?xml version="1.0" encoding="UTF-8"?>
<suite name="Suite_C" lib="ccppphys" ver="3.0.0">
  <group name="g1">
    <subcycle loop="1">
      <scheme>scheme_1</scheme>
      <scheme>scheme_2</scheme>
    </subcycle>
    <subcycle loop="2">
      <!-- block of schemes 3 and 4 is called twice -->
      <scheme>scheme_3</scheme>
      <scheme>scheme_4</scheme>
    </subcycle>
  </group>
</suite>

```

### 4.3.4 Operational GFS v15 Suite

Here is the *SDF* for the physics suite equivalent to the operational GFS v15 in the *UFS* Atmosphere, which employs various groups and subcycling:

```

<?xml version="1.0" encoding="UTF-8"?>
<suite name="FV3_GFS_v15" lib="ccppphys" ver="3.0.0">
  <!-- <init></init> -->
  <group name="fast_physics">
    <subcycle loop="1">
      <scheme>fv_sat_adj</scheme>
    </subcycle>
  </group>
  <group name="time_vary">
    <subcycle loop="1">
      <scheme>GFS_time_vary_pre</scheme>
      <scheme>GFS_rrtmg_setup</scheme>
      <scheme>GFS_rad_time_vary</scheme>
      <scheme>GFS_phys_time_vary</scheme>
      <scheme>stochastic_physics</scheme>
      <scheme>stochastic_physics_sfc</scheme>
    </subcycle>
  </group>
</suite>

```

(continues on next page)

(continued from previous page)

```

</group>
<group name="radiation">
  <subcycle loop="1">
    <scheme>GFS_suite_interstitial_rad_reset</scheme>
    <scheme>GFS_rrtmg_pre</scheme>
    <scheme>rrtmg_sw_pre</scheme>
    <scheme>rrtmg_sw</scheme>
    <scheme>rrtmg_sw_post</scheme>
    <scheme>rrtmg_lw_pre</scheme>
    <scheme>rrtmg_lw</scheme>
    <scheme>rrtmg_lw_post</scheme>
    <scheme>GFS_rrtmg_post</scheme>
  </subcycle>
</group>
<group name="physics">
  <subcycle loop="1">
    <scheme>GFS_suite_interstitial_phys_reset</scheme>
    <scheme>GFS_suite_stateout_reset</scheme>
    <scheme>get_prs_fv3</scheme>
    <scheme>GFS_suite_interstitial_1</scheme>
    <scheme>dcyc2t3</scheme>
    <scheme>GFS_surface_generic_pre</scheme>
    <scheme>GFS_surface_composites_pre
    <scheme>GFS_suite_interstitial_2</scheme>
  </subcycle>
  <!-- Surface iteration loop -->
  <subcycle loop="2">
    <scheme>sfc_diff</scheme>
    <scheme>GFS_surface_loop_control_part1</scheme>
    <scheme>sfc_nst_pre</scheme>
    <scheme>sfc_nst</scheme>
    <scheme>sfc_nst_post</scheme>
    <scheme>lsm_noah</scheme>
    <scheme>sfc_sice</scheme>
    <scheme>GFS_surface_loop_control_part2</scheme>
  </subcycle>
  <!-- End of surface iteration loop -->
  <subcycle loop="1">
    <scheme>GFS_surface_composites_post
    <scheme>dcyc2t3_post</scheme>
    <scheme>sfc_diag</scheme>
    <scheme>sfc_diag_post</scheme>
    <scheme>GFS_surface_generic_post</scheme>
    <scheme>GFS_PBL_generic_pre</scheme>
    <scheme>hedmf</scheme>
    <scheme>GFS_PBL_generic_post</scheme>
    <scheme>gwdps_pre</scheme>
    <scheme>gwdps</scheme>
    <scheme>gwdps_post</scheme>
    <scheme>rayleigh_damp</scheme>
    <scheme>GFS_suite_stateout_update</scheme>
    <scheme>ozphys_2015</scheme>
    <scheme>h2ophys</scheme>
    <scheme>GFS_DCNV_generic_pre</scheme>
    <scheme>get_phi_fv3</scheme>
    <scheme>GFS_suite_interstitial_3</scheme>
    <scheme>samfdeepcnv</scheme>
  </subcycle>

```

(continues on next page)

(continued from previous page)

```

    <scheme>GFS_DCNV_generic_post</scheme>
    <scheme>gwdc_pre</scheme>
    <scheme>gwdc</scheme>
    <scheme>gwdc_post</scheme>
    <scheme>GFS_SCNV_generic_pre</scheme>
    <scheme>samfshalcnv</scheme>
    <scheme>samfshalcnv_post</scheme>
    <scheme>GFS_SCNV_generic_post</scheme>
    <scheme>GFS_suite_interstitial_4</scheme>
    <scheme>cnvc90</scheme>
    <scheme>GFS_MP_generic_pre</scheme>
    <scheme>gfdl_cloud_microphys</scheme>
    <scheme>GFS_MP_generic_post</scheme>
    <scheme>sfc_sice_post</scheme>
    <scheme>maximum_hourly_diagnostics</scheme>
  </subcycle>
</group>
<group name="stochastics">
  <subcycle loop="1">
    <scheme>GFS_stochastics</scheme>
  </subcycle>
</group>
<!-- <finalize></finalize> -->
</suite>

```

The suite name is `FV3_GFS_v15`. Five groups (`fast_physics`, `time_vary`, `radiation`, `physics`, and `stochastics`) are used, because the physics needs to be called in different parts of the host model. The detailed explanation of each primary physics scheme can be found in scientific documentation. A short explanation of each scheme is below.

- `fv_sat_adj`: Saturation adjustment (for the UFS Atmosphere only)
- `GFS_time_vary_pre`: GFS physics suite time setup
- `GFS_rrtmg_setup`: Rapid Radiative Transfer Model for Global Circulation Models (RRTMG) setup
- `GFS_rad_time_vary`: GFS radiation time setup
- `GFS_phys_time_vary`: GFS physics suite time setup
- `stochastic_physics`: Stochastic physics
- `stochastic_physics_sfc`: Surface part of stochastic physics
- `GFS_suite_interstitial_rad_reset`: GFS suite interstitial radiation reset
- `GFS_rrtmg_pre`: Preprocessor for the GFS radiation schemes
- `rrtmg_sw_pre`: Preprocessor for the RRTMG shortwave radiation
- `rrtmg_sw`: RRTMG for shortwave radiation
- `rrtmg_sw_post`: Postprocessor for the RRTMG shortwave radiation
- `rrtmg_lw_pre`: Preprocessor for the RRTMG longwave radiation
- `rrtmg_lw`: RRTMG for longwave radiation
- `rrtmg_lw_post`: Postprocessor for the RRTMG longwave radiation
- `GFS_rrtmg_post`: Postprocessor for the GFS radiation schemes
- `GFS_suite_interstitial_phys_reset`: GFS suite interstitial physics reset

- `GFS_suite_stateout_reset`: GFS suite stateout reset
- `get_prs_fv3`: Adjustment of the geopotential height hydrostatically in a way consistent with FV3 discretization
- `GFS_suite_interstitial_1`: GFS suite interstitial 1
- `dcyc2t3`: Mapping of the radiative fluxes and heating rates from the coarser radiation timestep onto the model's more frequent time steps
- `GFS_surface_generic_pre`: Preprocessor for the surface schemes (land, sea ice)
- `GFS_surface_composites_pre`: Preprocessor for surface composites
- `GFS_suite_interstitial_2`: GFS suite interstitial 2
- `sfc_diff`: Calculation of the exchange coefficients in the GFS surface layer
- `GFS_surface_loop_control_part1`: GFS surface loop control part 1
- `sfc_nst_pre`: Preprocessor for the near-surface sea temperature
- `sfc_nst`: GFS Near-surface sea temperature
- `sfc_nst_post`: Postprocessor for the near-surface temperature
- `lsm_noah`: Noah land surface scheme driver
- `sfc_sice`: Simple sea ice scheme
- `GFS_surface_loop_control_part2`: GFS surface loop control part 2
- `GFS_surface_composites_post`: Postprocess for surface composites
- `Dcyc2t3_post`: Postprocessor for the mapping of the radiative fluxes and heating rates from the coarser radiation timestep onto the model's more frequent time steps
- `sfc_diag`: Land surface diagnostic calculation
- `sfc_diag_post`: Postprocessor for the land surface diagnostic calculation
- `GFS_surface_generic_post`: Postprocessor for the GFS surface process
- `GFS_PBL_generic_pre`: Preprocessor for all Planetary Boundary Layer (PBL) schemes (except MYNN)
- `hedmf`: Hybrid eddy-diffusivity mass-flux PBL
- `GFS_PBL_generic_post`: Postprocessor for all PBL schemes (except MYNN)
- `gw dps_pre`: Preprocessor for the orographic gravity wave drag
- `gw dps`: Orographic gravity wave drag
- `Gw dps_post`: Postprocessor for the orographic gravity wave drag
- `rayleigh_damp`: Rayleigh damping
- `GFS_suite_stateout_update`: GFS suite stateout update
- `ozphys`: Ozone photochemistry
- `GFS_DCNV_generic_pre`: Preprocessor for the GFS deep convective schemes
- `get_phi_fv3`: Hydrostatic adjustment to the height in a way consistent with FV3 discretization
- `GFS_suite_interstitial_3`: GFS suite interstitial 3
- `samfdeepcnv`: Simplified Arakawa Schubert (SAS) Mass Flux deep convection
- `GFS_DCNV_generic_post`: Postprocessor for all deep convective schemes

- `gwdc_pre`: Preprocessor for the convective gravity wave drag
- `gwdc`: Convective gravity wave drag
- `gwdc_post`: Postprocessor for the convective gravity wave drag
- `GFS_SCNV_generic_pre`: Preprocessor for the GFS shallow convective schemes
- `samfshalcnv`: SAS mass flux shallow convection
- `samfshalcnv_post`: Postprocessor for the SAS Mass Flux shallow convection
- `GFS_SCNV_generic_post`: Postprocessor for the GFS shallow convective scheme
- `GFS_suite_interstitial_4`: GFS suite interstitial 4
- `cnvc90`: Convective cloud cover
- `GFS_MP_generic_pre`: Preprocessor for all GFS microphysics
- `gfdl_cloud_microphys`: GFDL cloud microphysics
- `GFS_MP_generic_post`: Postprocessor for GFS microphysics
- `sfc_sice_post`: Postprocessor for the simple sea ice
- `Maximum_hourly_diagnostics`: Computation of the maximum of the selected diagnostics
- `GFS_stochastics`: GFS stochastics scheme: Stochastic Kinetic Energy Backscatter (SKEB), Perturbed boundary layer specific humidity (SHUM), or Stochastically Perturbed Physics Tendencies (SPPT)

## AUTOGENERATED PHYSICS CAPS

The connection between the host model and the physics schemes through the CCPP-Framework is realized with *caps* on both sides as illustrated in [Figure 1.1](#). The CCPP *prebuild* script discussed in [Chapter 3](#) generates the *caps* that connect the physics schemes to the CCPP-Framework, and a large fraction of code that can be included in the host model *cap*. The host model *cap* connects the framework (Physics Driver) to the host model and must be created manually incorporating the autogenerated code. This chapter describes the physics *caps*, while the host model *caps* are described in [Chapter 6](#). Note that the dynamic build, which can be used with the SCM and with the UFS Atmosphere, produces individual physics scheme *caps*, while the static build (for UFS Atmosphere only) produces group and suite *caps*. The physics *caps* autogenerated by `ccpp_prebuild.py` reside in the directory defined by the `CAPS_DIR` variable (see example in [Listing 8.1](#)).

### 5.1 Dynamic Build Caps

With the dynamic build, the CCPP-Framework and physics are dynamically linked to the executable to allow maximum runtime flexibility. A *cap* is generated using the metadata requested in each scheme argument table and the variables provided in the host model metadata tables (see left side of [Figure 3.2](#)). The CCPP *prebuild* step for the dynamic build performs the following tasks:

- Checks requested vs provided variables by `standard_name`.
- Checks units, rank, type.
- Creates Fortran code that adds pointers to the host model variables and stores them in the `ccpp-data` structure (`ccpp_fields_*.inc`). A hash table that is part of `cdata` is populated with `key = standard_name` of a variable and `value = location of that variable in memory` (i.e. a c-pointer).
- Creates one *cap* per physics scheme.
- Populates makefiles with schemes and *caps*.

The *prebuild* step will produce the following files for the UFS Atmosphere model:

- List of variables provided by host model and required by physics:

```
ccpp/framework/doc/DevelopersGuide/CCPP_VARIABLES_FV3.tex
```

- Makefile snippets that contain all *caps* to be compiled:

```
ccpp/physics/CCPP_CAPS.{cmake,mk}
```

- Makefile snippets that contain all schemes (and their dependencies) to be compiled:

```
ccpp/physics/CCPP_SCHEMES.{cmake,mk}
```

- List of variables provided by host model:

```
ccpp/physics/CCPP_VARIABLES_FV3.html
```

- One *cap* per physics scheme:

```
ccpp/physics/*_cap.F90
```

- \*.inc files that contain module use and ccpp\_field\_add statements that populate the ccpp data type (cdata) with the necessary information on where (in memory) to find required variables:

```
FV3/atmos_cubed_sphere/driver/fvGFS/ccpp_modules_{fast,slow}_physics.inc
FV3/atmos_cubed_sphere/driver/fvGFS/ccpp_fields_{fast,slow}_physics.inc
FV3/ipd/ccpp_modules_{fast,slow}_physics.inc
FV3/ipd/ccpp_fields_{fast,slow}_physics.inc
```

The variables added to \*\_fast\_physics.inc do not use GFS\_typedefs.F90 or CCPP\_data.F90.

- Autogenerated code to include in host model *caps* (called TARGET FILES) via CPP (C preprocessor) directives:

```
FV3/ipd/IPD_CCPP_driver.F90 for slow physics
FV3/atmos_cubed_sphere/driver/fvGFS/atmosphere.F90 for fast physics
```

For each *cap*, ccpp\_prebuild.py generates “use” statements based on the host model template. Only the public *caps* (init, run and finalize) are exposed (see code example below). Each *cap* consists of a module containing three functions. For example, scheme\_pre\_cap.F90 would contain the functions scheme\_pre\_init\_cap, scheme\_pre\_run\_cap and scheme\_pre\_finalize\_cap, which perform the functions below.

- Declare data types cptr, cdims and ckind.
- Create a pointer to the Fortran data type cdata.
- Call ccpp\_field\_get for each variable in metadata table and pulls data from the cdata structure.

The index defined in each call speeds up memory access by avoiding a binary search, since variables are no longer searched by name; the order of the data in cdata are known.

- Call the corresponding scheme entry-point at the end with an explicit argument list.

For example, the autogenerated scheme *cap* for rrtmg\_lw\_pre\_cap.F90 is shown in [Listing 5.1](#).

```
module rrtmg_lw_pre_cap
  use, intrinsic :: iso_c_binding, only: c_f_pointer, &
    c_ptr, c_int32_t
  use          :: ccpp_types, only: ccpp_t, CCPP_GENERIC_KIND
  use          :: ccpp_fields, only: ccpp_field_get
  use          :: ccpp_errors, only: ccpp_error, ccpp_debug
  use          :: rrtmg_lw_pre, only: rrtmg_lw_pre_run, &
    rrtmg_lw_pre_init, rrtmg_lw_pre_finalize
  ! Other modules required, e.g. type definitions
  use GFS_typedefs, only: GFS_control_type, GFS_grid_type, &
    GFS_sfcprop_type, GFS_radtend_type
  use machine, only: kind_phys
  implicit none
  private
  public :: rrtmg_lw_pre_run_cap, rrtmg_lw_pre_init_cap, &
    rrtmg_lw_pre_finalize_cap
  contains
  function rrtmg_lw_pre_init_cap(ptr) bind(c) result(ierr)
    integer(c_int32_t)          :: ierr
    type(c_ptr), intent(inout) :: ptr
```

(continues on next page)



(continued from previous page)

```

type(ccpp_t), pointer           :: cdata
type(c_ptr)                     :: cptr
integer, allocatable          :: cdims(:)
integer                        :: ckind
ierr = 0
call c_f_pointer(ptr, cdata)
call rrtmg_lw_pre_init()
end function rrtmg_lw_pre_init_cap

function rrtmg_lw_pre_run_cap(ptr) bind(c) result(ierr)
  integer(c_int32_t)           :: ierr
  type(c_ptr), intent(inout) :: ptr
  type(ccpp_t), pointer       :: cdata
  type(c_ptr)                 :: cptr
  integer, allocatable       :: cdims(:)
  integer                    :: ckind
  type(GFS_control_type), pointer :: Model
  type(GFS_grid_type), pointer   :: Grid
  type(GFS_sfcprop_type), pointer :: Sfcprop
  type(GFS_radtend_type), pointer :: Radtend
  integer, pointer :: im
  real(kind_phys), pointer :: tsfg(:)
  real(kind_phys), pointer :: tsfa(:)
  ierr = 0
  call c_f_pointer(ptr, cdata)
  call ccpp_field_get(cdata, 'GFS_control_type_instance', cptr, &
    ierr=ierr, kind=ckind, index=2)
  call c_f_pointer(cptr, Model)
  call ccpp_field_get(cdata, 'GFS_grid_type_instance', cptr, &
    ierr=ierr, kind=ckind, index=6)
  call c_f_pointer(cptr, Grid)
  call ccpp_field_get(cdata, 'GFS_sfcprop_type_instance', &
    cptr, ierr=ierr, kind=ckind, index=10)
  call c_f_pointer(cptr, Sfcprop)
  call ccpp_field_get(cdata, 'GFS_radtend_type_instance', &
    cptr, ierr=ierr, kind=ckind, index=9)
  call c_f_pointer(cptr, Radtend)
  call ccpp_field_get(cdata, 'horizontal_loop_extent', im, &
    ierr=ierr, kind=ckind, index=390)
  call ccpp_field_get(cdata, &
    'surface_ground_temperature_for_radiation', &
    tsfg, ierr=ierr, dims=cdims, kind=ckind, index=770)
  deallocate(cdims)
  call ccpp_field_get(cdata, &
    'surface_air_temperature_for_radiation', &
    tsfa, ierr=ierr, dims=cdims, kind=ckind, index=724)
  deallocate(cdims)
  call rrtmg_lw_pre_run(Model=Model, Grid=Grid, &
    Sfcprop=Sfcprop, Radtend=Radtend, im=im, &
    tsfg=tsfg, tsfa=tsfa, &
    errmsg=cdata%errmsg, errflg=cdata%errflg)
  ierr=cdata%errflg
end function rrtmg_lw_pre_run_cap
function rrtmg_lw_pre_finalize_cap(ptr) bind(c) result(ierr)
  integer(c_int32_t)           :: ierr
  type(c_ptr), intent(inout) :: ptr
  type(ccpp_t), pointer       :: cdata

```

(continues on next page)

(continued from previous page)

