



- **CMIP6 & co :**
 - **Production**
 - **Publication**
 - **Site web**
- **Ressources, stockage et portage**
- **Outils :**
 - **CliMAF**
 - **EcliS**
 - **Autres outils**

Document de travail :

<https://mypads2.framapad.org/mypads/?/mypads/group/gmgec-tech-cay793z/pad/view/document-partage-gmgec-tech-juillet-z5r8l9y8>

Réunion GMGEC-Tech – 3 juillet 2020

CMIP 6 & co

Production CMIP6

- Production des champs fixes (Marie-Pierre) :
 - Document préparé pour décrire la façon de faire.
 - Actuellement, un param pour une configuration type disponible pour générer les 33 champs fixes (tables fx, Ofx et Efx). Seuls 20 sortent :
 - Certains champs de surface doivent être récupérés dans le PGD (ex : fraction de sable) → voir avec Bertrand
 - Problème avec certains champs océaniques → voir avec Aurore
 - Pour la génération pour toutes les expériences, deux possibilités :
 - Faire une seule simulation puis modifier les attributs à la main
 - Faire un run de 1 jour par simulation et tirer parti de dr2xml pour générer les attributs
 - Plutôt sur la deuxième possibilité car beaucoup d'attributs à modifier. La liste des simulations avec param et settings a été faite par ailleurs, à transmettre.

Errata Service, known issues

- Lien Errata Service (liste des Errata déclarés) pour un modèle ou un institut donné dorénavant disponible (lien ajouté sur le site web CMIP6 du CNRM). Par ex pour CM6-1 :
<https://errata.es-doc.org/static/index.html?institute=cnrm-cerfacs&project=cmip6&source=cnrm-cm6-1>
- champs ***4co2 ESM-2-1** erronés (calculés 1xCO2) : déclaration faite sur l'ErrataService (456 datasets, 38 simulations concernées) et info en ligne sur site web CMIP6. Dépublication toujours en cours.....
- problème membre 29 historical CM6-1 et membre 3 de hist-GHG CM6-1 (continuité chaleur de l'océan) : info en ligne sur site web CMIP6, simulations dépubliées et déclarées sur Errata-Service. Rerun membre 29 historical CM6-1 publié, rerun hist-GHG en cours puis à republier quand disponible
- variables **seaice** mal pondérées : feu vert David pour publication variables P1 (siu, siv, sithick, siconc, sisnthick et sivol) avec message de warning pour siu et siv. Fait pour Slmon CM6-1, à faire encore pour ESM2-1 (Siday et Simon) et Slday CM6-1
- variables Slmon/**sivoln** et **sivols** (siarean et siareas à confirmer) bugguées (surestimation 1%) . Texte descriptif à écrire et insérer sur site web CMIP6 ([Aurore](#))
- Problème 'stratospheric aerosols' : instruit en réunion CMIP6-science 12/06. Texte explicatif inséré sur site web CMIP6
- Variables **nep** et **rh** à ne pas produire et publier pour CM6-1 et CM6-1-HR : texte en ligne sur site web CMIP6, datasets déjà publiés à déclarer sur l'Errata-service puis à dépublier ([les variables n'ont pas de sens pour les modèles CM6](#))
- Erreur d'unité variable **burntFractionAll** (Lmon) : texte en ligne sur site web CMIP6 ([pas de dépublication comme la correction est facile à faire et que les utilisateurs ont corrigé l'erreur eux-mêmes](#))
- [Un utilisateur a remonté un problème avec l'évaporation potentielle. Peut-être une mauvaise lecture ou une mauvaise interprétation liée au modèle de sol. En attente d'une description du problème par l'utilisateur.](#)
- (stand by) attribut **parent_variant_label** erroné pour certaines simulations ('variant_label' au lieu de 'parent_variant_label'). A faire encore : liste des simulations concernées ([en cours, Gaëlle](#)), texte pour site web (rubrique 'known issues')
- (stand by) Instruction à terminer suite à un retour utilisateur sur l'absence des champs lat_bnds / lon_bnds dans nos fichiers (attributs a priori présents pour grille native seulement, à confirmer encore)
- Rappel : tableau (usage interne) recensement erreurs sous /cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Production/Issues (InfoIssuesInterne.odt)

