

## Ensemble setup

---

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

### Table of Content

|  |          |
|--|----------|
| <b>Ensemble setup</b>                                    | <b>1</b> |
| <b>1. Prepare ensembles with <code>ins_job -e</code></b> | <b>2</b> |
| 1.1. Usage   | 2        |
| 1.2. Config.card   | 2        |
| 1.3. Ensemble.card                                       | 3        |
| 1.4. Configure a Date Restart ensemble                   | 3        |
| Periodic start dates                                     | 4        |
| Non-Periodic start dates                                 | 4        |
| 1.5. Configure a Perturbed ensemble                      | 5        |
| Periodic start dates                                     | 6        |
| Non-Periodic start dates                                 | 7        |
| List of members for a single start date                  | 8        |

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

## 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.

When `IPSLCM6.1.9-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:

```
modips1/config/IPSLCM6/EXPERIMENTS/IPSLCM/dcppAhindcast_CMIP6
```

into the directory:

```
modips1/config/IPSLCM6/
```

Several types of ensemble simulations can be prepared by filling [config.card](#) and more importantly [ensemble.card](#). All parameters for ensemble description are in [ensemble.card](#) and global simulation template is in the [config.card](#).

### 1.1. 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
- `Qsub.xxx.sh`: shell script to submit all jobs (`PeriodNb=5` for all simulations)
- `Qclean.PeriodLengt?`: shell script to submit all jobs (`PeriodNb=5` for all simulations)
- `Qclean.latestPackperiod`:

The `Qclean.PeriodLength?/year.xxx.sh` are bash files to use `clean_PeriodLength.job` or `clean_latestPackperiod.job` script for all simulations.

NOTE: If a directory exists, `ins_job` won't modify it. If only some directories of the ensemble are presents, it will create the missing ones and complete `Qsub.xxx.sh` shell file.

### 1.2. 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 overridden by variables present in [ensemble.card](#).

```
#D-- UserChoices -
[UserChoices]
#####
JobName=v3h4testB
#----- Short Name of Experiment
ExperimentName=v3h4testB
#----- DEVT TEST PROD
SpaceName=DEVT
LongName="IPSLCM5A CMIP5 DEVT phase decadal example with limited outputs."
TagName=IPSLCM5A
#D- Choice of experiment in EXPERIEMENTS directory
ExpType=IPSLCM5/decadal
```

```

#####
#-- leap, noleap, 360d
CalendarType=noleap
#-- Experiment dates : Beginning and ending
#-- "YYYY-MM-DD"
DateBegin=2013-01-01
DateEnd=2022-12-31
#####
#-- 1Y, 1M, 5D, 1D Period Length of one trunk of simulation
PeriodLength=1M
#####
#-- Total Number of Processors (minimum is 2 for a coupled configuration)
#JobNumProcTot=4
JobNumProcTot=32

```

A section [Ensemble] should also be present. It contains the information that we want to prepare an ensemble simulation with variable EnsembleRun set to y and three unset fields to be filled in the *config.card* of each member after `'ins_job -e` has run.

```

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

```

### 1.3. Ensemble.card

There are several sections in *ensemble.card*: [Ens\_PARAMETRIC], [Ens\_DATE] and [Ens\_PERTURB].

The choice of ensemble types is done by setting the variable *active* to *y* or *n*.

```

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

```

There are 3 types of ensembles :

- Parametric ensemble which is not implemented yet.
- Date restart ensemble which allows to configure simulations starting from different restart dates.
- Perturb ensemble which allows to generate members from an initial condition which is perturbed by different means.

### 1.4. Configure a Date Restart ensemble

We cover here the section which allows to generate identical simulations excepted the initial restart file. Indeed, the « Date Restart ensemble » was implemented to configure a set of simulations using several restart dates, generally chosen for a particular point (ex : randomly, particular climate oscillation phases, volcanic activity...).

In [ensemble.card](#) all configuration items of this ensemble are in [Ens\_DATE] section.

There are 2 types of possible configurations to define restarts dates : a periodic one (give year start / stop and periodicity) or non periodic one (give a list of desired restarts). The second one is recommended because it allows more options.

