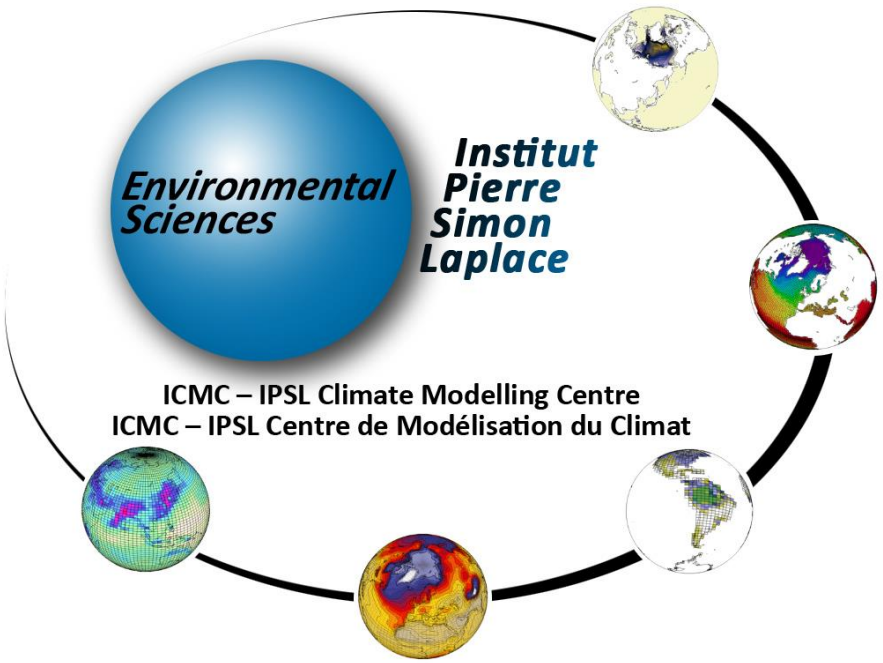
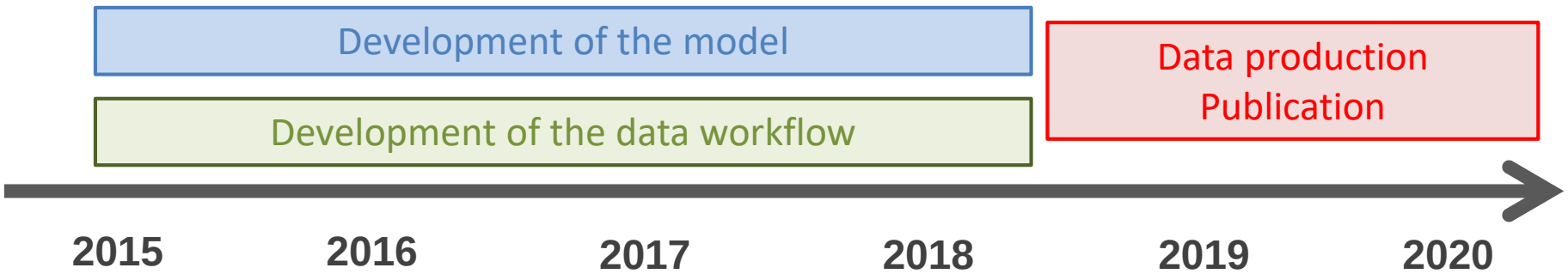



# CMIP6 production at IPSL : a 5-year story

Arnaud Caubel, Sébastien Denvil, M.Alice Foujols, Guillaume Levavasseur, Yann Meurdesoif, XIOS team, ICMC platform group (IPSL, France)
Gaëlle Rigoudy, Stéphane Sénési (Météo-France, France)
M.Pierre Moine, Sophie Valcke (Cerfacs, France)

*September, 24th 2020, Workshop on Coupling Technologies for Earth System Models*

