

Ressources nécessaires pour la modélisation du climat à l'IPSL

*Pôle de modélisation du climat de l'IPSL
Mai 2004*

1. La modélisation à l'IPSL

Les activités de modélisation du climat à l'IPSL sont centrées autour de projets ambitieux nécessitant des modèles globaux représentant les différents éléments du système climatique (atmosphère, océan, surfaces englacées, surfaces continentales, biogéochimie) et leurs couplages. De nombreuses avancées scientifiques dans ce domaine sont entièrement dépendantes des ressources en calcul intensif. En particulier notre capacité à prévoir les modifications climatiques majeures susceptibles de se produire au cours des prochaines années ou des prochains siècles est fortement conditionnée par notre capacité à intégrer dans des ensembles complexes des processus ayant des échelles de temps et d'espace très variés et par notre capacité à les évaluer dans des situations variées.

Ce document fait le bilan des ressources calcul nécessaires pour les projets de modélisation de l'IPSL qui s'appuie principalement sur les centres de calcul de l'IDRIS et du CEA. Il est destiné à servir de support à la prospective de calcul menée en 2004 par le groupe GEPCI2, en vue de renouveler la puissance de calcul du centre de l'IDRIS. Le document met l'accent sur les projets les plus critiques en terme de ressources de calcul à partir des demandes effectuées en 2004 et indiquent les principaux axes qui justifient une nette progression des ressources informatiques dans les années à venir.

Nous pouvons cependant noter que nous avons assisté à un décrochage de la puissance de calcul par rapport aux grands centres allemand et anglais (un facteur 3 à 7 suivant que l'on compare l'IPSL seul ou la communauté nationale dans son ensemble) ces dernières années, alors que nous étions dans une situation très favorable il y a 5 ans. Nos partenaires, eux-mêmes, sont loin derrière le Earth Simulator japonais.

Il faut donc réagir fortement pour éviter de sortir de la compétition scientifique internationale. **Les besoins de l'IPSL sont équivalents à 100 processeurs NEC SX-6 dès maintenant et le double dans deux ans.** Cette puissance de calcul doit, au moins pour les dix prochaines années, être fournie par des machines parallèles mais constituées de gros processeurs probablement vectoriels (parallélisme gros grains). Dans le même temps, nous devons nous préparer à la mise en œuvre d'un Earth Simulator européen qui accueillera les simulations de grande envergure.

La communauté directement concernée par la modélisation du climat de l'IPSL

L'IPSL (Institut Pierre Simon Laplace) est un observatoire des sciences de l'univers (OSU) et une fédération de 6 laboratoires de recherches sur le climat et l'environnement :

- CETP : Centre d'Etude des Environnements Terrestre et Planétaires,

- LMD : Laboratoire de Météorologie Dynamique,
- LODYC : Laboratoire d'Océanographie DYnamique et de Climatologie
- LBCM : Laboratoire de Biogéochimie Marine,
- LSCE : Laboratoire des Sciences du Climat et de l'Environnement,
- SA : Service d'Aéronomie

dont les tutelles sont : le CNRS, l'université Pierre et Marie Curie, l'Université de Versailles Saint-Quentin, le CEA, l'Ecole Normale Supérieure, l'Ecole Polytechnique, le CNES, l'IRD et le Muséum National d'Histoire Naturelle.

Le pôle de modélisation du climat de l'IPSL rassemble environ 80 chercheurs de l'IPSL. A ces personnels s'ajoutent des personnels du LOA(Lille), LGGE (Grenoble), UCL-ASTR (Louvain-la-Neuve) qui apportent leurs compétences respectivement sur la modélisation des aérosols, le climat des hautes latitudes et la glace de mer. Les résultats des simulations sont plus largement utilisés par des scientifiques de l'IPSL ou par des collaborateurs extérieurs, soit 200 personnes environ. Nous interagissons régulièrement avec des collègues européens et internationaux et participons à de nombreux exercices de comparaisons de modèles.

Le pôle de modélisation se structure autour d'un conseil scientifique présidé par Pascale Braconnot. Plusieurs groupes de travail (scientifiques et techniques) animent le pôle. Du point de vue du développement et de la mise en œuvre des modèles, nous distinguons les compétences entre les développeurs de modèles (une quinzaine de personnes), les utilisateurs avertis (qui adaptent et améliorent les modèles et lancent des simulations originales, soit une cinquantaine de personnes) et les scientifiques utilisateurs, IPSL ou non, qui analysent les résultats (environ 200).

Atouts et faiblesses actuelles

L'IPSL a la chance de regrouper dans le même institut des spécialistes des différents éléments du système climatique. L'activité de modélisation à l'IPSL se structure autour des différentes composantes du système climatique, en privilégiant les échelles globales. Les modèles atmosphériques, océaniques, chimiques et de surfaces continentales, sont développés et utilisés indépendamment dans des configurations variées pour étudier le fonctionnement des différents milieux interagissant du système climatique. Ces thématiques concernent aussi bien la variabilité du climat tropical, la caractérisation des puits et sources de CO₂ à partir de l'inversion du transport atmosphérique, la chimie des gaz à effet de serre, etc. ... A noter de plus que le modèle atmosphérique est utilisé pour d'autres planètes du système solaire (Mars, Titan). Les demandes de calcul spécifique à ces différents thèmes continueront dans le futur à constituer une fraction importante des demandes d'heures (cf. tableau 1).

Au-delà des études à partir des composantes individuelles, un effort important est mis sur le modèle couplé de l'IPSL. Les interactions entre composantes mettent en œuvre des processus complexes qui sont au cœur des études sur le changement climatique. C'est l'intersection entre ces développements sur le couplage et les compétences sur les cycles biogéochimiques qui ont permis à l'IPSL d'être l'un des deux premiers centres à réaliser des simulations couplées entre le climat et le cycle du carbone. Cette étude originale a, par exemple, permis d'aborder les questions suivantes : comment le climat réagit-il à l'augmentation des émissions de gaz carbonique ? Quelle est la perturbation sur le cycle du carbone du changement climatique ? Quel est l'effet en retour de la modification du cycle du carbone sur le climat ? Dans cette logique, le projet porté par l'IPSL, relayé par un ACI et le projet européen ENSEMBLES nécessite de

réaliser un ensemble de simulations couplées longues suivant les scénarios proposés par IPCC, auquel seront adjointes des séries d'expérience permettant de prendre en compte le couplage entre le climat et les cycles bio-géochimiques.