In both cases you must fill the following options : *active*, *NAME*, *LENGTH*, *INITFROM* and *INITPATH*.

```

[Ens_DATE]
# for using date ensemble, 'n' else.
active=y

# name of the ensemble (used to create root directory)

```

```

NAME= ENSTAMBORA

# default length of the simulation for non periodic and duration for all periodic (in Year or Month)
LENGTH=10Y

# Experiment name to find all restart files (and default one for non-periodic)
INITFROM=v3.historical6

# Restart root directory
INITPATH=/ccc/store/cont003/dsm/p86denv/dmf_import/IGCM_OUT/IPSLCM5A/PROD/historical

```

### Periodic start dates

In [ensemble.card](#), it is possible to specify a periodic list of start dates. Restart files will be generated for each member at each date starting from `BEGIN_INIT` to `END_INIT` with a periodicity of `PERIODICITY`, using `BEGIN_RESTART` as first restart. Leave all *NON\_PERIODIC* options empty (`NONPERIODIC`, `RESTART_NONPERIODIC`, `INITFROM_NONPERIODIC`, `LENGTH_NONPERIODIC`).

The following part of [ensemble.card](#) sets 10 years simulations from 1990-01-01 to 2000-01-01 every 2 years each with a restarts starting from 1814-12-31 every 2 years:

```

# start date of the first periodic simulation
BEGIN_INIT=19900101

# start date of the last periodic simulation
END_INIT=20000101

# duration between the start of 2 periodic simulations
PERIODICITY=2Y

# date for the first restart (next = first+periodicity). CAUTION of the calendar (use config.card one)!
BEGIN_RESTART=18141231

```

This will produce simulations starting at the dates : 1990-01-01, 1992-01-01, 1994-01-01, 1996-01-01, 1998-01-01, 2000-01-01. (`PERIODICITY` can be given in months for shorter periods).

The restart files are taken from `BEGIN_RESTART` every `PERIODICITY` step : 1814-12-31, 1816-12-31, 1818-12-31, etc...

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

To restart from experiment `v3.historical6` in directory `/ccc/store/cont003/dsm/p86denv/dmf_import/IGCM_OUT/IPSLCM5A/PROD/historical` fill:

```

# Restart name
INITFROM= v3.historical6

# Restart directory
/ccc/store/cont003/dsm/p86denv/dmf_import/IGCM_OUT/IPSLCM5A/PROD/historical

```

**CAUTION:** The variable `CalendarType` from `config.card` will be used to determine the next restart date. It should be consistent with the simulations from which you are initialising.

### Non-Periodic start dates

In [ensemble.card](#), it is also possible to specify manually all simulations running and restart dates, length, experiment names and directories to get restart files.

First, you need to left empty the periodic attributes `BEGIN_INIT`, `END_INIT`, `PERIODICITY` and `BEGIN_RESTART` in [ensemble.card](#). Then you can list the start date of all simulations with `NONPERIODIC` variable, all restart dates with `RESTART_NONPERIODIC` one, all experiments to get restart files in `INITFROM_NONPERIODIC`, all simulations restart path using `INITPATH_NONPERIODIC` and give the length of each simulation (`LENGTH_NONPERIODIC`).

Here is an example of a configuration :

```

# list of start dates for all simulations
NONPERIODIC=(18150101 19910101 19990101)

# list of corresponding restart dates
RESTART_NONPERIODIC=(18141230 19901230 19981231)

# simulation name to restart for each simulation. IF empty all simulations will use INITFROM one.
INITFROM_NONPERIODIC=( v3.historical6 v3.historical6 v5.historical1)

# directory of the restart for each simulation. IF empty all simulations will use INITPATH one.
INITPATH_NONPERIODIC= ( path/to/1st path/to/2nd path/to/3rd )

# length of each simulation. If empty all simulations duration will be the default LENGTH option.
LENGTH_NONPERIODIC=(10Y 10Y 50Y)

```

**WARNING:** For list variables, use blank between values (no coma).

This will produce 3 simulations which starting at the dates : 1815-01-01, 1990-01-01 and 1999-01-01 using respectively restarts from 1814-12-30, 1990-12-30, 1998-12-31 (note that the calendar should be different from the [config.card](#) one) taking into *v3.historical6* experiment for the 2 firsts and from *v5.historical1* for the last one (INITFROM is ignored when INITFROM\_NONPERIODIC is filled). Restarts will be taken respectively in the 3 directories specified with INITPATH\_NONPERIODIC (INITPATH is ignored when INITPATH\_NONPERIODIC is filled). Simulations length will be 10 years for the 2 firsts and 50 years for the last one. All restart experiments should be in the directory */ccc/store/cont003/dsm/p86denv/dmf\_import/IGCM\_OUT/IPSLCM5A/PROD/historical*.