# CMIP6 production at IPSL : a 5-year story



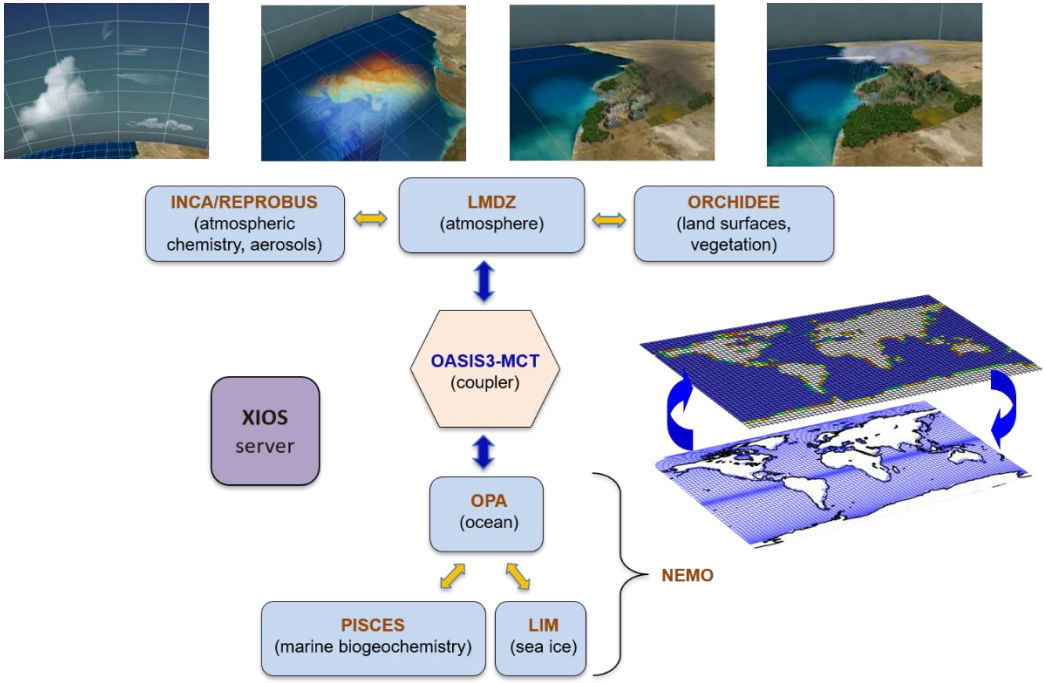
# Development of the model

Development of the model

Development of the data workflow

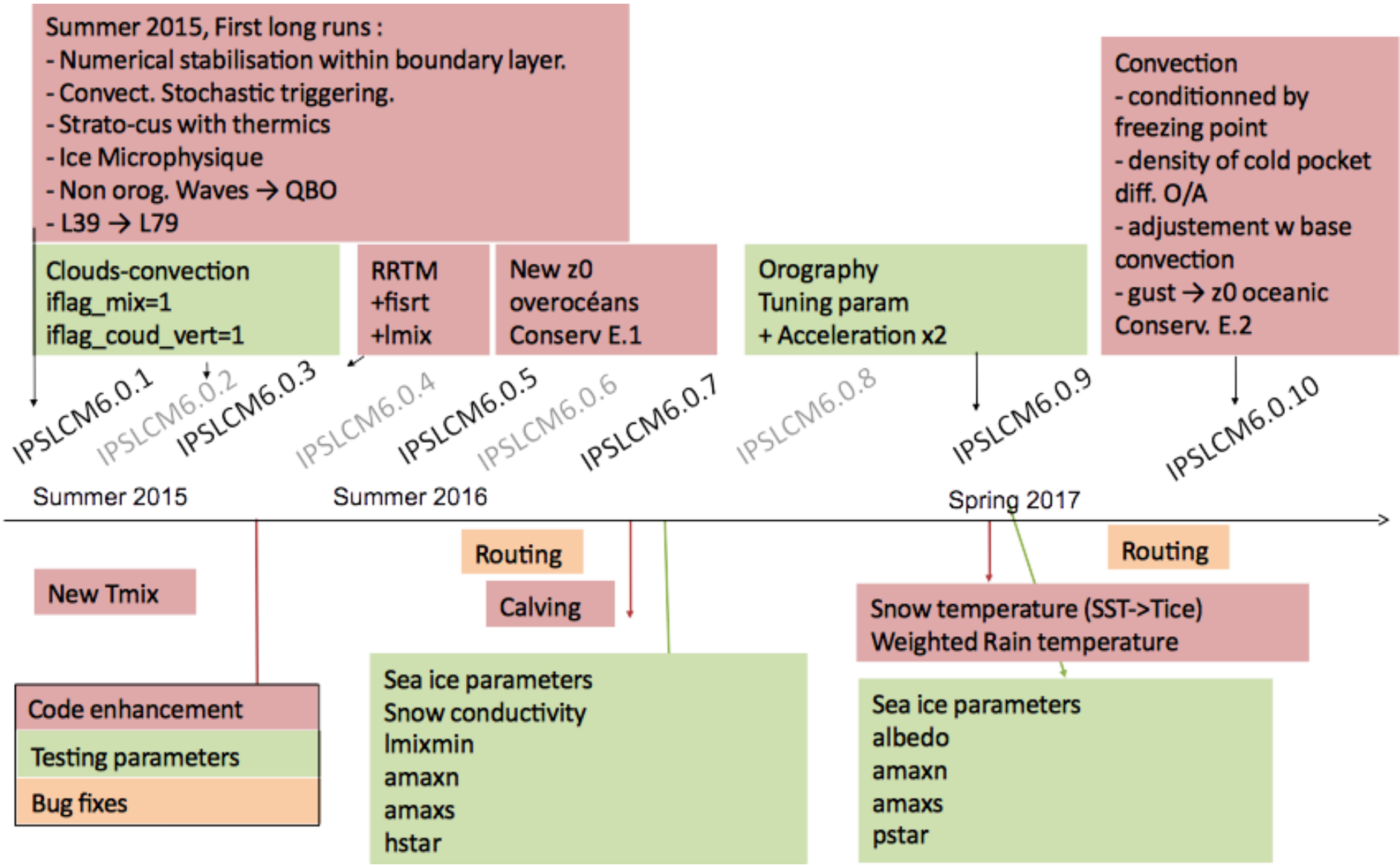
Data production
Publication

2015 2016 2017 2018 2019 2020

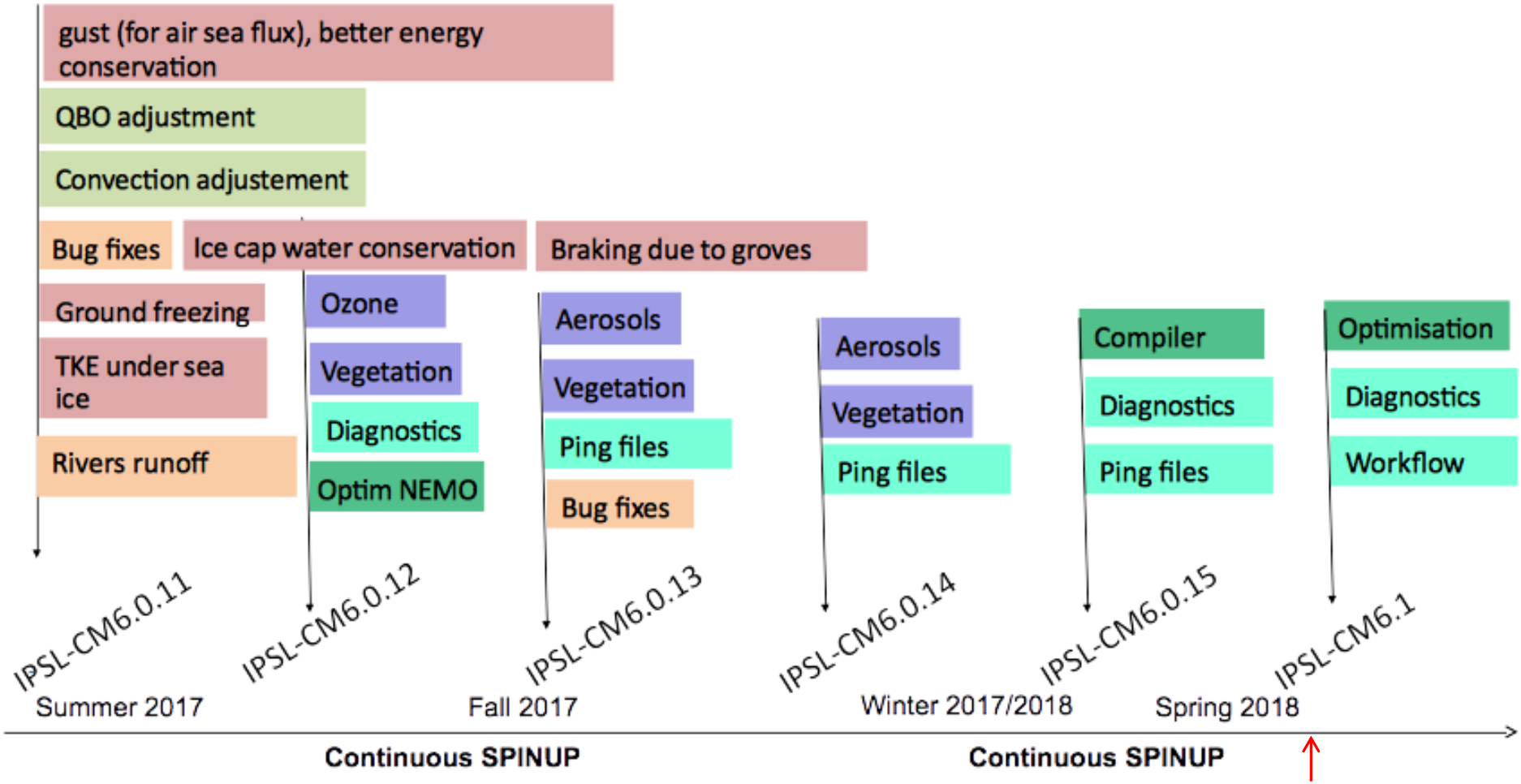



- **IPSL-CM6A-LR :** " the standard version"
  - LMDZ : Atmosphere 200 km, lon-lat grid
  - ORCHIDEE : Land Surface 200 km
  - NEMO : Ocean 100 km
  - LIM3 : Seaice 100 km
  - PISCES : Biogeochemistry 100 km
  - INCA : aerosols, atmospheric chemistry
  - OASIS3-MCT : ocean-atmosphere coupler
  - XIOS : Input/Output server
  - MPMD mode
    - Atmosphere : Hybrid parallelization MPI + OpenMP
    - Ocean : MPI only
    - XIOS server : MPI only

