

Demande d'allocation de temps de calcul pour 2016

22/10/2015

Dossier technique IPSL

Introduction

La plateforme de modélisation climatique de l'IPSL est un ensemble de composantes modélisant différentes parties du système climatique. Ces composantes peuvent être utilisées séparément, ou couplées ensemble dans une grande variété de configurations possibles en fonction des besoins scientifiques.

Les composantes sont le modèle de circulation atmosphérique LMDZ, le modèle de circulation océanique NEMO incluant la dynamique océanique, la glace de mer et la biogéochimie marine, le modèle de surfaces continentales ORCHIDEE et le modèle de chimie atmosphérique troposphérique INCA et stratosphérique REPROBUS. Les interactions entre les composantes sont décrites dans la Figure 1. Le système est conçu pour être évolutif et intégrer ou activer de nouvelles composantes, par exemple les calottes de glace. Il inclut également le coupleur OASIS. Il comprend aussi une bibliothèque d'entrées-sorties basée sur IOIPSL pour sa version historique et sur XIOS pour sa version massivement parallèle incluant aussi un serveur d'IO débranchable, un ensemble d'outils de gestion des versions (basés sur svn), de compilation MODIPSL et un ensemble de scripts d'exécution et de post-traitement libIGCM. Ces outils sont partagés par l'ensemble des composantes.

Le modèle couplé "Système Terre" de l'IPSL

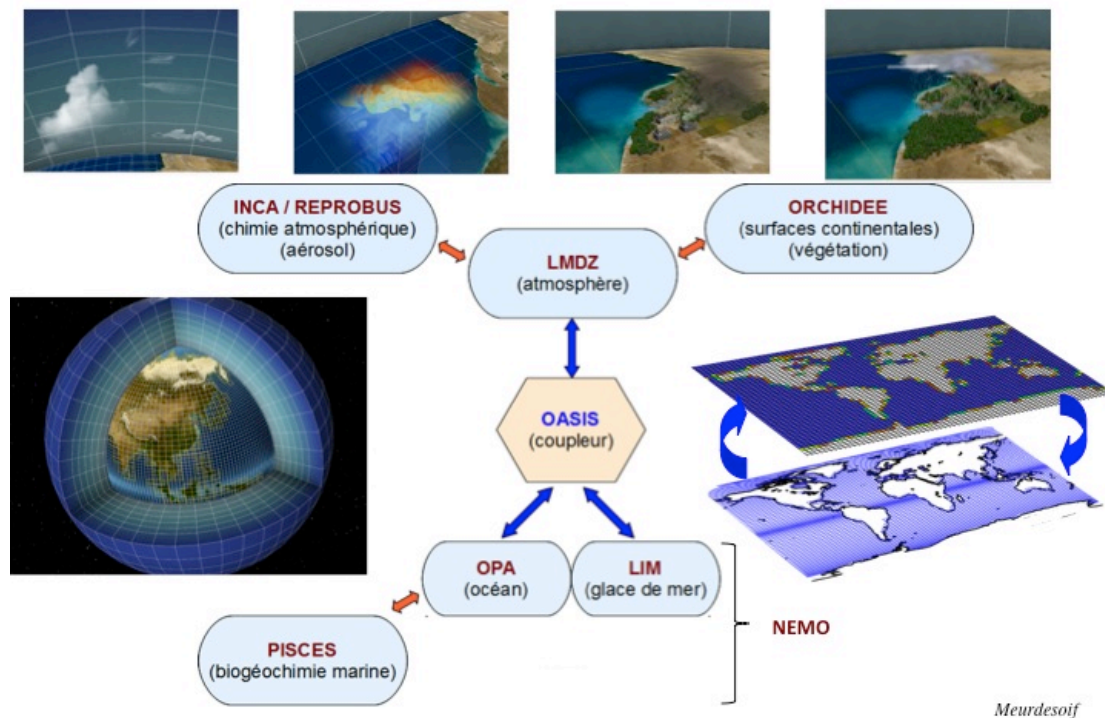


Figure 1. Le modèle système climatique de l'IPSL.