```

type(c_ptr)           :: cptr
integer, allocatable :: cdims(:)
integer              :: ckind
ierr = 0
call c_f_pointer(ptr, cdata)
call rrtmg_lw_pre_finalize()
end function rrtmg_lw_pre_finalize_cap
end module rrtmg_lw_pre_cap

```

*Listing 5.1: Condensed version of the autogenerated scheme cap rrtmg\_lw\_pre\_cap.F90 for the dynamic build. Note the calls to ccpp\_field\_get for each variable.*

The fields accessed from `cdata` are determined by the metadata in the scheme argument table(s). In this example, `rrtmg_lw_pre_init` and `rrtmg_lw_pre_finalize` are empty subroutines, i.e. they have no arguments passed in or out, no metadata table, and no calls to `ccpp_field_get`. However, `rrtmg_lw_pre_run` has a metadata table, so `ccpp_field_get` is called for each variable in the table and the value put into the call to `rrtmg_lw_pre_run`.

## 5.2 Static Build Caps

With a static build, the CCPP-Framework and physics are statically linked to the executable. This allows the best performance and efficient memory use. Similar to the dynamic build, the static build requires metadata provided by the host model and variables requested from the physics scheme. Unlike a dynamic build where all variables are kept and pulled multiple times for various parameterizations, a static build only keeps variables for specified suites, and therefore requires one or more SDFs (see left side of [Figure 3.3](#)) as arguments to the `ccpp_prebuild.py` script. The CCPP *prebuild* step for the static build performs the tasks below.

- Check requested vs provided variables by `standard_name`.
- Check units, rank, type.
- Filter unused schemes and variables.
- Create Fortran code for the static Application Programming Interface (API) that replaces the dynamic API (CCPP-Framework). The hash table used by the dynamic build to store variables in memory is left empty.
- Create *caps* for groups and suite(s).
- Populate makefiles with schemes and *caps*.

The *prebuild* step for the static build will produce the following files for the UFS Atmosphere:

- List of variables provided by host model and required by physics:

```
ccpp/framework/doc/DevelopersGuide/CCPP_VARIABLES_FV3.tex
```

- Makefile snippets that contain all *caps* to be compiled:

```
ccpp/physics/CCPP_CAPS.{cmake,mk}
```

- Makefile snippets that contain all schemes to be compiled:

```
ccpp/physics/CCPP_SCHEMES.{cmake,mk}
```

- List of variables provided by host model:

```
ccpp/physics/CCPP_VARIABLES_FV3.html
```

- One *cap* per physics group (fast\_physics, physics, radiation, time\_vary, stochastic, ...) for each suite:

```
ccpp/physics/ccpp_{suite_name}_{group_name}_cap.F90
```

- *Cap* for each suite:

```
ccpp/physics/ccpp_{suite_name}_cap.F90
```

- Autogenerated API for static build that replaces the dynamic API (aka CCPP-Framework), the interface is identical between the two APIs:

```
FV3/gfsphysics/CCPP_layer/ccpp_static_api.F90
```

- Same TARGET FILES as for the dynamic build

ccpp\_static\_api.F90 replaces the entire dynamic CCPP-Framework with an equivalent interface, which contains subroutines ccpp\_physics\_init, ccpp\_physics\_run and ccpp\_physics\_finalize. Each subroutine uses a suite\_name and an optional argument, group\_name, to call the groups of a specified suite (e.g. fast\_physics, physics, time\_vary, radiation, stochastic, etc.), or to call the entire suite. For example, ccpp\_static\_api.F90 would contain module ccpp\_static\_api with subroutines ccpp\_physics\_{init, run, finalize}. The subroutine ccpp\_physics\_init from the autogenerated code using suites FV3\_GFS\_v15 and FV3\_CPT\_v0 is shown in *Listing 5.2*.

```
subroutine ccpp_physics_init(cdata, suite_name, group_name, ierr)
  use ccpp_types, only : ccpp_t
  implicit none
  type(ccpp_t),          intent(inout) :: cdata
  character(len=*),     intent(in)   :: suite_name
  character(len=*), optional, intent(in) :: group_name
  integer,              intent(out)  :: ierr
  ierr = 0
  if (trim(suite_name)=="FV3_GFS_v15") then
    if (present(group_name)) then
      if (trim(group_name)=="fast_physics") then
        ierr = FV3_GFS_v15_fast_physics_init_cap(cdata=cdata, CCPP_interstitial=CCPP_
↪interstitial)
      else if (trim(group_name)=="time_vary") then
        ierr = FV3_GFS_v15_time_vary_init_cap(GFS_Interstitial=GFS_Interstitial, &
        cdata=cdata,GFS_Data=GFS_Data, GFS_Control=GFS_Control)
      else if (trim(group_name)=="radiation") then
        ierr = FV3_GFS_v15_radiation_init_cap()
      else if (trim(group_name)=="physics") then
        ierr = FV3_GFS_v15_physics_init_cap(cdata=cdata, GFS_Control=GFS_Control)
      else if (trim(group_name)=="stochastics") then
        ierr = FV3_GFS_v15_stochastics_init_cap()
      else
        write(cdata%errmsg, '(*(a))') "Group " // trim(group_name) // " not found"
        ierr = 1
      end if
    else
      ierr = FV3_GFS_v15_init_cap(GFS_Interstitial=GFS_Interstitial, cdata=cdata,GFS_
↪Control=GFS_Control, &
        GFS_Data=GFS_Data, CCPP_interstitial=CCPP_interstitial)
    end if
  else if (trim(suite_name)=="FV3_CPT_v0") then
```

(continues on next page)

(continued from previous page)

```

if (present(group_name)) then
  if (trim(group_name)=="time_vary") then
    ierr = FV3_CPT_v0_time_vary_init_cap(GFS_Interstitial=GFS_Interstitial, &
      cdata=cdata,GFS_Data=GFS_Data, GFS_Control=GFS_Control)
  else if (trim(group_name)=="radiation") then
    ierr = FV3_CPT_v0_radiation_init_cap()
  else if (trim(group_name)=="physics") then
    ierr = FV3_CPT_v0_physics_init_cap(con_hfus=con_hfus, &
      GFS_Control=GFS_Control,con_hvap=con_hvap, &
      con_rd=con_rd,con_rv=con_rv,con_g=con_g, &
      con_ttp=con_ttp,con_cp=con_cp,cdata=cdata)
  else if (trim(group_name)=="stochastics") then
    ierr = FV3_CPT_v0_stochastics_init_cap()
  else
    write(cdata%errmsg, '(*(a))') "Group " // trim(group_name) // " not found"
    ierr = 1
  end if
else
  ierr = FV3_CPT_v0_init_cap(con_g=con_g, GFS_Data=GFS_Data,GFS_Control=GFS_
↪Control, &
    con_hvap=con_hvap,GFS_Interstitial=GFS_Interstitial, con_rd=con_rd,con_
↪rv=con_rv, &
    con_hfus=con_hfus, con_ttp=con_ttp,con_cp=con_cp,cdata=cdata)
  end if
else
  write(cdata%errmsg,'(*(a))'), 'Invalid suite ' // trim(suite_name)
  ierr = 1
end if
cdata%errflg = ierr
end subroutine ccpp_physics_init

```

Listing 5.2: Code sample of subroutine `ccpp_physics_init` contained in the autogenerated file `ccpp_static_api.F90` for the multi-suite static build. This cap was generated using suites `FV3_GFS_v15` and `FV3_CPT_v0`. Examples of the highlighted functions are shown below in Listing 5.3 and Listing 5.4.

Note that if `group_name` is set, specified groups (i.e. `FV3_GFS_v15_physics_init_cap`) are called for the specified `suite_name`. These functions are defined in `ccpp_{suite_name}_{group_name}_cap.F90`, in this case `ccpp_FV3_GFS_v15_physics_cap.F90`. For example:

```

function FV3_GFS_v15_physics_init_cap(cdata,GFS_Control) &
  result (ierr)
  use ccpp_types, only: ccpp_t
  use GFS_typedefs, only: GFS_control_type
  implicit none
  integer                                :: ierr
  type(ccpp_t), intent(inout) :: cdata
  type(GFS_control_type), intent(in) :: GFS_Control
  ierr = 0
  if (initialized) return
  call lsm_noah_init (me=GFS_Control%me,isot=GFS_Control%isot,&
    ivegsrc=GFS_Control%ivegsrc,nlunit=GFS_Control%nlunit, &
    errmsg=cdata%errmsg,errflg=cdata%errflg)
  if (cdata%errflg/=0) then
    write(cdata%errmsg,'(a)') "An error occurred in lsm_noah_init"
    ierr=cdata%errflg
    return
  end if

```

(continues on next page)

(continued from previous page)

```

call gfdl_cloud_microphys_init(me=GFS_Control%me, &
    master=GFS_Control%master,nlunit=GFS_Control%nlunit, &
    input_nml_file=GFS_Control%input_nml_file, &
    logunit=GFS_Control%logunit,fn_nml=GFS_Control%fn_nml, &
    imp_physics=GFS_Control%imp_physics, &
    imp_physics_gfdl=GFS_Control%imp_physics_gfdl, &
    do_shoc=GFS_Control%do_shoc, &
    errmsg=cdata%errmsg,errflg=cdata%errflg)
if (cdata%errflg/=0) then
    write(cdata%errmsg,'(a)') "An error ocured in &
        gfdl_cloud_microphys_init"
    ierr=cdata%errflg
    return
end if
    initialized = .true.
end function FV3_GFS_v15_physics_init_cap

```

*Listing 5.3: The FV3\_GFS\_v15\_physics\_init\_cap contained in the autogenerated file ccpp\_FV3\_GFS\_v15\_physics\_cap.F90 showing calls to the lsm\_noah\_init, and gfdl\_cloud\_microphys\_init subroutines for the static build for suite 'FV3\_GFS\_v15' and group 'physics'.*

If the group\_name is not specified for a specified suite\_name, the suite is called from the autogenerated ccpp\_static\_api.F90, which calls the init, run and finalize routines for each group. *Listing 5.4* is an example of FV3\_GFS\_v15\_init\_cap.

```

function FV3_GFS_v15_init_cap(GFS_Interstitial, &
    cdata,GFS_Control,GFS_Data,CCPP_interstitial) result (ierr)
use GFS_typedefs, only: GFS_interstitial_type
use ccpp_types, only: ccpp_t
use GFS_typedefs, only: GFS_control_type
use GFS_typedefs, only: GFS_data_type
use CCPP_typedefs, only: CCPP_interstitial_type

implicit none

integer :: ierr
type(GFS_interstitial_type), intent(inout) :: GFS_Interstitial(:)
type(ccpp_t), intent(inout) :: cdata
type(GFS_control_type), intent(inout) :: GFS_Control
type(GFS_data_type), intent(inout) :: GFS_Data(:)
type(CCPP_interstitial_type), intent(in) :: CCPP_interstitial

    ierr = 0
    ierr = FV3_GFS_v15_fast_physics_init_cap(cdata=cdata, CCPP_interstitial=CCPP_
    ↪interstitial)
    if (ierr/=0) return

    ierr = FV3_GFS_v15_time_vary_init_cap (GFS_Interstitial=GFS_Interstitial,
    ↪cdata=cdata, &
        GFS_Data=GFS_Data,GFS_Control=GFS_Control)
    if (ierr/=0) return

    ierr = FV3_GFS_v15_radiation_init_cap()
    if (ierr/=0) return
    ierr = FV3_GFS_v15_physics_init_cap(cdata=cdata, &
        GFS_Control=GFS_Control)
    if (ierr/=0) return

```

(continues on next page)

(continued from previous page)

```
ierr = FV3_GFS_v15_stochastics_init_cap()  
if (ierr/=0) return  
end function FV3_GFS_v15_init_cap
```

*Listing 5.4: Condensed version of the FV3\_GFS\_v15\_init\_cap function contained in the autogenerated file ccpp\_FV3\_GFS\_v15\_cap.F90 showing calls to the group caps FV3\_GFS\_v15\_fast\_physics\_init\_cap, FV3\_GFS\_v15\_time\_vary\_init\_cap , etc. for the static build where a group name is not specified.*

---

## HOST SIDE CODING

This chapter describes the connection of a host model with the pool of *CCPP-Physics* schemes through the *CCPP-Framework*.

In several places, references are made to an Interoperable Physics Driver (IPD). The IPD was originally developed by EMC and later expanded by NOAA GFDL with the goal of connecting GFS physics to various models. A top motivation for its development was the dycore test that led to the selection of FV3 as the dycore for the *UFS*. Designed in a fundamentally different way than the *CCPP*, the IPD will be phased out in the near future in favor of the *CCPP* as a single way to interface with physics in the *UFS*. To enable a smooth transition, several of the *CCPP* components must interact with the IPD and, as such, parts of the *CCPP* code in the *UFS* currently carry the tag “IPD”.

### 6.1 Variable Requirements on the Host Model Side

All variables required to communicate between the host model and the physics, as well as to communicate between physics schemes, need to be allocated by the host model. An exception is variables `errflg`, `errmsg`, `loop_cnt`, `blk_no`, and `thrd_no`, which are allocated by the *CCPP-Framework*, as explained in [Section 6.4.1](#). A list of all variables required for the current pool of physics can be found in `ccpp/framework/doc/DevelopersGuide/CCPP_VARIABLES_XYZ.pdf` (XYZ: SCM, FV3).

At present, only two types of variable definitions are supported by the *CCPP-Framework*:

- Standard Fortran variables (character, integer, logical, real) defined in a module or in the main program. For character variables, a fixed length is required. All others can have a kind attribute of a kind type defined by the host model.
- Derived data types (DDTs) defined in a module or the main program. While the use of DDTs as arguments to physics schemes in general is discouraged (see [Section 2.2](#)), it is perfectly acceptable for the host model to define the variables requested by physics schemes as components of DDTs and pass these components to *CCPP* by using the correct `local_name` (e.g., `myddt%thecomponentIwant`; see [Section 6.2](#).)

### 6.2 Metadata Variable Tables in the Host Model

To establish the link between host model variables and physics scheme variables, the host model must provide metadata tables similar to those presented in [Section 2.1](#). The host model can have multiple metadata tables or just one. For each variable required by the pool of *CCPP-Physics* schemes, one and only one entry must exist on the host model side. The connection between a variable in the host model and in the physics scheme is made through its `standard_name`.

The following requirements must be met when defining variables in the host model metadata tables (see also [Listing 6.1](#) for examples of host model metadata tables).

- The `standard_name` must match that of the target variable in the physics scheme.

- The type, kind, shape and size of the variable (as defined in the host model Fortran code) must match that of the target variable.
- The attributes `units`, `rank`, `type` and `kind` in the host model metadata table must match those in the physics scheme table.
- The attributes `optional` and `intent` must be set to `F` and `none`, respectively.
- The `local_name` of the variable must be set to the name the host model cap uses to refer to the variable.
- The metadata table that exposes a DDT to the CCPP (as opposed to the table that describes the components of a DDT) must be in the same module where the memory for the DDT is allocated. If the DDT is a module variable, then it must be exposed via the module's metadata table, which must have the same name as the module.
- Metadata tables describing module variables must be placed inside the module.
- Metadata tables describing components of DDTs must be placed immediately before the type definition and have the same name as the DDT.

