

Réunion Tech

- **CliMAF**
 - **C-ESM-EP**
 - **Eclis**
 - **dr2xml**
 - **Retours sur la Data Request**
 - **Production CMIP6**
 - **Publications CMIP6**
 - **Ressources et site Web**
-

18 juin 2019

- Poursuite des baux CliMAF le jeudi avant pause estivale.
- Développements :
 - Préparation de la version V1.2.12 en cours (dernière version avant de figer les développements pour le passage à Python 3) :
 - Modifications pour le Cerfacs
 - Modifications relatives à la C-ESM-EP
 - Corrections de bogues
 - PEP8
 - **Intégration de corrections pour les bogues suivants : problème de lecture des champs fixes, lenteur dans les cas de gros fichiers**
 - Préparation de la V2.0 :
 - Compatibilité Python 3
 - Renommage de certaines fonctions
- Bogues connus :
 - Problème à la lecture des champs fixes
 - Problème dans certains cas avec sellonlat (cdo)
 - Lenteurs dans les cas où les fichiers sont gros

Atlas et C-ESM-EP

- C-ESM-EP :
 - Version V2.0 de la C-ESM-EP en cours de test à l'IPSL par Jérôme (devrait être disponible rapidement)
 - Atlas réalisés par Laure Coquart sur les simulations CMIP6 disponibles sous : <http://cerfacs.fr/giec6/>
 - Attente de l'installation d'une version fonctionnelle récente de la C-ESM-EP mais problème des lancements de jobs sur Aneto (un nombre important de job est lancé posant des problèmes de partage). Vérifier l'utilisation de la queue P8HOST sur le cluster et essayer de limiter le nombre de jobs lancés simultanément.
 - Installer une version de la C-ESM-EP récente et faire les modifications nécessaires dans CliMAF également.
- Atlas :
 - Le développement est toujours en cours.
 - Quelques problèmes soulevés (par exemple le bogue avec sellation)

Eclis : V6

- V6.42 :
 - Version disponible sur les calculateurs et sur le lustre du CNRM depuis le 16 mai
 - Embarque en particulier la compatibilité python 3
- V6.43 :
 - Contenu :
 - ▶ Augmentation du realization_index
 - ▶ Correction du plugin atm_lbc_nudging pour pouvoir commencer des run Aladin en milieu de mois
 - ▶ Corrections de bogues
 - ▶ **Déboguer cm_files (outil de suppression des simulations)**

Eclis : V7

- V7_beta :
 - Version utilisée pour le HR
 - Ajout de post-traitements pour quelques variables pondérées
 - Disponible sur Beaufix
- V7 (branche de développement) :
 - Contenu :
 - ▶ Rattrapages réguliers de la branche V7.0_beta
 - ▶ Rattrapages réguliers du master V6.*
 - ▶ Modifications liées à XIOS/XIAS
 - ▶ Changement de version de mtool
 - ▶ Développements envisagés (diapo suivante)
 - /\ Nécessite des binaires particuliers
 - Tests poussés à faire avant mise à disposition officielle de cette version. Sollicitation des personnes gérant des configurations particulières pour les tests.

Eclis : développements envisagés

- En mode développement, réduire le nombre de méta-données :
 - Fais dans la version V2.0 de dr2xml
 - Pilotage via le lab_and_model_settings (entrée print_variables)
 - Pas de pilotage via eclis pour l'instant (pas d'interfaçage avec eclis sur ce point)
- Correction de bogues sur job_postpros_new et assemble_and_QC_new
- Correction de cm_files (sera fait dans la V6.43).
- Suppression de LDR_EXPID
- Variables modifiées par Eclis dans les settings dr2xml : expid_in_filename, ...
- Ajouter un flag pour indiquer s'il s'agit de simulations de production et pour quel projet
- Liste des variables des settings de dr2xml spécifiques à CMIP6 : experiment, mip_era, activity_id... Gestion dr2xml et non eclis.
- Problème avec nctime pour des fichiers après l'an 4000. Exemple à envoyer à G. Levasseur (récupérer un fichier du control).
- Interfaçage de xiofilechecker (outil permettant de vérifier la cohérence entre ce qui est demandé dans les fichiers xml et les sorties effectivement produites).
- Interfaçage d'un débogueur