Notice that INITFROM\_NONPERIODIC, LENGTH\_NONPERIODIC and INITPATH\_NONPERIODIC are not mandatory for non-periodic configuration. If you don't fill one of them or all the INITFROM value and/or LENGTH value and/or INITPATH will be used for all simulations :

```

# default length of the simulation for non periodic and duration for all periodic (in Year or Month)
LENGTH=10Y
[...]

# list of start dates for all simulations
NONPERIODIC=(18150101 19910101 19990101)

# list of corresponding restart dates
RESTART_NONPERIODIC=(18141230 19901230 19901231)

# simulation name to restart for each simulation. IF empty all simulations will use INITFROM one.
INITFROM_NONPERIODIC=

# length of each simulation. IF left empty all simulations durations will be the default LENGTH option.
LENGTH_NONPERIODIC=

# directory of the restart for each simulation. IF empty all simulations will use INITPATH one.
INITPATH_NONPERIODIC=

# Restart name
INITFROM= v3.historical6

# Restart directory
/ccc/store/cont003/dsm/p86denv/dmf_import/IGCM_OUT/IPSLCM5A/PROD/historical

```

This will produce 3 simulations starting at the dates : 1815-01-01, 1990-01-01 and 1999-01-01 using respectively restarts from 1814-12-30, 1990-12-30, 1990-12-31.

All of them use *v3.historical6* experiment in */ccc/store/cont003/dsm/p86denv/dmf\_import/IGCM\_OUT/IPSLCM5A/PROD/historical* directory to get restart files and their duration is 10 years.

## 1.5. Configure a Perturbed ensemble

We cover here the section which allows 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

Each method applies only to the relevant type of ensemble generation available inside [Ens\_PERTURB] as will be explained later.

Before detailing the different functionalities available in [Ens\_PERTURB] let us discuss the NAME variable.

This variable will be both the global name of the ensemble (ie directory name) and the prefix for each member:

```
# ensemble name
NAME=v3h4testB
```

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.

### Periodic start dates

For this type of perturbed ensembles the following variables are left empty:

```
# member list (apply list of pattern to initial state)
PERTU_MAP_LIST=()

# member list of names corresponding to each member
MEMBER_NAMESLIST=()

# member pattern global name
MEMBER_INITFROM=

# member pattern global directory for name
MEMBER_INITPATH=
...
# start dates list
NONPERIODIC=()
# length list for non periodic simulation (NOTE: use length above if not fill)
LENGTH_NONPERIODIC=()
...
# Path of Mask file
MASKPATH=
```

In [ensemble.card](#), it is possible to specify a periodic list of start dates.

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.

The following part of [ensemble.card](#) sets 10 members from 19900101 to 20000101 every 2 years each lasting 10 years:

```
# 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=19900101
# start date of the last ensemble
END_INIT=20000101
# timestep between each periodic simulation
PERIODICITY=2Y
```

This will produce 10 members starting at the dates : 19900101, 19920101, 19940101, 19960101, 19980101, 20000101. (PERIODICITY can be given in months for shorter periods)

Each time the restart file to be perturbed in order to produce each member is taken from the previous day of the start date : 19893112, 19913112, etc...

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

To restart from experiment v3h4BTxx in directory /ccc/store/cont003/gen2211/nguyens/IGCM\_OUT/IPSLCM5A/PROD/historical fill:

```
# Restart name
INITFROM=v3h4BTxx
# Restart directory
INITPATH=/ccc/store/cont003/gen2211/nguyens/IGCM_OUT/IPSLCM5A/PROD/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

**!!NOTA!! 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 (in our example we have ten) a new restart file for the coupler will be generated using the executable *addnoise* to add some randomly generated temperature perturbation.

For the year 1990, the corresponding restart file of member 1 will be stored in

```
$WORKDIR/IGCM_IN/v3h4testB190/v3h4testB190A/CPL/Restart/
```

### Non-Periodic start dates

For this type of perturbed ensembles the following variables are left empty:

```
# member list (apply list of pattern to initial state)
PERTU_MAP_LIST=()

# member list of names corresponding to each member
MEMBER_NAMESLIST=()

# member pattern global name
MEMBER_INITFROM=

# member pattern global directory for name
MEMBER_INITPATH=
...
# start dates list
NONPERIODIC=()
# length list for non periodic simulation (NOTE: use length above if not fill)
LENGTH_NONPERIODIC=()
...
# start date of the first ensemble
BEGIN_INIT=
# start date of the last ensemble
END_INIT=
...
```

```
# Path of Mask file
MASKPATH=
```

The variable LENGTH must be set to something but is not used, PERIODICITY must be set to NONE:

```
# periodic and member list simulations length
LENGTH=10Y
...
# timestep between each periodic simulation (NONE for nonperiodic)
PERIODICITY=NONE
```

To set 10 members for the starting dates 1990 and 1992 for a duration of 10 years, set MEMBER, NONPERIODIC and LENGTH\_NONPERIODIC as follows:

```
# member nb (i.e nb of perturb initial restart for each date)
MEMBER=10
...
# start dates list
NONPERIODIC=(19900101 19920101)
# length list for non periodic simulation (NOTE: use length above if not fill)
LENGTH_NONPERIODIC=(10Y 10Y)
```

This results in 20 simulations in total.

The restart files to be perturbed to produce each member are sought in directory INITFROM which PATH is INITPATH.

```
# Restart name
INITFROM=v3h4BT00
# Restart directory
INITPATH=/ccc/store/cont003/gen0826/labetoul/dmf_import/IGCM_OUT/IPSLCM5A/PROD/historical
```

This will result in using restarts from experiment *v3h4BT00* located in directory */ccc/store/cont003/gen0826/labetoul/dmf\_import/IGCM\_OUT/IPSLCM5A/PROD/historical*.

The perturbation executable must be *AddNoise*.

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

### List of members for a single start date

For this type of perturbed ensembles the following variables are left empty:

```
# member nb (i.e nb of perturb initial restart for each date)
MEMBER=
# timestep between each periodic simulation (NONE for nonperiodic)
PERIODICITY=NONE
# start dates list
NONPERIODIC=( )
# length list for non periodic simulation (NOTE: use length above if not fill)
LENGTH_NONPERIODIC=( )
```

It is important to leave PERIODICITY set to NONE and LENGTH\_NONPERIODIC as an empty list: the list of member method only works for a single start date and neither with periodic start dates nor with non periodic start dates.

The variables BEGIN\_INIT and END\_INIT are set to the same date, only BEGIN\_INIT will be used to provide the start date of the simulation for each member.

```
# start date of the first ensemble
BEGIN_INIT=20560101
# start date of the last ensemble
```



```
END_INIT=20560101
```

The variable LENGTH is the computation time which is the same for all members.

```
# periodic and member list simulations length
LENGTH=10Y
```

MEMBER\_NAMESLIST is the list of names given to each member. It gives the names of the subdirectories from which the Job is submitted for each member as well as the subdirectories in which the results are stored for each member.

PERTU\_MAP\_LIST (previously named as *MEMBER\_LIST*) is the list of perturbation maps files names prefix to apply to the restart file. It is implied that the files are named *prefix.nc*.

MEMBER\_INITFROM is the directory in which the perturbations maps are stored.

MEMBER\_INITPATH is the path to this directory.

```
# member list of names corresponding to each member
MEMBER_NAMESLIST=(OWN3DTA, OWN3DTB, OWN3DTC, OWN3DTD)

# member list (apply list of pattern to initial state)
PERTU_MAP_LIST=(OWN3DT_A, OWN3DT_B, OWN3DT_C, OWN3DT_D)

# member pattern global directory name
MEMBER_INITFROM=OWN3DTpf

# member pattern global directory for name
MEMBER_INITPATH=/ccc/work/cont003/gen2211/nguyens/PERTU/VECTORS
```

The variables INITFROM and INITPATH are still used to point to the directory where the restart files including the one to be perturbed are available.

```
# Restart name
INITFROM=piControl2

# Restart directory
INITPATH=/ccc/store/cont003/dsm/p86caub/dmf_import/IGCM_OUT/IPSLCM5A/PROD/piControl
```

For the member list perturbation type we use the executable AddPertu3DOCE and set PERTURB\_BIN this way:

```
# perturbation type
PERTURB_BIN=(AddPertu3DOCE, OCE, restart, tn, ORCA2_mesh_mask.nc)
```

The elements of the list mean:

- the executable to be called to generate the perturbation is *AddPertu3DOCE*
- the component is the Ocean (*OCE*)
- the restart file to perturb is *\*restart\*.nc*
- the field to perturb in the restart file is *tn*
- the meshmask file to tell if the gridcell is land or sea is *ORCA2\_mesh\_mask.nc*

The path to the mesh mask file is given in MASKPATH.

```
# Path of Mask file
MASKPATH=/ccc/cont003/home/gen2211/nguyens/addpertu
```

Once [config.card](#) and [ensemble.card](#) properly filled the directories containing the jobs to launch the simulations are created by issuing the command:

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