```

module example_vardefs

  implicit none

  !> \section arg_table_example_vardefs
  !! | local_name | standard_name | long_name | units | rank | type      | kind |
  ↪intent | optional |
  !! |-----|-----|-----|-----|-----|-----|-----|-----|
  ↪-----|-----|
  !! | ex_int      | example_int   | ex. int   | none | 0 | integer |      |
  ↪none   | F           |
  !! | ex_reall    | example_reall | ex. real  | m     | 2 | real    | kind=8 |
  ↪none   | F           |
  !! | ex_ddt      | ex_ddt        | ex. ddt type | DDT   | 2 | ex_ddt  |      |
  ↪none   | F           |
  !! | ext         | ex_ddt_instance | ex. ddt inst | DDT   | 2 | ex_ddt  |      |
  ↪none   | F           |
  !! | errmsg      | error_message | err. msg.  | none  | 0 | character | len=64 |
  ↪none   | F           |
  !! | errflg      | error_flag    | err. flg.  | flag  | 0 | logical  |      |
  ↪none   | F           |
  !!
  integer, parameter      :: r15 = selected_real_kind(15)
  integer                 :: ex_int
  real(kind=8), dimension(:, :) :: ex_reall
  character(len=64)       :: errmsg
  logical                 :: errflg

  ! Derived data types

  !> \section arg_table_ex_ddt
  !! | local_name | standard_name | long_name | units | rank | type      | kind |
  ↪intent | optional |
  !! |-----|-----|-----|-----|-----|-----|-----|-----|
  ↪---|-----|
  !! | ext%l      | example_flag  | ex. flag  | flag  | 0 | logical  |      |
  ↪none   | F           |
  !! | ext%r      | example_real3 | ex. real  | kg    | 2 | real     | r15   |
  ↪none   | F           |
  !! | ext%r(:,1) | example_slice | ex. slice | kg    | 1 | real     | r15   |
  ↪none   | F           |

```

(continues on next page)



(continued from previous page)

```

!!
  type ex_ddt
    logical          :: l
    real, dimension(:, :) :: r
  end type ex_ddt

  type(ex_ddt) :: ext

end module example_vardefs

```

Listing 6.1: Example Host Model Metadata Table. In this example, both the definition and the declaration (memory allocation) of a DDT `ext` (of type `ex_ddt`) are in the same module.

### 6.3 CCPP Variables in the SCM and UFS Atmosphere Host Models

While the use of standard Fortran variables is preferred, in the current implementation of the CCPP in the UFS Atmosphere and in the SCM almost all data is contained in DDTs for organizational purposes. In the case of the SCM, DDTs are defined in `gmtb_scm_type_defs.f90` and `GFS_typedefs.F90`, and in the case of the UFS Atmosphere, they are defined in both `GFS_typedefs.F90` and `CCPP_typedefs.F90`. The current implementation of the CCPP in both host models uses the following set of DDTs:

- `GFS_init_type` variables to allow proper initialization of GFS physics
- `GFS_statein_type` prognostic state data provided by dycore to physics
- `GFS_stateout_type` prognostic state after physical parameterizations
- `GFS_sfcprop_type` surface properties read in and/or updated by climatology, obs, physics
- `GFS_coupling_type` fields from/to coupling with other components, e.g., land/ice/ocean
- `GFS_control_type` control parameters input from a namelist and/or derived from others
- `GFS_grid_type` grid data needed for interpolations and length-scale calculations
- `GFS_tbd_type` data not yet assigned to a defined container
- `GFS_cldprop_type` cloud properties and tendencies needed by radiation from physics
- `GFS_radtend_type` radiation tendencies needed by physics
- `GFS_diag_type` fields targeted for diagnostic output to disk
- `GFS_interstitial_type` fields used to communicate variables among schemes in the slow physics group required to replace interstitial code in `GFS_{physics, radiation}_driver.F90` in CCPP
- `GFS_data_type` combined type of all of the above except `GFS_control_type` and `GFS_interstitial_type`
- `CCPP_interstitial_type` fields used to communicate variables among schemes in the fast physics group

The DDT descriptions provide an idea of what physics variables go into which data type. `GFS_diag_type` can contain variables that accumulate over a certain amount of time and are then zeroed out. Variables that require persistence from one timestep to another should not be included in the `GFS_diag_type` nor the `GFS_interstitial_type` DDTs. Similarly, variables that need to be shared between groups cannot be included in the `GFS_interstitial_type` DDT. Although this memory management is somewhat arbitrary, new variables provided by the host model or derived in an interstitial scheme should be put in a DDT with other similar variables.

Each DDT contains a create method that allocates the data defined in the metadata table. For example, the `GFS_stateout_type` contains:

```

type GFS_stateout_type

  !-- Out (physics only)
  real (kind=kind_phys), pointer :: gu0 (:,:) => null() !< updated zonal wind
  real (kind=kind_phys), pointer :: gv0 (:,:) => null() !< updated meridional wind
  real (kind=kind_phys), pointer :: gt0 (:,:) => null() !< updated temperature
  real (kind=kind_phys), pointer :: gq0 (:,:,) => null() !< updated tracers

  contains
    procedure :: create => stateout_create !< allocate array data
end type GFS_stateout_type

```

In this example, `gu0`, `gv0`, `gt0`, and `gq0` are defined in the host-side metadata table, and when the subroutine `stateout_create` is called, these arrays are allocated and initialized to zero. With the CCPP, it is possible to not only refer to components of DDTs, but also to slices of arrays in the metadata table as long as these are contiguous in memory. An example of an array slice from the `GFS_stateout_type` looks like:

```

!! | GFS_Data(cdata%blk_no)%Stateout%gq0(:, :, GFS_Control%ntsw) | snow_water_mixing_
↪ ratio_updated_by_physics | moist (dry+vapor, no_
↪ condensates) mixing ratio of snow water updated by physics | kg kg-1 | _
↪ 2 | real | kind_phys | none | F

```

Array slices can be used by physics schemes that only require certain values from an array.

## 6.4 CCPP API

The CCPP Application Programming Interface (API) is comprised of a set of clearly defined methods used to communicate variables between the host model and the physics and to run the physics. The bulk of the CCPP API is located in the CCPP-Framework, and is described in file `ccpp_api.F90`. Some aspects of the API differ between the dynamic and static build. In particular, subroutines `ccpp_physics_init`, `ccpp_physics_finalize`, and `ccpp_physics_run` (described below) are made public from `ccpp_api.F90` for the dynamic build, and are contained in `ccpp_static_api.F90` for the static build. Moreover, these subroutines take an additional argument (`suite_name`) for the static build. File `ccpp_static_api.F90` is auto-generated when the script `ccpp_prebuild.py` is run for the static build.

### 6.4.1 Data Structure to Transfer Variables between Dynamics and Physics

The roles of `cdata` structure in dealing with data exchange are not the same between the dynamic and the static builds of the CCPP. For the dynamic build, the `cdata` structure handles the data exchange between the host model and the physics schemes. `cdata` is a DDT containing a list of pointers to variables and their metadata and is persistent in memory.

For both the dynamic and static builds, the `cdata` structure is used for holding five variables that must always be available to the physics schemes. These variables are listed in a metadata table in `ccpp/framework/src/ccpp_types.F90` ([Listing 6.2](#)).

- Error flag for handling in CCPP (`errmsg`).
- Error message associated with the error flag (`errflg`).
- Loop counter for subcycling loops (`loop_cnt`).
- Number of block for explicit data blocking in CCPP (`blk_no`).
- Number of thread for threading in CCPP (`thrd_no`).

```

!! | local_name                | standard_name                | long_name                |
↪ | units | rank | type                | kind |
↪ intent | optional |
!! |-----|-----|-----|-----|
↪-----|-----|-----|-----|
↪--|-----|
!! | cdata%errflg              | ccpp_error_flag              | error flag for
↪ error handling in CCPP      | flag | 0 | integer |
↪ none | F |
!! | cdata%errmsg              | ccpp_error_message           | error message
↪ for error handling in CCPP  | none | 0 | character | len=512 |
↪ none | F |
!! | cdata%loop_cnt            | ccpp_loop_counter            | loop counter for
↪ subcycling loops in CCPP    | index | 0 | integer | | none
↪ | F |
!! | cdata%blk_no              | ccpp_block_number            | number of block
↪ for explicit data blocking in CCPP | index | 0 | integer | |
↪ none | F |
!! | cdata%thrd_no             | ccpp_thread_number           | number of thread
↪ for threading in CCPP      | index | 0 | integer | | none
↪ | F |
!!

```

*Listing 6.2: Mandatory variables provided by the CCPP-Framework from `ccpp/framework/src/ccpp_types.F90`. These variables must not be defined by the host model.*

Two of the variables are mandatory and must be passed to every physics scheme: `errmsg` and `errflg`. The variables `loop_cnt`, `blk_no`, and `thrd_no` can be passed to the schemes if required, but are not mandatory. For the static build of the CCPP, the `cdata` structure is only used to hold these five variables, since the host model variables are directly passed to the physics without the need for an intermediate data structure.

Note that `cdata` is not restricted to being a scalar but can be a multidimensional array, depending on the needs of the host model. For example, a model that uses a one-dimensional array of blocks for better cache-reuse may require `cdata` to be a one-dimensional array of the same size. Another example of a multi-dimensional array of `cdata` is in the SCM, which uses a one-dimensional `cdata` array for  $N$  independent columns.

Due to a restriction in the Fortran language, there are no standard pointers that are generic pointers, such as the C language allows. The CCPP system therefore has an underlying set of pointers in the C language that are used to point to the original data within the host application cap. The user does not see this C data structure, but deals only with the public face of the Fortran `cdata` DDT. The type `ccpp_t` is defined in `ccpp/framework/src/ccpp_types.F90`.

## 6.4.2 Adding and Retrieving Information from `cdata` (dynamic build option)

Subroutines `ccpp_field_add` and `ccpp_field_get` are part of the CCPP-Framework and are used (in the dynamic build only) to load and retrieve information to and from `cdata`. The calls to `ccpp_field_add` are auto-generated by the script `ccpp_prebuild.py` and inserted onto the host model code via include files (i.e. `FV3/CCPP_layer/ccpp_fields_slow_physics.inc`) before it is compiled.

A typical call to `ccpp_field_add` is below, where the first argument is the instance of `cdata` to which the information should be added, the second argument is the `standard_name` of the variable, the third argument is the corresponding host model variable, the fourth argument is an error flag, the fifth argument is the units of the variable, and the last (optional) argument is the position within `cdata` in which the variable is expected to be stored.

```

call ccpp_field_add(cdata, 'y_wind_updated_by_physics', GFS_Data(cdata%blk_no)
↪ %Stateout%v0, ierr=ierr, units='m s-1', index=886)

```

For DDTs, the interface to `CCPP_field_add` is slightly different:

```
call ccpp_field_add(cdata, 'GFS_cldprop_type_instance', '', c_loc(GFS_Data(cdata%blk_
↳no)%Cldprop), ierr=ierr, index=1)
```

where the first argument and second arguments bear the same meaning as in the first example, the third argument is the units (can be left empty or set to “DDT”), the fourth argument is the C pointer to the variable in memory, the fifth argument is an error flag, and the last (optional) argument is the position within `cdata` as in the first example.

Each new variable added to `cdata` is always placed at the next free position, and a check is performed to confirm that this position corresponds to the expected one, which in this example is 886. A mismatch will occur if a developer manually adds a call to `ccpp_field_add`, in which case a costly binary search is applied every time a variable is retrieved from memory. Adding calls manually is not recommended as all calls to `ccpp_fields_add` should be auto-generated.

The individual physics *caps* used in the dynamic build, which are auto-generated using the script `ccpp_prebuild.py`, contain calls to `ccpp_field_get` to pull data from the `cdata` DDT as a Fortran pointer to a variable that will be passed to the individual physics scheme.

### 6.4.3 Initializing and Finalizing the CCPP

At the beginning of each run, the `cdata` structure needs to be set up. Similarly, at the end of each run, it needs to be terminated. This is done with subroutines `ccpp_init` and `ccpp_finalize`. These subroutines should not be confused with `ccpp_physics_init` and `ccpp_physics_finalize`, which were described in [Chapter 5](#).

Note that optional arguments are denoted with square brackets.

#### Suite Initialization Subroutine

The suite initialization subroutine, `ccpp_init`, takes three mandatory and two optional arguments. The mandatory arguments are the name of the suite (of type character), the name of the `cdata` variable that must be allocated at this point, and an integer used for the error status. Note that the suite initialization routine `ccpp_init` parses the SDF corresponding to the given suite name and initializes the state of the suite and its schemes. This process must be repeated for every element of a multi-dimensional `cdata`. For performance reasons, it is possible to avoid repeated reads of the SDF and to have a single state of the suite shared between the elements of `cdata`. To do so, specify an optional argument variable called `cdata_target = X` in the call to `ccpp_init`, where `X` refers to the instance of `cdata` that has already been initialized.

For a given suite name `XYZ`, the name of the suite definition file is inferred as `suite_XYZ.xml`, and the file is expected to be present in the current run directory. It is possible to specify the optional argument `is_filename=.true.` to `ccpp_init`, which will treat the suite name as an actual file name (with or without the path to it).

Typical calls to `ccpp_init` are below, where `ccpp_suite` is the name of the suite, and `ccpp_sdf_filepath` the actual SDF filename, with or without a path to it.

```
call ccpp_init(trim(ccpp_suite), cdata, ierr)
call ccpp_init(trim(ccpp_suite), cdata2, ierr, [cdata_target=cdata])
call ccpp_init(trim(ccpp_sdf_filepath), cdata, ierr, [is_filename=.true.])
```

#### Suite Finalization Subroutine

The suite finalization subroutine, `ccpp_finalize`, takes two arguments, the name of the `cdata` variable that must be de-allocated at this point, and an integer used for the error status. A typical call to `ccpp_finalize` is below:

```
call ccpp_finalize(cdata, ierr)
```

If a specific data instance was used in a call to `ccpp_init`, as in the above example in [Section 6.4.3](#), then this data instance must be finalized last:

```
call ccpp_finalize(cdata2, ierr)
call ccpp_finalize(cdata, ierr)
```

## 6.4.4 Running the physics

The physics is invoked by calling subroutine `ccpp_physics_run`. This subroutine is part of the CCPP API and is included with the CCPP-Framework (for the dynamic build) or auto-generated (for the static build). This subroutine is capable of executing the physics with varying granularity, that is, a single scheme (dynamic build only), a single group, or an entire suite can be run with a single subroutine call. Typical calls to `ccpp_physics_run` are below, where `scheme_name` and `group_name` are optional and mutually exclusive (dynamic build), and where `suite_name` is mandatory and `group_name` is optional (static build).

Dynamic build:

```
call ccpp_physics_run(cdata, [group_name], [scheme_name], ierr=ierr)
```

Static build:

```
call ccpp_physics_run(cdata, suite_name, [group_name], ierr=ierr)
```

## 6.4.5 Initializing and Finalizing the Physics

Many (but not all) physical parameterizations need to be initialized, which includes functions such as reading lookup tables, reading input datasets, computing derived quantities, broadcasting information to all MPI ranks, etc. Initialization procedures are typically done for the entire domain, that is, they are not subdivided by blocks. Similarly, many (but not all) parameterizations need to be finalized, which includes functions such as deallocating variables, resetting flags from *initialized* to *non-initialized*, etc. Initialization and finalization functions are each performed once per run, before the first call to the physics and after the last call to the physics, respectively.

The initialization and finalization can be invoked for a single parameterization (only in dynamic build), for a single group, or for the entire suite. In all cases, subroutines `ccpp_physics_init` and `ccpp_physics_finalize` are used and the arguments passed to those subroutines determine the type of initialization.

These subroutines should not be confused with `ccpp_init` and `ccpp_finalize`, which were explained previously.

### Subroutine `ccpp_physics_init`

This subroutine is part of the CCPP API and is included with the CCPP-Framework (for the dynamic build) or auto-generated (for the static build). It cannot contain thread-dependent information but can have block-dependent information. Typical calls to `ccpp_physics_init` are below.

Dynamic build:

```
call ccpp_physics_init(cdata, [group_name], [scheme_name], ierr=ierr)
```

Static build:

```
call ccpp_physics_init(cdata, suite_name, [group_name], ierr=ierr)
```

### Subroutine `ccpp_physics_finalize`

This subroutine is part of the CCPP API and is included with the CCPP-Framework (for the dynamic build) or auto-generated (for the static build). Typical calls to `ccpp_physics_finalize` are below.

Dynamic build:

```
call ccpp_physics_finalize(cdata, [group_name], [scheme_name], ierr=ierr)
```

Static build:

```
call ccpp_physics_finalize(cdata, suite_name, [group_name], ierr=ierr)
```

## 6.5 Host Caps

The purpose of the host model *cap* is to abstract away the communication between the host model and the CCPP-Physics schemes. While CCPP calls can be placed directly inside the host model code (as is done for the relatively simple SCM), it is recommended to separate the *cap* in its own module for clarity and simplicity (as is done for the UFS Atmosphere). While the details of implementation will be specific to each host model, the host model *cap* is responsible for the following general functions:

- Allocating memory for variables needed by physics
  - All variables needed to communicate between the host model and the physics, and all variables needed to communicate among physics schemes, need to be allocated by the host model. The latter, for example for interstitial variables used exclusively for communication between the physics schemes, are typically allocated in the *cap*.
- Allocating the `cdata` structure(s)
  - For the dynamic build, the `cdata` structure handles the data exchange between the host model and the physics schemes, while for the static build, `cdata` is utilized in a reduced capacity.
- Calling the suite initialization subroutine
  - The suite must be initialized using `ccpp_init`.
- Populating the `cdata` structure(s)
  - For the dynamic build, each variable required by the physics schemes must be added to the `cdata` structure (or to each element of a multi-dimensional `cdata`) on the host model side using subroutine `ccpp_field_add`. This is an automated task accomplished by inserting a preprocessor directive

```
#include ccpp_modules.inc
```

at the top of the *cap* (before `implicit none`) to load the required modules and a second preprocessor directive

```
#include ccpp_fields.inc
```

after the `cdata` variable and the variables required by the physics schemes are allocated and after the call to `ccpp_init` for this `cdata` variable. For the static build, this step can be skipped because the autogenerated *caps* for the physics (groups and suite *caps*) are automatically given memory access to the host model variables and they can be used directly, without the need for a data structure containing pointers to the actual variables (which is what `cdata` is).

**Note:** The CCPP-Framework supports splitting physics schemes into different sets that are used in different parts of the host model. An example is the separation between slow and fast physics processes for the GFDL microphysics implemented in the UFS Atmosphere: while the slow physics are called as part of the usual model physics, the fast physics are integrated in the dynamical core. The separation of physics into different sets is determined in the CCPP *prebuild* configuration for each host model (see [Chapter 5.1](#), and [Figure 8.1](#)), which allows to create multiple include files (e.g. `ccpp_fields_slow_physics.inc` and `ccpp_fields_fast_physics.inc` that can be used by different `cdata` structures in different parts of the model). This is a highly advanced feature and developers seeking to take further advantage of it should consult with GMTB first.

- Providing interfaces to call the CCPP
  - The *cap* must provide functions or subroutines that can be called at the appropriate places in the host model time integration loop and that internally call `ccpp_init`, `ccpp_physics_init`, `ccpp_physics_run`, `ccpp_physics_finalize` and `ccpp_finalize`, and handle any errors returned See [Listing 6.3](#).