# Development of the model : 25 releases over 3 years



# Development of the model : 25 releases over 3 years




**2018 : Ready to start**

**… still continuing**

**IPSLCM6.1.11 in September 2020**

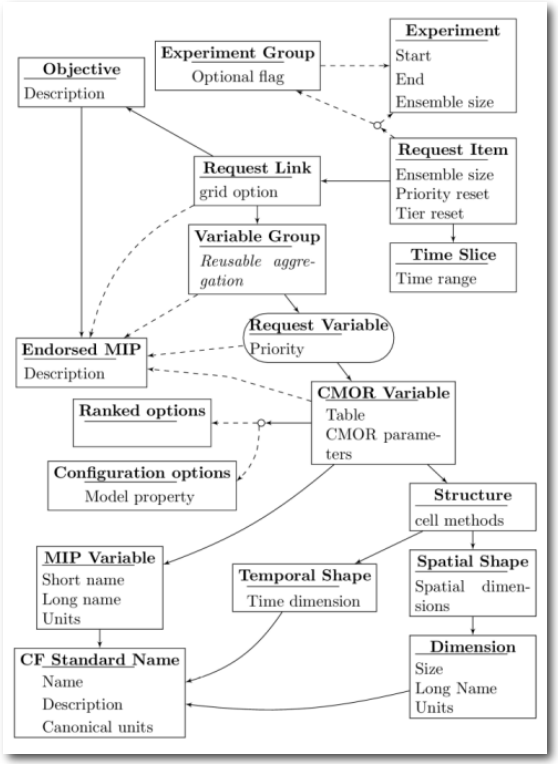
# Development of the data workflow

Development of the model

Development of the data workflow

Data production
Publication

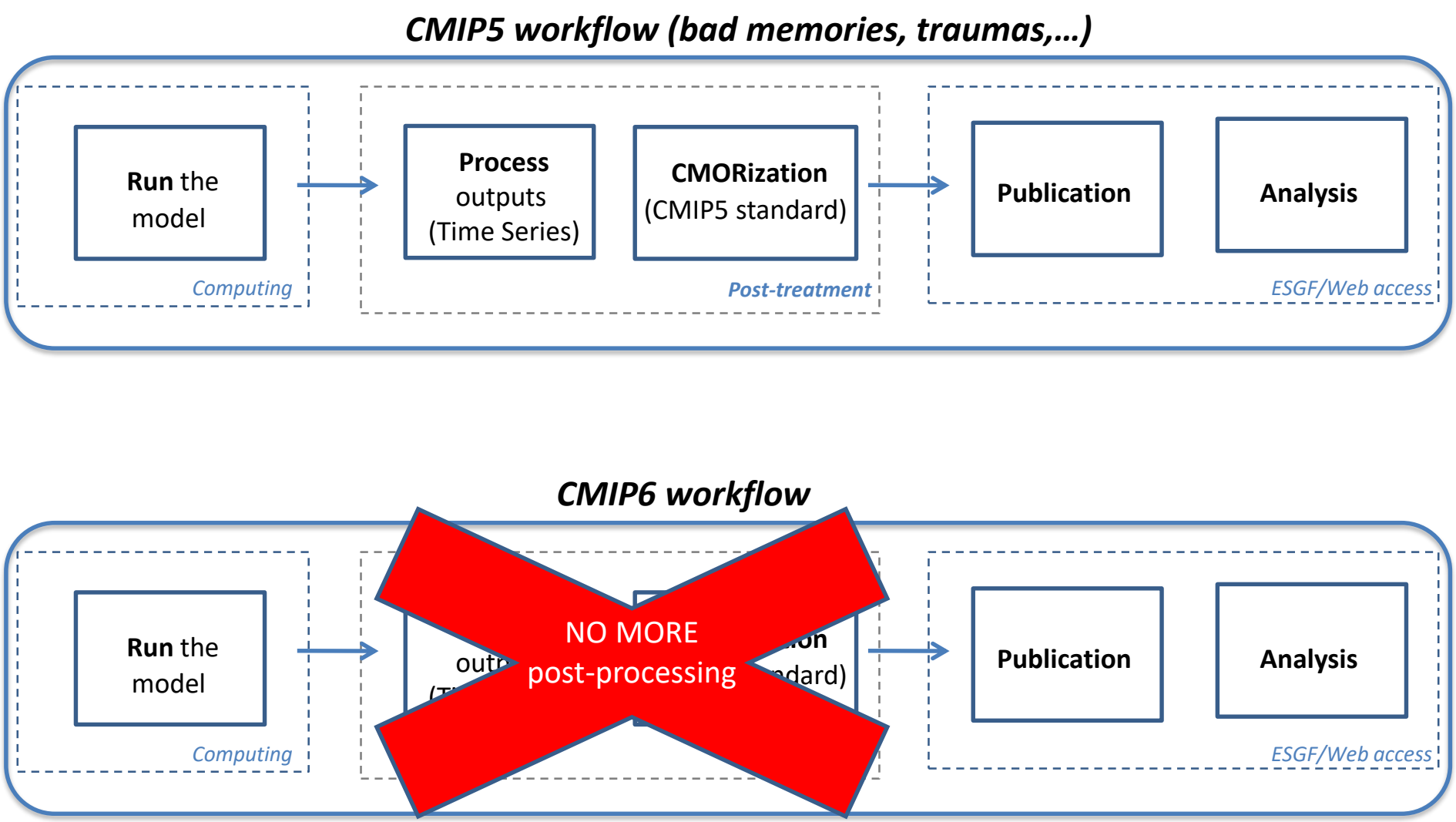
2015 2016 2017 2018 2019 2020



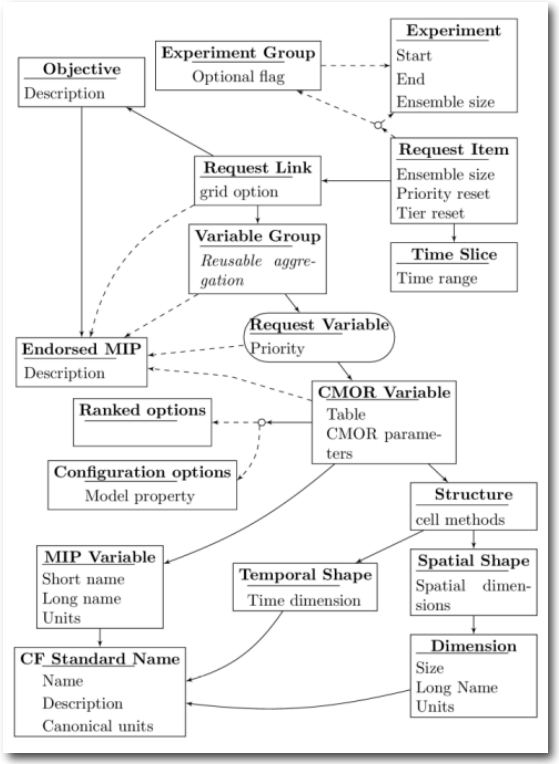

## The challenging CMIP6 Data Request

- What ? Required specific variables directly given by the model or to be processed from model variable : vertical interpolation (pressure levels), sites, …
- When ? Specifies variables (and sampling) needed for each experiments with high variability : from one experiment to the other, from one simulated year to the next one, from a modelling group to an other depending on the MIPs it is engaged in,…
- How ? Requires a specific format for the data to be published : name of the variable, name of the file, file attributes, variables attributes,…