Les dimensions des projets actuels correspondent ainsi à plusieurs types de configurations :

- le modèle couplé de l'IPSL en résolution typique sur une centaine d'années
- les simulations d'ensembles à partir d'un ensemble de simulations de même durée
- le modèle couplé incluant la chimie (transport de cent ou cent cinquante espèces et plusieurs centaines de réactions chimiques)
- le modèle océan ou atmosphère à haute résolution

Pour chacun de ces domaines, le tableau joint reprend les caractéristiques essentielles en terme de durée de simulation, de résolution, de ressources CPU, de mémoire, de taille des sorties.

Modèle	Expérience type	Résolution	CPU nécessaire	Mémoire	Sorties
Couplé	100 ans	Atm. : 96x72x19 Océan : 182x149x31	10h / an	3 Go	5 Go/an
Ensemble	Nx10 ans	idem	idem	Nx3 Go	idem
Couplé-chimie	Nx10ans	Atm 96x72x19 100 traceurs et 200 réactions chimiques	30h / an	5Go	40 Go/an
Couplé-chimie avec stratosphère	Nx10ans	Atm 96x72x50 100 traceurs et 200 réactions chimiques	50h / an	5Go	00 Go/an
Océan haute résolution avec biogéochimie		722x511x31 + 26 traceurs			
Etude de processus haute résolution avec biogéochimie	10 ans	1000x1500x200 + 9 traceurs pas de temps 5mn			

La saturation progressive des moyens de calcul à l'IDRIS s'est traduite ces dernières années par une forte limitation des ressources accessibles à la communauté et une auto-censure dans le dimensionnement des projets scientifiques. Concrètement, la participation de l'IPSL au prochain rapport de l'IPCC se fera sur la base de simulations avec une résolution spatiale nettement plus grossière que nos partenaires, l'accent ayant été principalement porté sur le développement des couplages. Cette résolution risque de diminuer d'autant le poids accordé à notre contribution. Outre la résolution, les limitations actuelles conditionnent la complexité du modèle de climat. Dans la configuration retenue, des simulations longues (de type IPCC) incluant la chimie atmosphérique correctement représentée et la biogéochimie océaniques ne sont pas possibles. Pour l'année à venir, seuls des *snapshots* sur des périodes courtes sont envisagés. A cela, s'ajoute le risque de limiter les développements (paramétrisation, complexification, résolution) effectués par l'ensemble de la communauté.

1.3 Les modèles utilisés et leurs spécificités

Les composantes du système climatique

Les modèles développés et manipulés à l'IPSL concernent les différents éléments du système climatiques:

- l'atmosphère (LMDZ y compris sa partie transport des espèces chimiques),
- l'océan (y compris le transport des traceurs) et la glace de mer (OPA dans sa configuration ORCA-LIM),
- la biogéochimie océanique (PISCES, LOBSTER),
- les processus de surface incluant le carbone et l'utilisation des sols (ORCHIDEE),
- la chimie atmosphérique (INCA).

Ce sont des modèles en points de grille. La programmation de ce type de modèles s'appuie sur des vecteurs longs, ce qui leur permet d'être efficaces sur les calculateurs vectoriels. Cette connaissance de base est acquise à l'IPSL depuis de longues années grâce à l'existence successive des Cray, Fujitsu et NEC du CCVR, de l'IDRIS et du CEA. Ces modèles sont également adaptables au calcul parallèle. Certains ont tourné de manière très efficace sur les T3E. Un point clé reste la performance monoprocesseur qui doit être la plus efficace possible afin de permettre la mise en œuvre rapide de nouvelles études sur des prototypes qui peuvent ne pas être parallèles.

Plusieurs des composantes sont utilisées aussi bien sur les super-calculateurs de l'IDRIS ou du CEA que, dans des configurations plus légères, sur des stations de travail ou PC-Linux. Ceci est extrêmement important et doit être maintenu en œuvre notamment pour des collaborations avec des pays qui n'ont pas les moyens de calcul équivalents.

Ces modèles sont développés depuis de nombreuses années et restent d'actualité pendant plusieurs années. Une nouvelle version majeure de ces modèles sort, typiquement, tous les 3 à 5 ans. **Les études scientifiques prévues s'appuieront largement sur ces modèles et toute évolution des moyens de calcul aura un impact immédiat sur nos études.**

Les activités de modélisation ne consistent pas seulement à lancer des simulations à partir d'un outil unique. Le travail de modélisation est central dans la thématique de recherche. Développer un modèle et le confronter aux observations, c'est progresser dans la compréhension de processus complexes. Du coup, les modèles sont extrêmement évolutifs et doivent pouvoir rester, en partie, entre les mains des chercheurs non-spécialistes des aspects informatiques et numériques comme des étudiants en thèse et en stage. Pour conclure, les modèles doivent rester souples et faciles à utiliser et à faire évoluer.

Le système couplé

Pour aborder le climat dans toutes ses composantes, l'IPSL a entrepris le développement d'un modèle du système terre complet où les interactions entre composantes sont activées. Ce modèle couplé s'appuie sur l'ensemble des différentes composantes du système climatique: Certaines des composantes échangent leurs informations par le coupleur OASIS, développé au CERFACS. D'autres comme le couplage entre le modèle de chimie-aérosols INCA et l'atmosphère font l'objet de procédures internes.