```

module example_ccpp_host_cap

  use ccpp_api,           only: ccpp_t, ccpp_init, ccpp_finalize
  use ccpp_static_api,   only: ccpp_physics_init, ccpp_physics_run,   &
                                ccpp_physics_finalize

  implicit none
  ! CCPP data structure
  type(ccpp_t), save, target :: cdata
  public :: physics_init, physics_run, physics_finalize
contains

  subroutine physics_init(ccpp_suite_name)
    character(len=*), intent(in) :: ccpp_suite_name
    integer :: ierr
    ierr = 0

    ! Initialize the CCPP framework, parse SDF
    call ccpp_init(trim(ccpp_suite_name), cdata, ierr=ierr)
    if (ierr/=0) then
      write(*, '(a)') "An error occurred in ccpp_init"
      stop
    end if

    ! Initialize CCPP physics (run all _init routines)
    call ccpp_physics_init(cdata, suite_name=trim(ccpp_suite_name),   &
                          ierr=ierr)
    ! error handling as above

  end subroutine physics_init

  subroutine physics_run(ccpp_suite_name, group)
    ! Optional argument group can be used to run a group of schemes   &
    ! defined in the SDF. Otherwise, run entire suite.
    character(len=*),           intent(in) :: ccpp_suite_name
    character(len=*), optional, intent(in) :: group

    integer :: ierr
    ierr = 0

```

(continues on next page)

(continued from previous page)

```

    if (present(group)) then
        call ccpp_physics_run(cdata, suite_name=trim(ccpp_suite_name), &
                             group_name=group, ierr=ierr)
    else
        call ccpp_physics_run(cdata, suite_name=trim(ccpp_suite_name), &
                             ierr=ierr)
    end if
    ! error handling as above

end subroutine physics_run

subroutine physics_finalize(ccpp_suite_name)
    character(len=*), intent(in) :: ccpp_suite_name
    integer :: ierr
    ierr = 0

    ! Finalize CCPP physics (run all _finalize routines)
    call ccpp_physics_finalize(cdata, suite_name=trim(ccpp_suite_name), &
                               ierr=ierr)

    ! error handling as above
    call ccpp_finalize(cdata, ierr=ierr)
    ! error handling as above

end subroutine physics_finalize

end module example_ccpp_host_cap

```

*Listing 6.3: Fortran template for a CCPP host model cap from `ccpp/framework/doc/DevelopersGuide/host_cap_template.F90`.*

The following sections describe two implementations of host model caps to serve as examples. For each of the functions listed above, a description for how it is implemented in each host model is included.

## 6.5.1 SCM Host Cap

The only build type available for the SCM is the dynamic build. The cap functions are mainly implemented in:

`gmtb-scm/scm/src/gmtb_scm.F90`

With smaller parts in:

`gmtb-scm/scm/src/gmtb_scm_type_defs.f90`

`gmtb-scm/scm/src/gmtb_scm_setup.f90`

`gmtb-scm/scm/src/gmtb_scm_time_integration.f90`

The host model *cap* is responsible for:

- Allocating memory for variables needed by physics

All variables and constants required by the physics are in the host-side metadata tables, `arg_table_physics_type` and `arg_table_gmtb_scm_physical_constants`, which are implemented in `gmtb_scm_type_defs.f90` and `gmtb_scm_physical_constants.f90`. To mimic the UFS Atmosphere and to hopefully reduce code maintenance, currently, the SCM uses GFS DDTs as sub-types within the physics DDT.



In `gmtb_scm_type_defs.f90`, the physics DDT has a create type-bound procedure (see subroutine `physics_create` and `type physics_type`), which allocates GFS sub-DDTs and other physics variables and initializes them with zeros. `physics%create` is called from `gmtb_scm.F90` after the initial SCM state has been set up.

- Allocating the `cdata` structure

The SCM uses a one-dimensional `cdata` array for `N` independent columns, i.e. in `gmtb_scm.F90`:

```
allocate(cdata(scm_state%n_cols))
```

- Calling the suite initialization subroutine

Within `scm_state%n_cols` loop in `gmtb_scm.F90` after initial SCM state setup and before first timestep, the suite initialization subroutine `ccpp_init` is called for each column with own instance of `cdata`, and takes three arguments, the name of the runtime SDF, the name of the `cdata` variable that must be allocated at this point, and `ierr`.

- Populating the `cdata` structure

Within the same `scm_state%n_cols` loop, but after the `ccpp_init` call, the `cdata` structure is filled in with real initialized values:

- `physics%Init_parm` (GFS DDT for setting up suite) are filled in from `scm_state%`
- call `GFS_suite_setup()`: similar to `GFS_initialize()` in the UFS Atmosphere, is called and includes:
- `%init/%create` calls for GFS DDTs
- initialization for other variables in physics DDT
- `ini` calls for legacy non-ccpp schemes
- call `physics%associate()`: to associate pointers in physics DDT with targets in `scm_state`, which contains variables that are modified by the SCM “dycore” (i.e. forcing).
- Actual `cdata` fill in through `ccpp_field_add` calls:

```
#include "ccpp_fields.inc"
```

This include file is auto-generated from `ccpp/scripts/ccpp_prebuild.py`, which parses tables in `gmtb_scm_type_defs.f90`.

- Providing interfaces to call the CCPP

- Calling `ccpp_physics_init()`

Within the same `scm_state%n_cols` loop but after `cdata` is filled, the physics initialization routines (`*_init()`) associated with the physics suite, group, and/or schemes are called at each column.

- Calling `ccpp_physics_run()`

At the first timestep, if the forward scheme is selected (i.e. `scm_state%time_scheme == 1`), call `do_time_step()` to apply forcing and `ccpp_physics_run()` calls at each column; if the leapfrog scheme is selected (i.e. `scm_state%time_scheme == 2`), call `ccpp_physics_run()` directly at each column.

At a later time integration, call `do_time_step()` to apply forcing and `ccpp_physics_run()` calls at each column. Since there is no need to execute anything between physics groups in the SCM, the `ccpp_physics_run` call is only given `cdata` and an error flag as arguments.

- Calling `ccpp_physics_finalize()` and `ccpp_finalize()`

`ccpp_physics_finalize()` and `ccpp_finalize()` are called after the time loop at each column.

## 6.5.2 UFS Atmosphere Host Cap

For the UFS Atmosphere, there are slightly different versions of the host cap implementation depending on the desired build type (dynamic or static). As discussed in [Chapter 8](#), these modes are controlled via appropriate strings included in the `MAKEOPTS` build-time argument. Within the source code, the three modes are executed within appropriate pre-processor directive blocks:

For any build that uses CCPP (dynamic or static):

```
#ifdef CCPP
#endif
```

For static (often nested within `#ifdef CCPP`):

```
#ifdef STATIC
#endif
```

The following text describes how the host cap functions listed above are implemented for the dynamic build only. Where the other modes of operation differ in their implementation, it will be called out.

- Allocating memory for variables needed by physics
  - Within the `atmos_model_init` subroutine of `atmos_model.F90`, the following statement is executed

```
allocate(IPD_Data)
```

`IPD_Data` is of `IPD_data_type`, which is defined in `IPD_typedefs.F90` as a synonym for `GFS_data_type` defined in `GFS_typedefs.F90`. This data type contains GFS-related DDTs (`GFS_statein_type`, `GFS_stateout_type`, `GFS_sfcprop_type`, etc.) as sub-types, which are defined in `GFS_typedefs.F90`.
- Allocating the `cdata` structures
- For the current implementation of the UFS Atmosphere, which uses a subset of fast physics processes tightly coupled to the dynamical core, three instances of `cdata` exist within the host model: `cdata_tile` to hold data for the fast physics, `cdata_domain` to hold data needed for all UFS Atmosphere blocks for the slow physics, and `cdata_block`, an array of `cdata` DDTs with dimensions of (number of blocks, number of threads) to contain data for individual block/thread combinations for the slow physics. All are defined as module-level variables in the `CCPP_data` module of `CCPP_data.F90`. The `cdata_block` array is allocated (since the number of blocks and threads is unknown at compile-time) as part of the 'init' step of the `CCPP_step` subroutine in `CCPP_driver.F90`. Note: Although the `cdata` containers are not used to hold the pointers to the physics variables for the static mode, they are still used to hold other CCPP-related information for that mode.
- Calling the suite initialization subroutine
  - Corresponding to the three instances of `cdata` described above, the `ccpp_init` subroutine is called within three different contexts, all originating from the `atmos_model_init` subroutine of `atmos_model.F90`:
  - For `cdata_tile` (used for the fast physics), the `ccpp_init` call is made from the `atmosphere_init` subroutine of `atmosphere.F90`. Note: when fast physics is used, this is the *first* call to `ccpp_init`, so it reads in the SDF and initializes the suite in addition to setting up the fields for `cdata_tile`.

- For `cdata_domain` and `cdata_block` used in the rest of the physics, the ‘init’ step of the `CCPP_step` subroutine in `CCPP_driver.F90` is called. Within that subroutine, `ccpp_init` is called once to set up `cdata_domain` and within a loop for every block/thread combination to set up the components of the `cdata_block` array. Note: as mentioned in the [CCPP API Section 6.4](#), when fast physics is used, the SDF has already been read and the suite is already setup, so this step is skipped and the suite information is simply copied from what was already initialized (`cdata_tile`) using the `cdata_target` optional argument.
- Populating the `cdata` structures
- When the dynamic mode is used, the `cdata` structures are filled with pointers to variables that are used by physics and whose memory is allocated by the host model. This is done using `ccpp_field_add` statements contained in the autogenerated include files. For the fast physics, this include file is named `ccpp_fields_fast_physics.inc` and is placed after the call to `ccpp_init` for `cdata_tile` in the `atmosphere_init` subroutine of `atmosphere.F90`. For populating `cdata_domain` and `cdata_block`, IPD data types are initialized in the `atmos_model_init` subroutine of `atmos_model.F90`. The `Init_parm` DDT is filled directly in this routine and `IPD_initialize` (pointing to `GFS_initialize` and for populating diagnostics and restart DDTs) is called in order to fill the GFS DDTs that are used in the physics. Once the IPD data types are filled, they are passed to the ‘init’ step of the `CCPP_step` subroutine in `CCPP_driver.F90` where `ccpp_field_add` statements are included in `ccpp_fields_slow_physics.inc` after the calls to `ccpp_init` for the `cdata_domain` and `cdata_block` containers.
- Note: for the static mode, filling of the `cdata` containers with pointers to physics variables is not necessary. This is because the autogenerated `caps` for the physics groups (that contain calls to the member schemes) can fill in the argument variables without having to retrieve pointers to the actual data. This is possible because the host model metadata tables (that are known at `ccpp_prebuild` time) contain all the information needed about the location (DDTs and local names) to pass into the autogenerated `caps` for their direct use.
- Providing interfaces to call the CCPP
  - Calling `ccpp_physics_init`
  - In order to call the initialization routines for the physics, `ccpp_physics_init` is called in the `atmosphere_init` subroutine of `atmosphere.F90` after the included `ccpp_field_add` calls for the fast physics. For the slow physics, the ‘physics\_init’ step of the `CCPP_step` subroutine in `CCPP_driver.F90` is invoked immediately after the call to the ‘init’ step in the `atmos_model_init` subroutine of `atmos_model.F90`. Within the ‘physics\_init’ step, calls to `ccpp_physics_init` for all blocks are executed.
  - Note: for the static mode, `ccpp_physics_init` is autogenerated and contained within `ccpp_static_api.F90`. As mentioned in the [CCPP API Section 6.4](#), it can be called to initialize groups as defined in the SDFs or the suite as a whole, depending on whether a group name is passed in as an optional argument.
  - Calling `ccpp_physics_run`
  - For actually running the physics within the FV3 time loop, `ccpp_physics_run` is called from a couple of different places in the FV3 source code. For the fast physics, `ccpp_physics_run` is called for the fast physics group from the `Lagrangian_to_Eulerian` subroutine of `fv_mapz.F90` within the dynamical core. For the rest of the physics, the subroutine `update_atmos_radiation_physics` in `atmos_model.F90` is called as part of the FV3 time loop. Within that subroutine, the various physics steps (defined as groups within a SDF) are called one after the other. The ‘time\_vary’ step of the `CCPP_step` subroutine within `CCPP_driver.F90` is called. Since this step is called over the entire domain, the call to `ccpp_physics_run` is done once using `cdata_domain` and the `time_vary` group. The ‘radiation’, ‘physics’, and ‘stochastics’ steps of the `CCPP_step` subroutine are called next. For each of these steps within `CCPP_step`, there is a loop over the number of blocks for calling

`ccpp_physics_run` with the appropriate group and component of the `cdata_block` array for the current block and thread.

- Note: The execution of calls to `ccpp_physics_run` is different for the three build types. For the static mode, `ccpp_physics_run` is called from `ccpp_static_api.F90` and contains autogenerated caps for groups and the suite as a whole as defined in the SDFs.
- calling `ccpp_physics_finalize` and `ccpp_finalize`
- At the conclusion of the FV3 time loop, calls to finalize the physics are executed. For the fast physics, `ccpp_physics_finalize` is called from the `atmosphere_end` subroutine of `atmosphere.F90`. For the rest of the physics, the ‘finalize’ step of the `CCPP_step` subroutine in `CCPP_driver.F90` is called from the `atmos_model_end` subroutine in `atmos_model.F90`. Within the ‘finalize’ step of `CCPP_step`, calls for `ccpp_physics_finalize` and `ccpp_finalize` are executed for every thread and block for `cdata_block`. Afterward, `ccpp_finalize` is called for `cdata_domain` and lastly, `cdata_tile`. (That is, the calls to `ccpp_finalize` are in reverse order than the calls to `ccpp_initialize`.) In addition, `cdata_block` is also deallocated in the ‘finalize’ step of `CCPP_step`.
- Note: for the static mode, `ccpp_physics_finalize` is autogenerated and contained within `ccpp_static_api.F90`. As mentioned in the [CCPP API Section 6.4](#), it can be called to finalize groups as defined in the current SDFs or the suite as a whole, depending on whether a group name is passed in as an optional argument.

## CCPP CODE MANAGEMENT

### 7.1 Organization of the Code

This chapter describes the organization of the code, provides instruction on the GitHub workflow and the code review process, and outlines the release procedure. It is assumed that the reader is familiar with using basic GitHub features and has a GitHub account set up.

#### 7.1.1 UFS Atmosphere

The *UFS* Atmosphere source code is contained in the NEMSFv3gfs code repository. CCPP users and developers should use the NEMSFv3gfs code and submodules maintained by GMTB in GitHub. For codes whose authoritative repository is in VLab, GMTB synchronizes these with VLab at periodic intervals.

<https://github.com/NCAR/NEMSFv3gfs>

There are five submodules referenced in the NEMSFv3gfs repository

<https://github.com/NCAR/ccpp-framework>

<https://github.com/NCAR/ccpp-physics>

<https://github.com/NCAR/FV3>

<https://github.com/NCAR/NEMS>

<https://github.com/NCAR/FMS>

Users have read-only access to these repositories and as such cannot accidentally destroy any important (shared) branches of these authoritative repositories.

Some of these repositories are public (no GitHub account required) and some are private. The public repositories (ccpp-framework, ccpp-physics, and Flexible Modeling System - FMS) may be used directly to read or create forks. Write permission is generally restricted, however. The private repositories require access - please send a request and your GitHub username to [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu).

The primary development by GMTB, including the latest CCPP developments, are maintained in the following branches:

Repository (GMTB development version)	Branch name
<a href="https://github.com/NCAR/NEMSFv3gfs">https://github.com/NCAR/NEMSFv3gfs</a>	gmtb/ccpp
<a href="https://github.com/NCAR/FV3">https://github.com/NCAR/FV3</a>	gmtb/ccpp
<a href="https://github.com/NCAR/ccpp-physics">https://github.com/NCAR/ccpp-physics</a>	master
<a href="https://github.com/NCAR/ccpp-framework">https://github.com/NCAR/ccpp-framework</a>	master
<a href="https://github.com/NCAR/NEMS">https://github.com/NCAR/NEMS</a>	gmtb/ccpp
<a href="https://github.com/NCAR/FMS">https://github.com/NCAR/FMS</a>	GFS-FMS

## 7.1.2 SCM

CCPP developers should use the SCM code and submodules maintained by GMTB in GitHub. The `ccpp-framework` and `ccpp-physics` repositories are used by the two models, creating a two-way dependency between them (see *Committing Changes to your Fork* for implications on the code review and testing procedure).

<https://github.com/NCAR/gmtb-scm>

As for NEMSFv3gfs, there are two submodules referenced in the `gmtb-scm` repository:

<https://github.com/NCAR/ccpp-framework>

<https://github.com/NCAR/ccpp-physics>

Users have read-only access to these repositories and as such cannot accidentally destroy any important (shared) branches of these authoritative repositories. Both CCPP repositories are public (no GitHub account required) and may be used directly to read or create forks. Write permission is generally restricted, however. The SCM repository is private, to request access please send a message and your GitHub username to [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu).

The primary development by GMTB, including the latest CCPP developments, are maintained in the following branches:

Repository (GMTB development version)	Branch name
<a href="https://github.com/NCAR/gmtb-scm">https://github.com/NCAR/gmtb-scm</a>	master
<a href="https://github.com/NCAR/ccpp-physics">https://github.com/NCAR/ccpp-physics</a>	master
<a href="https://github.com/NCAR/ccpp-framework">https://github.com/NCAR/ccpp-framework</a>	master

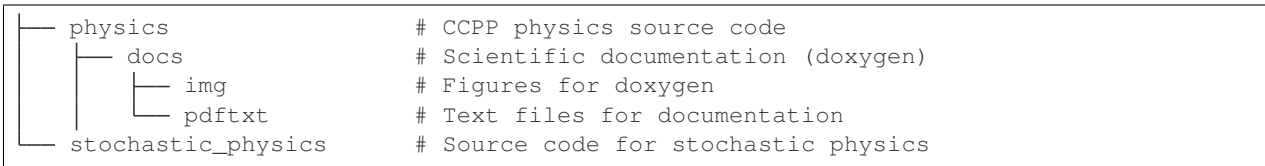
## 7.1.3 Directory Structure of `ccpp/framework`

The following is the directory structure for the `ccpp/framework` (condensed version):

—	cmake	# cmake files for building
—	doc	# Documentation for design/implementation # and developers guide
	CCPPtechnical	# CCPP Technical documentation
	build	
	source	
	_static	
	_templates	
	common	# Latex style file used by various documentation efforts
	DevelopersGuide	
	images	
	img	
—	schemes	# Example <code>ccpp_prebuild_config.py</code>
	check	
—	scripts	# Scripts for <code>ccpp_prebuild.py</code> , metadata # parser, etc.
	fortran_tools	
	parse_tools	
—	src	# CCPP framework source code
	tests	# SDFs and code for testing
—	test	
	nemsfv3gfs	# NEMSFv3gfs regression test scripts
—	tests	# Development for framework upgrades

### 7.1.4 Directory Structure of ccpp/physics

The following is the directory structure for the ccpp/physics (condensed version):



## 7.2 GitHub Workflow (setting up development repositories)

The CCPP development practices make use of the GitHub forking workflow. For users not familiar with this concept, this website provides some background information and a tutorial.

### 7.2.1 Creating Forks

The GitHub forking workflow relies on forks (personal copies) of the shared repositories on GitHub. These forks need to be created only once, and only for directories that users will contribute changes to. The following steps describe how to create a fork for the example of the ccpp-physics submodule/repository:

Go to <https://github.com/NCAR/ccpp-physics> and make sure you are signed in as your GitHub user.

Select the “fork” button in the upper right corner.

- If you have already created a fork, this will take you to your fork.
- If you have not yet created a fork, this will create one for you.

Note that the repo name in the upper left (blue) will be either “NCAR” or “your GitHub name” which tells you which fork you are looking at.

Note that personal forks are not required until a user wishes to make code contributions. The procedure for how to check out the code laid out below can be followed without having created any forks beforehand.

### 7.2.2 Checking out the Code (NEMSfv3gfs)

Because of the large number of submodules of the central repository NEMSfv3gfs, we strongly recommend setting up passwordless access to GitHub via https or ssh (see <https://help.github.com/categories/authenticating-to-github>). In the following, we assume access is via https rather than ssh.

Start with checking out the main repository from the NCAR GitHub

