Ensemble setup

This chapter describes how to setup your simulation ensemble once you have compiled your configuration at a choosen resolution.

Table of Content

In this chapter we describe how to generate/use **ensemble** simulation.

Here **ensemble** simulation define a group of climate model (coupled) simulations used for climate projections. There are multiple simulations that have same model version but slightly different initial conditions. Ensemble simulations can be useful to have multiple projections, to analyse prediction of model status and to use spread for projections and scenarios. There differents type of ensemble, for IPSLCM6 cuopled model it is possible to have "Perturbation of initial condition ensemble" or "ensemble with different starting dates".

1. Prepare ensembles with ins_job -e

To create an ensemble configuration you need to create an `ensemble.card` file.

NOTE: a template of `ensemble.card` is given in `IPSLCM6/EXPERIMENTS/IPSLCM/dcppAhindcast_CMIP6` for IPSLCM6 model. This is a specific protocol for experiment DCPP-A CMIP6.
A simpler template to configure ensemble (without CMIP6 constraint) is given in `IPSLCM6/EXPERIMENTS/IPSLCM/decadal`

Since IPSLCM6.1.9-LR, when IPSLCM6-LR is downloaded with `./model IPSLCM6.1.9-LR` it will offer the possibility to launch experiments of the decadal type.
To prepare an ensemble of simulations copy the config.card and ensemble.card files from the directory:

```
modipsl/config/IPSLCM6/EXPERIMENTS/IPSLCM/dcppAhindcast_CMIP6
```

into the directory:

```
modipsl/config/IPSLCM6/
```

Two types of ensemble are allowed with this template:

1. Ens_DATE : allows to configure simulations starting from different restart dates (choosen non-periodicaly).
1. Ens_PERTURB : allows to configure simulations of differents members from an initial condition which is perturbed, using PERIODIC restart dates (it is possible to create different initial states perturbing Sea Surface Temperature with a with noise).

All parameters for ensemble description are in ensemble.card and in config.card.

2. Usage

Check that COMP, POST, PARAM and DRIVER directories are present in the experiment folder.
Once ensemble.card and config.card are correctly filled, to create an ensemble simply type:

    ../../libIGCM/ins_job -e

This will create :

- All the directories of the ensemble (by default PeriodNb=5 for all simulations)
- Qsub.*xxx*.sh : shell script to submit all jobs
- Qclean.PeriodLengt.*xxx*.sh : shell script to clean (if necessary) all simulations after an error, and to re-launch them
- Qclean.latestPackperiod.*xxx*.sh : shell script to clean (if necessary) last packed period after an error, and to re-launch them

NOTE:

- *xxx* it will be the *JobName* configured in config.card
- If a directory exists, ins_job won't modify it, the creation of ensemble will be stopped.

2.1. Config.card

The file *config.card* is filled as a regular *config.card* (`ins_job` without the -e option). It will be used as a template for all simulations that will be created.

The important lines for the ensemble set up are in the [UserChoices] section. Make sure that JobName and ExperimentName are filled with proper values.
The variables DateBegin and DateEnd will be overwritten by variables present in ensemble.card.

```
#
# This is config.card file for IPSLCM6 configuration
#
#==================================================================
#D-- Compatibility -
[Compatibility]
libIGCM=1.0
#D-- UserChoices -
[UserChoices]
#==========================
JobName=CM619-LR-dcppA-hindcast

#----- Short Name of Experiment
ExperimentName=dcppA-hindcast

#----- DEVT TEST PROD
SpaceName=DEVT

LongName="IPSLCM6.1.9-LR"

TagName=IPSLCM6

ModelName=IPSL-CM6A-LR

Member=r1i1p1f1

#D- Choice of experiment in EXPERIMENTS directory
ExpType=IPSLCM/dcppAhindcast_CMIP6
#==========================

#-- leap, noleap, 360d
CalendarType=leap

#-- Experiment dates : Beginning and ending
#-- "YYYY-MM-DD"
DateBegin=1961-01-01
DateEnd=1980-12-31

#==========================
ORCA_version=eORCA1.2

#==========================
#-- 1Y, 1M, 5D, 1D Period Length of one trunk of simulation
PeriodLength=1Y
```

A section [Ensemble] should also be present.
It contains the informations that we want to store for an ensemble simulation. The variable EnsembleRun must be set to y and three other fields will to be filled in the *config.card* of each member after **'ins_job -e** has ran.

```
#==========================
[Ensemble]
#D- Ensemble run ? 'y' or 'n'
#D- If 'y', fill in ensemble.card !!
EnsembleRun=y

EnsembleName=

EnsembleDate=

EnsembleType=
#
```

NOTE: Notice that if you have a `Member=` parameter in the `[UserChoices]` section, it will be automatically incremented for each member.