La version actuelle du modèle couplé océan - glace de mer – atmosphère - surfaces continentales est arrivée récemment à maturité. Dans sa configuration actuelle, le modèle climatique de l'IPSL (couplage entre les différentes composantes ou seuls les échanges de chaleur et d'eau sont activés) simule un an en 10 heures CPU. Ce modèle couplé est l'outil de base pour effectuer des simulations dites de scénarios dans le cadre de la préparation du prochain rapport de l'IPCC. Elles constituent à la fois un enjeu scientifique important et un point de passage obligé pour la communauté concernée. La qualité des résultats simulés conditionnera notre crédibilité dans les années qui viennent. Une part significative (30-50%) de la consommation d'heures de calcul sur les super-calculateurs sera consacrée à ces simulations.

Les modèles de chimie atmosphérique ou de végétation peuvent déjà être couplés à l'atmosphère. De même le modèle de biogéochimie marine peut être activé en ligne dans le modèle d'océan. Les prochaines étapes de développement concerneront le couplage dans le système global des espèces chimiques atmosphériques et biogéochimiques océaniques.

Un élément important de la réussite de cette entreprise repose sur un ensemble d'outils informatiques génériques utilisés par le modèle couplé et ses composantes. Ce sont les mêmes outils qui permettent de relancer une simulation de référence que le modèle utilisé soit une composante seule, ou le modèle couplé. Cet aller-retour permanent entre simulations couplées et découplées est au cœur de notre démarche de compréhension des rétroactions internes. Ces méthodes de travail similaires permettent de partager les efforts de développement et de les pérenniser.

Quelques éléments de positionnement des modèles de l'IPSL dans la communauté internationale

Les modèles européens sont de qualité scientifique plus que respectables et la communauté française est très présente dans la course à la qualité des modèles. Lors d'un exercice d'intercomparaison aux données satellitaires ISCCP, les propriétés des nuages simulés par LMDZ se sont révélées être parmi les plus proches des observations. Dans sa version transport, LMDZ a été utilisé par Olivier Boucher pour calculer un des champs de concentration de sulfate recommandé par l'IPCC. Le modèle d'océan OPA est également utilisé dans tous les autres grands centres européens (Hadley Center – Royaume Uni, ECMWF, MPI-M – Allemagne, MPI- Kiel Allemagne, Bologne Italie, KNMI – Pays Bas, ...).

Un travail important a été effectué au cours des trois dernières années pour améliorer le climat simulé. Les comparaisons du cycle saisonnier simulé par le modèle en différents endroits des océans tropicaux montrent que le cycle saisonnier est bien représenté. Ce résultat résulte en partie de la très bonne qualité des ajustements radiatifs au sommet de l'atmosphère comme l'atteste les comparaisons avec les données satellites. Des comparaisons préliminaires des simulations couplées de différents modèles d'atmosphère avec OPA comme modèle de circulation océanique indiquent également que le modèle produit des résultats très satisfaisants pour la variabilité tropicale malgré notre résolution plus grossière.

Le modèle couplé de l'IPSL est tout à fait dans la course en terme de modélisation intégrée des composantes du système terre.

2. Enjeux scientifiques de la modélisation du climat global

Les projets scientifiques menés à l'IPSL demandent d'être capable d'accroître nos systèmes dans plusieurs directions à la fois. En effet, les enjeux actuels conduisent à disposer de système de modélisation capable de représenter le monde qui nous entoure de la façon la plus réaliste possible. Cet objectif ne peut être atteint que par un ensemble de modélisation modulaire représentant les différents aspects du système climatique, ainsi que les aspects spécifiques de l'activité humaine, comme le traitement des espèces chimiques anthropiques ou de l'agriculture. Les grandes thématiques perdureront dans les années qui viennent. L'évolution sera guidée par un besoin de mieux comprendre les interactions d'échelles intervenant dans le système, des études de prédictibilité de plus en plus poussées pour toutes les échelles de temps, l'accroissement de la complexité du système pour pouvoir intégrer dans un ensemble cohérent climat global, climat régional et problèmes environnementaux ou de sociétés. L'accroissement des ressources informatiques a un impact direct sur les études possibles et sur leurs qualités.

Les évolutions souhaitables vont toutes accroître les ressources nécessaires (calcul, mémoire, espace fichiers, outils d'analyse, ...). Elles se caractérisent par :

- la variété des simulations pour pouvoir prendre en compte le système dans son ensemble ou le rôle de l'une ou l'autre composante ou paramétrisation,
- l'augmentation de la résolution horizontale et verticale,
- l'augmentation des longueurs des simulations ,
- la généralisation des simulations d'ensemble et l'augmentation du nombre de membres de ces ensembles,
- la complexification des composantes (plus d'espèces chimiques, plus de processus physiques, ...),
- des couplages de plus en plus poussés et offrant des possibilités d'analyses croisées avec nos partenaires

Des simulations longues et multiples :

Pour pouvoir étudier le comportement du système climatique, nous devons pouvoir mener de front des simulations de type très différent. Ces simulations doivent permettre de :

- valider les modèles,
- étudier les mécanismes mis en jeu en forte interaction avec les observations,
- mener des simulations longues du climat actuel pour étudier la variabilité naturelle,
- mener des simulations longues des climats passés pour étudier les mécanismes de changement de climat et évaluer les modèles utilisés pour les scénarios futurs,
- mener des études à l'échelle régionale (en zoomé ou haute résolution),
- quantifier les erreurs des projections climatiques liées aux paramétrisations,
- intégrer les processus non représentés, comme, par exemple, les couplages biogéochimiques,
- aborder les analyses statistiques par la réalisation d'ensemble de simulations.

Si l'on s'intéresse au climat et aux différentes échelles de temps qui le composent, il faut réaliser des simulations de :

- 1000 ans au minimum pour les simulations de contrôle,
- plusieurs fois 500 ans pour les tests autour simulations de changement climatique,
- plusieurs fois 100 ans pour les études de processus.