```

git clone -b gmtb/ccpp https://github.com/NCAR/NEMSfv3gfs
cd NEMSfv3gfs
git submodule init
git submodule update

```

At this point, you have checked out the correct branches of all six repositories. Each repository is connected to a remote destination (simply called remote in GitHub language), which points to the NCAR GitHub pages and is labeled as origin. For consistency with the CCPP workflow, we recommend renaming the NCAR remote destination to upstream for all six repositories in order to guarantee that your development will be pushed to your fork and not directly to the NCAR repository. For example:

```
cd ccpp/framework
git remote rename origin upstream
cd ../../
```

Checking out remote branches as submodules means that your local branches are in a detached head state, since the submodule points to a specific commit. As long as you are not making any modifications in a certain repository, this is not a problem. If during your development changes are made to the corresponding upstream branch, you can simply navigate to this repository and check out the updated version (example NEMS):

```
cd NEMS
git remote update
git checkout upstream/gmtb/ccpp
cd ..
```

However, if you are making changes in a repository (submodule or main repository), you must create a local branch, for example in NEMSfv3gfs:

```
git checkout -b my_local_development_branch
```

Once you are ready to contribute the code to the upstream repository, you need to create a pull request (PR; see next section). In order to do so, you will use your own fork of this repository (see previous section) and configure your fork as an additional remote destination, which we typically label as origin. For the same example NEMSfv3gfs:

```
git remote add origin https://github.com/YOUR_GITHUB_USER/NEMSfv3gfs
git remote update
```

Then, push your local branch to your fork:

```
git push origin my_local_development_branch
```

For each repository/submodule, you can check the configured remote destinations and all existing branches (remote and local):

```
git remote -v show
git remote update
git branch -a
```

As opposed to branches without modifications described in step 3, changes to the upstream repository can be brought into the local branch by pulling them down. For the example of submodule FV3 (where a local branch is checked out):

```
cd FV3
git remote update
git pull upstream gmtb/ccpp
```

## 7.2.3 Checking out the Code (SCM)

The process for checking out SCM is described in the following, assuming access via https rather than ssh. We strongly recommend setting up passwordless access to GitHub (see <https://help.github.com/categories/authenticating-to-github>).

Start with checking out the main repository from the NCAR GitHub

```
git clone https://github.com/NCAR/gmtb-scm
cd gmtb-scm
```

(continues on next page)



(continued from previous page)

```
git submodule init
git submodule update
```

At this point, you have checked out the correct branches of all three repositories. Each repository is connected to a remote destination (simply called remote in GitHub language), which points to the NCAR GitHub pages and is labeled as origin. For consistency with the CCPP workflow, we recommend renaming the NCAR remote destination to upstream for all repositories. For example:

```
cd ccpp/framework
git remote rename origin upstream
cd ../../
```

Checking out remote branches means that your local branches are in a detached state, since you cannot commit directly to a remote branch. As long as you are not making any modifications in a certain repository, this is not a problem. If during your development work changes are made to the corresponding upstream branch, you can simply navigate to this repository and check out the updated version (example ccpp-physics):

```
cd ccpp/physics
git remote update
git checkout upstream/master
cd ../../
```

However, if you are making changes in a repository (submodule or main repository), you must create a local branch, for example in gmtb-scm:

```
git checkout -b my_local_development_branch
```

Once you are ready to contribute the code to the upstream repository, you need to create a PR (see next section). In order to do so, you first need to create your own fork of this repository (see previous section) and configure your fork as an additional remote destination, which we typically label as origin. For the same example gmtb-scm:

```
git remote add origin https://github.com/YOUR_GITHUB_USER/gmtb-scm
git remote update
```

Then, push your local branch to your fork:

```
git push origin my_local_development_branch
```

For each repository/submodule, you can check the configured remote destinations and all existing branches (remote and local):

```
git remote -v show
git remote update
git branch -a
```

As opposed to branches without modifications described in step 3, changes to the upstream repository can be brought into the local branch by pulling them down. For the example of submodule ccpp-physics (where a local branch is checked out):

```
cd ccpp/physics
git remote update
git pull upstream master
```

## 7.3 Committing Changes to your Fork

Once you have your fork set up to begin code modifications, you should check that the cloned repositories upstream and origin are set correctly:

```
git remote -v
```

This should point to your fork as origin and the repository you cloned as upstream:

```
origin          https://github.com/YOUR_GITHUB_USER/ccpp-physics (fetch)
origin          https://github.com/YOUR_GIRHUB_USER/ccpp-physics (push)
upstream       https://github.com/NCAR/ccpp-physics (fetch)
upstream       https://github.com/NCAR/ccpp-physics (push)
```

Also check what branch you are working on:

```
git branch
```

This command will show what branch you have checked out on your fork:

```
* features/my_local_development_branch
  master
```

After making modifications and testing, you can commit the changes to your fork. First check what files have been modified:

```
git status
```

This git command will provide some guidance on what files need to be added and what files are “untracked”. To add new files or stage modified files to be committed:

```
git add filename1 filename2
```

At this point it is helpful to have a description of your changes to these files documented somewhere, since when you commit the changes, you will be prompted for this information. To commit these changes to your local repository and push them to the development branch on your fork:

```
git commit
git push origin features/my_local_development_branch
```

When this is done, you can check the status again:

```
git status
```

This should show that your working copy is up to date with what is in the repository:

```
On branch features/my_local_development_branch
Your branch is up to date with 'origin/features/my_local_development_branch'.
nothing to commit, working tree clean
```

At this point you can continue development or create a PR as discussed in the next section.

## 7.4 Contributing Code, Code Review Process

Once your development is mature, and the testing has been completed (see next section), you are ready to create a PR using GitHub to propose your changes for review.

### 7.4.1 Creating a PR

Go to the [github.com](https://github.com) web interface, and navigate to your repository fork and branch. In most cases, this will be in the `ccpp-physics` repository, hence the following example:

Navigate to: <https://github.com/<yourusername>/ccpp-physics>

Use the drop-down menu on the left-side to select a branch to view your development branch

Use the button just right of the branch menu, to start a “New Pull Request”

Fill in a short title (one line)

Fill in a detailed description, including reporting on any testing you did

Click on “Create pull request”

Several people (aka CODEOWNERS) are automatically added to the list of reviewers on the right hand side. If others should be reviewing the code, click on the “reviewers” item on the right hand side and enter their GitHub usernames

Once the PR has been approved, the change is merged to master by one of the code owners. If there are pending conflicts, this means that the code is not up to date with the trunk. To resolve those, pull the target branch from upstream as described above, solve the conflicts and push the changes to the branch on your fork (this also updates the PR).

Note. GitHub offers a draft pull request feature that allows users to push their code to GitHub and create a draft PR. Draft PRs cannot be merged and do not automatically initiate notifications to the CODEOWNERS, but allow users to prepare the PR and flag it as “ready for review” once they feel comfortable with it.



## TECHNICAL ASPECTS OF THE *CCPP PREBUILD*

### 8.1 *Prebuild* Script Function

The *CCPP prebuild* script `ccpp/framework/scripts/ccpp_prebuild.py` is the central piece of code that connects the host model with the *CCPP-Physics* schemes (see Figures 3.2 and 3.3). This script must be run before compiling the *CCPP-Physics* library and the host model cap. This may be done manually or as part of a host model build-time script. In the case of the SCM, `ccpp_prebuild.py` must be run manually, as it not incorporated in that model's build system. In the case of NEMSfv3gfs, `ccpp_prebuild.py` can be run manually or automatically as a step in the build system.

The *CCPP prebuild* script automates several tasks based on the information collected from the metadata tables on the host model side and from the individual physics schemes (see Figure 8.1):

- Compiles a list of variables required to run all schemes in the *CCPP-Physics* pool.
- Compiles a list of variables provided by the host model.
- Matches these variables by their `standard_name`, checks for missing variables and mismatches of their attributes (e.g., units, rank, type, kind) and processes information on optional variables.
- For the static build only, filters out unused variables for a given suite.
- For the dynamic build only, creates Fortran code (`ccpp_modules_*.inc`, `ccpp_fields_*.inc`) that stores pointers to the host model variables in the `cdata` structure.
- Autogenerates software caps as appropriate, depending on the build type.
  - If dynamic, the script generates individual caps for all physics schemes.
  - If static, the script generates caps for the suite as a whole and physics groups as defined in the input *SDFs*; in addition, the *CCPP* API for the static build is generated.
- Populates makefiles with schemes, scheme dependencies, and caps. For the static build, statements to compile the static *CCPP* API are included as well.

### 8.2 Script Configuration

To connect the *CCPP* with a host model XYZ, a Python-based configuration file for this model must be created in the host model's repository. The easiest way is to copy an existing configuration file for the TEST model in sub-directory `schemes/check` of the *ccpp-Framework* repository. The configuration in `ccpp_prebuild_config.py` depends largely on (a) the directory structure of the host model itself, (b) where the `ccpp-framework` and the `ccpp-physics` directories are located relative to the directory structure of the host model, and (c) from which directory the `ccpp_prebuild.py` script is executed before/during the build process (this is referred to as `basedir` in `ccpp_prebuild_config_XYZ.py`).

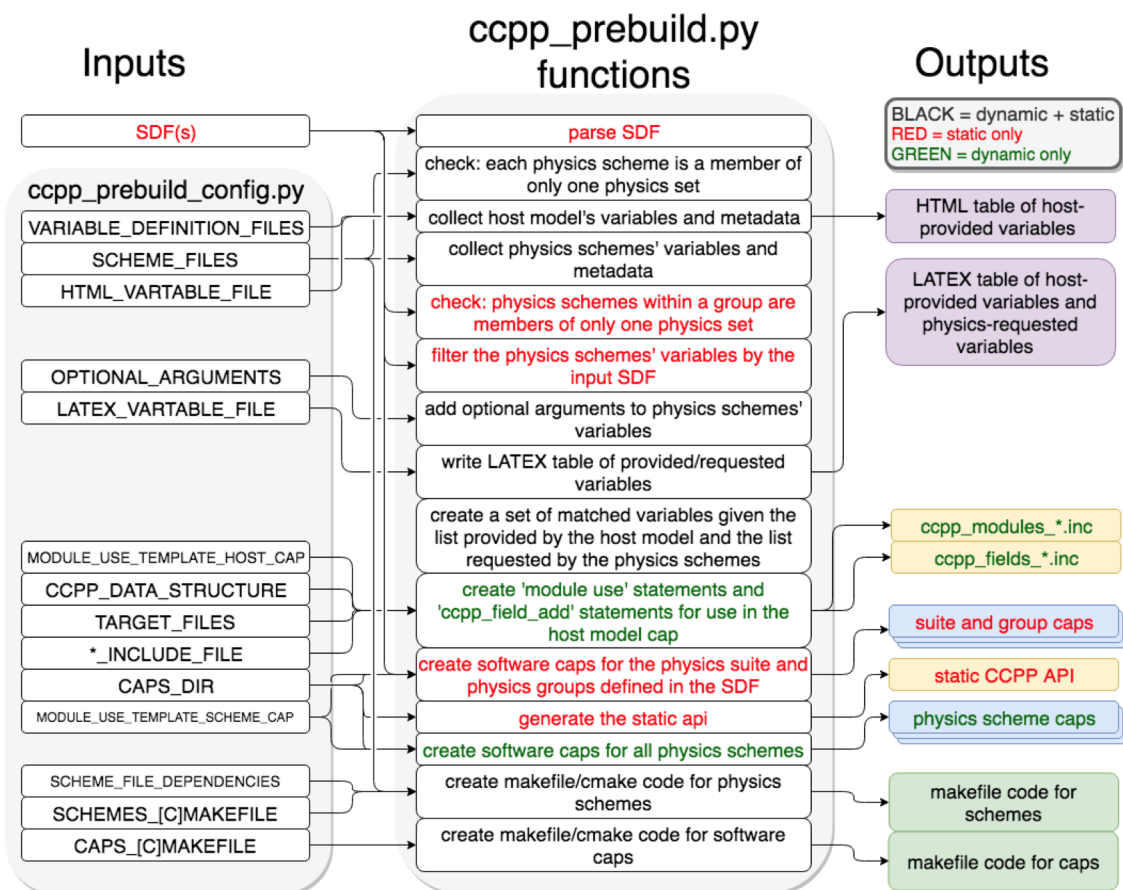


Fig. 8.1: Schematic of tasks automated by the `ccpp_prebuild.py` script and associated inputs and outputs. Red text denotes entries that are valid for the static build only, green text entries that are valid for the dynamic build only, and black text denotes entries valid for both the dynamic and static builds. The majority of the input is controlled through the host-model dependent `ccpp_prebuild_config.py`, whose user-editable variables are included as all-caps within the `ccpp_prebuild_config.py` bubble. Outputs are color-coded according to their utility: purple outputs are informational only (useful for developers, but not necessary to run the code), yellow outputs are used within the host model, blue outputs are connected to the physics, and green outputs are used in the model build.

*Listing 8.1* contains an example for the CCPP-SCM prebuild config. Here, it is assumed that both `ccpp-framework` and `ccpp-physics` are located in directories `ccpp/framework` and `ccpp/physics` of the top-level directory of the host model, and that `ccpp_prebuild.py` is executed from the same top-level directory.