Publications CMIP6 ESGF

- Publication CMIP6 réalisée à ce jour sur 30 nœuds (38 institutions, 86 modèles). Détails sur l'ESGF et sur la page de monitoring :
https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- Nouvelle version de la pile logicielle ESGF (v4.0.4) : installée sur le nœud de test esg2 du CNRM et validée (publication ok pour tous projets). A déployer sur nœud de prod esg1 à l'issue des dépublications en cours....
- Informations CDNOT (CMIP Datanode Operations Team) : poursuite réunion mensuelle, composant de publication porté sur Python3 à tester prochainement au CNRM, discussions sur stratégie d'affectation numéro de version pour les cas de dépublication/republication (simulations dépubliées encore non déclarées sur Errata-Service par de nombreux centres, ok côté CNRM-CERFACS) et en cas de complément de la période temporelle d'une simulation donnée
- URL générique permettant la recherche du DOI/citation pour un jeu CMIP6 donné :
http://bit.ly/CMIP6_Citation_Search qui renvoie vers
<https://cera-www.dkrz.de/ords/f?p=127:2>

Publication ESGF au CNRM

■ Bilan publication CNRM au 01/07 :

- détails sous /cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Publication/SuiviPubli/SuiviPubliGlobal.ods
- **295** bons à publier émis au CNRM
- **288** simulations publiées soit **~94 %** du nombre total de simulations à publier sur l'exercice (**305** prévues à ce jour....). 15 MIPs couverts côté CNRM-CERFACS, volume données CMIP6 d'ores et déjà publiées **520 To**
- depuis la dernière réunion Tech du 05/06 :
 - * publication membres r5 à r10 simulation CNRM-ESM2-1_1pctCO2
 - * publication simulation C4MIP CNRM-ESM2-1_esm-ssp585_r1i1p1f3
 - * publication rerun CNRM-CM6-1_historical_r29i1p1f2

■ A venir côté CNRM :

* dépublications :

- variables erronées (cf diapo Errata)

* publications :

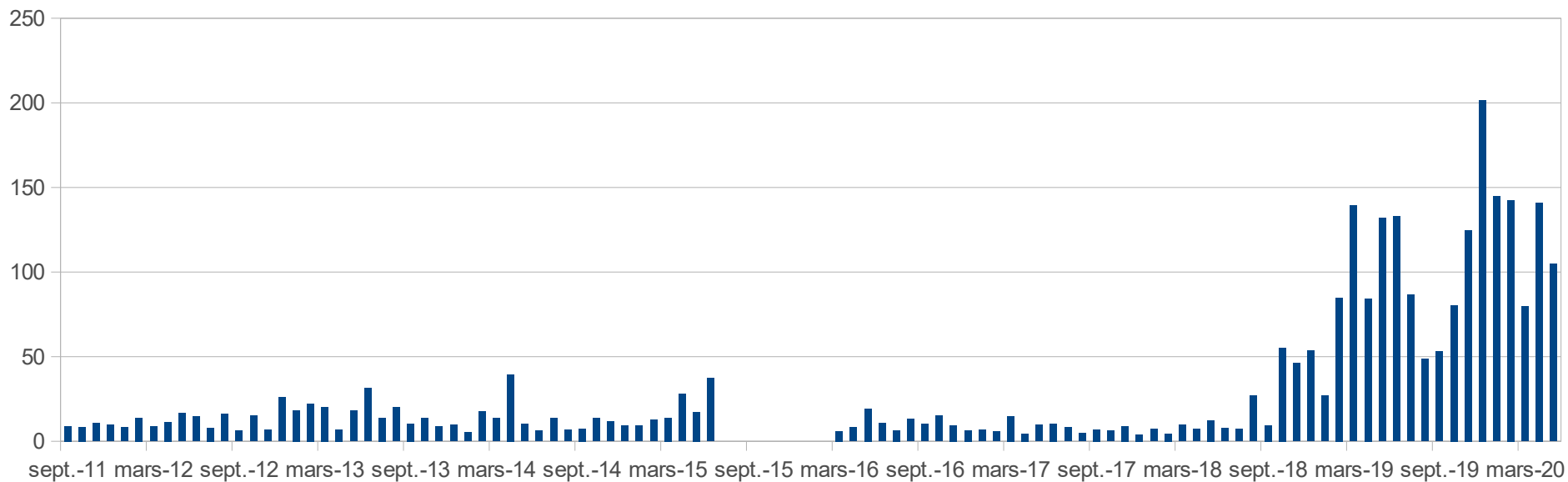
- rerun CNRM-CM6-1_hist-GHG_r3i1p1f2 quand disponible
- variables P1 tables Slday et Simon pour ESM2-1 (et Slday pour CM6-1)
- complément DECK piControl LR 2350-4850
- variables complémentaires 'Cordex': bons à publier reçus pour ESM2-1, ssp126 et ssp585 membres r2 et r3
- extension année 2009 simulation omip1 CNRM-CM6-1
- extension '2300' 3 membres r1 ScenarioMIP ESM2-1
- ESM2-1_esm-hist r1i1p1f3 (modèle en émission)
- simulations ESM2-1 CDRMIP (2) et LUMIP (membres r2 et r3 hist-noLu)
- puis ordre chronologique d'arrivée des futurs bons à publier....

