

Réunion Tech

- production CMIP6
 - publications CMIP6
 - ressources et site Web
 - Xios / Xias
 - Eclis
 - dr2xml
 - CliMAF
-

09 mai 2019

Production CMIP6

- vt100 :
 - Variable demandée par la DR :
northward_heat_flux_in_air_due_to_eddy_advection (W.m-2)
 - Variable publiée : $v \cdot T$ à 100 hPa (K.m.s-1)
 - Dépublication et errata [à suivre](#)
 - A priori, une modification de la DR sera faite pour cette variable (sortie de la variable $v' \cdot T'$)
- ialb : mal masquée (variable maison, ne doit pas être publiée)
- Fichiers incorrects présents sur /scratch (non accessibles via Eclis) :
 - PGD.tl127.AESM.fa
 - PGD.tl127.AOESMemis.fa
 - PGD.tl127.AOESM.fa

Qu'en fait-on ?

Remplacement par des fichiers corrects ?

Suppression des fichiers présents sur le scratch et modification des param de référence (les PGD seront pris dans le répertoire FORLUC pour l'initialisation).

Production CMIP6

- Concernant les variables de glace non pondérées, il n'y a rien de particulier, quelques errata à faire (liste des variables concernées à figer encore). Problèmes de cohérence dans les demandes de la DR. Les données sorties sont cohérentes et permettent de fermer les bilans d'énergie (cf. liste de variables seaice de la slide Errata Service & Known Issues)
- Quelques variables de glace sont boguées (liste des variables à faire).
- Présence des champs fixes à vérifier pour les différentes simulations (LS3mip par exemple). Si absence, les sorties ne sont pas utilisées par certains groupes (ESMValTools notamment).
- Pour le HR, il manque la définition des bassins. Roland regarde s'il est possible de récupérer un fichier.
- Pour les simulations Offline, problème avec les run quand on demande via dr2xml à XIOS de changer le contenu du fichier areacella. A voir au passage à une version plus récente de XIOS, post-traitement ajouté à la version V6.42 d'Eclis.

Publication ESGF : informations générales

- Publication CMIP6 réalisée à ce jour sur 13 nœuds. Détails sur l'ESGF et sur la page de monitoring : https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- Réplication CMIP6 : 110 To d'ores et déjà répliqués à l'IPSL, 100 To au CEDA, 500 To au DKRZ, 35 To à NCI, 580 To au LLNL
- Pas de réunion du groupe CDNOT depuis le 10/04
- Livraison d'une nouvelle version majeure de la pile logicielle ESGF (upgrade à réaliser avec un nouveau mode opératoire). Montée de niveau prévue à l'IPSL cette semaine ou semaine 20. A réaliser ensuite au CNRM

Data Node	
<input type="checkbox"/>	aims3.llnl.gov (29)
<input type="checkbox"/>	cmip.bcc.cma.cn (2467)
<input type="checkbox"/>	crd-esgf-drc.ec.gc.ca (24716)
<input type="checkbox"/>	dist.nmlab.snu.ac.kr (1)
<input type="checkbox"/>	esg.lasg.ac.cn (1)
<input type="checkbox"/>	esg1.umr-cnrm.fr (31049)
<input type="checkbox"/>	esgdata.gfdl.noaa.gov (547)
<input type="checkbox"/>	esgf-data.ucar.edu (30600)
<input type="checkbox"/>	esgf-data2.diasjp.net (5633)
<input type="checkbox"/>	esgf-data3.ceda.ac.uk (2257)
<input type="checkbox"/>	esgf.bsc.es (336)
<input type="checkbox"/>	esgf.nccs.nasa.gov (10041)
<input type="checkbox"/>	vesg.ipsl.upmc.fr (199232)

Publication ESGF au CNRM

■ Bilan publication au 09/05 :

- 202 bons à publier émis (+ 3 extensions 2100-2300)

- 100 simulations publiées (~200 To). Depuis la dernière réunion :

* fin DAMIP (10 membres hist-nat CM6-1), ESM2-1 1pctCO2 (membre r4), membres r1 ScenarioMIP ESM2-1 (s119, ssp126, ssp245, ssp370, ssp434, ssp460, ssp585), autres membres ScenarioMIP CM6-1 (membres r2 à r6 ssp126)