```
# Host model identifier
HOST_MODEL_IDENTIFIER = "TEST"
# Add all files with metadata tables on the host model side,
# relative to basedir = top-level directory of host model
VARIABLE_DEFINITION_FILES = [
    'scm/src/gmtb_scm_type_defs.f90',
    'scm/src/gmtb_scm_physical_constants.f90'
]
# Add all physics scheme dependencies relative to basedir - note that the CCPP
# rules stipulate that dependencies are not shared between the schemes!
SCHEME_FILES_DEPENDENCIES = [] # can be empty
# Add all physics scheme files relative to basedir
SCHEME_FILES = {
# Relative path : [ list of sets in which scheme may be called ]
'ccpp/physics/physics/GFS_DCNV_generic.f90' : ['physics'],
'ccpp/physics/physics/sfc_sice.f' : ['physics'],
}
# Auto-generated makefile/cmakefile snippets that contains all schemes
SCHEMES_MAKEFILE = 'ccpp/physics/CCPP_SCHEMES.mk'
SCHEMES_CMAKEFILE = 'ccpp/physics/CCPP_SCHEMES.cmake'
# CCPP host cap in which to insert the ccpp_field_add statements;
# determines the directory to place ccpp_{modules,fields}.inc
TARGET_FILES = ['scm/src/gmtb_scm.f90', ]
# Auto-generated makefile/cmakefile snippets that contains all caps
CAPS_MAKEFILE = 'ccpp/physics/CCPP_CAPS.mk'
CAPS_CMAKEFILE = 'ccpp/physics/CCPP_CAPS.cmake'
# Directory where to put all auto-generated physics caps
CAPS_DIR = 'ccpp/physics/physics'
# Directory where the suite definition files are stored
SUITES_DIR = 'ccpp/suites'

# Optional arguments - only required for schemes that use optional arguments.
# ccpp_prebuild.py will throw an exception if it encounters a scheme subroutine with
↳ optional arguments if no entry is made here. Possible values are:
OPTIONAL_ARGUMENTS = {
# 'subroutine_name_1' : 'all', #'subroutine_name_2' : 'none', #'subroutine_name_3' : [
↳ 'var1', 'var2'],}
# Names of Fortran include files in the host model cap (do not change);
# both files will be written to the directory of each target file
MODULE_INCLUDE_FILE = 'ccpp_modules.inc'
FIELDS_INCLUDE_FILE = 'ccpp_fields.inc'
# Directory where to write static API to
STATIC_API_DIR = 'scm/src'

# HTML document containing the model-defined CCPP variables
HTML_VARIABLE_FILE = 'ccpp/physics/CCPP_VARIABLES.html'
# LaTeX document containing the provided vs requested CCPP variables
LATEX_VARIABLE_FILE = 'ccpp/framework/doc/DevelopersGuide/CCPP_VARIABLES.tex'
##### Template code to generate include files #####
# Name of the CCPP data structure in the host model cap;
# in the case of SCM, this is a vector with loop index i
CCPP_DATA_STRUCTURE = 'cdata(i)'

# EOF
```

*Listing 8.1: CCPP prebuild config for SCM (shortened)*

Although most of the variables in the `ccpp_prebuild_config.py` script are described by in-line comments in the code listing above and their use is described in [Figure 8.1](#), some clarifying comments are in order regarding the `SCHEME_FILES` variable. This is a list of CCPP-compliant physics scheme entry/exit point source files. For each item in this list, a list of physics “sets” in which the scheme may be executed is included. A physics set refers to a collection of physics schemes that are able to be called together and executed in one software domain of a host model that do not share variables with schemes from another physics set. This feature was included to cater to the needs of the FV3GFS, which provides a clear-cut example of this concept. In this model, part of the microphysics scheme needed to be coupled more tightly with the dynamics, so this part of the microphysics code was put into a physics set labeled “fast\_physics” which is executed within the dycore code. The variables in this physics set are distinct (in memory, due to a lack of a model variable registry) from variables used in the rest of the physics, which are part of the “slow\_physics” set. In the future, it may be necessary to have additional sets, e.g. for chemistry or separate surface model components that do not share data/memory with other model components. For simpler models such as the GMTB SCM, only one physics set (labeled “physics”) is necessary. The concept of physics sets is different from physics “groups”, which are capable of sharing variables among their members and between groups but are used to organize schemes into sequential, callable units.

### 8.3 Running `ccpp_prebuild.py`

Once the configuration in `ccpp_prebuild_config.py` is complete, the `ccpp_prebuild.py` script can be run from the top level directory. For the SCM, this script must be run (in dynamic build mode only) to reconcile data provided by the SCM with data required by the physics schemes before compilation and to generate physics caps and makefile segments. For the *UFS* Atmosphere host model, the `ccpp_prebuild.py` script is called automatically by the NEMSFv3gfs build system when the *CCPP* build is requested (by running the *CCPP* regression tests or by passing the option `CCPP=Y` and others to the `compile.sh` script; see the `compile` commands defined in the *CCPP* regression test configurations for further details).

For developers adding a CCPP-compliant physics scheme, running `ccpp_prebuild.py` periodically is recommended to check that the metadata in the argument tables match between the host model and the physics scheme. For the *UFS* Atmosphere, running `ccpp_prebuild.py` manually is identical to running it for the SCM (since the relative paths to their respective `ccpp_prebuild_config.py` files are identical), except it may be necessary to add the `--static` and `--suites` command-line arguments for the static option.

As alluded to above, the `ccpp_prebuild.py` script has six command line options, with the path to a host-model specific configuration file (`--config`) being the only necessary input option:

```
-h, --help show this help message and exit
--config PATH_TO_CONFIG/config_file path to CCPP prebuild configuration file
--clean remove files created by this script, then exit
--debug enable debugging output
--static enable a static build for a given suite definition file
--suites SUITES SDF(s) to use (comma-separated, for static build only, without path)
```

So, the simplest possible invocation of the script (called from the host model’s top level directory) would be:

```
./ccpp/framework/scripts/ccpp_prebuild.py --config ./ccpp/config/ccpp_prebuild_config.py
↪py [--debug]
```

which assumes a dynamic build with a configuration script located at the specified path. The `debug` option can be used for more verbose output from the script.

For a static build (described above), where the *CCPP-Framework* and the physics libraries are statically linked to the executable and a set of one or more suites are defined at build-time, the `--suites` and `--static` options must be included. The *SDF*(s) should be specified using the `--suites` command-line argument. Such files are included



with the SCM and NEMSFV3GFS repositories, and must be included with the code of any host model to use the *CCPP*. Unless the `--static` command-line argument is used with the script, it will assume dynamically linked libraries. An example of a static build using two *SDFs* is:

```
./ccpp/framework/scripts/ccpp_prebuild.py --config=./ccpp/config/ccpp_prebuild_config.  
↪py --static \  
--suites=FV3_GFS_v15,FV3_CPT_v0
```

If the *CCPP* *prebuild* step is successful, the last output line will be:

```
INFO: CCPP prebuild step completed successfully.
```

To remove all files created by `ccpp_prebuild.py`, for example as part of a host model's `make clean` functionality, execute the same command as before, but with `--clean` appended:

```
./ccpp/framework/scripts/ccpp_prebuild.py --config=./ccpp/config/ccpp_prebuild_config.  
↪py --static \  
--suites=FV3_GFS_v15,FV3_CPT_v0 --clean
```

## 8.4 Troubleshooting

If invoking the `ccpp_prebuild.py` script fails, some message other than the success message will be written to the terminal output. Specifically, the terminal output will include informational logging messages generated from the script, any error messages written to the python logging utility, and a Python traceback that pinpoints the line within the script where the error caused failure. Some common errors (minus the typical logging output and traceback output) and solutions are described below, with non-bold font used to denote aspects of the message that will differ depending on the problem encountered. This is not an exhaustive list of possible errors, however. For example, in this version of the code, there is no cross-checking that the information provided in the metadata tables corresponds to the actual Fortran code, so even though `ccpp_prebuild.py` may complete successfully, there may be related compilation errors later in the build process. For further help with an undescribed error, please contact [gmtb-help@ucar.edu](mailto:gmtb-help@ucar.edu).

### 1. **ERROR: Configuration file erroneous/path/to/config/file not found**

- Check that the path entered for the `--config` command line option points to a readable configuration file.

### 2. **KeyError: 'erroneous\_scheme\_name' when using the --static and --suites options**

- This error indicates that a scheme within the supplied *SDFs* does not match any scheme names found in the `SCHEME_FILES` variable of the supplied configuration file that lists scheme source files. Double check that the scheme's source file is included in the `SCHEME_FILES` list and that the scheme name that causes the error is spelled correctly in the supplied *SDFs* and matches what is in the source file (minus any `*_init`, `*_run`, `*_finalize` suffixes).

### 3. **CRITICAL: Suite definition file erroneous/path/to/SDF.xml not found.**

#### **Exception: Parsing suite definition file erroneous/path/to/SDF.xml failed.**

- Check that the path `SUITES_DIR` in the *CCPP* *prebuild* config and the names entered for the `--suites` command line option are correct.

### 4. **ERROR: Scheme file path/to/offending/scheme/source/file belongs to multiple physics sets: set1, set2**

#### **Exception: Call to check\_unique\_pset\_per\_scheme failed.**

- This error indicates that a scheme defined in the `SCHEME_FILES` variable of the supplied configuration file belongs to more than one set. Currently, a scheme can only belong to one physics set.

5. **ERROR:** Group `group1` contains schemes that belong to multiple physics sets: `set1,set2`

**Exception:** `Call to check_unique_pset_per_group failed.`

- This error indicates that one of the groups defined in the supplied *SDF*(s) contains schemes that belong to more than one physics set. Make sure that the group is defined correctly in the *SDF*(s) and that the schemes within the group belong to the same physics set (only one set per scheme is allowed at this time).

6. **INFO:** Parsing metadata tables for variables provided by host model...

**IOError:** `[Errno 2] No such file or directory: 'erroneous_file.f90'`

- Check that the paths specified in the `VARIABLE_DEFINITION_FILES` of the supplied configuration file are valid and contain CCPP-compliant host model variable metadata tables.

7. **Exception:** `Encountered invalid line "some fortran" in argument table variable_metadata_table_name`

- This is likely the result of not ending a variable metadata table with a line containing only `!!`. Check that the formatting of the offending variable metadata table is correct.

8. **Exception:** `Error parsing variable entry "erroneous variable metadata table entry data" in argument`

- Check that the formatting of the metadata entry described in the error message is OK. The number of metadata columns must match the table header and each entry's columns must be separated by a `'` character (be sure that the `'` character was not used accidentally).

9. **Exception:** `New entry for variable var_name in argument table variable_metadata_table_name is i`

Existing: Contents of `<mkcap.Var object at 0x10299a290>` (\* = mandatory for compatibility):

```

standard_name = var_name *
long_name =
units = various *
local_name =
type = real *
rank = (,,:) *
kind = kind_phys *
intent = none
optional = F
target = None
container = MODULE_X TYPE_Y

```

vs. new: Contents of `<mkcap.Var object at 0x10299a310>` (\* = mandatory for compatibility):

```

standard_name = var_name *
long_name =
units = frac *
local_name =
type = real *
rank = (,,:) *
kind = kind_phys *
intent = none

```

```

optional = F
target = None
container = MODULE_X TYPE_Y

```

- This error is associated with a variable that is defined more than once (with the same standard name) on the host model side. Information on the offending variables is provided so that one can provide different standard names to the different variables.

**10. Exception: Scheme name differs from module name: module\_name= “X” vs. scheme\_name= “Y”**

- Make sure that each scheme in the errored module begins with the module name and ends in either \*\_init, \*\_run, or \*\_finalize.

**11. Exception: Encountered closing statement "end" without descriptor (subroutine, module**

- This script expects that subroutines and modules end with descriptor and name, e.g. ‘end subroutine subroutine\_name’.

**12. Exception: New entry for variable var\_name in argument table of subroutine scheme\_subroutine**

existing: Contents of <mkcap.Var object at 0x10299a290> (\* = mandatory for compatibility):

```

standard_name = var_name *
long_name =
units = various *
local_name =
type = real *
rank = (:,:) *
kind = kind_phys *
intent = none
optional = F
target = None
container = MODULE_X TYPE_Y

```

vs. new: Contents of <mkcap.Var object at 0x10299a310> (\* = mandatory for compatibility):

```

standard_name = var_name *
long_name =
units = frac *
local_name =
type = real *
rank = (:,) *
kind = kind_phys *
intent = none
optional = F
target = None
container = MODULE_X TYPE_Y

```

- This error is associated with physics scheme variable metadata entries that have the same standard name with different mandatory properties (either units, type, rank, or kind currently – those attributes denoted with a \*). This error is distinguished from the error described in 9 above, because the error message mentions “in argument table of subroutine” instead of just “in argument table”.

13. **ERROR: Check that all subroutines in module `module_name` have the same root name:**  
i.e. `scheme_A_init`, `scheme_A_run`, `scheme_A_finalize` Here is a list of the subroutine names for scheme `scheme_name`: `scheme_name_finalize`, `scheme_name_run`  
\* All schemes must have `*_init`, `*_run`, `*_finalize` subroutines contained within its entry/exit point module.
14. **ERROR: Variable `X` requested by `MODULE_Y SCHEME_Z SUBROUTINE_A` not provided by the model**  
Exception: Call to `compare_metadata` failed.
- A variable requested by one or more physics schemes is not being provided by the host model. If the variable exists in the host model but is not being made available for the *CCPP*, an entry must be added to one of the host model variable metadata tables.
15. **ERROR: error, variable `X` requested by `MODULE_Y SCHEME_Z SUBROUTINE_A` cannot be identified**
- A variable is defined in the host model variable metadata tables more than once (with the same standard name). Remove the offending entry or provide a different standard name for one of the duplicates.
16. **ERROR: incompatible entries in metadata for variable `var_name`:**
- provided: Contents of `<mkcap.Var object at 0x104883210>` (\* = mandatory for compatibility):
- ```
standard_name = var_name *
long_name =
units = K *
local_name =
type = real *
rank = *
kind = kind_phys *
intent = none
optional = F
target = None
container =
```
- requested: Contents of `<mkcap.Var object at 0x10488ca90>` (\* = mandatory for compatibility):
- ```
standard_name = var_name *
long_name =
units = none *
local_name =
type = real *
rank = *
kind = kind_phys *
intent = in
optional = F
target = None
container =
```
17. **Exception: Call to `compare_metadata` failed.**
- This error indicates a mismatch between the attributes of a variable provided by the host model and what is requested by the physics. Specifically, the units, type, rank, or kind don't match for a given variable standard name. Double-check that the attributes for the provided and requested mismatched variable are accurate. If after checking the attributes are indeed mismatched, reconcile as appropriate

(by adopting the correct variable attributes either on the host or physics side).

Note: One error that the `ccpp_prebuild.py` script will not catch is if a physics scheme lists a variable in its actual (Fortran) argument list without a corresponding entry in the subroutine's variable metadata table. This will lead to a compilation error when the autogenerated scheme cap is compiled:

```
Error: Missing actual argument for argument 'X' at (1)
```



## BUILDING AND RUNNING HOST MODELS

The following instructions describe how to compile and run the CCPP code with the SCM (Section 9.1) and with the UFS Atmosphere (Section 9.2). Instructions are for the *Theia*, *Jet* and *Cheyenne* computational platforms, with examples on how to run the code on *Theia*.

### 9.1 SCM

One option for a CCPP host model is the SCM. This can be a valuable tool for diagnosing the performance of a physics suite, from validating that schemes have been integrated into a suite correctly to deep dives into how physical processes are being represented by the approximating code. In fact, this SCM likely serves as the simplest example for using the CCPP and its framework in an atmospheric model.

#### 9.1.1 System Requirements, Libraries, and Tools

The source code for the SCM and CCPP component is in the form of programs written in FORTRAN, FORTRAN 90, and C. In addition, the I/O relies on the netCDF libraries. Beyond the standard scripts, the build system relies on the use of the Python scripting language, along with cmake, GNU make and date.

The basic requirements for building and running the CCPP and SCM bundle are listed below. The versions listed reflect successful tests and there is no guarantee that the code will work with different versions.

- **FORTRAN 90+ compiler versions**
  - ifort 18.0.1.163 and 19.0.2
  - gfortran 6.2, 8.1, and 9.1
  - pgf90 17.7 and 17.9
- **C compiler versions**
  - icc v18.0.1.163 and 19.0.2
  - gcc 6.2 and 8.1
  - AppleClang 10.0.0.10001145
  - pgcc 17.7 and 17.9
- cmake versions 2.8.12.1, 2.8.12.2, and 3.6.2
- netCDF with HDF5, ZLIB and SZIP versions 4.3.0, 4.4.0, 4.4.1.1, 4.5.0, 4.6.1, and 4.6.3 (not 3.x)
- Python versions 2.7.5, 2.7.9, and 2.7.13 (not 3.x)
- Libxml2 versions 2.2 and 2.9.7 (not 2.9.9)

Because these tools and libraries are typically the purview of system administrators to install and maintain, they are considered part of the basic system requirements. Further, there are several utility libraries as part of the NCEPLibs package that must be installed prior to building the SCM.

- bacio v2.0.1 - Binary I/O library
- sp v2.0.2 - Spectral Transformation Library
- w3nco v2.0.6 - GRIB decoder and encoder library

These libraries are prebuilt on most NOAA machines using the Intel compiler. For those needing to build the libraries themselves, GMTB recommends using the source code from GitHub at <https://github.com/NCAR/NCEPLibs.git>, which includes build files for various compilers and machines using OpenMP flags and which are thread-safe. Instructions for installing NCEPLibs are included on the GitHub repository webpage, but for the sake of example, execute the following for obtaining and building from source in `/usr/local/NCEPLibs` on a Mac:

```
mkdir /usr/local/NCEPLibs
cd /usr/local/src
git clone https://github.com/NCAR/NCEPLibs.git
cd NCEPLibs
./make_ncep_libs.sh -s macosx -c gnu -d /usr/local/NCEPLibs -o 1
```

Once NCEPLibs is built, the `NCEPLIBS_DIR` environment variable must be set to the location of the installation. For example, if NCEPLibs was installed in `/usr/local/NCEPLibs`, one would execute

```
export NCEPLIB_DIR=/usr/local/NCEPLibs
```

If using *Theia* or *Cheyenne* HPC systems, this environment variable is automatically set to an appropriate installation of NCEPLibs on those machines through use of one of the setup scripts described below.

## 9.1.2 Building and Running the SCM

Instructions for downloading the code are provided in [Chapter 7](#). Here are the steps to compile and run SCM:

- Run the CCPP *prebuild* script to match required physics variables with those available from the dycore (SCM) and to generate physics *caps* and makefile segments.

```
./ccpp/framework/scripts/ccpp_prebuild.py --config=./ccpp/config/ccpp_
↳prebuild_config.py [ -- debug ]
```

- Change directory to the top-level SCM directory.

```
cd scm
```

- (Optional) Run the machine setup script if necessary. This script loads compiler modules (Fortran 2003-compliant), netCDF module, etc. and sets compiler environment variables.

- source `etc/Theia_setup_intel.csh` (for csh) or `. etc/Theia_setup_intel.sh` (for bash)
- source `etc/Theia_setup_gnu.csh` (for csh) or `. etc/Theia_setup_gnu.sh` (for bash)
- source `etc/Theia_setup_pgi.csh` (for csh) or `. etc/Theia_setup_pgi.sh` (for bash)
- source `etc/Cheyenne_setup_intel.csh` (for csh) or `. etc/Cheyenne_setup_intel.sh` (for bash)
- source `etc/Cheyenne_setup_gnu.csh` (for csh) or `. etc/Cheyenne_setup_gnu.sh` (for bash)



- source etc/Cheyenne\_setup\_pgi.csh (for csh) or . etc/Cheyenne\_setup\_pgi.sh (for bash)
- source etc/UBUNTU\_setup.csh (for csh) or . etc/UBUNTU\_setup.sh (for bash) if following the instructions in doc/README\_UBUNTU.txt
- source etc/CENTOS\_setup.csh (for csh) or . etc/CENTOS\_setup.sh (for bash) if following the instructions in doc/README\_CENTOS.txt
- source etc/MACOSX\_setup.csh (for csh) or . etc/MACOSX\_setup.sh (for bash) if following the instructions in doc/README\_MACOSX.txt

---

**Note:** If using a local Linux or Mac system, we provide instructions for how to set up your development system (compilers and libraries) in doc/README\_{MACOSX, UBUNTU, CENTOS}.txt. If following these, you will need to run the respective setup script listed above. If your computing environment was previously set up to use modern compilers with an associated netCDF installation, it may not be necessary, although we recommend setting environment variables such as CC and FC. For version 3.0 and above, it is required to have the NETCDF environment variable set to the path of the netCDF installation that was compiled with the same compiler used in the following steps. Otherwise, the cmake step will not complete successfully.

---

- Make a build directory and change into it.