Les ensembles de simulation :

Le système climatique est un système chaotique. Une légère perturbation de l'un des éléments peut ainsi le conduire sur une trajectoire différente. L'étude de la prévisibilité climatique et les projections climatiques doivent donc être abordées de façon probabiliste. Le système peut également bifurquer rapidement d'un état dans un autre, comme cela a été montré à partir de modèles simplifiés. Il n'existe actuellement pas de cadre théorique général pour ce problème. L'approche est empirique et passe par la réalisation de nombreuses simulations.

Les modèles incorporent également des paramétrisations dont certains paramètres sont définis dans une plage d'incertitude à partir de données disponibles. De légères variations de ces paramètres peuvent modifier la réponse d'un modèle à une perturbation. Ces effets doivent être pris en compte dans les incertitudes des simulations climatiques. Là encore l'approche actuelle consiste à réaliser de nombreuses simulations en variant les paramètres dans un domaine acceptable ou en réalisant des ensembles de simulations en remplaçant une paramétrisation par une paramétrisation équivalente dans son niveau de complexité.

Les différents projets de comparaison de modèles et les études de prédictibilité ont montré qu'actuellement aucun modèle n'était capable de représenter correctement toutes les facettes du climat et que la moyenne d'ensemble de simulations réalisées avec des modèles différents était le meilleur estimateur d'un climat ou d'un changement de climat. C'est encore autre ce qui justifie que de nombreux groupes de modélisation réalisent les simulations IPCC et que les études du changement climatique se basent sur des projections réalisées par différents modèles.

En conclusion, le recourt à des ensembles de simulations doit se généraliser pour tous les types d'étude de l'échelle de la prévision saisonnière au changement de climat à long terme pour :

- étudier la sensibilité à l'état initial,
- étudier la sensibilité à la représentation des processus physiques,
- étudier les rétroactions dans le système climatique.

Importance de la résolution spatiale horizontale :

L'accroissement de la résolution horizontale est un facteur important pour mieux représenter les caractéristiques de la circulation de grande échelle, avoir accès aux détails régionaux, rapprochant ainsi plus naturellement l'échelle de la simulation à l'échelle des observations, une meilleure représentation de la variabilité climatique et des extrêmes.

Ainsi par exemple :

- La résolution par le modèle atmosphérique de la dynamique ondulatoire aux moyennes latitudes est importante pour bien simuler la variabilité synoptique dans ces régions et les transports de vapeur d'eau et d'énergie vers les hautes latitudes. Il faudrait atteindre une résolution d'environ $1^\circ \times 1^\circ$, soit augmenter la résolution actuelle d'un facteur 9 (3×3) environ. Cette amélioration permettra d'étudier les modifications des régimes de temps et des fréquences des tempêtes en Europe, les changements d'hydrologie (accroissement des précipitations et des risques d'inondation) en Europe du Nord...
- La résolution des tourbillons océaniques est un élément important pour la dynamique de

l'océan et pour la précision des modèles biogéochimiques. Bien résoudre ces phénomènes demande de travailler à une résolution de 2 Km.

- L'étude du risque climatique associé par exemple à la modification brusque de la circulation thermo haline nécessite de représenter correctement la topographie complexe des différents seuils et bassins des moyennes et hautes latitudes. Une résolution de $0,5^\circ$ dans le modèle d'océan est souhaitable.
- La prise en compte des cyclones dans l'étude du changement climatique nécessite d'accroître considérablement la résolution du modèle atmosphérique dans les régions tropicales. Actuellement une version avec un zoom sur l'océan Indien permet de simuler de façon réaliste les cyclones dans cette région. Pour prendre en compte l'impact éventuel de ces cyclones sur l'océan, il faudrait étendre ce zoom à toute la région tropicale. La résolution visée est de $0,5^\circ \times 0,5^\circ$ environ, soit un accroissement par 40 de la résolution dans les régions tropicales.
- L'interprétation au niveau régional et local des scénarii d'évolution du climat ouvre la question d'avoir des simulations très haute résolution, en particulier pour évaluer et caractériser les extrêmes climatiques.
- Pour la chimie et les aérosols atmosphériques, la résolution des sources des différentes espèces nécessite une résolution d'au moins 1° degré pour les représenter correctement à l'échelle de l'Europe ou de l'Asie. Plusieurs groupes envisagent de réaliser les simulations IPCC à cette résolution (T106) et les Japonais produiront des scénarios à 50 Km.

Augmenter la résolution horizontale d'un facteur 2 ajoute un facteur 4 à 8 en ressources mémoire et CPU.

Importance de la résolution verticale :

La résolution verticale est un facteur contraignant de nombreux processus. Augmenter la résolution verticale est nécessaire mieux prendre en compte les gradients verticaux et les échanges entre différents milieux ou régions de l'océan ou de l'atmosphère. C'est le cas par exemple de la couche limite atmosphérique, de l'interface air-mer, des échanges entre la stratosphère et la troposphère.

Les évolutions envisagées sont les suivantes:

- pour la couche limite atmosphérique, une nouvelle paramétrisation en cours de développement permettra une meilleure estimation des flux à la surface ainsi qu'une meilleure simulation des nuages de sommet de couche limite et des nuages bas. Un accroissement de résolution d'un facteur 2 à 4 dans la couche limite est envisagé.
- avoir une bonne résolution verticale est nécessaire pour simuler correctement les processus dans la stratosphère et les échanges entre la stratosphère et la troposphère. Une résolution de l'ordre du Km est nécessaire pour représenter ces phénomènes. Cela aura pour principales conséquences d'améliorer la variabilité basse fréquence de l'atmosphère, de mieux simuler la perturbation (le refroidissement) de la haute atmosphère en réponse à un accroissement de CO₂, de mieux simuler les échanges de chaleurs, des masses, de composés chimiques à l'interface stratosphère-troposphère.

Le passage de 19 à 50 niveaux dans le modèle d'atmosphère accroît le coût de calcul d'un facteur 2,5.