* rappel fichier suivi publication (utilisé pour tableau de bord site web CMIP6) : fichier SuiviPubliGlobal.ods sous /cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Publication/SuiviPubli

■ A venir (par ordre de priorité décroissant, **à valider**) :

* autres membres ScenarioMIP CM6-1 : membres r2 à r6 ssp245, ssp370, ssp585

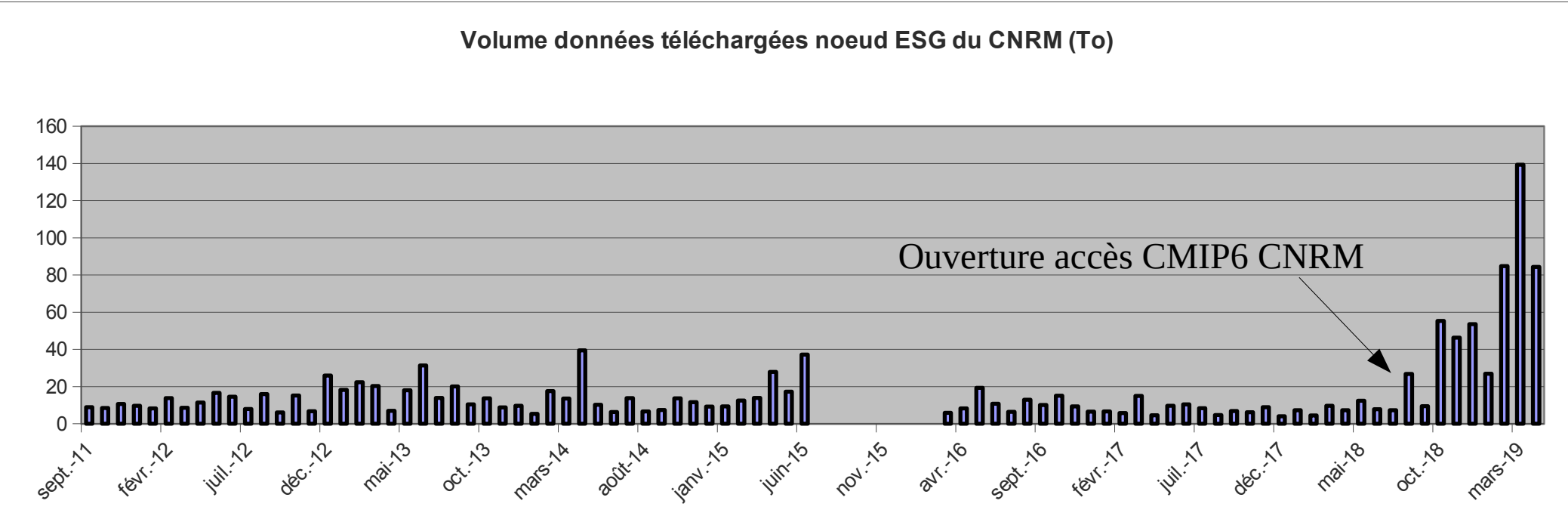
* simulations omip1 CM6-1 et ESM2-1

* autres membres ScenarioMIP ESM2-1

* simulations Crescendo (jalon été)

* puis ordre chronologique d'arrivée des bons à publier....

Trafic noeud ESGF au CNRM



Errata Service, known issues

- Errata déclarés sur l'Errata Service pour :
 - * ozone **o3** (8 datasets). Dépublication puis republication nouveaux champs
 - * albiscpp/pctisccp (82 datasets). Dépublication datasets concernés
- variables **seaice** mal pondérées (cf mail Stéphane 23/04 : siv, siu, sisnconc, sisnmass, sisnhc, sistrxdtop, sipr, sistrydtop, sistryubot, sistrxubot, sidragbot, siflwdrain, siflwbob, siflswutop) : décision de dépublication (738 datasets déjà publiés) / publication à suivre ? [Pas de dépublication](#).
- sous-échantillonnage champs 3D à 3h pour moyenne temporelle : texte site web à valider avant mise en ligne
- **vt100** (27 datasets publiés) : errata à déclarer sur Errata Service, datasets à dépublier et texte à site web à valider avant mise en ligne
- suivi 'known issues' CMIP6 :
 - * fichier 'InfolssuesInterne.odt' récapitulatif des erreurs à usage interne ss /cnrm/est/USERS/franchisteguy/CMIP6/CNRM/Production/Issues
 - * mise en place page dédiée 'Data informations and known issues' sur site web CMIP6, <http://www.umr-cnrm.fr/cmip6/spip.php?article23> . Remarques et avis bienvenus.