2.2. Ensemble.card

It is possile to choice between 2 type of ensembles in *ensemble.card*:

- Date restart ensemble which allows to configure simulations starting from different restart dates: need to fill section `[Ens_DATE]`
- Perturb ensemble which allows to generate members from an initial condition which is perturbed by different means: need to fill section`[Ens_PERTURB]`

To choose between the 2 types of ensembles, set the variable `active` to *y* of section that you want activate:
To activate **Date Restart** (non-Periodic) Ensembles :

```
[Ens_DATE]
active=y

[Ens_PERTURB]
active=n
```

To activate **Perturbed Restart** (Periodic) Ensembles :

```
[Ens_DATE]
active=n

[Ens_PERTURB]
active=y
```

NOTE: Period can be only 1Y (to be able to run a kind of "NON-periodic perturbed ensemble")

3. Configure a Date Restart ensemble experience

The « Date Restart ensemble » is implemented to configure a set of simulations using several restart dates, generally chosen following a particular protocol (ex : randomly, particular climate oscillation phases, volcanic activity…).

In ensemble.card as mentionned previously you need to activate `[Ens_DATE]` and desactivate the other ensemble types and configure the `[Ens_DATE]` section.
This is a non-periodic definition of the start dates (ie you'll need to specify all of them manually). Here is an example of a basic ensemble configuration:

```
[Ens_DATE]
# active=y to use date ensemble, 'n' for no DO NOT use.
active=y

#--- Default values for all members ---
# name of the ensemble (used to create root directory)
NAME= CM618-LR-volc-pinatubo-full

# default length for all simulations
LENGTH=10Y

# Default start date for all simulations
STARTDATE=19900601

# Experiment name to find all restart files (and default one for non-periodic)
INITFROM=CM61-LR-pi-03

# Restart root directory
INITPATH=$PATH_TO_RESTART/piControl

#--- Specific values for each member (overule default) ---
# list of corresponding restart dates
```

```
RESTART_NONPERIODIC=(18700531 18810531)
```

In this example, you'll generate a 2-members ensemble starting from 1870-05-31 and 1881-05-31 (as defined in last line). Other parameters filled are "global ones", so it means that they are the same for all members. In all members' `config.card` files you'll get:

- `DateBegin= 1990-06-01` due to `STARTDATE=19900601`
- `DateEnd= 2000-05-31` due to `LENGTH=10Y`
- In all components' restarts section `RestartJobName=CM61-LR-pi-03` due to `INITFROM=CM61-LR-pi-03` option and `RestartPath$PATH_TO_RESTART/piControl` according to `INITPATH=$PATH_TO_RESTART/piControl`.

If you don't want to use the same value in all members for those parameters, you could use optional parameters `*_NONPERIODIC` to overrule them. It is not necessary to fill all of them if only few are required for your configuration (in this case you'll mix general and specific ones). All of those optional parameters must be filled with the same format that for the `RESTART_NONPERIODIC` one. Here is an example with all the possibilities:

```
[Ens_DATE]
# active=y to use date ensemble, 'n' else.
active=y

#--- Default values for all members ---
# name of the ensemble (used to create root directory)
NAME= CM618-LR-volc-pinatubo-full
# default length for all simulations
LENGTH=10Y
# Default start date for all simulations
STARTDATE=19900601
# Experiment name to find all restart files (and default one for non-periodic)
INITFROM=CM61-LR-pi-03
# Restart root directory
INITPATH=$PATH_TO_RESTART/piControl

#--- Specific values for each member (overrule default) ---
# list of corresponding restart dates
RESTART_NONPERIODIC=(18700531 18810531)

### Optional options (use default values if not filled) ###
# list of start dates for all simulations
NONPERIODIC=(19910601 19910601)
# simulation name to restart for each simulation. IF empty all simulations will use INITFROM one.
INITFROM_NONPERIODIC=(CM61-LR-pi-03-REDO.MAY0A CM61-LR-pi-03-REDO.MAY1A CM61-LR-pi-03-REDO.MAY2A)
# directory of the restart for each simulation. IF empty all simulations will use INITPATH one.
INITPATH_NONPERIODIC=(path/to/simu1 path/to/simu2)
# length of each simulation. If empty all simulations duration will be the default LENGTH option.
LENGTH_NONPERIODIC=(10Y 50Y)
```

For list-variables: use blank between values (no coma).

4. Configure a Perturbed Restart ensemble experience

In this section it is explained how to generate members from an initial condition which is perturbed by different means.

There are two ways to perturb the initial condition:

- apply some random white noise of defined amplitude to the temperature field of the coupler component (CPL) restart file
- apply some previously generated 3D temperature perturbation map to the temperature field of the ocean component (OCE) restart file

NOTE: It is MANDATORY that "NAME" in ensemble.card is the same of JobName in config.card JobName variable in config.card will be the name of the root directory that would be created containing all config and script files and the ensemble.