Les configurations supportées aujourd'hui sont :

- IPSLCM5_v5 : le couplé de référence pour les simulations de type CMIP5 incluant l'atmosphère, les surfaces continentales, l'océan, la glace de mer et le cycle du carbone
- LMDZOR_v5 : la version atmosphère-continentes correspondante.
- LMDZORINCA_v5 : la version atmosphère-continentes-chimie troposphérique correspondante.
- LMDZREPR_v5 : la version atmosphère-chimie stratosphérique correspondante.
- IPSLCM5CHT_v5 : le couplé de référence avec chimie troposphérique.
- IPSLCM5CHS_v5 : le couplé de référence avec chimie stratosphérique.
- NEMO : la version océanique correspondante incluant la glace de mer et la biogéochimie marine.
- ORCHIDEE_OL :
 - ORCHIDEE_SVN_AR5 : la version surfaces continentales correspondante.
 - ORCHIDEE_TAG : la version surfaces continentales utilisant la dernière version de référence d'Orchidee

Depuis 2013, la configuration IPSLCM6 est en chantier. Une première version est disponible : IPSLCM6_VLR_rc0. La version intégrant les résolutions retenues pour CMIP6 est en cours de développement : IPSLCM6_LR_rc1. Elle inclut les développements techniques les plus récents et va servir de base aux prochaines versions qui incorporeront les versions les plus à jour scientifiquement. La configuration IPSLCM6_rc0 inclut en particulier :

- NEMO v3-6_STABLE et LMDZ6
- La bibliothèque OASIS3-MCT
- XIOS pour toutes les composantes
- La parallélisation hybride MPI/OpenMP pour l'atmosphère, les surfaces continentales et la chimie
- Les résolutions CM6-VLR et CM6-LR

Algorithmes utilisés

LMDZ et NEMO/OPA

Ce sont les composantes dynamiques du système. Les deux modèles résolvent les équations dites primitives des fluides géophysiques :

- Equations de Navier-Stokes en repère tournant pour la vitesse horizontale, avec diffusion horizontale pour représenter les phénomènes sous-maille.
- Pour la verticale : approximation hydrostatique et diffusion turbulente verticale dans la couche limite.
- Conservation de la masse, de l'énergie, du sel. Conservation de la vorticité pour l'écoulement horizontal.

- Pour l'atmosphère : équation de conservation de la vapeur d'eau avec paramétrisation des phénomènes de condensation; également paramétrisations des transferts radiatifs.

Dans les configurations couplées NEMO/OPA-LMDZ, l'essentiel du temps de calcul est consommé par LMDZ.

NEMO/PISCES

NEMO/PISCES traite des traceurs océaniques. Il est composé d'un module de transport de traceurs océaniques, qui consomme l'essentiel du CPU, et un module de biogéochimie océanique. La parallélisation de NEMO/PISCES suit celle de la partie dynamique NEMO/OPA. NEMO/PISCES peut être utilisé off-line (la dynamique est lue dans des fichiers) ou in-line (un seul exécutable pour océan et biogéochimie marine qui interagissent directement).

NEMO/LIM

LIM traite de la dynamique (rhéologie et cinématique) et de la thermodynamique de la glace de mer. NEMO_LIM peut être utilisé off-line (la dynamique est lue dans des fichiers via OASIS et le Standalone Surface module) ou in-line (un seul exécutable pour océan et biogéochimie marine qui interagissent directement).

INCA

INCA traite de la chimie et des aérosols atmosphériques. Il calcule les réactions chimiques (modèle Eulérien 3D, choix entre méthode explicite et/ou implicite), l'effet des aérosols sur le système radiatif ainsi que les émissions, les transformations photochimiques et les dépôts secs et humides des espèces chimiques et des aérosols. Il est couplé au module de transport du modèle LMDZ. La parallélisation de INCA suit celle de la partie dynamique LMDZ. Un niveau de parallélisation supplémentaire, qui permettra de déporter le transport de traceurs sur des processeurs dédiés à cette tâche, a été étudié par le passé.

REPROBUS

REPROBUS traite de la chimie et des aérosols stratosphériques.

ORCHIDEE

ORCHIDEE rassemble trois composantes : SECHIBA traite des transferts d'eau et d'énergie entre l'atmosphère, la végétation et le sol ; STOMATE traite du cycle du carbone dans la végétation et les sols ; LPJ traite de la dynamique de la végétation. ORCHIDEE consomme peu de temps CPU par rapport aux autres composantes. ORCHIDEE peut aussi être utilisé en offline, sur un seul point, sur un ensemble de points ou sur l'ensemble de la planète, en étant piloté par des champs climatologiques précalculés.