■ Rappel approche proposée :

* renseignement des 9 fichiers realm/coupling/toplevel pour la configuration la plus complète (esm2-1) puis déclinaison pour les 2 autres configurations cm6-1 et cm6-1-hr par élagage/modifications. realm ocean : mise en commun éventuelle/réutilisation informations déjà renseignées par IPSL (cf GitHub)

* coordination de l'alimentation de chacun des 9 fichiers par un référent puis dépôt final sur GitHub par Marie-Pierre et/ou Laurent

* consignes systématiques pour le remplissage fournies sur le 1^{er} onglet de chaque fichier, champs obligatoires marqués d'un *

■ Point d'étape :

* fichiers à renseigner (esm2-1) envoyés à chaque référent le 12/04

* réponses reçues pour les 2 realm aerosol et atmoschem

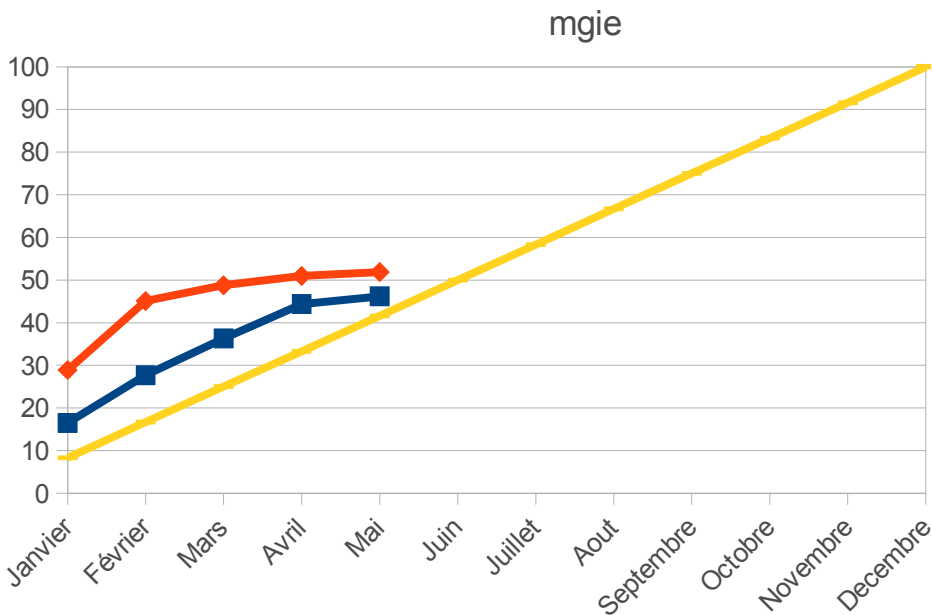
* à suivre : récolte puis vérification des fichiers, dépôt sur GitHub ESDOC et envoi à l'équipe ESDOC de vos remarques éventuelles sur les fichiers à renseigner

■ **Se renseigner auprès d'Es-Doc pour savoir si l'alimentation de la base de donnée CIM pour accès aux services Es-Doc search/view est prévue (Marie-Pierre/Laurent)**

