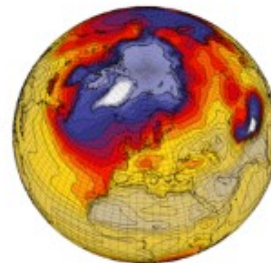
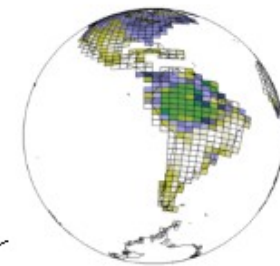
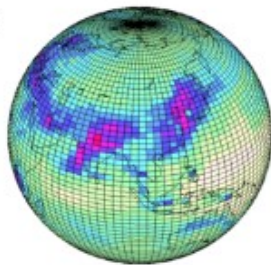
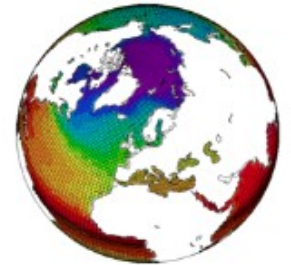
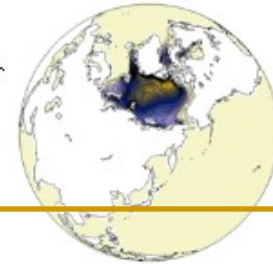


# Diffusion et accès aux données CMIP5

Une vingtaine de personnes et al. entre Paris et Toulouse



# Etat des lieux distribution CMIP5 : CMOR2

- The **GOOD**
  - Dès octobre 2010 nous avons des fichiers «cmorizés»
  - Le recensement des variables CMIP5 s'est bien passé
  - Echange IPSL/CNRM/CERFACS sur les codes cmor2
- The **BAD**
  - Tables figées en février 2010 seulement. Lourd à gérer
  - Quelques changements de table ont impliqués des réécritures
- The **UGLY**
  - Bug cmor2 → mauvais vertices dans les fichiers océan
  - Bug fixé dans cmor 2.7.0 hier → réécriture

# Etat des lieux distribution CMIP5 : ESG

## ■ The GOOD

- Collaboration pour installation du nœud de données.
- ✓ CNRM/CERFACS et IPSL : parmi les 5 premiers groupes présents sur ESG (mars/avril 2011).
- ✓ ~ 62% des 95To de l'archive CMIP5 viennent de la communauté ENES (nous principalement).

## ■ The BAD

- Procédures de mise à jour des données très lourdes.

## ■ The UGLY

- Installation et upgrade du data node cauchemardesques.
- Pas de «release cycle» ni de tests unitaires.

# Etat des lieux distribution CMIP5 : accès aux données via ESG

- The **GOOD**
  - DOI pour les données
  - ESG Ca marche
- The **BAD**
  - Pas très bien d'après moi
  - Couche sécurité (cf licence) : c'est louuurd
- The **UGLY**
  - Les portails CMIP5 n'exposent pas les mêmes données
  - Téléchargement complexe (authentification/autorisation)
  - Incohérence des arborescences des data nodes

# Etat des lieux distribution CMIP5 : accès aux données via ciclad/prodiguer

- Ciclad
  - Cluster de calcul (20 nœuds dual quad-core AMD)
  - « Climate analysis enabled » netcdf, ferret, matlab...
- Prodiguer
  - Sous ensemble CMIP5 (160 To → 300 To fin 2011)
  - Un portail «beau et intuitif» pour naviguer dans le sous ensemble CMIP5
  - Un data node «maison» (accès opendap sur agrégation) (année 1→10 + 11→20 = 1→20)
  - Un outil simple de téléchargement CMIP5 (fonctionnalité proche du script type wget CMIP3)

# Etat des lieux distribution CMIP5 : financement/projet actuel

- Source **européenne**: METAFOR/IS-ENES/EGI
- Source **internationale** : G8 : projet ExArch
- Source **nationale** : aucune. Future : ANR?
  