# Development of the data workflow



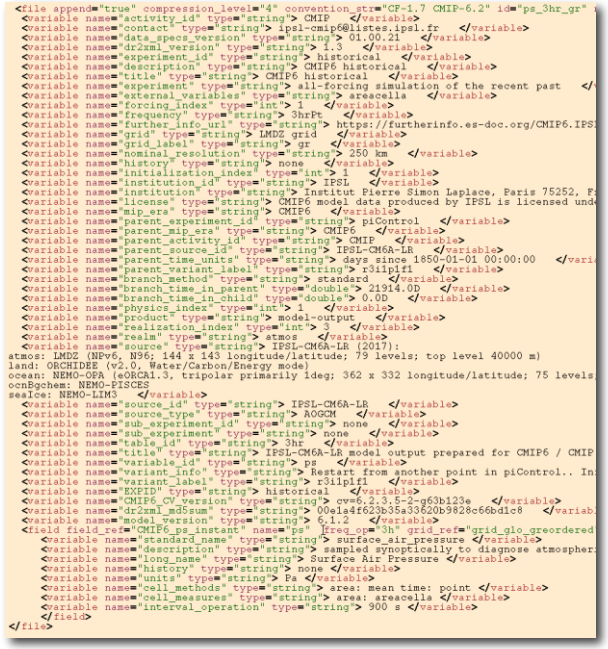
# Development of the data workflow




Tool for translating a CMIP Data Request to publishable datafiles using XIOS

dr2xml

*Developped at CNRM*



XIOS

*Developped at IPSL*

- publication ready data files thanks to workflow in-situ functionalities
  - CF and CMIP6 compliant
  - Units rescaling
  - Time integration (averaging, minimum, maximum)
  - Vertical interpolation on pressure levels
  - Horizontal remapping (ex : NEMO grid to regular grid, Icosaedral grid to regular grid)
  - Transects (flux across ocean straight)
- asynchroneous parallel writing in server mode