dr2xml : versions disponibles

- V1_CMIP6_dev (sur les calculateurs) :
 - Version parallèle qui embarque toutes les nouveautés des branches récentes sans les modifications pouvant impacter le contenu ou les métadonnées des fichiers.
- V2.0 disponible depuis le 3 juin :
 - Compatibilité python 3 (mais correction à apporter pour gérer la présence d'accents dans les fichiers de settings)
 - Axe de pression configurable dans les settings
 - Fichiers contenant les pattern des variables dev et perso
 - Possibilité de spécifier les méta-données à ajouter
- Bogues connus (développements en cours) :
 - Mauvaise gestion des accents dans les fichiers de settings (lié à la compatibilité python 2 / python 3)
 - Vérifier pourquoi il manque l'attribut global `branch_method` pour certaines simulations (CNRM-CM6-1_omip1_r1i1p1f2 par exemple)
- Emplacement sur les calculateurs : sous `~rigoudyg/dr2xml`

dr2xml : développements

- Développements en cours :
 - Adaptation de dr2xml pour Cordex (tests en cours) :
 - ▶ Modifications finalisées mais devant être intégrées dans le master
 - ▶ Il faudrait ajouter un objet de type char des les fichiers NetCDF via XIOS. A priori, quelque chose de similaire a été fait pour le modèle océanique.
 - Écriture des fichiers xml via un module spécifique en gardant la reproductibilité (en particulier dans l'ordre des attributs).
 - Gestion des accents

Retours sur la Data Request CMIP6

- Demande de Martin Juckes.
- Préparation du retour :
 - Brainstorming CNRM/Cerfacs via le document partagé :
<https://docs.google.com/document/d/18UHMgzumX5a6ZIWM SqRXoTntwaJohbfwb1AxWfGANu0/edit>
 - Questionnaire plus ciblé CNRM/Cerfacs
 - Synthèse CNRM/Cerfacs
 - Synthèse IPSL/CNRM/Cerfacs
- Trois thématiques :
 - Contenu de la Data Request
 - Utilisation de la Data Request et des interfaces
 - Pérennité et adaptation de la structure à d'autres projets

Retours sur la Data Request CMIP6

- **Contenu de la Data Request : définition et cohérence des sorties demandées**
 - **Variabilité/Cohérence de la spécification (inter-membres, inter-expériences, inter-MIPs, etc)**
 - **Quantité de variables demandées**
 - **Pertinence des diagnostics demandés**
 - **Manques éventuels**
- **Remarques :**
 - Hiérarchisation des données nécessaires avec des arbitrages, des homogénéisations et des retours vers les MIP sur les demandes :
 - Variables essentielles à produire sur toutes les simulations
 - Quelques variables supplémentaires selon les MIP/expériences
 - Supprimer les différenciations inter-membres et inter-expériences (dans un même MIP)
 - Rationaliser les périodes où des diagnostics supplémentaires sont sortis.
 - Uniformiser et expliciter les noms des tables et des variables ainsi que les inclusions.
 - Cycles diurne et climatologies ne devraient pas être dans la DR.

Retours sur la Data Request CMIP6

- **Utilisation de la Data Request et des interfaces :**
 - **Mise à disposition des versions successives (fréquence, délai)**
 - **Visualisation du contenu de la Data Request (interface html, générateur de fichier excel, etc)**
 - **Mécanisme d'interrogation de la DR (python)**
 - **Correspondance/cohérence avec les standards (CMIP6_CV, CMOR, CF)**
 - **Recherche d'erreurs/d'incohérences**
 - **Vérification que les sorties demandées par les groupes sont bien produites**
 - **Impact de la DR sur les tables PrePARE maison utilisées pour la publication**
- **Remarques :**
 - **Mise à disposition des versions successives à améliorer (trop tardives, trop nombreuses)**
 - **Outils de visualisation incomplets : quelles données sont produites pour chaque simulation à résolution donnée ?**

Retours sur la Data Request CMIP6

- **Pérennité, adaptation de la structure à d'autres projets :**
 - **Dépendance du workflow (dr2xml) à la structure de la DR**
 - **Arguments pour/contre une DR-CMIP7 (point de vue structure, pas contenu)**
- **Remarques :**
 - **Bonne idée sur le principe mais mise en œuvre de la Data Request à améliorer**

Retours sur la Data Request CMIP6

- La Data Request a été utilisée via son interface Python via dr2xml (IPSL/CNRM/Cerfacs) uniquement. Les autres groupes ont utilisé CMOR (et certainement le générateur de documents excel pour savoir quelles variables produire).
- Dans dr2xml, une partie exploite la Data Request, une autre permet d'ajouter des tables maison. Possibilité d'ajouter des tables ou des variables à la Data Request officielle ?
- Vu l'utilisation faite de la Data Request, pourquoi est-elle si compliquée ?
- La structure de la DR est intéressante mais son contenu doit être rationalisé.