XIOS

XIOS (XML-IO-server) est une nouvelle bibliothèque capable de gérer les entrées-sorties des modèles climat avec souplesse et performances. Le pilotage des sorties se fait grâce à un fichier XML. Les performances sont atteintes grâce aux processeurs supplémentaires dédiés. Depuis mi-2012 elle est disponible en production et en cours de déploiement dans les différentes composantes du modèle IPSL. Voir <http://forge.ipsl.jussieu.fr/ioserver>

Structure des programmes

Pour les deux modèles dynamiques NEMO et LMDZ, les équations sont discrétisées en différences finies, sur une grille C (Arakawa 1972) et l'intégration temporelle est réalisée par un schéma saute-mouton. Pour éviter la divergence des pas de temps pairs et impairs due au schéma saute-mouton, un filtre Asselin est utilisé dans le cas de l'océan alors que pour l'atmosphère un schéma Matsuno (avant-arrière) est appliqué tous les cinq pas de temps. Pour satisfaire le critère de stabilité CFL, le modèle d'atmosphère a un pas de temps de six minutes et le modèle d'océan de une heure trente aux résolutions dans sa configuration IPSLCM5A-LR. Dans le modèle d'atmosphère, les termes de diffusion et de sources d'énergie sont calculés toutes les trente minutes et toutes les six heures pour le rayonnement.

Parallélisation

LMDZ

LMDZ est parallélisé par découpage de domaine, chaque partie étant découpée de façon particulière.

Dynamique

Le découpage se fait par bande de latitude et par niveaux verticaux. Le découpage en latitude est une parallélisation en mémoire distribuée, avec l'utilisation de la bibliothèque *MPI*. Le découpage vertical se fait grâce à une parallélisation en mémoire partagée, avec l'utilisation d'*OpenMP*. Toutefois, cette parallélisation possède des limites intrinsèques de l'ordre de la centaine de cœurs de calcul aux résolutions usuelles.

Transport

La méthode est identique à celle de la dynamique, mais les bandes de latitudes sont réparties de façon égale entre les processeurs.

Physique

Découpage mono-dimensionnel : on répartit les points également sur les processeurs, en mélangeant les parallélisations *MPI* et *OpenMP* pour conserver le même nombre de tâches *MPI* et de thread *OpenMP* que dans les parties dynamique et transport. Il faut noter que cette partie est *embarassingly* parallèle et s'adaptera aisément à toute autre module dynamique comme DYNAMICO.

LMDz et DYNAMICO

Le maillage latitude-longitude utilisé par LMDZ possède deux points singuliers aux pôles qui limitent, de fait, la scalabilité du modèle. DYNAMICO, le nouveau cœur de calcul de la dynamique, a été conçu en utilisant un maillage icosaédrique sur lequel les équations de la dynamique ont été discrétisées en utilisant des schémas numériques nouveaux conservant l'énergie et la vorticit  potentielle. Le d veloppement du code DYNAMICO a d marr  en 2012 et, en 2013, les premi res simulations avec DYNAMICO ont  t  r alis es. Elles ont montr  une scalabilit  sur des cas acad miques, sans physique et avec une physique simplifi e de type aquaplan te,   la r solution usuelle (comparable   celle de IPSLCM5A-LR) jusqu'  3000 c eurs, l  o  LMDZ stagnait   partir de 256 processeurs.

ORCHIDEE

Lorsqu'il est couplé à LMDZ (via une interface Fortran), la parallélisation d'ORCHIDEE est identique à celle de la partie physique de LMDZ à l'exception des tâches associées à des domaines sans aucun point de terre qui ne font que participer aux barrières globales.

NEMO

NEMO est parallélisé par découpage de domaine dans les directions X et Y. C'est une parallélisation à mémoire distribuée utilisant la bibliothèque *MPI*.

Précisions informatiques

Temps CPU, mémoire utilisée et volumes de fichiers

Voir les performances plus loin.

Sur le serveur de fichiers (Ergon et CCCSTOREDIR) une simulation de 150 ans avec IPSLCM5A, toutes sorties activées (comme historical) représente 400 Go et 8000 inodes. Les ATLAS représentent également un volume de 20 000 inodes, stockés sur Ergon ou sur CCCWORKDIR.

Langages de programmation utilisés

Tous les modèles sont programmés en *Fortran 90*. Cependant, pour des raisons de performances, peu des spécificités nouvelles du Fortran 90 sont utilisées. La bibliothèque d'entrée-sortie XIOS et son serveur détachable sont écrits en C++.