Accroissement de la complexité du système :

Un enjeu important pour l'étude du climat est de prendre en compte les interactions entre le climat et les cycles bio-géochimiques, ce qui demande d'évoluer vers un modèle du Système Terre (couplage du climat avec les cycles du carbone, le cycle du méthane, la chimie de l'ozone et la physique des aérosols). Inclure la chimie atmosphérique multiplie par 4 le temps de calcul du modèle atmosphérique, inclure la biogéochimie marine multiplie par 4 le temps de calcul du modèle océanique. L'accroissement significatif des ressources permettrait d'effectuer les simulations longues et en particulier les scénarios IPCC avec le système complet intégrant chimie et biogéochimie. C'est un objectif que nous poursuivons.

Un facteur particulièrement contraignant concerne la chimie et les aérosols atmosphériques. Réaliser des simulations avec une chimie complète est très pénalisant à cause du nombre de réactions à prendre en compte mais aussi du schéma de transport utilisé. Prendre en compte simultanément la chimie troposphérique et stratosphérique nécessite de considérer 200 traceurs (soit 150 traceurs pour la troposphère en incluant les différents hydrocarbures, 20 à 30 traceurs pour la stratosphère et 20 à 30 traceurs pour les aérosols). Cette étape doit pouvoir être franchie si l'on veut quantifier les effets relatifs des différents constituants à effet de serre et leur impact sur les caractéristiques globales et régionales du climat.

Par rapport à l'état actuel, intégrer toutes les composantes du système représente un facteur 5 en ressources calcul et mémoire. Ajouter les niveaux nécessaires pour aborder la stratosphère ajoute encore un facteur 3, soit un facteur 15 en tout, pour traiter le système intégré.

3. Moyens de calcul: un bilan mitigé

Etat des lieux en France : 3 centres

En France, le pôle toulousain utilise les calculateurs Fujitsu VPP de Météo-France dont la configuration a été récemment doublée. L'IPSL utilise les calculateurs de l'IDRIS et du CEA.

Les projets de l'IPSL ont reçu comme allocation en 2004 100 000 heures NEC SX-5 à l'IDRIS et 83 000 heures NEC SX-6 au CEA. Les demandes correspondantes ne reflètent pas la pression de la demande car nous souhaitons rester dans l'épure réalisable. Nous nous sommes autocensurés fortement en limitant principalement la résolution spatiale.

Centre	Calculateur	Détail	Perf crête soutenue	Part climat	Perf soutenue pour climat
IDRIS	NEC SX-5 depuis 2000	3 noeuds 40 proc	320 Gflops 102 Gflops	30% 100 000 h	30 Gflops
CEA	NEC SX-6 depuis 2003	6 noeuds 44 proc	350 Gflops 113 Gflops	35% 83 000 h	40 Gflops
Météo-France	VPP 5000 depuis 2003	124 proc	1190 Gflops 381 Gflops	30%	114 Gflops

L'existence de ces trois centres est un fait. La vie des utilisateurs n'est pas simplifiée car nous devons gérer les projets et les comptes personnels séparément. C'est pour cette raison que l'IPSL s'est concentrée sur 2 sites seulement. Les avantages sont que cela nous oblige à vérifier constamment la portabilité sur différents calculateurs de nos modèles. Les changements et éventuelles pannes de calculateurs sont décalés dans le temps. Ceci nous garantit presque un accès permanent à des ressources. Cela a néanmoins un coût en temps de portage et d'optimisation spécifique, en temps d'adaptation aux spécificités des centres (serveurs de fichiers, machines dédiées ou non aux compilations, post-traitement, logiciels spécifiques, etc. ...). Notons que l'utilisation de deux centres est rendu possible grâce à des ressources réseau existantes bien adaptées. Enfin, dans un contexte multi-sites, toute coopération est à favoriser, comme le déploiement d'outils type OpenDAP (ex-DODS) ou Unicore.

Coordination des centres ou centre unique

Actuellement nous travaillons sur deux centres différents au moins. Cette situation se traduit par des avantages et des inconvénients que nous assumons. Le but à atteindre est de bénéficier de ressources calcul accessibles aisément à la hauteur de nos partenaires. Il ne nous appartient pas de trancher sur la meilleure façon d'atteindre ce but (site unique dédié ou coordination des différents centres avec un suivi dans le temps pour maintenir une capacité de calcul à la hauteur au cours du temps) mais il en va de la capacité du domaine à se maintenir sur la scène internationale.

Etat des lieux en Europe : des centres dédiés

En Europe, l'Allemagne s'est équipée depuis une vingtaine d'années d'un centre de calcul dédié au climat : le DKRZ (Deutsches Klimarechenzentrum). Le Royaume-Uni concentre ses recherches climatiques au sein du Hadley Center qui utilise une part des ressources calcul de la météorologie anglaise : *UK Met Office*. L'effet structurant de telles structures thématiques est visiblement bénéfique pour ces deux pays.

Centre	Calculateur	Détail	Perf (crête) soutenue	Part climat	Perf sout. pour climat
IDRIS	NEC SX-5 depuis 2000	3 noeuds 40 proc	320 Gflops 102 Gflops	30%	30 Gflops
CEA	NEC SX-6 depuis 2003	6 noeuds 44 proc	350 Gflops 113 Gflops	35%	40 Gflops
Météo-France	VPP 5000 x2 en 2003	124 noeuds 124 proc	1190 Gflops 381 Gflops	30%	114 Gflops
DKRZ	NEC SX-6 depuis 2003	24 noeuds 192 proc	1536 Gflops 492 Gflops	100 %	492 Gflops

UK Met Office	NEC SX-6 depuis 2004	30 noeuds 240 proc	1920 Gflops 614 Gflops	50%	307 Gflops
----------------------	----------------------	-----------------------	----------------------------------	-----	-------------------