Production CMIP6

- vt100 :
 - Variable demandée par la DR :
northward_heat_flux_in_air_due_to_eddy_advection (W.m-2)
 - Variable publiée : v^*T à 100 hPa (K.m.s-1)
 - Dépublication en cours et errata déclaré
 - A priori, une modification de la DR sera faite pour cette variable (sortie de la variable v^*T')
- Liste des variables de glace non pondérées et éventuellement boguées en attente d'instruction (variables des tables Simon et Slday ne sont plus publiées dans l'attente de cette instruction).
- Présence des champs fixes pour les différentes simulations ?
Pas d'informations supplémentaires. Faire la liste des champs fixes et la liste des simulations pour lesquelles des champs fixes ont été publiés (et lesquels). On pourra ensuite régénérer les champs fixes pour les différentes simulations.
- Définition des bassins pour le HR : en attente

Production CMIP6

- Problème avec la simulation piClim-control appartenant à deux MIP : seules les variables d'un des deux MIP ont été produites. Vérifier que les autres simulations à faire pour deux MIP ont bien sorti les bonnes variables.
Les simulations sortant des variables pour deux MIP sont les suivantes :
 - **land-hist**
 - **piClim-aer** (ok pour AerChemMIP car identique à piClim-NTCF)
 - **piClim-control** (ok RFMIP, manque variables AerChemMIP)
 - **ssp370** (ok pour AerChemMIP)
 - **dcppC-forecast-addPinatubo**

Publication ESGF : informations générales

- Publication CMIP6 réalisée à ce jour sur 15 nœuds. Détails sur l'ESGF et sur la page de monitoring : https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings/
- Nouvelle version majeure de la pile logicielle ESGF disponible (upgrade à réaliser selon un nouveau mode d'installation). Montée de niveau toujours en stand-by à l'IPSL et à réaliser ensuite au CNRM
- Réplication CMIP6 :
 - * augmentation des volumes de réplication au sein de la fédération ESGF (110 To d'ores et déjà répliqués à l'IPSL, 150 To au CEDA, 600 To au DKRZ, 35 To à NCI, 580 To au LLNL)
 - * courrier WIP au groupes de modélisation concernant la liste des variables prioritaires à répliquer selon auteurs WG1 rapport IPCC (liste de variables par chapitres établie fin avril) : cf document :

<https://docs.google.com/spreadsheets/d/1H2qXofwjOpCospDcj9LRokpz0HGysRcDK2HmNSJoqyk/edit#gid=0>