Bibliothèques et logiciels tiers utilisés

Pour l'exécution des codes de calculs

- NetCDF

En post-traitement

- evince (remplace acroread ou xpdf)
- cdo
- NetCDF
- ferret
- Firefox
- ghostscript
- imagemagick
- nco
- netpbm
- Python 2.6.2, avec les modules NetCDF4 et numpy
- subversion
- tetex-latex

Etat de préparation des programmes

Le modèle couplé et toutes ses composantes sont exploités de façon intensive depuis plus de vingt ans sur les différents calculateurs du CEA et du CNRS. Ils ont régulièrement été optimisés sur ces machines.

La production scientifique se fait actuellement en utilisant quelques dizaines à quelques centaines de cœurs.

Les développements autour du nouveau cœur dynamique DYNAMICO vont permettre l'utilisation, tout d'abord en développement puis en production, de milliers de cœurs de calcul.

Checkpoint restart

Tout les codes composant le modèle couplé disposent d'un système de checkpoint-restart. Un ensemble de scripts permet de décomposer les simulations longues (plusieurs dizaines ou centaines d'années). Généralement, l'exécutable est relancé chaque année de simulation, avec écriture et sauvegarde d'un checkpoint. Chaque job enchaîne plusieurs années, de façon à avoir des travaux en machine environ 24h. Les jobs sont enchaînés sur plusieurs semaines pour atteindre la longueur de simulation voulue.

Performances

Performances du modèle couplé et des configurations forcées associées

Ces chiffres sont issus de programmes en production scientifique, avec en particulier des conditions réalistes d'entrées/sorties. Les calculs et les post-traitements se font sur curie nœuds fins. L'expérience montre qu'il faut demander **10% de temps calcul supplémentaires** pour les post-traitements **au TGCC**. Pour simplifier, comme les performances sont globalement similaires sur ada à l'IDRIS, un seul tableau est indiqué. De même pour la machine occigen au CINES. Néanmoins la chaîne de calcul n'est pas opérationnelle au CINES. Seules les compilations et les exécutions simples sont supportées par l'IPSL à ce jour.

Mesure consommation 10 ans IPSLCM6-VLR Résolution 96x95x39 - ORCA2

machine	Nb cœurs	Temps réel/10 ans	Temps à demander
curie	160 (31 océan + 32*4 atm. + 1 xios)	8 h	1 300 h

Mesure consommation 10 ans IPSLCM6-LR Résolution 144x142x79 – ORCA1

machine	Nb cœurs	Temps réel/10 ans	Temps à demander
curie	608 (211 océan + 47*8 atm. + 1 xios)	43 h	28 750 h

Mesure consommation 10 ans LMDZOR6-LR Résolution 144x142x79

machine	Nb cœurs	Temps réel/10 ans	Temps à demander
curie	377 (47*8 atm. + 1 xios)	43 h	17 800 h

Estimation consommation 10 ans IPSLCM6-MR Résolution 280x280x39 – ORCA1

machine	Nb cœurs	Temps réel/10 ans	Temps à demander
curie	803 (67 océan + 92*8 atm. + 1 xios)	80 h	64 300 h

1 mois configuration forcée LMDZORINCA (LMD 96x95x39–INCA-NMHC-AER) :

Real time : 6000s sur 32 cœurs

A noter : la nouvelle physique de LMDZ (runs de type IPSLCM5B) prend 2,5 fois plus de temps que l'ancienne (runs de type IPSLCM5A).

Influence de la charge des noeuds sur les performances :

Comparatif d'un nœud avec 16 cores running (plusieurs jobs LMDZ par nœuds) vs non chargé (1 job LMDZ par nœud) sur Curie, en mode Newtonien 96x95x93 (sans physique) 480 pas de temps par jour, simulation de 5 jours, en mode MPI seul (= 1 thread openmp).

#core	Nœuds chargés	Nœuds non chargé
1	386 s	205 s
2	197 s	110 s
4	98.18 s	72 s
8	50.70 s	54 s

- ⇒ L'influence de la charge du nœud est importante notamment lorsqu'on a peu de cœurs running par nœud : presque un facteur pour une charge à $\frac{1}{4}$, équivalent à une charge à $\frac{1}{2}$.
- ⇒ La limitation des performances est essentiellement liée à la bande passante mémoire du nœud de calcul.