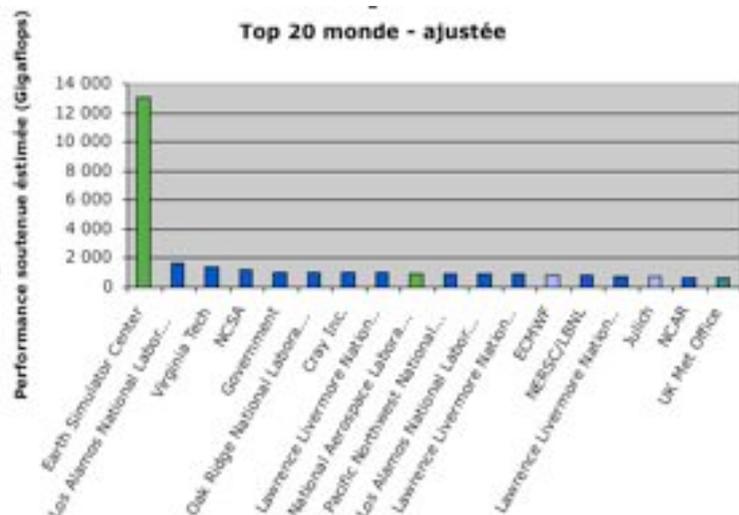
La péréquation entre les parts des centres dédiés au climat entre les pays permet une certaine comparaison entre les ressources européennes de notre communauté. Le retard que nous accusons par rapport à nos partenaires européens est de l'ordre de 3 à 7 selon les points de comparaison.

Le numéro 1 mondial : le Earth Simulator est loin devant depuis 2 ans

Le Earth Simulator est composé de 640 nœuds de 8 processeurs vectoriels NEC chacun, soit un total de 5120 processeurs c'est à dire plus de 25 fois la puissance du DKRZ. Lors de notre semaine de travail au Japon, le premier test de l'océan à 0,5° de résolution est passé en 4 mn (à comparer à plusieurs jours d'attente à l'IDRIS). Ainsi, le modèle d'océan à 0,5° et 300 niveaux, au cœur de la coopération avec le Japon, est impossible à faire tourner en France ne serait-ce que sur quelques jours de simulation.

Même nos partenaires européens sont ainsi loin derrière cette machine unique au monde qui tient la tête du Top 500 (www.top500.org) et ce, depuis sa mise en route en 2002.

Une seule figure résumera cet état de fait : l'histogramme des calculateurs selon leur performance estimée grâce à une formule ad hoc sur nos types d'applications (32% performances crêtes sur processeurs vectoriels et 8% sur processeurs scalaires).



La faiblesse actuelle des moyens de calcul disponibles, au regard de ce qui existe par exemple en Angleterre et en Allemagne, s'est traduite ces dernières années par une limitation en termes de résolution spatiale des simulations. Pour se retrouver au niveau de ces pays, il faudrait multiplier par 1,5 à 2 la résolution horizontale ce qui se traduit illico par une augmentation d'un facteur 4 à 8 du coût de calcul. C'est exactement ce ratio qui nous sépare des concurrents en termes de moyens

Importance du stockage et du post-traitement

Nos études nécessitent des simulations longues, des ensembles de simulation à différentes échelles de temps. Y sont associés des fichiers d'entrée et des fichiers de résultats. Les résultats sont basés sur des séries temporelles d'état du système qui approchent la variabilité du système simulé. Ils sont donc constitués d'un grand nombre de fichiers de taille importante. Elles comportent également une part d'analyse systématique ou post-processing à faire dans le flot des simulations. Nos travaux ont un profil complexe associant des parties de calcul intensif, des parties d'accès à des fichiers et des parties d'analyses systématiques.

Sur l'IDRIS actuellement, le facteur bloquant concerne l'accès aux fichiers résultats et montre, encore une fois, qu'il est impératif de faire évoluer les ressources calcul et les ressources de stockage et d'analyses de façon équilibrée.

Architecture : Vectoriel/parallèle

Un enjeu pour notre communauté sera d'évoluer vers une parallélisation progressive des outils. Ce travail est soit abouti soit bien engagé sur la plupart des composantes. Cependant, une évolution vers du parallélisme massif à base de processeurs scalaires serait un handicap majeur. Les configurations adaptées à nos besoins sont des grappes de processeurs vectoriels. C'est d'ailleurs ce choix qui a été retenu par l'ensemble des grands centres de modélisation climatique. Les Américains qui avaient fait le choix d'abandonner le vectoriel ont été contraints de faire machine arrière, entre autre sous une forte pression de la communauté climatique. Cet aspect de parallélisation deviendra essentiel dès lors qu'on disposera de plus de la douzaine actuelle de processeurs vectoriels sur chacun des centres de calcul.

L'effort de parallélisation engagé autour de nos codes s'appuie sur la mise en œuvre, simultanée ou non, de l'ensemble des techniques parallèles qui permettent d'accélérer la production de résultats : vectorisation, parallélisation par directives OPENMP, parallélisation par échanges de messages. Nous nous devons d'intégrer progressivement dans nos codes les possibilités de parallélisation efficaces sur tout type de processeurs scalaires et vectoriels. Ceci est un processus nécessairement itératif pour maintenir dans le même temps l'évolution scientifique des codes et leur garantir les performances sur tout type d'architecture.

D'autre part, le coût qu'a représenté le passage du Fujitsu à l'IBM à l'ECMWF est de l'ordre de plusieurs années/hommes. Cet effort d'optimisation extrême est envisageable dans un centre opérationnel dont la tâche principale est de fournir une prévision de la meilleure qualité possible dans un temps limité et au meilleur coût. Il n'est pas réalisable dans le cadre recherche qui est le notre où la créativité et la publication de résultats originaux sont les moteurs centraux. De plus les coopérations internationales effectives imposent à nos modèles de tourner sur tout type de

calculateur : du PC au Earth Simulator. La souplesse, la lisibilité et la facilité d'évolution des codes sont primordiales par rapport à un effort ponctuel d'optimisation pointue.

Ceci nous conduit à la conclusion qu'il est impératif de maintenir une forte capacité de calcul sur machine vectorielle avec du parallélisme (parallélisme gros-grain pour les 10 années qui viennent).