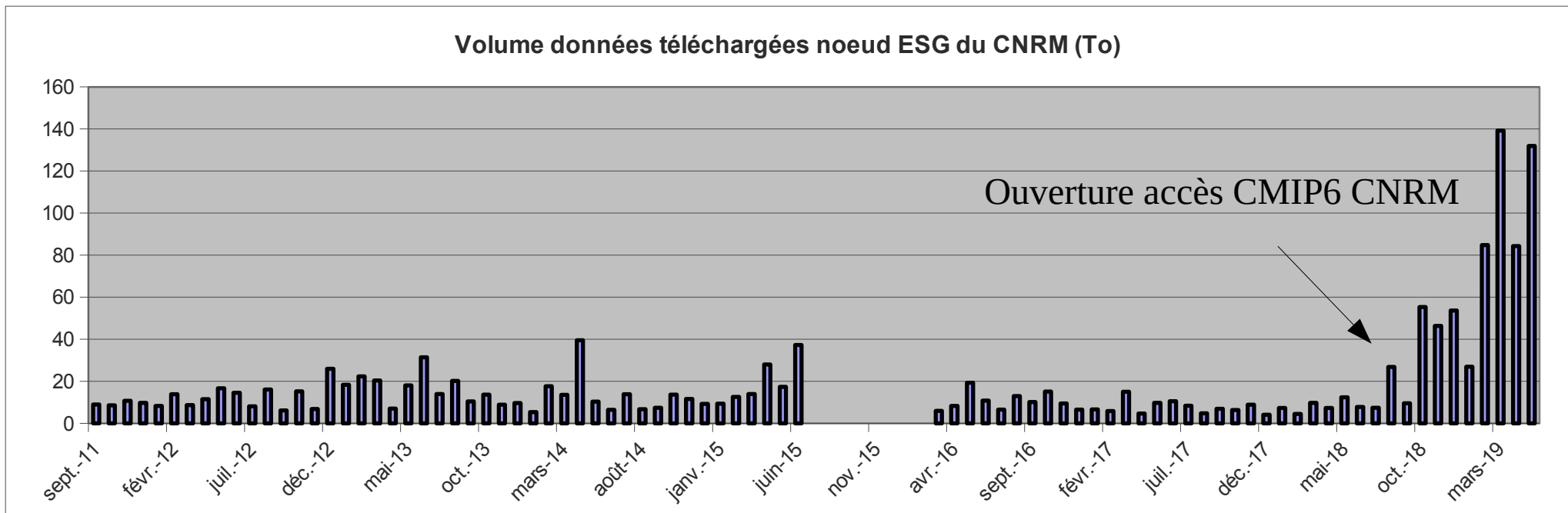
| Data Node | |
|--------------------------|-------------------------------|
| <input type="checkbox"/> | aims3.llnl.gov (29) |
| <input type="checkbox"/> | cmip.bcc.cma.cn (6517) |
| <input type="checkbox"/> | crd-esgf-drc.ec.gc.ca (43763) |
| <input type="checkbox"/> | dist.nmlab.snu.ac.kr (840) |
| <input type="checkbox"/> | esg-dn1.nsc.liu.se (129) |
| <input type="checkbox"/> | esg.lasg.ac.cn (717) |
| <input type="checkbox"/> | esg1.umr-cnrm.fr (38916) |
| <input type="checkbox"/> | esgdata.gfdl.noaa.gov (2586) |
| <input type="checkbox"/> | esgf-data.ucar.edu (65079) |
| <input type="checkbox"/> | esgf-data2.diasjp.net (8291) |
| <input type="checkbox"/> | esgf-data3.ceda.ac.uk (7574) |
| <input type="checkbox"/> | esgf.bsc.es (336) |
| <input type="checkbox"/> | esgf.nccs.nasa.gov (11317) |
| <input type="checkbox"/> | esgf3.dkrz.de (250) |
| <input type="checkbox"/> | vesg.ipsl.upmc.fr (201358) |

Publication ESGF au CNRM

- Quelques soucis d'infrastructure pénalisants pour la publication : DMZ CNRM saturée pendant 2 semaines, transfert Hendrix – nœud impossible durant une semaine (débit fortement dégradé depuis cet incident)
- **Bilan publication au 17/05** :
 - 208 bons à publier émis (+ 3 extensions 2100-2300 membres r1 ESM2-1)
 - 128 simulations publiées (~210 To). Depuis la dernière réunion Tech :
 - * publication de tous les membres ScenarioMIP CM6-1 (membres r2 à r6 simulations ssp126, ssp245, ssp370 et ssp585), simulation omip1 ESM2-1 et 12 simulations ScenarioMIP ESM2-1
 - * 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**) :
 - * simulations omip1 CM6-1 (reste table Omon, pb attribut 'branch_method')
 - * autres membres ScenarioMIP ESM2-1 (19 simulations restantes ssp119, ssp126, ssp245, ssp370, ssp434, ssp460, ssp534-over, ssp585)
 - * fin simulations Crescendo (jalón été)
 - * **simulations RFMIP (liste simulations prioritaires à préciser)**
 - * extension '2300' membres r1 ESM2-1
 - * puis ordre chronologique d'arrivée des bons à publier....