- Labellisation INSU PRODIGUER : refusé
  - **Dossiers classés A et considérés comme de bons projets pouvant sans conteste être labellisés, mais qui mériteraient certainement une ouverture nationale plus grande ou dont l'implémentation doit être revue.**
  - **PRODIGUER système de diffusion des données de modélisation, issues des exercices du GIEC, est un projet tout à fait intéressant mais dont le soutien comme SOERE peut faire débat. L'avis du comité est que cette initiative mérite d'être soutenue dans un cadre qu'Allenvi pourrait déterminer.**



# https://prodiguer.ipsl.fr/



## Prodiguer Portal v1\_0a4

Diffusion internationale des données scientifiques

Experiments

Simulations

Output

### EXPERIMENTS - ATMOSPHERE

ATMOSPHERE

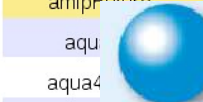
Long-Term

Decadal

1 to 12 of 12 entries

Filter:

Name	Description	Number	Tier	Years / run	Ensembles		
amip	AMIP	3.3	CORE	>=30	>=1		
sstClim	control SST climatology	6.2a	CORE	30	1		
sstClim4xCO2	CO2 forcing	6.2b	CORE	30	1		
amip4K	AMIP plus 4K anomaly	6.8	TIER 1	30	1		
amip4xCO2	4xCO2 AMIP	6.5	TIER 1	30	1		
amipFuture	AMIP plus patterned anomaly	6.6	TIER 1	30	1		



## Prodiguer Portal v1\_0a4

Diffusion internationale des données scientifiques

Experiments

Simulations

Output

### OUTPUT - CMIP5 - KEYWORD

DRS

KEYWORD

Data Node

Keywords

50

CICLAD

sublimation sea ice i1p1 historical

Search



1 to 6 of 6 entries

Filter:

Node	Institute	Model	Experiment	Freq.	Realm	Ensemble	Version	MIP	Name	Units	Start	End	GB	
CICLAD	IPSL	ipsl-cm5a-lr	historicalGHG	mon	seaIce	r1i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>
CICLAD	IPSL	ipsl-cm5a-lr	historical	mon	seaIce	r1i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>
CICLAD	IPSL	ipsl-cm5a-lr	historical	mon	seaIce	r2i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>
CICLAD	IPSL	ipsl-cm5a-lr	historical	mon	seaIce	r3i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>
CICLAD	IPSL	ipsl-cm5a-lr	historical	mon	seaIce	r4i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>
CICLAD	IPSL	ipsl-cm5a-lr	historicalNat	mon	seaIce	r1i1p1	v1	Olmon	sbsi	kg m-2 s-1	1850-01-01	2005-12-01	1.52	<input type="checkbox"/>

50 per page

First Previous 1 Next Last



# MEMO 1)

## 1) Mode d'accès CMIP5 officiel:

- Le portail web du pcmdi est le point d'entrée officiel.

<http://pcmdi3.llnl.gov/esgcet/home.htm>

- L'ensemble de la base CMIP5 est accessible depuis ce point d'entrée.
- Dans la grande majorité des cas ; les données sont physiquement localisées sur des serveurs hébergées par les groupes de modélisation.
- Nous vous recommandons de vous enregistrer sur ce portail si vous comptez analyser les sorties CMIP5 multi-modèles

[http://cmip-pcmdi.llnl.gov/cmip5/docs/data\\_getting\\_started.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/data_getting_started.pdf)

ou [http://cmip-pcmdi.llnl.gov/cmip5/data\\_getting\\_started.html](http://cmip-pcmdi.llnl.gov/cmip5/data_getting_started.html)



# MEMO 2)

2) Étant donné les volumes des résultats il faut éviter les rapatriements dispersés. Voici ce que l'IPSL a fait pour faciliter le travail d'analyse.

- Le cluster ciclad a accès à un espace disque conséquent sur lequel sera rapatrié un sous ensemble de la base CMIP5 (espace de 160 To actuellement, 300 To courant 2011)
- Le portail prodiguer permet d'effectuer des recherches sur ce qui est disponible dans cet espace commun. Les fonctionnalités de prodiguer évolueront régulièrement. Tout retour d'expérience est le bienvenu:

<https://prodiguer.ipsl.fr> (onglet simulation "protégé". login unix IPSL).

- La réplication partielle des données CMIP5 des autres modèles a démarré (CNRM-CERFACS, CCCma, MOHC).
- En attendant d'avoir figé la procédure de réplication de la base cmip5 et d'ingestion dans prodiguer elles sont temporairement stockées ici:

`/data/jripsl/data/cmip5`

# MEMO 3)

3) Nous vous invitons à demander dès maintenant un compte sur le cluster CICLAD en suivant les indications du site

<http://ciclad-web.ipsl.jussieu.fr/sections/informations/par-commencer>

Trois documents incontournables pour pouvoir appréhender cette base de données.

1) Le protocole expérimental CMIP5:

[http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor\\_CMIP5\\_design.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf)

2) L'arborescence utilisée pour stocker les résultats:

[http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5\\_data\\_reference\\_syntax.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf)

3) La liste des variables demandées pour cmip5:

[http://cmip-pcmdi.llnl.gov/cmip5/docs/standard\\_output.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.pdf)