```
mkdir bin && cd bin
```

- Invoke cmake on the source code to build using one of the commands below.
- Without threading / OpenMP

```
cmake ../src
```

- With threading / OpenMP

```
cmake -DOPENMP=ON ../ src
```

- Debug mode

```
cmake -DCMAKE_BUILD_TYPE=Debug ../ src
```

- If cmake cannot find libxml2 because it is installed in a non-standard location, add the following to the cmake command.

```
-DPC_LIBXML_INCLUDEDIR=...
-DPC_LIBXML_LIBDIR=...
```

- Compile with make command. Add VERBOSE=1 to obtain more information on the build process.

```
make
```

Note that this will produce executable `gmtb_scm` and library `libccppphys.so.X.Y.Z` (where X is a major version number; Y is a minor version number, and Z is a patchlevel) and `libccppphys.so`, which is a link to `libccppphys.so.X.Y.Z`. The library, which is located in `ccpp/lib`, will be dynamically linked to the executable at runtime.

If compilation successfully completes, a working executable named `gmtb_scm` will have been created in the `bin` directory.

Although `make clean` is not currently implemented, an out-of-source build is used, so all that is required to clean the `build/run` directory is (from the `bin` directory)

```
pwd #confirm that you are in the build/run directory before deleting files
rm -rfd *
```

**Warning:** This command can be dangerous (deletes files without confirming), so make sure that you're in the right directory before executing!

There are several test cases provided with this version of the SCM. For all cases, the SCM will go through the time steps, applying forcing and calling the physics defined in the chosen SDF using physics configuration options from an associated namelist. The model is executed through one of two Python run scripts that are pre-staged into the `bin` directory: `run_gmtb_scm.py` or `multi_run_gmtb_scm.py`. The former sets up and runs one integration while the latter sets up and runs several integrations serially.

### Single Run Script Usage

Running a case requires three pieces of information: the case to run (consisting of initial conditions, geolocation, forcing data, etc.), the physics suite to use (through a CCPP SDF), and a physics namelist (that specifies configurable physics options to use). Cases are set up via their own namelists in `./etc/case_config`. A default physics suite is provided as a user-editable variable in the script and default namelists are associated with each physics suite (through `./src/default_namelists.py`), so, technically, one must only specify a case to run with the SCM. The single run script's interface is described below.

```
./run_gmtb_scm.py -c CASE_NAME [-s SUITE_NAME] [-n PHYSICS_NAMELIST_PATH] [-g]
```

When invoking the run script, the only required argument is the name of the case to run. The case name used must match one of the case configuration files located in `./etc/case_config` (*without the .nml extension!*). If specifying a suite other than the default, the suite name used must match the value of the suite name in one of the SDFs located in `./../ccpp/suites` (Note: not the filename of the SDF). As part of the third CCPP release, the following suite names are valid:

- `SCM_GFS_v15`
- `SCM_GFS_v15plus`
- `SCM_csawmg`
- `SCM_GSD_v0`

**Note that using the Thompson microphysics scheme (as in `SCM_GSD_v0`) requires the existence of lookup tables during its initialization.**

- <https://dtcenter.org/GMTB/freezeH2O.dat> (243 M)
- [https://dtcenter.org/GMTB/qr\\_acr\\_qg.dat](https://dtcenter.org/GMTB/qr_acr_qg.dat) (49 M)
- [https://dtcenter.org/GMTB/qr\\_acr\\_qs.dat](https://dtcenter.org/GMTB/qr_acr_qs.dat) (32 M)

These files should be staged in `gmtb-scm/scm/data/physics_input_data` prior to executing the run script. Since binary files can be system-dependent (due to endianness), it is possible that these files will not be read correctly on your system. For reference, the linked files were generated on *Theia* using the Intel v18 compiler.

Also note that some cases require specified surface fluxes. Special SDFs that correspond to the suites listed above have been created and use the `*_prescribed_surface` decoration. It is not necessary to specify this filename decoration when specifying the suite name. If the `spec_sfc_flux` variable in the configuration file of the case being run is set to `.true.`, the run script will automatically use the special SDF that corresponds to the chosen suite from the list above.

If specifying a namelist other than the default, the value must be an entire filename that exists in `./../ccpp/physics_namelists`. Caution should be exercised when modifying physics namelists since some redundancy

between flags to control some physics parameterizations and scheme entries in the SDFs currently exists. Values of numerical parameters are typically OK to change without fear of inconsistencies. Lastly, the `-g` flag can be used to run the executable through the `gdb` debugger (assuming it is installed on the system).

If the run aborts with the error message

```
gmtb_scm: libccppphys.so.X.X.X: cannot open shared object file: No such file or_
↳directory
```

the environment variable `LD_LIBRARY_PATH` must be set to

```
export LD_LIBRARY_PATH=$PWD/ccpp/physics:$LD_LIBRARY_PATH
```

before running the model.

A netCDF output file is generated in the location specified in the case configuration file, if the `output_dir` variable exists in that file. Otherwise an output directory is constructed from the case, suite, and namelist used (if different from the default). All output directories are placed in the `bin` directory. Any standard netCDF file viewing or analysis tools may be used to examine the output file (`ncdump`, `ncview`, `NCL`, etc).

### Multiple Run Script Usage

A second Python script is provided for automating the execution of multiple integrations through repeated calling of the single run script. From the run directory, one may use this script through the following interface.

```
./multi_run_gmtb_scm.py {[-c CASE_NAME] [-s SUITE_NAME] [-f PATH_TO_FILE]} [-v{v}] [-
↳t]
```

No arguments are required for this script. The `-c` or `--case`, `-s` or `--suite`, or `-f` or `--file` options form a mutually-exclusive group, so exactly one of these is allowed at one time. If `-c` is specified with a case name, the script will run a set of integrations for all supported suites (defined in `./src/supported_suites.py`) for that case. If `-s` is specified with a suite name, the script will run a set of integrations for all supported cases (defined in `./src/supported_cases.py`) for that suite. If `-f` is specified with the path to a filename, it will read in lists of cases, suites, and namelists to use from that file. If multiple namelists are specified in the file, there either must be one suite specified or the number of suites must match the number of namelists. If none of the `-c` or `--case`, `-s` or `--suite`, or `-f` or `--file` options group is specified, the script will run through all permutations of supported cases and suites (as defined in the files previously mentioned).

In addition to the main options, some helper options can also be used with any of those above. The `-vv` or `--verbose` option can be used to output more information from the script to the console and to a log file. If this option is not used, only completion progress messages are written out. If `-v` is used, the script will write out completion progress messages and all messages and output from the single run script. If `-vv` is used, the script will also write out all messages and single run script output to a log file (`multi_run_gmtb_scm.log`) in the `bin` directory. The final option, `-t` or `--timer`, can be used to output the elapsed time for each integration executed by the script. Note that the execution time includes file operations performed by the single run script in addition to the execution of the underlying (Fortran) SCM executable. By default, this option will execute one integration of each subprocess. Since some variability is expected for each model run, if greater precision is required, the number of integrations for timing averaging can be set through an internal script variable. This option can be useful, for example, for getting a rough idea of relative computational expense of different physics suites.

### Batch Run Script

If using the model on HPC resources and significant amounts of processor time is anticipated for the experiments, it will likely be necessary to submit a job through the HPC's batch system. An example script has been included in the repository for running the model on *Theia's* batch system (SLURM). It is located in `gmtb-scm/scm/etc/gmtb_scm_slurm_example.py`. Edit the `job_name`, `account`, etc. to suit your needs and copy to the `bin` directory. The case name to be run is included in the command variable. To use, invoke

```
./gmtb_scm_slurm_example.py
```

from the `bin` directory.

Additional information on the SCM can be found at [https://dtcenter.org/gmtb/users/ccpp/docs/SCM-CCPP-Guide\\_v3.0.pdf](https://dtcenter.org/gmtb/users/ccpp/docs/SCM-CCPP-Guide_v3.0.pdf)

## 9.2 UFS Atmosphere

Another option for a CCPP host model is the UFS Atmosphere, located in the umbrella repository NEMSFv3gfs.

### 9.2.1 System Requirements, Libraries, and Compilers

The build system for the UFS with CCPP relies on the use of the Python scripting language, along with `cmake`.

The basic requirements for building and running the UFS with CCPP are listed below. The versions listed reflect successful tests and there is no guarantee that the code will work with different versions.

- **FORTRAN 90+ compiler versions**
  - ifort 15.1.133, 18.0.1.163 and 19.0.2
  - gfortran 6.2, 8.1, and 9.1
- **C compiler versions**
  - icc v18.0.1.163 and 19.0.2
  - gcc 6.2.0 and 8.1
  - AppleClang 10.0
- **MPI job scheduler versions**
  - mpt 2.19
  - impi 5.1.1.109 and 5.1.2.150
  - mpich 3.2.1
- `cmake` versions 2.8.12.1, 2.8.12.2, and 3.6.2
- netCDF with HDF5, ZLIB and SZIP versions 4.3.0, 4.4.0, 4.4.1.1, 4.5.0, 4.6.1, and 4.6.3 (not 3.x)
- Python versions 2.7.5, 2.7.9, and 2.7.13 (not 3.x)

A number of NCEP libraries are required to build and run FV3 and are listed in [Table 9.1](#).

Table 9.1: *NCEP libraries required to build the UFS Atmosphere*

Library	Version	Description
bacio	2.0.1	NCEP binary I/O library
ip	2.0.0/3.0.0	NCEP general interpolation library
nemsio	2.2.3	NEMS I/O routines
sp	2.0.2	NCEP spectral grid transforms
w3emc	2.2.0	NCEP/EMC library for decoding data in GRIB1 format
w3nco/v2.0.6	2.0.6	NCEP/NC0 library for decoding data in GRIB1 format

These libraries are prebuilt on most NOAA machines using the Intel compiler. For those needing to build the libraries themselves, GMTB recommends using the source code from GitHub at <https://github.com/NCAR/NCEPLibs.git>, which includes build files for various compilers and machines using OpenMP flags and which are thread-safe.

In addition to the NCEP libraries, some additional external libraries are needed (Table 9.2).

Table 9.2: External libraries necessary to build the UFS Atmosphere

Library	Version	Description
ESMF	V7.1.0r and v8.0.0_bs21	Earth System Modeling Framework for coupling applications
netCDF	4.3.0 and 4.6.1	Interface to data access functions for storing and retrieving data arrays
SIONlib (optional)	v1.7.2	Parallel I/O library (link) that can be used to read precomputed lookup tables instead of computing them on the fly (or using traditional Fortran binary data files)

The Earth System Modeling Framework (ESMF), the SIONlib, the NCEPLibs, and the netCDF libraries must be built with the same compiler as the other components of the UFS Atmosphere.

## 9.2.2 Building the UFS Atmosphere

A complete listing and description of the FV3 build options were discussed in Chapter 3 and are shown in Figure 3.1. This section will describe the commands needed to build the different options using the script `compile.sh` provided in the NEMSfv3gfs distribution. This script calls `ccpp_prebuild.py`, so users do not need to run the *prebuild* step manually. All builds using `compile.sh` are made from the `./tests` directory of NEMSfv3gfs and follow the basic command:

```
./compile.sh $PWD/./FV3 system.compiler 'MAKEOPTS'
```

Here, `system` stands for the machine on which the code is compiled and can be any of the following machines and compilers: *theia*, *jet*, *cheyenne*, *gaea*, *stampede*, *wcoss\_cray*, *wcoss\_dell\_p3*, *supermuc\_phase2*, *macosx*, or *linux*.

`compiler` stands for the compiler to use and depends on the system. For *theia* and *cheyenne*, the available options are `intel` and `gnu`. For *macosx* and *linux*, the only tested compiler is `gnu`. For all other platforms, `intel` is the only option at this time.

The `MAKEOPTS` string, enclosed in single or double quotes, allows to specify options for compiling the code. The following options are of interest for building the CCPP version of NEMSfv3gfs:

- **CCPP=Y** - enables *CCPP* (default is N)
- **STATIC=Y** - enables the CCPP static mode; requires `CCPP=Y` (default is N) and `SUITES=...` (see below)
- **SUITES=XYZ, ABC, DEF, ...** - specify SDF(s) to use when compiling the code in CCPP static mode; SDFs are located in `ccpp/suites/`, omit the path in the argument; requires `CCPP=Y` `STATIC=Y` (default is '')
- **SION=Y** - enables support for the SIONlib I/O library (used by CCPP to read precomputed lookup tables instead of computing them on the fly); available on *Theia*, *Cheyenne*, *Jet*; also available on *Mac OS X* and *Linux* if instructions in `doc/README_{macosx,linux}.txt` are followed (default is N)
- **32BIT=Y** - compiles FV3 dynamical core in single precision; note that physics are always compiled in double precision; this option is only available on *Theia*, *Cheyenne*, and *Jet* (default is N)
- **REPRO=Y** - compiles code in *REPRO* mode, i.e. removes certain compiler optimization flags used in the default *PROD* mode to obtain bit-for-bit (b4b) identical results between CCPP and non-CCPP code (default is N)
- **DEBUG=Y** - compiles code in *DEBUG* mode, i.e. removes all optimization of *PROD* mode and add bound checks; mutually exclusive with `REPRO=Y` (default is N)

- **INTEL18=Y** - available on *Theia* and *Jet* only, compiles code with Intel 18 compiler instead of the default Intel 15 compiler (default is N); note that Intel 18 is the only supported compiler on *Cheyenne*.
- **TRANSITION=Y** - applies selective lowering of optimization for selected files to obtain b4b with non-CCPP code in PROD mode (only when using Intel 15 on *Theia*)

Examples:

- Compile non-CCPP code with 32-bit dynamics on *Theia* with the Intel compiler

```
./compile.sh $PWD/../FV3 theia.intel `32BIT=Y`
```

- Compile dynamic CCPP code in DEBUG mode on *Jet* with Intel 18

```
./compile.sh $PWD/../FV3 jet.intel `CCPP=Y DEBUG=Y INTEL18=Y`
```

- Compile static CCPP code for the CPT suite on *Linux* with the GNU compiler, enable support for the SIONlib I/O library (requires that the library to be installed)

```
./compile.sh $PWD/../FV3 linux.gnu `SION=Y CCPP=Y STATIC=Y SUITES=FV3_CPT_  
↪v0`
```

- *Cheyenne* static build with multiple suites:

```
./compile.sh $PWD/../FV3 cheyenne.intel `CCPP=Y STATIC=Y SUITES=FV3_GFS_  
↪v15,FV3_CPT_v0`
```

## 9.2.3 Running the UFS Atmosphere Using the Regression Tests (RTs)

Regression testing is the process of testing changes to the programs to make sure that the existing functionalities still work when changes are introduced. By running the RTs (or a subset of them by copying a RT configuration file and editing it), the code is compiled, the run directories are set up, and the code is executed. The results are typically compared against a pre-existing baseline, but in certain occasions it is necessary to first create a new baseline (for example, in a new platform where a baseline does not exist or when it is expected that a new development will change the answer). Because the RTs set up the run directories, this is a useful and easy way to get started, since all the model configuration files and necessary input data (initial conditions, fixed data) are copied into the right place.

### Overview of the RTs

The RT configuration files are located in `./tests` relative to the top-level directory of NEMSfv3gfs and have names `rt*.conf`. The default RT configuration file, supplied with the NEMSfv3gfs master, compares the results from the non-CCPP code to the *official baseline* and is called `rt.conf`. Before running the RT script `rt.sh` in the same directory, the user has to set one or more environment variables and potentially modify the script to change the location of the automatically created run directories. The environment variables are ACCNR (mandatory unless the user is a member of the default project *nems*; sets the account to be charged for running the RTs), NEMS\_COMPILER (optional for the `intel` compiler option, set to `gnu` to switch), and potentially RUNDIR\_ROOT. RUNDIR\_ROOT allows the user to specify an alternative location for the RT run directories underneath which directories called `rt_$PID` are created (`$PID` is the process identifier of the `rt.sh` invocation). This may be required on systems where the user does not have write permissions in the default run directory tree.

```
export ACCNR=...  
export NEMS_COMPILER=intel  
export RUNDIR_ROOT=/full/path/under/which/rt_$PID/will/be/created
```

Running the full default RT suite defined in `rt.conf` using the script `rt.sh`:

```
./rt.sh -f
```

This command can only be used on a NOAA machine using the Intel compiler, where the output of a non-CCPP build using the default Intel version is compared against the *official baseline*. For information on testing the CCPP code, or using alternate computational platforms, see the following sections.

This command and all others below produce log output in `./tests/log_machine.compiler`. These log files contain information on the location of the run directories that can be used as templates for the user. Each `rt*.conf` contains one or more compile commands preceding a number of tests.

## Baselines

Regression testing is only possible on machines for which baselines exist. EMC maintains *official baselines* of non-CCPP runs on *Jet* and *Wcoss* created with the Intel compiler. GMTB maintains additional baselines on *Theia*, *Jet*, *Cheyenne*, and *Gaea*. While GMTB is trying to keep up with changes to the official repositories, baselines maintained by GMTB are not guaranteed to be up-to-date.

When porting the code to a new machine, it is useful to start by establishing a *personal baseline*. Future runs of the RT can then be compared against the *personal baseline* to ascertain that the results have not been inadvertently affected by code developments. The `rt.sh -c` option is used to create a *personal baseline*.