4. Conclusion

Le bilan est net, la communauté est très en retard par rapport à nos principaux partenaires européens pour ce qui concerne les ressources calcul alors même que nos modèles sont dans la course en terme de qualité scientifique. Ce retard est visible dans les études menées à l'IPSL et il est urgent de combler ce retard en changeant le NEC SX-5 de l'IDRIS âgé de 5 ans.

Dans le même temps, nous devons maintenir à jour notre expertise pour être partie prenante des discussions prospectives. Dans ce dessein, nous avons engagé une coopération avec les Japonais pour porter et faire tourner sur le Earth Simulator nos codes prêts pour ce type d'architecture (vectoriel et parallèle à deux niveaux) même s'il est quasi impossible de les préparer actuellement.

Nos partenaires européens ont décidé depuis longtemps d'équiper de moyens calculs importants leurs instituts travaillant sur le climat. Sans remettre en cause le choix différent fait en France, il est primordial de mettre en œuvre un suivi spécifique dans la durée des besoins de la communauté et de faciliter le travail inter-centres nationaux et européens.

Un Earth Simulator européen permettrait la réalisation de simulations de grande envergure. Par nature non généraliste, un calculateur de ce type doit être entouré de ressources dimensionnées pour la mise en œuvre d'un vaste ensemble d'études plus standards et le développement des outils de base des simulations extrêmes.

Pour toutes ces raisons, le besoin de l'IPSL est de bénéficier maintenant de l'équivalent de 100 processeurs NEC, de les doubler dans 2 ans et de retrouver et garder un juste équilibre entre les ressources calcul et les ressources de stockage et de post-traitement des résultats.

Annexe 1 Objectifs scientifiques et ressources associées

Objectif	Exigence, contrainte	Résolution	Temps CPU/mémoire
Améliorer la connaissance de la réponse globale du climat aux perturbations anthropiques en améliorant les rétroactions nuageuses	Bonne représentation des nuages bas et de leur sensibilité aux changements de circulation, de température et d'humidité.	Résolution verticale x2 à x4 dans la couche limite Résolution horizontale x4	CPU x 10
Simuler les changements climatiques et les couplages climat-carbone à l'échelle globale.	Idem ci-dessus avec en plus les modèles de carbone des surface continentales et de l'océan	Idem ci-dessus	Idem ci-dessus
Quantifier les changements climatiques en Europe (tempêtes, inondations, sécheresse)	Bien simuler les régimes dépressionnaires des moyennes latitudes.	Résolution horizontale 1°x1° (300x225) Facteur 9	CPU x 15 mémoire x 10
Quantifier l'impact régional du changement climatique sur la chimie dans la troposphère et la stratosphère	Modèle couplé avec chimie-aérosols. Ajouter 200 traceurs et encore plus de réactions chimiques	Amélioration résolution verticale à la tropopause pour atteindre 100 niveaux Facteur 4 en résolution	CPU x 100 Mémoire x 40
Simuler les phénomènes extrêmes dans les régions tropicales et leur évolution	Simuler les cyclones tropicaux	Résolution horizontale 0,5°x0,5°(150x110)	CPU x 90 mémoire x40

Annexe 2 Tableau comparatif des résolutions des modèles couplés du projet européen ENSEMBLES

Scientific/technical questions:

- What is the current estimate of uncertainties in climate projections using state-of-the-art global coupled models ?
- Can this estimate of uncertainties be improved by using the earth-system models and ensemble prediction system which will be delivered by RT1?
- How to combine the best available climate simulations with different global coupled models to produce probabilistic estimates of climate change projections for use in other applications?

Table 2A. Combination of atmosphere-ocean models used to produce the coupled simulations in RT2A.

The 'resolution' column gives the triangular truncation in the case of a spectral model, or the horizontal grid resolution in degrees over latitude x longitude in the case of a grid point model. For ocean models a range of resolutions is indicated since these models have generally irregular grids with higher resolution near the equator. The number of levels in the vertical is given in column 'Lev'. The workpackages of RT2A in which the models are applied are specified in column 'WP'.

Partners	Atmosphere	resolution	Lev	Ocean	resolution	Lev	WP
METO-HC	HC-AGCM	1.25°x1.875°	38	HC-OGCM	0.33-1°	40	1, 2,3
UREADMM METO-HC	UK-HiGEM	1°x1°	38	HC-OGCM	0.33°	40	3
IPSL UCL-ASTR	LMDZ-4	2.5°x3.75°	19	ORCA	0.5-2°	31	2, 3
MPI DMI	ECHAM5	T63	31	MPI-OM	1.5°	40	2, 3
INVG	ECHAM4.6	T106	19	OPA 8.2	0.5-2°		2, 3
FUB	EGMAM	T31	39	EGMAM	0.5-2.8°	20	2, 3
ECMWF	IFS	T95	60	ORCA	0.5-2°	29	1
CERFACS	ARPEGE	T63	31	ORCA	0.5-2°	31	1
CNRM	ARPEGE	T63	31	ORCA	0.5-2°	31	1
			45	OPA8	0.5-2°	31	2,3
NERSC	ARPEGE	T63	31	MICOM	1.2°	36	2, 3

Annexe 3

Les enjeux de modélisation pour la communauté climatique

Le climat de la Terre est le résultat d'interactions complexes entre de nombreux processus faisant intervenir l'atmosphère, l'océan et les surfaces continentales. Comment fonctionne ce système ? Peut-on prévoir son évolution à l'échelle d'une saison ou à plus long terme ? Les activités humaines sont-elles en train de modifier les grands équilibres climatiques ? Quelles en sont et seront les conséquences pour l'humanité ? Autant de questions qui rendent particulièrement sensibles les recherches sur le climat et son évolution, et mettent sous le feu des projecteurs les résultats des modèles de climat.