Publications ESGF au Cerfac

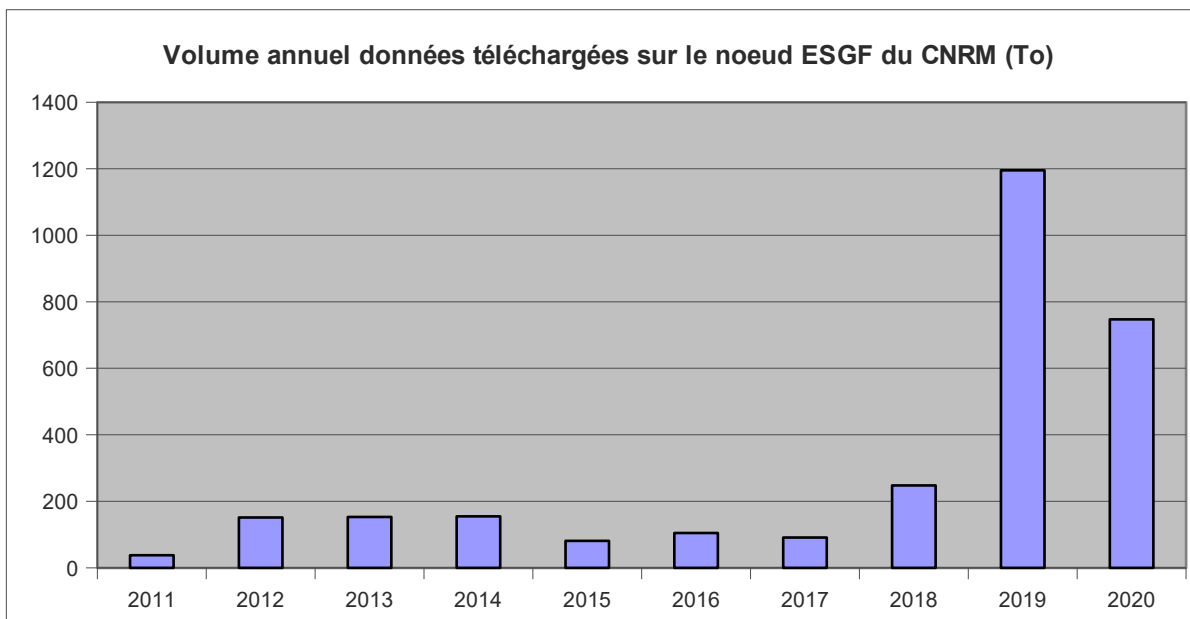
- Publication en cours des membres additionnels réalisés dans le cadre HighResMIP (les premiers membres sont déjà publiés car obligatoires dans le cadre de CMIP6).
- Environ 80 % des données sont dispo sur l'ESGF.