```
./rt.sh -l rt.conf -c fv3 # create own reg. test baseline
```

Once the *personal baseline* has been created, future runs of the RT should be compared against the *personal baseline* using the `-m` option.

```
./rt.sh -l rt.conf -m # compare against own baseline
```

## The script `rt.sh`

`rt.sh` is a bash shell file to run the RT and has the following options:

```
Usage: $0 -c <model> | -f | -s | -l <file> | -m | -r | -e | -h
-c create new baseline results for <model>
-f run full suite of regression tests
-s run standard suite of regression tests
-l run test specified in <file>
-m compare against new baseline results
-r use Rocoto workflow manager
-e use ecFlow workflow manager
-h display this help
```

The location of the run directories and *personal baseline* directories is controlled in `rt.sh` on a per-machine basis. The user is strongly advised to NOT modify the path to the *official baseline* directories.

The *official baseline* directory is defined as:

```
RTPWD=${DISKNM}/trunk-yyyyymmdd/${COMPILER} # on Cheyenne
RTPWD=${DISKNM}/trunk-yyyyymmdd           # elsewhere
```

Note that `yyyyymmdd` is the year, month and day the RT was created.

**Warning:** Modifying `$DISKNM` will break the RTs!

*Personal baseline* results (see below) are stored in

```
NEW_BASELINE=${STMP}/${USER}/FV3_RT/REGRESSION_TEST
```

and RTs are run in \$RUNDIR\_ROOT.

Example: *Theia*

```
...
dprefix=/scratch4/NCEPDEV
DISKNM=${dprefix}/nems/noscrub/emc.nemspara/RT
STMP=${dprefix}/stmp4
PTMP=${dprefix}/stmp3
..
```

In case a user does not have write permissions to \$STMP (/scratch4/NCEPDEV/stmp4/), \$STMP must be modified without modifying \$DISKNM (i.e. dprefix). Similarly, if the user does not have write permissions to \$PTMP, the user can set the \$RUNDIR\_ROOT environment variable to change the location of the run directories as described below.

```
# Overwrite default RUNDIR_ROOT if environment variable RUNDIR_ROOT is set
RUNDIR_ROOT=${RUNDIR_ROOT:-${PTMP}/${USER}/FV3_RT}/rt_$$
```

### Non-CCPP vs CCPP Tests

While the official EMC RTs do not execute the CCPP code, GMTB provides RTs to exercise the CCPP in its various modes: `rt_ccpp_standalone.conf` tests the CCPP with dynamic build and `rt_ccpp_static.conf` tests the CCPP with static build. These tests compare the results of runs done using the CCPP against a previously generated *personal baseline* created without the CCPP by running `rt_ccpp_ref.conf`. For this comparison, both the non-CCPP *personal baseline* and the tests using the CCPP are performed with code built with the *REPRO* compiler options.

The command below should be used to create a *personal baseline* using non-CCPP code compiled in *REPRO* mode.

```
./rt.sh -l rt_ccpp_ref.conf -c fv3 # create own reg. test baseline
```

Once the *personal baseline* in *REPRO* mode has been created, the CCPP tests can be run to compare against it. Use the `-l` option to select the test suite and the `-m` option to compare against the *personal baseline*.

```
./rt.sh -l rt_ccpp_standalone.conf -m # dynamic build
./rt.sh -l rt_ccpp_static.conf -m # static build
```

### Compatibility between the Code Base, the SDF, and the Namelist in the UFS Atmosphere

The variable `suite_name` within the `namelist.input` file used in the UFS Atmosphere determines which suite will be employed at run time (e.g., `suite_name=FV3_GFS_v15`). It is the user's responsibility to ascertain that the other variables in `namelist.input` are compatible with the chosen suite. When runs are executed using the RT framework described in the preceding sections, compatibility is assured. For new experiments, users are responsible for modifying the two files (SDF and `namelist.input`) consistently, since limited checks are in place.

Information about the UFS Atmosphere physics namelist can be found with the CCPP Scientific Documentation at [https://dtcenter.org/GMTB/v3.0/sci\\_doc/](https://dtcenter.org/GMTB/v3.0/sci_doc/).



## TIPS FOR ADDING A NEW SCHEME

This chapter contains a brief description on how to add a new scheme to the *CCPP-Physics* pool.

- Identify the variables required for the new scheme and check if they are already available for use in the CCPP by checking the metadata tables in `GFS_typedefs.F90` or by perusing file `ccpp-framework/doc/DevelopersGuide/CCPP_VARIABLES_{FV3, SCM}.pdf` generated by `ccpp_prebuild.py`.
  - If the variables are already available, they can be invoked in the scheme’s metadata table and one can skip the rest of this subsection. If the variable required is not available, consider if it can be calculated from the existing variables in the CCPP. If so, an interstitial scheme (such as `scheme_pre`; see more in [Chapter 2](#)) can be created to calculate the variable. However, the variable must be defined but not initialized in the host model as the memory for this variable must be allocated on the host model side. Instructions for how to add variables to the host model side is described in [Chapter 6](#).
  - It is important to note that not all data types are persistent in memory. Most variables in the interstitial data type are reset (to zero or other initial values) at the beginning of a physics group and do not persist from one set to another or from one group to another. The diagnostic data type is periodically reset because it is used to accumulate variables for given time intervals. However, there is a small subset of interstitial variables that are set at creation time and are not reset; these are typically dimensions used in other interstitial variables.

---

**Note:** If the value of a variable must be remembered from one call to the next, it should not be in the interstitial or diagnostic data types.

---

- If information from the previous timestep is needed, it is important to identify if the host model readily provides this information. For example, in the Model for Prediction Across Scales (MPAS), variables containing the values of several quantities in the preceding timesteps are available. When that is not the case, as in the UFS Atmosphere, interstitial schemes are needed to compute these variables. As an example, the reader is referred to the GF convective scheme, which makes use of interstitials to obtain the previous timestep information.
- Examine scheme-specific and suite interstitials to see what needs to be replaced/changed; then check existing scheme interstitial and determine what needs to be replicated. Identify if your new scheme requires additional interstitial code that must be run before or after the scheme and that cannot be part of the scheme itself, for example because of dependencies on other schemes and/or the order the scheme is run in the SDF.
- Follow the guidelines outlined in [Chapter 2](#) to make your scheme CCPP-compliant. Make sure to use an upper-case suffix `.F90` to enable C preprocessing.
- Locate the CCPP *prebuild* configuration files for the target host model, for example:
  - `NEMSFV3GFS/ccpp/config/ccpp_prebuild_config.py` for the UFS Atmosphere
  - `gmtb-scm/ccpp/config/ccpp_prebuild_config.py` for the SCM

- Add the new scheme to the Python dictionary in `ccpp_prebuild_config.py` using the same path as the existing schemes:

```
SCHEME_FILES = [ ...
'../some_relative_path/existing_scheme.F90',
'../some_relative_path/new_scheme.F90',
...]
```

- If the new scheme uses optional arguments, add information on which ones to use further down in the configuration file. See existing entries and documentation in the configuration file for the possible options:

```
OPTIONAL_ARGUMENTS = {
    'SCHEME_NAME' : {
        'SCHEME_NAME_run' : [
            # list of all optional arguments in use for this
            # model, by standard_name ],
            # instead of list [...], can also say 'all' or 'none'
        ],
    },
}
```

- Place new scheme in the same location as existing schemes in the CCPP directory structure, e.g., `../some_relative_path/new_scheme.F90`.
- Edit the SDF and add the new scheme at the place it should be run. SDFs are located in
  - `NEMSFV3GFS/ccpp/suites` for the UFS Atmosphere
  - `gmtb-scm/ccpp/suites` for the SCM
- Before running, check for consistency between the namelist and the SDF. There is no default consistency check between the SDF and the namelist unless the developer adds one. Errors may result in segmentation faults in running something you did not intend to run if the arrays are not allocated.
- Test and debug the new scheme:
  - Typical problems include segment faults related to variables and array allocation.
  - Make sure SDF and namelist are compatible. Inconsistencies may result in segmentation faults because arrays are not allocated or in unintended scheme(s) being executed.
  - A scheme called `GFS_debug` (`GFS_debug.F90`) may be added to the SDF where needed to print state variables and interstitial variables. If needed, edit the scheme beforehand to add new variables that need to be printed.
  - Check *prebuild* script for success/failure and associated messages.
  - Compile code in `DEBUG` mode, run through debugger if necessary (`gdb`, `Allinea DDT`, `totalview`, ...). See [Chapter 9](#) for information on debugging.
  - Use memory check utilities such as `valgrind`.
  - Double-check the metadata table in your scheme to make sure that the standard names correspond to the correct local variables.
- Done. Note that no further modifications of the build system are required, since the *CCPP-Framework* will autogenerate the necessary makefiles that allow the host model to compile the scheme.

## ACRONYMS

Abbreviation	Explanation
aa	Aerosol-aware
API	Application Programming Interface
b4b	Bit-for-bit
CCPP	Common Community Physics Package
CF conventions	Climate and Forecast Metadata Conventions
CPP	C preprocessor
CPT	Climate Process Team
CSAW	Chikira-Sugiyama convection with Arakawa-Wu extension
DDT	Derived Data Type
dycore	Dynamical core
EDMF	Eddy-Diffusivity Mass Flux
EMC	Environmental Modeling Center
eps	Encapsulated PostScript
ESMF	The Earth System Modeling Framework
ESPC	Earth System Prediction Capability
FMS	Flexible Modeling System
FV3	Finite-Volume Cubed Sphere
GF	Grell-Freitas convective scheme
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GMTB	Global Model Test Bed
GSD	Global Systems Division
HEDMF	Hybrid eddy-diffusivity mass-flux
HTML	Hypertext Markup Language
IPD	Interoperable Physics Driver
LSM	Land Surface Model
MG	Morrison-Gottelman
MP	Microphysics
MPAS	Model for Prediction Across Scales
MPI	Message Passing Interface
MYNN	Mellor-Yamada-Nakanishi-Niino
NCAR	National Center for Atmospheric Research
NEMS	National Oceanic and Atmospheric Administration (NOAA) Environmental Modeling System
NEMSFv3gfs	National Oceanic and Atmospheric Administration (NOAA) Environmental Modeling System using FV3 dynamic core and GFS physics
NGGPS	Next Generation Global Prediction System

Continued on next page

Table 11.1 – continued from previous page

Abbreviation	Explanation
NOAA	National Oceanic and Atmospheric Administration
NRL	Naval Research Laboratory
NUOPC	National Unified Operational Prediction Capability
OpenMP	Open Multi-Processing
PBL	Planetary Boundary Layer
png	Portable Network Graphic
PR	Pull request
PROD	Compiler optimization flags for production mode
REPRO	Compiler optimization flags for reproduction mode (bit-for-bit testing)
RRTMG	Rapid Radiative Transfer Model for Global Circulation Models
RT	Regression test
RUC	Rapid Update Cycle
sa	Scale-aware
SAS	Simplified Arakawa-Schubert
SCM	Single Column Model
SDF	Suite Definition File
sfc	Surface
SHUM	Perturbed boundary layer specific humidity
SKEB	Stochastic Kinetic Energy Backscatter
SPPT	Stochastically Perturbed Physics Tendencies
TKE	Turbulent Kinetic Energy
TWP-ICE	Tropical Warm Pool International Cloud Experiment
UML	Unified Modeling Language
UFS	Unified Forecast System
VLab	Virtual Laboratory
WRF	Weather Research and Forecasting

## GLOSSARY

**CCPP** Model agnostic, vetted, collection of codes containing atmospheric physical parameterizations and suites for use in NWP along with a framework that connects the physics to host models

**CCPP-Framework** The infrastructure that connects physics schemes with a host model; also refers to a software repository of the same name

**CCPP-Physics** The pool of CCPP-compliant physics schemes; also refers to a software repository of the same name

**Dynamic CCPP build** A CCPP build type in which the *CCPP-Framework* and *CCPP-physics* libraries are dynamically linked to the executable and all CCPP-compliant schemes available in the library can be invoked at runtime. This build type is available for both the SCM and the UFS Atmosphere

**“Fast” physics** Physical parameterizations that require tighter coupling with the dynamical core than “slow” physics (due to the approximated processes within the parameterization acting on a shorter timescale) and that benefit from a smaller time step. The distinction is useful for greater accuracy, numerical stability, or both. In the UFS Atmosphere, a saturation adjustment is used in some suites and is called directly from the dynamical core for tighter coupling

**Group** A set of physics schemes within a suite definition file (SDF) that are called together without intervening computations from the host application

**Group cap** Autogenerated interface between a group of physics schemes and the host model. They are used only in the static CCPP build, and effectively replace the code from *CCPP-Framework* that provides the flexibility of the dynamic CCPP build

**Host model/application** An atmospheric model that allocates memory, provides metadata for the variables passed into and out of the physics, and controls time-stepping

**Interstitial scheme** A modularized piece of code to perform data preparation, diagnostics, or other “glue” functions that allows primary schemes to work together as a suite. They can be categorized as “scheme-specific” or “suite-level”. Scheme-specific interstitial schemes augment a specific primary scheme (to provide additional functionality). Suite-level interstitial schemes provide additional functionality on top of a class of primary schemes, connect two or more schemes together, or provide code for conversions, initializing sums, or applying tendencies, for example.

**Multi-suite static CCPP build** A static CCPP build type in which a set of physics suites is specified at compile time from which one can be chosen at runtime.

**NEMS** The NOAA Environmental Modeling System - a software infrastructure that supports NCEP/EMC’s forecast products. The coupling software is based on ESMF and the NUOPC layer.

**NUOPC** The National Unified Operational Prediction Capability is a consortium of Navy, NOAA, and Air Force modelers and their research partners. It aims to advance the weather modeling systems used by meteorologists, mission planners, and decision makers. NUOPC partners are working toward a common model architecture - a standard way of building models - in order to make it easier to collaboratively build modeling systems.

**Parameterization** The representation, in a dynamic model, of physical effects in terms of admittedly oversimplified parameters, rather than realistically requiring such effects to be consequences of the dynamics of the system (AMS Glossary)

**Physics *cap*** Autogenerated interface between an individual physics scheme and the host model. Used only in the dynamic CCPP build

**Primary scheme** A parameterization, such as PBL, microphysics, convection, and radiation, that fits the traditionally-accepted definition, as opposed to an interstitial scheme

**PROD** Compiler flags used by NCEP for operational runs of the UFS Atmosphere and by EMC for regression tests of the code

**REPRO** Compiler flags used by EMC to guarantee reproducibility of the UFS Atmosphere code

**Scheme** A CCPP-compliant parameterization (primary scheme) or auxiliary code (interstitial scheme)

**SDF** Suite Definition File (SDF) is an external file containing information about the construction of a physics suite. It describes the schemes that are called, in which order they are called, whether they are subcycled, and whether they are assembled into groups to be called together

**Set** A collection of physics schemes that do not share memory (e.g. fast and slow physics)

**“Slow” physics** Physical parameterizations that can tolerate looser coupling with the dynamical core than “fast” physics (due to the approximated processes within the parameterization acting on a longer timescale) and that often use a longer time step. Such parameterizations are typically grouped and calculated together (through a combination of process- and time-splitting) in a section of an atmospheric model that is distinct from the dynamical core in the code organization

**Standard\_name** Variable names based on CF conventions (<http://cfconventions.org>) that are uniquely identified by the *CCPP-compliant* schemes and provided by a host model

**Static CCPP build** A build type in which the *CCPP-Framework* and the *CCPP-physics* libraries are statically linked to the executable and only the suites determined at compile time can be invoked at runtime. This build type is only available for the UFS Atmosphere

**Subcycling** Executing a physics scheme more frequently (with a shorter timestep) than the rest of the model physics or dynamics

**Suite** A collection of primary physics schemes and interstitial schemes that are known to work well together

**Suite *cap*** Autogenerated interface between an entire suite of physics schemes and the host model. They are used only in the static CCPP build and consist of calls to autogenerated group caps. They may be used to call an entire suite at once or to call a specific group within a physics suite

**UFS** A Unified Forecast System (UFS) is a community-based, coupled comprehensive Earth system modeling system. The UFS numerical applications span local to global domains and predictive time scales from sub-hourly analyses to seasonal predictions. It is designed to support the Weather Enterprise and to be the source system for NOAA’s operational numerical weather prediction applications

**UFS Atmosphere** The atmospheric model component of the UFS. Its fundamental parts are the dynamical core and the physics

**VLab** Virtual Laboratory - a service and information technology framework, that enables NOAA employees and their partners to share ideas, collaborate, engage in software development, and conduct applied research (<https://www.nws.noaa.gov/mdl/vlab/>)

**.xsd file extension** XML schema definition

## Symbols

"Fast" physics, [89](#)  
"Slow" physics, [90](#)  
.xsd file extension, [90](#)

## C

CCPP, [89](#)  
CCPP-Framework, [89](#)  
CCPP-Physics, [89](#)

## D

Dynamic CCPP build, [89](#)

## G

Group, [89](#)  
Group cap, [89](#)

## H

Host model/application, [89](#)

## I

Interstitial scheme, [89](#)

## M

Multi-suite static CCPP build, [89](#)

## N

NEMS, [89](#)  
NUOPC, [89](#)

## P

Parameterization, [90](#)  
Physics cap, [90](#)  
Primary scheme, [90](#)  
PROD, [90](#)

## R

REPRO, [90](#)

## S

Scheme, [90](#)

SDF, [90](#)  
Set, [90](#)  
Standard\_name, [90](#)  
Static CCPP build, [90](#)  
Subcycling, [90](#)  
Suite, [90](#)  
Suite cap, [90](#)

## U

UFS, [90](#)  
UFS Atmosphere, [90](#)

## V

VLab, [90](#)