CMIP6-publication-ready data files

# Assembling model and workflow

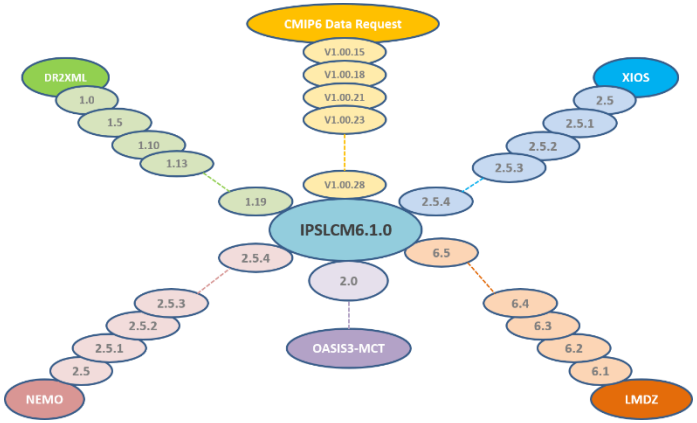
Development of the model

Development of the data workflow

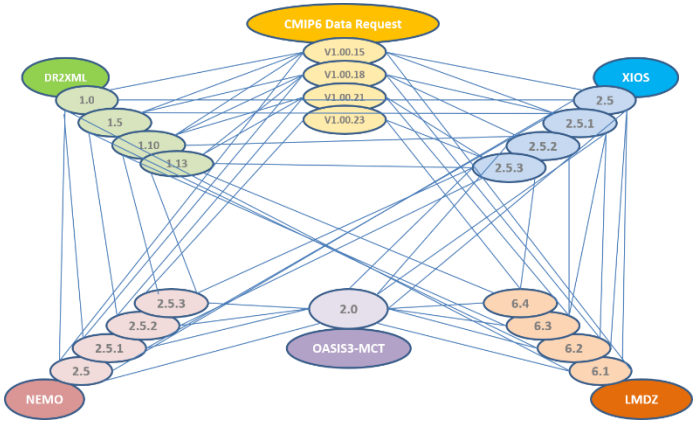
Data production
Publication

2015 2016 2017 2018 2019 2020

- Many tools evolved simultaneously = not easy to assemble Data Request, dr2xml, XIOS, NEMO, LMDZ, Oasis3-MCT into 25 releases
- A rigorous quality control was needed to validate each release
  - Technical validation : basic quality tests on short simulation, reproductibility tests during production phase (on the same machine)
  - Scientific validation : long simulation in development phase (and also to validate the use of a new machine)




*Evolution of components during development phase*




*Assembling of components during development phase*

# Model configurations and computing performances

Development of the model

Development of the data workflow

Data production
Publication

2015 2016 2017 2018 2019 2020

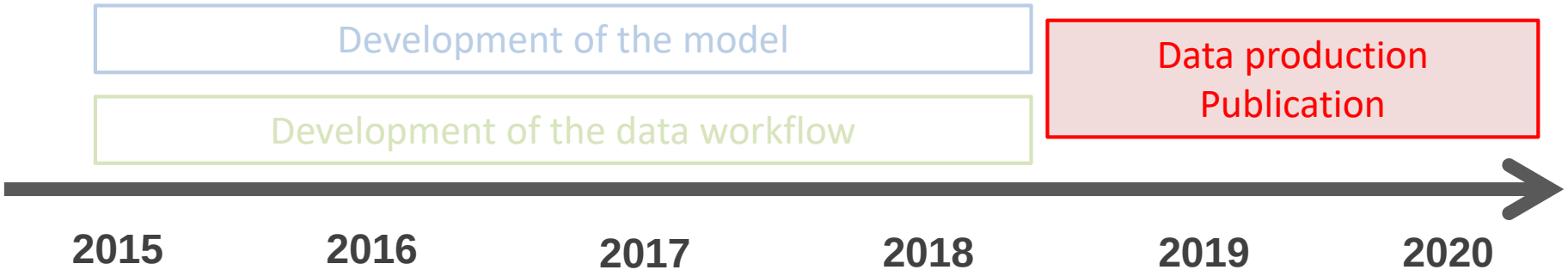
+ 20 % elapsed time

- **IPSL-CM6A-LR :** " the standard version"
  - Atmosphere 200 km, Ocean 100 km
  - Lon-lat atmospheric grid
- IPSL-CM6A-INCA-LR (Interactive aerosols)
- IPSL-CM5A2-INCA (Atmospheric chemistry)
  - Atmosphere 300 km, Ocean 200 km
- IPSL-CM6A-ATM-HR (High resolution)
  - Atmosphere 50 km, lon-lat grid
- **IPSL-CM7A-ATM-HR**
  - DYNAMICO icosaedric grid 50km (and 25 km soon)

| | No IO | IOs CMIP6 |
|---|---|---|
| **IPSL-CM6A-LR**<br>Lon-lat grid<br>Atmosphere 200 km<br>Ocean 100 km | **23 SYPD**<br>*928 cores* | **19 SYPD**<br>*940 cores*<br>*(928 + 12 XIOS servers)* |
| **IPSL-CM5A2-INCA**<br>Lon-lat grid<br>Atmospheric chemistry<br>Atmosphere 300 km<br>Ocean 200 km | **13 SYPD**<br>*609 cores* | **11 SYPD**<br>*621 cores*<br>*(609 + 12 XIOS servers)* |
| **IPSL-CM7A-ATM-HR**<br>DYNAMICO icosaedric grid, 50km<br>Data output both on native and regular grid | **8 SYPD**<br>*2560 cores* | **6 SYPD**<br>*2848 cores*<br>*(2560 + 12 XIOS servers, 24 cores/XIOS server)* |