Trafic noeud ESGF au CNRM (volume téléchargé)

Volume données téléchargées noeud ESG du CNRM (To)



Volume annuel données téléchargées sur le noeud ESGF du CNRM (To)



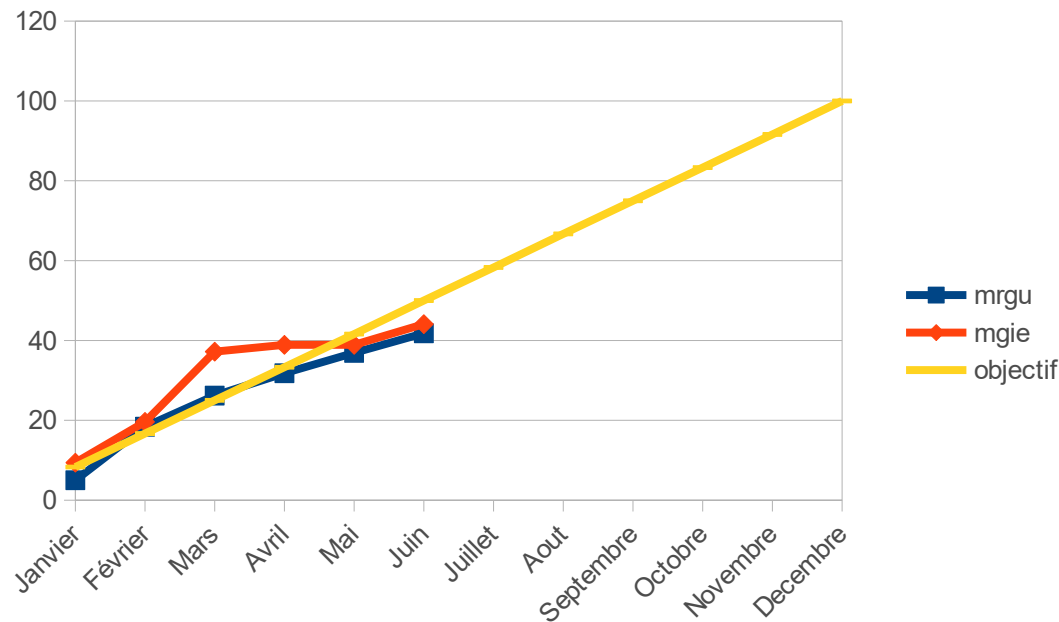
(01/07)

- 'further info URL' mis en œuvre pour modèles décrits sur ESDOC, regroupe toutes les informations sur une expérience donnée : liens esgf search, errata, citation Institute, citation dataset. Par exemple :
<https://explore.es-doc.org/cmip6/further-info?target=CMIP6.CNRM-CERFACS.CNRM-ESM2-1.historical.none.r1i1p1f1>
- ESDOC alimenté par 14 centres seulement (au 30/06), voir <https://explore.es-doc.org/cmip6/models>
- Côté CNRM :
 - * ESM2-1 : seaice, toplevel et citation à documenter encore (nouvelle version du fichier ocean disponible à pousser sur ESDOC). Depuis la dernière réunion : échanges avec l'équipe ESDOC pour que nos inputs déposés sur le GitHub du projet soient enfin correctement pris en compte sur le portail : ok le 08/06 mais de nouveau incomplet le 30/06 pour les realm atmos et land....(ok pour les autres realm)
 - * CM6-1 : fichiers initialisés fournis aux responsables de chaque realm. Retours d'ores et déjà reçus pour coupling, atmos et ocean (à pousser sur ESDOC). Autres realm à suivre
 - * CM6-1-HR : initialisation fichiers à partir des fichiers CM6-1 finaux
- Point à voir sur la façon de remplir les fichiers pour le HR comme il y a deux versions un peu différentes utilisées (une pour HighResMIP et une pour les simulations CM6-HR) avec en particulier un nombre de catégories de glace différent (Aurore et Marie-Pierre).

Ressources calcul, stockage, portage

Ressources calcul et stockage

- Renouvellement calculateurs : fin VSR Belenos a priori le 02/07 (03/07, pas de nouvelles de la fin de la VSR), début VSR Taranis le 28/09 (fin VSR au plus tard le 20/11). Arrêt Beaufix/Prolix prévu pour l'instant mi-novembre 2020
- Nouvelles règles : quotas (50 Go) par user et par groupe unix sur /home. Fichiers les plus anciens non utilisés sur /scratch seront supprimés (tous les 2-3 mois), données \$TMPDIR supprimées à J+1, données \$WORKDIR supprimées si non accédées depuis 30 j. Demande /scratch GMGEC sur Belenos envoyée à la DSI le 01/07
- Consommation mgie 2020 sur Beaufix : 2.2 Mh sur queue classique (quota 5 Mh), 9.3 Mh sur queue Libre. Rappel : quota 2020 prévu pour OMIP-HR et quelques simulations LR. OMIP-HR (~3Mh) pas encore lancé (lié au passage de l'oper sur Beaufix) donc il ne pourra pas être terminé sur Beaufix (changement de machine entre deux cycles pour limiter les impacts). Donc ce n'est pas grave si ce quota est entamé.
- Réajustement quotas 2020 demandé à la DSI (CPCC le 26/06) pour prise en compte décalage calendrier Calcul2020 (7 Mh sur Beaufix, 600 000 h sur Prolix, 700 000 h sur Belenos256) : quota accordé.