Filtre FFT Vs filtre classique

Le filtrage zonal des variables dans les régions polaires du modèle d'atmosphère LMDZ est nécessaire (critère CFL) pour permettre des simulations utilisant des pas de temps raisonnables. On peut soit utiliser un filtre « classique » (basé sur des produits matrice-vecteur) ou un filtre « FFT ».

Configuration 96x95x39 : 16 processus MPI

Filtre FFT : 31.96s, coût du filtre 2.87 s

Filtre classique : 32.37s, coût du filtre : 4.11 s

- ⇒ Coût du filtre assez marginal à cette résolution, les progrès en performance des SandyBridge (AVX) ont nettement amélioré les performances de la MKL (FFT et BLAS) alors que LMDZ n'a pas profité du même gain.

Configuration 360x180x39 : 40 processus MPI

Filtre FFT : 92.54 s, coût du filtre 22 s

Filtre classique : 132s, coût du filtre : 62 s

- ⇒ Le poids du filtre augmente avec la résolution.
- ⇒ Le filtre FFT ($\sim n \log n$) est 3x plus efficace que le filtre classique ($\sim n^2$)
- ⇒ Gains de l'ordre de 40%

Post-traitement

La chaîne de calcul climat est composée de nombreux travaux qui s'enchaînent. Pour une simulation d'une centaine d'années, il y a quelques dizaines de jobs de calcul et

près d'un millier de jobs de post-traitement. Un audit de cette chaîne de calcul a été réalisé en 2012-2013 et est disponible à : http://forge.ipsl.jussieu.fr/igcmg/attachment/wiki/BasculeCCRTTGCC/Prestation_CC_RT_TGCC/Audit_chaine_calcul_IPSL.pdf

A noter : l'utilisation de XIOS simplifie et fiabilise cette chaîne en supprimant les travaux de rebuild/mask puisque les fichiers sont alors directement générés depuis les jobs de calcul. Ce qui supprime 150 jobs de post-traitement, soit 15%, par rapport à l'ancienne méthode.

TGCC

Les jobs de calcul utilisent les nœuds fins. Le passage des jobs se fait selon une priorité dont le calcul tient compte de l'ancienneté du job, de la consommation récente et du type de projet (PRACE 80 % ou GENCI 20%). Il n'est pas possible de prévenir qu'un job appartient à une longue chaîne de calcul qui va rester plusieurs semaines en machine. Dans certains cas, nous pourrions être amenés à demander des priorités afin de garantir des temps de passage compatibles avec les échéances des projets.

Les fichiers intermédiaires sont stockés en SCRATCHDIR dont le quota à 20 To s'est avéré suffisant, sauf exception.

Certains fichiers (produits de simulations pour la communauté) sont destinés à être visibles depuis les serveurs dods : <http://dods.extra.cea.fr/work/login> et <http://dods.extra.cea.fr/store/login/>

Cette fonctionnalité devrait passer sur le datanode ESGF du TGCC dès sa remise en route opérationnelle.

IDRIS

Les jobs de calcul utilisent Ada. Les jobs de post-traitements utilisent Adapp.

Les fichiers intermédiaires sont stockés WORKDIR, et l'archivage pérenne sur Ergon.

Certains fichiers sont également destinés à être visibles depuis le serveur dods : <http://dodsp.idris.fr/login>

Cette fonctionnalité devrait passer sur le datanode ESGF de l'IDRIS dès sa remise en route opérationnelle.

Logiciels de post-traitements nécessaires

- evince (remplace acroread ou xpdf)
- cdo
- NetCDF
- ferret
- Firefox
- ghostscript
- imagemagick
- nco
- netpbm
- Python 2.6.2 au moins, avec les modules NetCDF4 et numpy
- subversion
- tetex-latex