*Performances on TGCC Irene Supercomputer (Intel Skylake)*

# Data production and publication

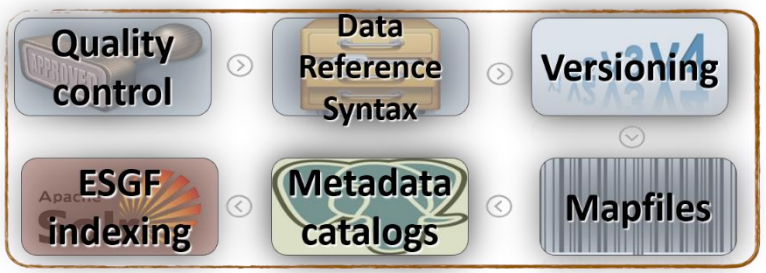



## CMIP6 production at IPSL

- 28 MIPs, 228 experiments, 850 simulations
- 55 000 years of simulation performed
- 20 logins (means different users) for the production campaign
- 200 logins for the climate model development and CMIP6 analysis
- in average 20 000 cores, with peak at 80 000 cores (one week at the end of Curie)
- ~400 millions computing hours (development + production) on TGCC (Curie Intel SandyBridge, Irene Intel Skylake, Irene AMD Rome)
- 4 Pb of data produced

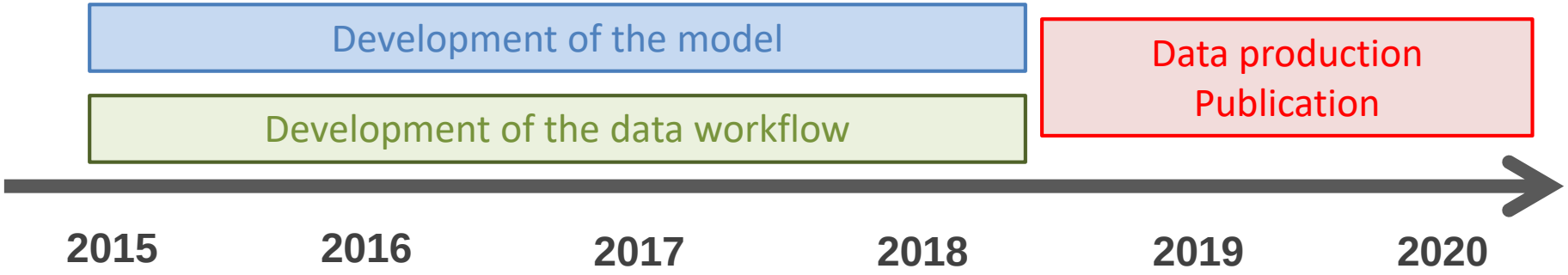
## CMIP6 publication at IPSL

***Publication process***

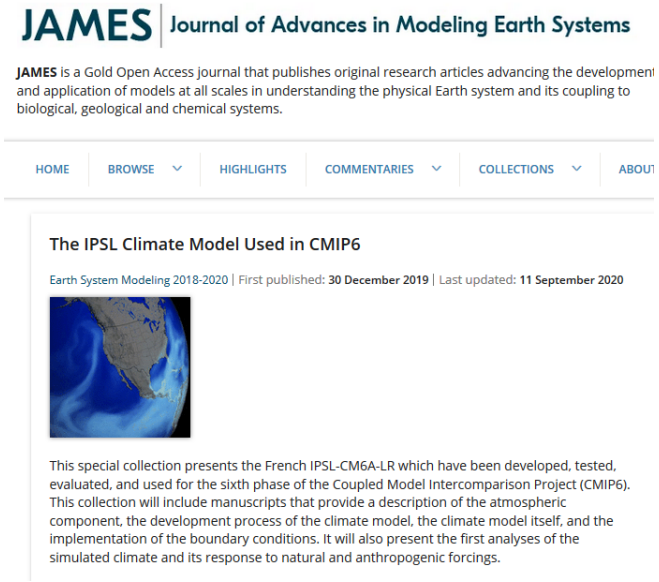



- 1st group to publish CMIP6 datasets
  *(2018-07-17T18:17:18.913Z)*
- 655 488 datasets published = 1.2 Pb

# Conclusion




- CMIP6 exercise at IPSL was a scientific, technical and human challenge
- Production on the fly of data ready to publication
  - work to do in development phase to set up the workflow
  - nothing to do in post-treatment step
- CMIP6 production is still in progress : HighResMIP 25 km,…
- Special Collection of articles in the JAMES AGU journal (2019/2020)



LSCE
Laboratoire des sciences du climat & de l'environnement

Institut Pierre Simon Laplace
Sciences de l'environnement

# Thank you !