Trafic noeud ESGF au CNRM

- Record volume maximal téléchargé en une journée battu sur le noeud le 31/05 (10.96 To)
- ~ 132 To téléchargé au mois de mai (2ème volume mensuel le plus important depuis 2011)
- ~ 356 To sur la période mars-mai (à comparer au volume moyen mensuel sur la période 2011 à 2017 qui s'élève à 14 To/mois)



Errata Service, known issues

- **vt100** :

- errata déclaré sur l'Errata Service :

- <https://errata.es-doc.org/static/view.html?uid=ff701570-b976-b9eb-8b88-348836e14b09>

- dépublication 27 datasets concernés en cours, texte site web mis en ligne

- variables **seaice** mal pondérées (cf mail Stéphane 23/04 : siv, siu, sisnconc, sisnmass, sisnhc, sistrxdtop, sipr, sistrydtop, sistryubot, sistrxubot, sidragbot, siflfwdrain, siflfwbot, siflswutop). Attente verdict final pour décision de dépublication (738 datasets déjà publiés) et publication jeux à suivre

- sous-échantillonnage champs 3D à 3h pour moyenne temporelle : texte site web mis en ligne

■ Point d'étape :

- * fichiers à renseigner (esm2-1) envoyés à chaque référent le 12/04, rappel le 06/05
- * fichiers déposés sur GitHub ESDOC <https://github.com/ES-DOC-INSTITUTIONAL/cnrm-cerfacs/> et visibles sur <https://explore.es-doc.org/cmip6/models> pour les 4 realm aerosol, atmoschem, land, landice
- * fichiers realm coupling (Sophie, Aurore) en cours de finalisation
- * remarques émises sur les fichiers proposés (Pierre, Bertrand et Christine) transmises à l'équipe ESDOC
- * **Relance à faire auprès des référents n'ayant pas répondu.**

■ A suivre :

1) récolte puis vérification des fichiers restants pour esm2-1 (atmos, ocean, ocnbgchem, seaice, toplevel et citation), dépôt sur GitHub ESDOC et envoi à l'équipe ESDOC de vos remarques éventuelles sur les fichiers à renseigner

2) extension à cm6-1 puis cm6-1-hr

■ Question landice : landice déclaré dans CMIP6_CV seulement pour CM6-1 :

```
"CNRM-CM6-1":{  
  [...] "landIce":{  
    "description":"GRISLI: Collaborating with l'Institut des Geosciences de l'Environnement in  
Grenoble on offline ice-sheet simulations",  
    "native_nominal_resolution":"10 km"  
  }  
}
```

➡ CMIP6_CV à modifier ? **Modèle non utilisé, en attente informations**

Prise en compte fichiers esdoc aerosol et atmoschem GitHub sur <https://explore.es-doc.org/cmip6/models>

Institute CNRM-CERFACS

Model CNRM-ESM2-1

Realm > Process

- Top Level
 - Radiative Forcings
 - Aerosol**
 - Grid
 - Transport
 - Emissions
 - Concentrations
 - Optical Radiative Properties
 - Model
- Atmosphere
 - Grid
 - Dynamical Core
 - Radiation
 - Turbulence Convection
 - Microphysics Precipitation
 - Cloud Scheme
 - Observation Simulation
 - Gravity Waves
 - Natural Forcing
 - Atmospheric Chemistry**
 - Grid
 - Transport
 - Emissions Concentrations

CNRM-CERFACS > CNRM-ESM2-1 :: Aerosol

Top Level Properties

Aerosol > Name

| | |
|-------------|----------------------------|
| Description | Name of aerosol model code |
| Value | TACTIC_v2 |

Aerosol > Keywords

| | |
|-------------|---|
| Description | Keywords associated with aerosol model code |
| Value | Aerosol, aerosol-radiation interactions, deposition, emission, transport, sedimentation |

Aerosol > Overview

| | |
|-------------|---|
| Description | Overview of aerosol model. |
| Value | TACTIC (Tropospheric Aerosols for ClimaTe In CNRM) is an interactive tropospheric aerosol scheme, able to represent the main anthropogenic and natural aerosol types in the troposphere. Originally developed in the GEMS/MACC project (Morcrette et al., 2009), this scheme has been adapted to the ARPEGE/ALADIN-Climat models (Michou et al., 2015 and Nabat et al., 2015). Aerosols are included through sectional bins, separating desert dust (3 size bins), sea-salt (3 size bins), sulphate (1 bin, as well as 1 additional variable for sulfate precursors considered as SO2), organic matter (2 bins: hydrophobic and hydrophilic particles) and black carbon (2 bins: hydrophobic and hydrophilic particles) particles. All these 12 species are prognostic variables in the model, submitted to transport (semi-lagrangian advection, and convective transport), dry deposition, in-cloud and below-cloud scavenging. The interaction with shortwave and longwave radiation, is also taken into account through optical properties (extinction coefficient, single scattering albedo and asymmetry parameter) calculated using the Mie theory. Sulfate, organic matter and sea salt concentrations are used to determine the cloud droplet number concentration following Menon et al. (2002), thus representing the cloud-albedo effect (1st indirect aerosol effect). |