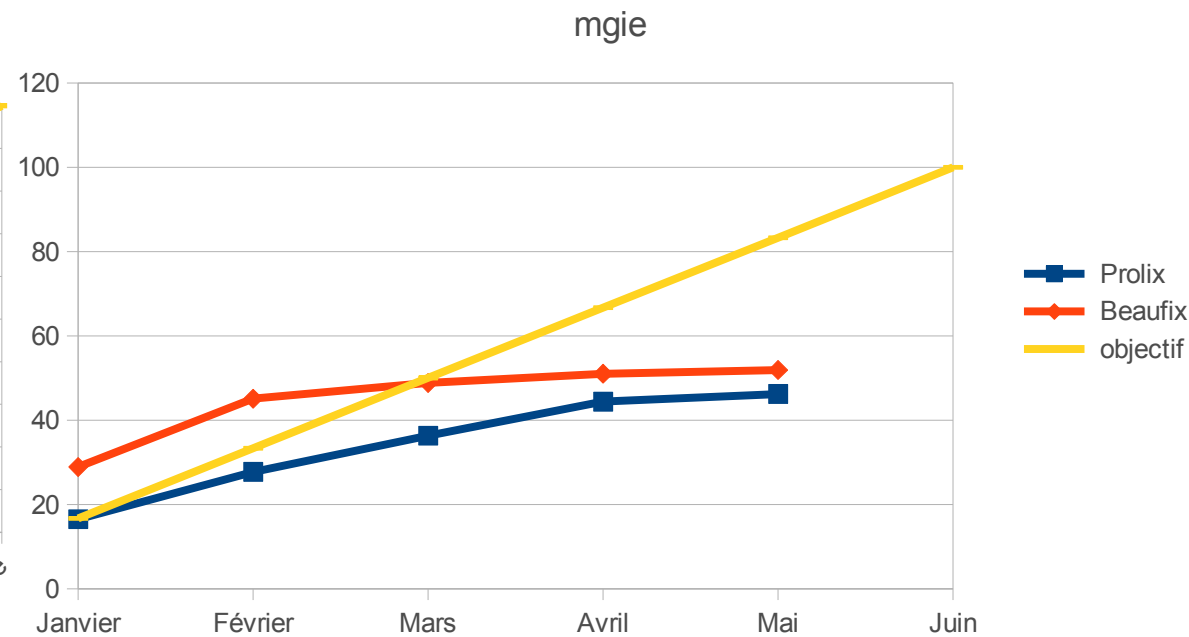
Ressources calcul et stockage, site web

- Demande DSI réactualisation ressources 2019 (DL 23/05).Maj besoins calcul HR(?)
- Consommation calcul mgie (au 09/05) : il reste 13.6 Mh sur Beaufix et 7.5 Mh sur Prolix

Utilisation quota sur l'année



Utilisation quota sur le 1^{er} semestre 2019



Ressources calcul et stockage, site web

- Réalisation de scénarii, avec et sans rallonge de temps de calcul pour voir ce qu'il est possible de faire.
 - Pour la simulation PIControl HR, il n'y a pas trop d'augmentation du temps de calcul quand on augmente le nombre de diagnostics.
 - Question de sortir les champs à utiliser pour forcer des modèles régionaux ou des simulations Cordex, mais problème au niveau du stockage. A voir maintenant comme on ne refera pas tourner. Actuellement, forçage d'Aladin à partir des membres ESM puis utilisation par Arome. Intérêt de forcer Aladin par le HR avant de forcer Arome ? S'il faut les sortir, quels scénarii si plusieurs sont faits ? A discuter avec Samuel.
 - Utilisation de la table avec les sorties en niveaux pression ? Pas assez de niveau pression pour Cordex et pas les bonnes périodes.
 - Pour le HR, on ne fait qu'un membre. Au niveau calcul, il vaut mieux prévoir plus de scénarii que plusieurs membres en cas de rallonge. Question des scénarii à faire. Qu'est-ce qui est fait dans les autres centres ?
 - Dans Primavera (HighResMIP), 3 membres sont faits pour control+hist, mais sur 100 ans (1950-2050). Le plus intéressant serait d'augmenter ce nombre de membres (jusqu'à 10 ?).

Refonte Xias – symétrisation Arpege/Surfex

- Basé sur la 6.4.1
 - après la simplification d'abandon d'auto-déclaration des champs
 - bcp de sources impactés (mais bcp par des changements de noms de routines, ou abandon d'option)
- Activation : Arpege par variable d'env. / Surfex par namelist
- Contexte Xios partagé entre Arpege et Surfex
- Quelques routines partagées : xias_*
 - xias_send_block, xias_setups(date, tstep, nblock), xias_set_axis ...
 - aujourd'hui hébergées dans le code Surfex,
 - à terme ailleurs (e.g. sous-rép de Surfex ou avec le code Xios)
- Utilisation par Arpege d'une routine Surfex (latitudes_gauss) ; à rectifier
- Ajouts dans Arpege : init_xios (master), suxiosb (avant sucospsim), suxiosf
 - Dont : déclaration de la grille de Gauss,
- Légères modifs dans arpsfx.xml :
 - Les domaines Surfex (SEA, NATURE, WATER) héritent de FULL (Surfex déclare seulement les masques des tiles)
 - La dimension des axes est initialisée (et cohérente avec DR CMIP6 quand applicable)
- Tests:
 - Fonctionnels : Arpège avec Xios / Surfex avec Xios / les deux.
 - Reproductibilité OK (diags et NORMS)
- Encore à traiter: Offline ([en cours](#)), Aladin