The example generates :

- Ensemble in directory named CM619-LR-dcppA-hindcast ensemble of 10 members, of 10 years duration, starting from year 1961 every 2 years (i.e. 1961, 163, 1965,... 1979) till 1980.
- Tree of subdirectories of ensemble will be:

CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast1961
CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast1963
CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast1965
CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast1967
....
CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast/CM619-LR-dcppA-hindcast1979

```
################################################################################
[Ens_PERTURB]
# active=y to use this ensemble type
active=y

# ensemble name (must be equal to JobName in config.card)
NAME=CM619-LR-dcppA-hindcast

# member nb (i.e nb of perturb initial restart for each date)
MEMBER=10

# periodic and member list simulations length
LENGTH=10Y

# start date of the first ensemble
BEGIN_INIT=19610101

# start date of the last ensemble
END_INIT=19801231

# timestep between each periodic simulation
PERIODICITY=2Y

# Restart name
INITFROM=CM61-LR-nudgSSTSire-r1-2DERS-m

# Restart directory
INITPATH=directory/IGCM_OUT/IPSLCM6/DEVT/historical
```

Restart files will be generated for each member at each date starting from BEGIN_INIT to END_INIT with a periodicity of PERIODICITY.
The variable MEMBER sets the number of members for each start date.
In this example members will be:

- For year 1961 :

CM619-LR-dcppA-hindcast1961/CM619-LR-dcppA-hindcast1961-01
CM619-LR-dcppA-hindcast1961/CM619-LR-dcppA-hindcast1961-02
CM619-LR-dcppA-hindcast1961/CM619-LR-dcppA-hindcast1961-03
...
CM619-LR-dcppA-hindcast1961/CM619-LR-dcppA-hindcast1961-10

- For year 1963 :

CM619-LR-dcppA-hindcast1963/CM619-LR-dcppA-hindcast1963-01
CM619-LR-dcppA-hindcast1963/CM619-LR-dcppA-hindcast1963-02
CM619-LR-dcppA-hindcast1963/CM619-LR-dcppA-hindcast1963-03
...
CM619-LR-dcppA-hindcast1963/CM619-LR-dcppA-hindcast1963-10

The directory in which the start date is retrieved is given by INITPATH and INITFROM.

For this example restart is from experiment **CM61-LR-nudgSSTSire-r1-2DERS-m** in directory $directory/IGCM_OUT/IPSLCM6/DEVT/historical

The way the perturbed member is generated depends on PERTURB_BIN array. The first two elements are the most important. The first one is the executable to be used to produce the members, the second one is the component from which the restart is perturbed.

In the Periodic Case it is only possible to build the members by applying a randomly generated temperature pattern on the restart file of the coupler. PERTURB_BIN should look like this :

```
PERTURB_BIN=(AddNoise, CPL, sstoc, O_SSTSST, 0.1)
```

The list is interpreted as follows:

- the used executable is *AddNoise*,
- the component is the coupler (*CPL*),
- the restart file to perturb contains *sstoc* in its name,
- the variable to perturb in the restart file is *O_SSTSST*,
- the randomly generated perturbation is in [-.05;+0.05] degrees

NOTE: The perturbation is not applied to grid points located under the sea ice.

This condition is "hard-written" in the AddNoise code. Because of a change of the name of the sea ice cover variable from IPSL-CM5A (OIceFrac) and IPSL-CM6 (OIceFrc), a modification of the code has been made by Olivier Marti in June 2016 to allow the code to search for both names

For each member a new restart file for the coupler will be generated using the executable *AddNoise?* to add some randomly generated temperature perturbation.

The corresponding restart file of each member will be stored in

```
directory/IGCM_IN/dcppA-hindcast/dcppA-hindcast/CPL/Restart/
```

The perturbation executable must be *AddNoise*.

```
PERTURB_BIN=(AddNoise, CPL, sstoc, O_SSTSST, 0.1)
```

Once config.card and ensemble.card will be properly filled launch ins_job -e to generate all ensemble directories.

```
ins_job -e # Check and complet job's header
```

5. Start your ensemble experience

Once your configurations files are properly filled and ins_job -e ends successfully you'll probably want to run your new ensemble. As mentionned upper, some scripts to run and clean all ensemble members will becreated.

You can use Qsub.*xxx*.sh (at the JobName directory level) to launch at same time ALL jobs:

```
sh Qsub.''xxx''.sh
```

If an error occured in all the runs, you could use *Qclean.*.sh* scripts.

Before starting all the ensemble we recommanded you to test only with one member to avoid massive cleaning for basic configuration troubles