Aerosol > Scheme Scope

| | |
|-------------|--|
| Description | Atmospheric domains covered by the aerosol model |
| Values | Troposphere |

Aerosol > Basic Approximations

| | |
|-------------|---|
| Description | Basic approximations made in the aerosol model |
| Value | The conversion from sulphate precursors into sulphate aerosols is calculated from an exponential decay based on a time constant depending only on the latitude. |

Aerosol > Prognostic Variables Form

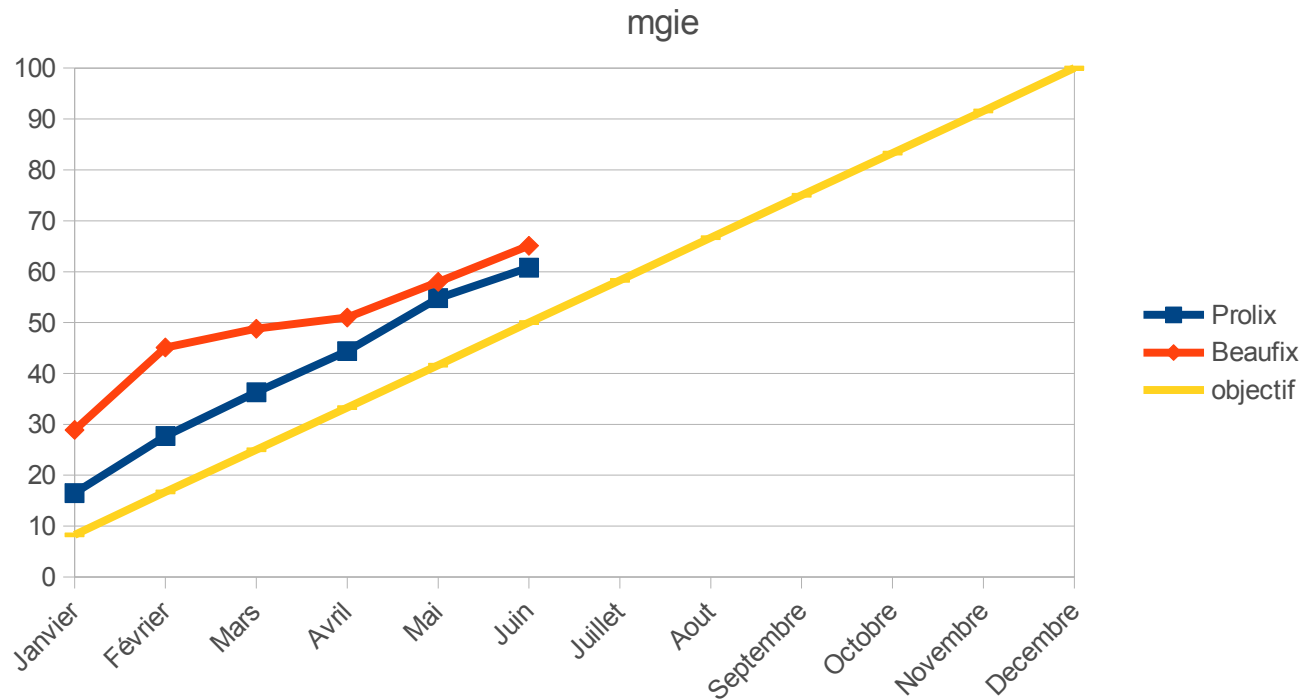
| | |
|-------------|---|
| Description | Prognostic variables in the aerosol model |
| Values | 3D mass/volume ratio for aerosols |

Aerosol > Number Of Tracers

| | |
|-------------|--|
| Description | Number of tracers in the aerosol model |
| Value | 12 |

Ressources calcul et stockage, site web

- Demande DSI réactualisation ressources 2019 : +5 Mh demandé sur Beaufix (HR). Arbitrage lors du CPCC du 27/06.
- Consommation calcul mgie (au 17/06) : il reste 9.9 Mh sur Beaufix et 5.5 Mh sur Prolix.



Suite...

- Prochaine réunion ? Pendant l'été si suffisamment de participants, cf sondage. Puis reprise à l'automne.