Eclis (V6.42)

- Problème avec nctime pour des fichiers après l'an 4000. Exemple à envoyer à G. Levavasseur.
- Développements intégrés:
 - Ajout de post-traitements pour certains fichiers de l'ESM
 - Modification des archivages dans assemble_and_QC (pas de ssh direct sur les nœuds de transfert depuis les nœuds de calcul)
 - Corrections de bogues dans cm_files
- En cours :
 - Post-traitement pour modifier les fichiers areacella (plantage Offline quand on demande la variable en mode data request)
 - Augmentation du realization_index ([report à la prochaine version](#))
 - Passage à nctime via Eclis d'un fichier contenant les pattern des noms de fichiers à ne pas vérifier (modification de do_prepare associée)
 - [Récupérer les modifications Python 3 et PEP8 dans cette version pour pouvoir faire tourner des simulations en version V6.* après le changement de calculateur.](#)
 - [Faire en sorte que si on est en mode développement, les méta-données des fichiers netCDF soient plus réduits \(nécessite aussi une modif dr2xml\).](#)
- Mise à disposition : mi-mai

Eclis (V7.0)

- V7_beta :
 - Version utilisée pour le HR
 - Ajout de post-traitements pour quelques variables pondérées
 - Disponible sur Beaufix
- V7 :
 - Discussion sur la suppression de LDR_EXPID

Modification du comportement de LDR_EXPID pour la v $\bar{7}$ d'ECLIS ?

- Etat des lieux
- Proposition

Fonctionnement actuel (rappel du manuel simulanceur)

- **Postulat de départ:** les fichiers produits sont nommés avec EXPID
 - Sauf si expid_in_filename est mis dans experiment_settings.py
- EXPID dépend de LDR_EXPID
 - **0 : EXPID = suffixe du param_EXPID**
 - Pour les expé de test, mais alors ECLIS impose un n° de réalisation > 90
 - Pour les expés de rerun ou prod hors CMIP6 avec r>90 mais en précisant le expid_in_filename dans experiment_settings.py → compatible avec la production
 - **1 : EXPID =<nom_model>_<experiment>_<ripf>**
 - **Production** Ex: CNRM-CM6-1_piControl_r1i1p1f2
 - **2 : concaténation de 1 et 2 :**
 - Utile si on veut que le rerun garde le nom de production mais avec un suffixe ex _rr1 → compatible avec la production

Enfinement tous les LDR_EXPID sont compatibles avec la production!

Modification du comportement de LDR_EPID pour la v7 d'ECLIS

Pb de base : intrication nom de simu et nom donné par DR2XML aux fichiers de sortie → complexifie bcp le comportement d'ECLIS

- ne permet pas l'usage de EXPID dans le fichier param
- on peut finalement faire de la prod avec n'importe quel LDR_EXPID, intérêt pas évident?

Doit-on revenir à param_EXPID donne le nom de l'expé EXPID = EXPID unique?

La raison d'être de LDR_EXPID était de vouloir distinguer le mode "production CMIP6" et le mode "standard"

Qu'implique le fait qu'une simulation soit en mode "production"?

- le nom de la simulation (ie EXPID) colle au nom CMIP6

 - model_experimentid_r*i*p*f* → CNRM-CM6-1_piControl_r1i1p1f2

 - ✓ Vu le nombre de reruns fait qui n'ont pas suivi ce format → pas nécessaire

- les sorties nommées selon standard CMIP6

 - ✓ c'est vrai qqsoit le LDR_EXPID → c'est à gérer dans experiment_setting.py

- les restarts nommés selon standard CMIP6

 - ✓ Vu le nombre de reruns faits qui n'ont pas suivi ce format → pas nécessaire

- simulation entre dans la liste de production (fichier html de Stéphane)

 - ✓ Non en fait c'est LDR et le numéro de membre, on pourrait avoir une clef ECLIS pour identifier le run de prod?

