diag_run – wrapper script for NorESM diagnostics

09.04.2018

AUTHOR
Johan Liakka, NERSC (johan.liakka@nersc.no).

LOCATION
At the moment diag_run is only available on NIRD:
/projects/NS2345K/noresm_diagnostics/bin/diag_run

SYNTAX
`diag_run [-m model] [-c case_name] [-s start_yr] [-e end_yr] [-i input_dir] [-c2 case_name2] [-s2 start_yr2] [-e2 end_yr2] [-i2 input_dir2] [-o output_dir] [-w web_dir] [-t type] [-p]`

DESCRIPTION
diag_run is a wrapper script, which is used to run the diagnostics for each NorESM component (cam, clm, cice, micom, and hamocc). These diagnostic tools can be used to either plot model results with respect to observations (so-called model-obs diagnostics), or to another simulation (model1-model2 diagnostics). The diagnostics packages for the atmosphere (cam), land (clm) and sea-ice (cice) are based on the NCAR packages, but has undergone some major improvements, particularly the climatology and time-series computations. The ocean (micom) and its geo-bio-chemistry (hamocc) have been developed in-house by the author.

diag_run has two modes: (i) an “active-mode”, for which diag_run runs the diagnostic scripts; and (ii) a “passive-mode”, for which diag_run only configures the scripts. In the passive-mode the diagnostic scripts have to be run manually by the user. By default, diag_run is always in the active-mode, but switches into passive-mode if at least one of these two criteria are fulfilled:
1. The user invokes the option -p (see below), or
2. The user does not give enough information needed to run the diagnostics (next subsection).

Active-mode:
If you want to use diag_run to run the diagnostic scripts, the minimum requirement is to specify the options *model*, *case_name*, *start_yr* and *end_yr*
(-m, -c, -s and -e), e.g.:

```
Example 1:
diag_run -m cam -c N1850_f19_tn14_191017 -s 21 -e 50
```

This command runs atmospheric model-obs diagnostics of the case N1850_f19_tn14_191017 using a climatology between model years 21 and 50. It is assumed that the N1850_f19_tn14_191017 history files are located in /projects/NS2345K/noresm/cases. The resulting plots and html will be stored in /projects/NS2345K/www/noresm_diagnostics/N1850_f19_tn14_191017 /CAM_DIAG, which links to the following URL:
http://ns2345k.web.sigma2.no/noresm_diagnostics/N1850_f19_tn14_191017/CAM_DIAG/yrs21to50-obs.html.
The climatology and time-series files in /projects/NS2345K/noresm_diagnostics/out/$USER/CAM_DIAG (where $USER is your NIRD username).

If you want to run model1-model2 diagnostics, you also need to specify *case_name2*, *start_yr2* and *end_yr2* (-c2, -s2, -e2) in addition to those in `Example 1`, i.e.:

```
Example 2:
```

```
diag_run -m cam -c N1850_f19_tn14_191017 -s 21 -e 50 -c2 B1850MICOM_f09_tn14_01
-s2 21 -e2 50
```

would be the same as in `Example 1` above, except for comparing N1850_f19_tn14_191017 to B1850MICOM_f09_tn14_01 instead of observations.

diag_run uses some template scripts for each of the model components. When diag_run is executed, these scripts are changed according to the settings you have invoked, and renamed with a time stamp. For example, if you run the micom diagnostics, the run script template (micom_diag_template.sh) will be renamed with a time-stamp as micom_diag_YYMMDD_HHMMSS.

diag_run also creates a config and output file with the same time stamp (config_YYMMDD_HHMMSS and out_YYMMDD_HHMMSS, respectively). The config file stores information about changes in the diagnostics scripts invoked by the user, and the output file contains the standard output and error (i.e. what is shown in your terminal during runtime).

When the diagnostics a component is finished the run scripts are copied to:
output_dir/$USER/*model_diag*/config/*case_name*/run_scripts
and the config and output files to:
output_dir/$USER/*model_diag*/config/*case_name*/logs

Hence, for Example 1 above, the run scripts are saved in:
/projects/NS2345K/noresm_diagnostics/out/
$USER/CAM_DIAG/config/N1850_f19_tn14_191017/run_scripts
and the config and out files in:
/projects/NS2345K/noresm_diagnostics/out/
$USER/CAM_DIAG/config/N1850_f19_tn14_191017/logs

passive-mode:
Another important property of diag_run is that it will only run the diagnostics if sufficient information has been provided by the user; otherwise it switches into passive-mode. diag_run will then configure the diagnostics scripts as much as possible (based on the information provided by the user), and add information to the config file about which variables are still required to be modified by the user in order to run diag_run. This option is particularly useful if you want to do some development work on the diagnostics scripts, or if you want to change any variables in the diagnostics scripts that are not included in diag_run. Hence, if you run the following command:

Example 3:
```
diag_run -m clm
```

you will get the following information on the screen:

```
------------------------------------------------
Program:
/projects/NS2345K/noresm_diagnostics/bin/diag_run
Version: 3.0
------------------------------------------------
-CHANGING DIAGNOSTICS DIRECTORY to
/projects/NS2345K/noresm_diagnostics/out/johiak/CLM_DIAG in lnd_template.csh
-CHANGING ROOT DIRECTORY FOR CODE AND DATA to
/projects/NS2345K/noresm_diagnostics/packages/CLM_DIAG in lnd_template.csh
-CHANGING INPUT DIR 1 to /projects/NS2345K/noresm/cases in lnd_template.csh
```

-CHANGING publish_html_root to /projects/NS2345K/www/noresm_diagnostics in lnd_template.csh
CLM DIAGNOSTICS SUCCESSFULLY CONFIGURED in /projects/NS2345K/noresm_diagnostics/out/johiak/CLM_DIAG
---
lnd_template.csh IS NOT RUNNING: NOT ALL REQUIRED VARIABLES HAVE BEEN CONFIGURED (see /projects/NS2345K/noresm_diagnostics/out/johiak/CLM_DIAG/config.log).

The (semi-configured) run script has then been copied to /projects/NS2345K/noresm_diagnostics/out/johiak/CLM_DIAG/lnd_template.csh, and all information about the configuration is contained in /projects/NS2345K/noresm_diagnostics/out/johiak/CLM_DIAG/config.log

diag_run options (flags) come in both short (single-letter) and long forms. A complete description of all options is given below in alphabetical order of the short option letter. When invoked without options, diag_run prints a table of all options along with some examples.

-c *case_name* (-c1, --case, --case1)

Name of the test case experiment that you want to run diagnostics for. This option is required if you want to use diag_run in active-mode.

-c2 *case_name2* (--case2)

Name of the control case experiment. This option is required if you want to run model1-model2 diagnostics in active-mode.

-e *end_year* (-e1,--end_yr,--end_yr1)

End year of climatology for *case_name*. This is required if you want to run in active-mode.

-e2 *end_year* (--end_yr2)

End year of climatology for *case_name2*. This is required if you want to run model1-model2 diagnostics in active-mode.

-i *input_dir* (-i1, --input-dir, --input-dir1)

Name of the root directory of the monthly history files for *case_name*. For example, if your micom history files are located in /this/is/a/directory/case1/ocn/hist, this option should be set to input_dir=/this/is/a/directory. Default is input_dir=/projects/NS2345K/noresm/cases

-i2 *input-dir2* (--input-dir2)

Name of the root directory of the monthly history files for *case_name2*. Also here, default is input_dir2=/projects/NS2345K/noresm/cases

-m *model* (--model)

Name of the model you want to run the diagnostics for. Valid options are cam, clm, cice, micom, hamocc and all. This is the only option that is required for both the active and passive mode. If you invoke the “all” option, the cam, clm, cice, micom and hamocc diagnostics will be run

subsequently. It is also possible to combine different models as you wish within this option: for example, if you only want to run cam and clm diagnostics, you can simply add the names of those models and separate them with a comma (-m cam,clm).

-o *output_dir* (--output_dir)

Root directory where you want to store the output from the diagnostics (i.e. the climatology and time-series files). For example, if you set output_dir=/just/another/directory, the climatology and time-series files from the micom diagnostics will be stored in /just/another/directory/MICOM_DIAG/. Default is output_dir=/projects/NS2345K/noresm_diagnostics/out/$USER, where $USER is your user name on NIRD.

-p, --passive-mode

This option, which takes no argument, forces diag_run into passive-mode. This means, even if you have given sufficient information to run in active-mode, the diagnostic scripts will not be executed.

-s *start_year* (-s1,--start_yr,--start_yr1)

Start year of climatology for *case_name*. This is required if you want to run in active-mode.

-s2 *start_year2* (--start_yr2)

Start year of climatology for *case_name2*. This is required if you want to run model1-model2 diagnostics in active mode.

-t *type* (--type)

Specifies if you only run climatology or time-series diagnostics: valid options are --type=climo and --type=time_series. Default is to run both. Note that the time series are computed over the entire simulation, so if wish to have quick results from a long simulation, you can use --type=climo.

-w *webdir* (--web-dir)

Specifies the directory where the html should be stored. This directory should preferably be linked to a web server so that one can look at the results with a web browser. Default is --web-dir=/projects/NS2345K/www/noresm_diagnostics/.