Référence sur le modèle d'atmosphère LMDz

- Bony S., H. Le Treut, J. P. Duvel and R. Kandel, Satellite Validation of GCM simulated annual cycle of the Earth Radiation Budget and Cloud Forcing, *J. Geophys. Res.*, 97, 18 061-18 081, 1992.
- Harzallah A., and R. Sadourny : Internal versus SST-forced atmospheric variability as simulated by an atmospheric general circulation model, *J. of Climate*, 8, 474-495, 1995.
- Hourdin, F. and A. Armengaud (1999). "The use of finite-volume methods for atmospheric advection of trace species. Part I: Test of various formulations in a general circulation model." *Monthly Weather Review* 127(5): 822-837.
- Krinner, G., Ch. Genton, Z.X. Li, et P. Le Van, 1997, Studies of the Antarctic climate with a stretched grid general circulation model, *J. Geophys. Res.*, 102D, 13,731-13,745.
- Le Treut H., Z. X. Li, and M. Forichon, Sensitivity of the LMD general circulation model to greenhouse forcing associated with two different cloud water parameterizations, *J. of Climate*, 7, 1827-1841, 1994.
- Le Treut, H. and ZX. Li, Sensitivity of an atmospheric general circulation model to prescribed SST changes : feedback effects associated with the simulation of cloud optical properties, *Climate. Dynamics*, 175-187, 1991.
- Li X., 1999 : Ensemble atmospheric GCM simulations of climate interannual variability from 1979 to 1994, *J. of Clim*, 12, 986-1000.
- Frédéric Hourdin, Jean-Yves Grandpeix, Catherine Rio, Sandrine Bony, Arnaud Jam, Frédérique Cheruy, Nicolas Rochetin, Laurent Fairhead, Abderrahmane Idelkadi, Ionela Musat, Jean-Louis Dufresne, Alain Lahellec, Marie-Pierre Lefebvre, Romain Roehrig (2013). LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, 40, Issue 9-10.

Références de base sur le modèle NEMO

L'ensemble des références est disponible ici : <http://www.nemo-ocean.eu/About-NEMO/Reference-manuals>

On citera en particulier :

- Madec G. 2014: "NEMO ocean engine" (Draft edition r5171). Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No 1288-1619.
- PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies O. Aumont, C. Ethé, A. Tagliabue, L. Bopp, and M. Gehlen *Geosci. Model Dev.*, 8, 2465-2513, 2015
- The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities C. Rousset, M. Vancoppenolle, G. Madec, T. Fichefet, S. Flavoni, A. Barthélemy, R. Benshila, J. Chanut, C. Levy, S. Masson, and F. Vivier *Geosci. Model Dev.*, 8, 2991-3005, 2015

Référence de base sur le coupleur et le modèle couplé

- Marti, O., P. Braconnot, et al. (2005). The new IPSL climate system model: IPSL-CM4. Paris, Institut Pierre Simon Laplace, Note du Pôle de modélisation n° 26, 84 pages, <http://igcmg.ipsl.jussieu.fr/Doc/IPSLCM4>.
- Terray, L., E. Sevault, et al. (1995). The OASIS Coupler User Guide Version 2.0. Toulouse, Cerfacs: 95-46
- Marti, O., P. Braconnot, J. L. Dufresne, J. Bellier, R. Benshila, S. Bony, P. Brockmann, P. Cadule, A. Caubel, F. Codron, N. de Noblet, S. Denvil, L. Fairhead, T. Fichefet, M. A. Foujols, P. Friedlingstein, H. Goosse, J. Y. Grandpeix, E. Guilyardi, F. Hourdin, A. Idelkadi, M. Kageyama, G. Krinner, C. Levy, G. Madec, J. Mignot, I. Musat, D. Swingedouw, and C. Talandier, (2010). Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dynamics*, 34, 1-26.
- J.-L. Dufresne, M.-A. Foujols, S. Denvil, A. Caubel, O. Marti, O. Aumont, Y. Balkanski, S. Bekki, H. Bellenger, R. Benshila, S. Bony, L. Bopp, P. Braconnot, P. Brockmann, P. Cadule, F. Cheruy, F. Codron, A. Cozic, D. Cugnet, N. de Noblet, J.-P. Duvel, C. Ethé, L. Fairhead, T. Fichefet, S. Flavoni, P. Friedlingstein, J.-Y. Grandpeix, L. Guez, E. Guilyardi, D. Hauglustaine, F. Hourdin, A. Idelkadi, J. Ghattas, S. Joussaume, M. Kageyama, G. Krinner, S. Labetoulle, A. Lahellec, M.-P. Lefebvre, F. Lefevre, C. Levy, Z. X. Li, J. Lloyd, F. Lott, G. Madec, M. Mancip, M. Marchand, S. Masson, Y. Meurdesoif, J. Mignot, I. Musat, S. Parouty, J. Polcher, C. Rio, M. Schulz, D. Swingedouw, S. Szopa, C. Talandier, P. Terray, N. Viovy, N. Vuichard, (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40, Issue 9-10.