Proposition

- Supprime LDR_EXPID et on met un flag qui dit si la simulation est de la prod ou pas, voire même le nom de la prod pour permettre d'avoir un suivi général de la prod d'un projet donné.

- On utilise la propriété expid_in_filename de experiment_settings.py pour gérer le format des sorties, si on ne veut pas utiliser le expid qui est dans le experiment_settings.py

NB: J'ai expérimenté de changer expid_in_filename pour des expés hors prod et ca me permettait de conserver LDR_EXPID=0, ie d'avoir un nom de simu "court". Je ne vois pas l'intérêt a priori de ne pas utiliser expid. Au contraire en utiliser un autre demande une gestion fine entre avec get et esg-drs, etc... Finalement, autant se rappeler que DR2XML a travaillé comme si c'était "expid", même si notre simu n'a rien à voir. Pour d'autres projets, de toute façon il faudra un flag pour DR2XML et les formats de fichiers risquent d'avoir leur propre spécificité. → en bref, empêcher de simplement utiliser expid n'était pas vraiment utile, ni souhaitable.

- On laisse les restarts suivre la logique ECLIS

Eclis (V7.0)

- V7 :
 - Discussion sur la suppression de LDR_EXPID
 - ▶ A priori, on peut s'en passer
 - ▶ Faire en sorte qu'une variable du param modifie expid_in_filename dans le fichier experiment_settings
 - ▶ Ajout d'un flag pour indiquer s'il s'agit de production et le projet associé (pour utilisation hors CMIP6). Pourra gérer les post-traitements lancés.
 - ▶ Discuter lors de la prochaine réunion des autres paramètres qui ne sont pas utiles pour les simulations non CMIP6 ou dont les valeurs pourraient poser problème si elles ne sont pas spécifiées (experiment, mip_era, activity_id, end_year...).

Eclis (V7.0)

- V7 (suite) :
 - Ajout de la compatibilité python 3 pour les quelques fichiers concernés
 - Mise à jour des plugin suivants : aladin, nudging, esm (en cours). Ils nécessitent des modifications de binaires. Il s'agit juste d'une préparation et une phase de test (non possible) devra être prévue.
 - Changement de version mtool (en test)
 - Modification d'Eclis relative à la mise à jour d'XIOS à prévoir.
 - Rattrapage des modifications de la V6.42
 - Les modes de fonctionnement avec et sans boucle interne seront utilisable dans cette version. Pour Surfex, pas d'intérêt à passer au nouveau mode de fonctionnement pour l'instant (fichiers très lourds). Par contre, ce mode de fonctionnement pourra être intéressant si on utilise XIOS pour la lecture des fichiers au fil de l'eau.

dr2xml (V2.0)

- Nouveautés :
 - Compatibilité python 3
 - Axe de pression configurable au niveau des settings (shape XY-perso) :
 - Description complète à mettre dans le dictionnaire perso_sdims_description (simulation_settings)
 - Création d'un fichier contenant les pattern des fichiers correspondants aux variables perso et dev
 - Ajout d'un arbre d'appel de dr2xml (pas très lisible mais en progression...)
- A prévoir :
 - Possibilité, en mode développement, d'avoir un nombre réduit de méta-données dans les fichiers netCDF produits par les simulations (entre tout et rien).
 - Pas d'utilisation et de développement de dr2xml pour la prévision saisonnière pour le moment.
- Mise à disposition d'ici la fin mai
- Emplacement sur les calculateurs ? Sous `~rigoudyg/dr2xml`

- Poursuite des bords CliMAF.
- Développement :
 - Compatibilité Python3 et PEP8 bien avancées
 - Récupération en cours des branches à intégrer à la nouvelle version
 - ▶ Modifications pour le Cerfacs
 - ▶ Modifications pour faire tourner C-ESM-EP
 - ▶ Corrections de bogues
- Problème à la lecture des champs fixes ?

Suite...

- Prochaine réunion **mardi 18 juin à 14h30 en salle 151**