(au 30/06)

Portage

- Difficultés à avoir une visibilité sur la manière de s'organiser pour le portage (modifications de l'environnement non annoncées) → La compilation telle qu'elle a été faite par Pastel ne fonctionne plus en l'état (faite en début de VSR) en lien avec un changement d'environnement.
- Certains comptes viennent juste d'être créés
- Tout n'aura pas pu être testé mais on va prendre le temps pour le faire et que tout fonctionne.

- État des versions « officielles » (hors versions compilées côté Pastel) :
 - Oasis est disponible (Aurore)
 - XIOS est disponible (Gaëlle)
 - NEMO compile (Aurore)
 - Arpège à tester
- Réaliser des points réguliers par mail sur le portage, pour faire des itérations sur les versions à utiliser.

Outils



Eclis

- Documentation :
 - En cours de rédaction sous LaTeX (mais pas prêt...) : <https://fr.overleaf.com/read/dzngchnvckbr>
- Base de tests :
 - En cours de mise en place
- Modifications en cours (V6.45) :
 - Intégration des dernières corrections pour tourner sur Belenos
 - Ajout de la récupération des forçages sur Hendrix s'ils ne sont pas disponibles sur le scratch
 - Modification du post-traitement (utilisation d'un système de plugins pour le rendre plus générique, création d'un fichier de configuration particulier pour faciliter la relance) :
 - Philosophie des modifications prévues sur le post-traitement :
 - Passer d'une terminologie CMIP6 à une terminologie plus générale permettant d'utiliser le post-traitement dans d'autres contextes que CMIP6.
 - En pratique, utilisation d'un système de plugins comme c'est le cas pour les run.
 - La documentation sur le post-traitement s'adaptera à cette nouvelle version du post-traitement.
- Modifications à court terme :
 - Modification du répertoire de relance et de l'archivage des fichiers intermédiaires pour s'adapter aux changements d'architecture
 - Vérifier les runs sur aneto
- Suivi des soucis d'accès à Hendrix au moment de l'install pour voir s'il faut modifier quelque chose.

- Prochaines versions :
 - V1.2.13 (en cours de test à l'IPSL, avec la nouvelle version de la C-ESM-EP) :
 - Mise à jour de la documentation et des exemples
 - Correction de bogues
 - Modification de l'appel à cdo
 - Amélioration des figures
 - Possibilité d'appel des opérateurs sans fichiers d'entrée
 - Modification du cache
 - Les modifications apportées pour la gestion des gros fichiers netcdf sont les suivantes : modifications du cache et des sélections réalisées (Stéphane), modification du script de sélection de données (changement de l'ordre des opérations et création de fichiers intermédiaire). Vérifications de la diminution du temps de calcul à faire. Pour la suite, on envisage de remplacer le script de sélection utilisant cdo par un script utilisant python (dask et xarray) pour aller plus vite mais les tests sont toujours en cours.
 - V2.0 (pendant l'été ?) :
 - Compatibilité python 3
 - Écriture de tests unitaires

Autres outils

- dr2xml :
 - Documentation :
 - En cours : architecture prête, il reste à remplir... (disponible dans le code)
 - Développement court terme :
 - Dans la sélection des attributs, séparer la sélection au niveau des fichiers de la sélection au niveau des champs
- Get :
 - Correction en cours (« -n » ne fonctionne pas)
 - Réflexion sur l'outil en cours
- XIOS :
 - Problèmes de compilation sur mageia7 :
 - Ok version 1852 en mode debug
 - Ok version 1442 + modifs __1442++__ en mode prod (David)
 - Pas ok pour les autres
 - Compilation ok sur Belenos en 1442 (avec les modifs __1442++_) et en 1637

Suite...

- Prochaine réunion ? (début septembre)