Il est donc plus que jamais nécessaire de comprendre les mécanismes essentiels qui gouverneront l'évolution du climat. Une difficulté majeure de ces études est que la perturbation anthropique, qui apparaît comme une dérive du système, se superpose et interfère avec la variabilité naturelle dont de nombreux aspects sont encore mal connus. L'enjeu de la recherche climatique est donc de déterminer les éléments clés du système en fonction des échelles de temps abordées, d'évaluer l'impact de l'homme (émissions de gaz à effet de serre et d'aérosols,

utilisation des sols, déforestation, etc.) sur le changement climatique aux échelles globales et régionales et d'évaluer le risque de "surprises climatiques" dans le futur.

Seule une approche multidisciplinaire et intégrée du Système Terre peut relever ce défi, en associant les études du climat actuel, des climats passés et les projections climatiques. Elle doit faire intervenir à la fois les interactions entre l'océan, l'atmosphère, la biosphère et la cryosphère sous l'ensemble de leurs aspects physiques, chimiques et biologiques. Les avancées scientifiques attendues dans ce domaine sont très fortement dépendantes des ressources en calcul intensif. Notre capacité à prévoir les modifications climatiques majeures susceptibles de se produire au cours des prochaines années ou des prochains siècles est fortement conditionnée par notre capacité à intégrer dans des ensembles complexes des processus ayant des échelles de temps et d'espace très variés et par notre capacité à les évaluer par différents types de simulations. Elle repose aussi sur différentes études de prévisibilité climatiques à l'échelle saisonnière, pluri-annuelle, voire à l'échelle du siècle.

Au niveau international, le grand défi pour la communauté scientifique à l'échelle de 10 ans, est de disposer de systèmes intégrés très complets du système Terre, intégrant la physique et les cycles bio-géochimiques, et de se diriger pour une part vers la prévision climatique. Cet enjeu est particulièrement dimensionnant pour les ressources de calcul à l'heure où les possibilités offertes par des calculateurs comme le "Earth Simulator" ouvrent de nouvelles perspectives de recherches dans le domaine du climat et de l'étude des processus intervenant dans les différents milieux qui le composent. La communauté est aussi fortement mobilisée pour le prochain rapport du Groupe Intergouvernemental sur l'Etude du changement climatique qui rendra ses conclusions en 2007 (www.ipcc.ch). D'ici là, les différents groupes de modélisation sont fortement sollicités pour produire dès la fin 2004 un ensemble coordonné de simulations des 20^{ème}, 21^{ème} et 22^{ème} siècles. Au niveau européen ces défis sont repris par le Projet Intégré du FP6 ENSEMBLES (coordinateur D. Grigs, Hadley Centre, UK). Ce projet, qui devrait démarrer au cours de l'année 2004 pour une période de 5 ans, fait le lien entre les scénarios climatiques, la prévisibilité climatique, la régionalisation, les études d'impacts et l'économie. Il regroupe 70 partenaires européens. D'autres projets internationaux ou européens viendront renforcer cet effort. C'est en particulier le cas pour les simulations paléoclimatiques avec le projet international PMP (Paleoclimate Modeling Intercomparaison Project) qui regroupe des simulations couplées pour des périodes extrêmes dans le passé pour lesquelles il existe de nombreuses données.

La communauté française est fortement intégrée à ces différents projets. Les aspects les plus dimensionnants en terme de ressource de calcul concernent les études nécessitant le couplage de différentes composantes du système terre. Dans la suite logique des simulations climat-carbone effectuées précédemment, l'IPSL développe un ensemble de thématiques permettant d'analyser et de quantifier le rôle des interactions entre le système climatique et les cycles biogéochimiques. Cette activité comprend: l'analyse du couplage entre le climat et le cycle du carbone, l'étude de l'impact du changement d'utilisation des sols sur le climat et le cycle du carbone, et la quantification du rôle des différents gaz à effets de serre et des aérosols. Les analyses s'appuieront sur des expériences numériques où les différentes composantes du modèle système Terre de l'IPSL sont successivement activées, ainsi que sur des simulations avec les composantes individuelles. Cette approche permet de quantifier les échanges entre les différents milieux et d'évaluer le rôle des rétroactions. Un objectif est de disposer d'un modèle complet (ensembles des couplages pouvant être activés simultanément) d'ici fin 2006.

Les volets d'analyse du changement de climat simulé mettent l'accent sur les changements du cycle hydrologique, au travers des rétroactions liées aux variations de la circulation océanique et aux rétroactions nuageuses. Ces études demandent de pouvoir augmenter rapidement la résolution spatiale des modèles d'océan et d'atmosphère utilisés aujourd'hui pour avoir une meilleure représentation des échelles synoptiques dans l'atmosphère et de la topographie dans les régions de l'atlantique Nord ou de nombreux seuils conditionnent le cheminement des masses d'eau. La validation du modèle et de ses différents modules et l'évaluation des résultats s'appuient sur l'ensemble des recherches menées dans les différents laboratoires de l'IPSL pour le climat actuel et sa variabilité et les climats passés. Les études directement menées à l'IPSL partir de scénarios climatiques globaux sont donc essentiellement axées sur l'étude des rétroactions dans le système climatique et l'analyse de la sensibilité climatique.

Météo-France est également fortement impliqué dans les études du changement climatique. Bien que s'intéressant aux rétroactions du système climatique dans son ensemble en liaison avec le cycle hydrologique et les processus de surface, le projet de Météo-France met d'avantage l'accent sur des études aux échelles régionales. Cela concerne à la fois des études de variabilité climatique (Atlantique Nord-Europe, régions de moussons, régions polaires), le développement d'une méthodologie de régionalisation dynamique des changements climatiques, et des études de détection et d'attribution des changements climatiques aux échelles régionales en collaboration avec le CERFACS.

A partir des configurations de modèles utilisées aujourd'hui pour l'ensembles de ces études (cf tableau), les éléments les plus contraignants et dimensionnants pour continuer à faire une recherche de qualité dans le domaine et continuer à être reconnus au niveau international concernent :

- l'augmentation de la résolution des modèles pour avoir une meilleure description des échelles synoptiques, représenter les modes de variabilité de l'atmosphère et inclure les niveaux stratosphériques,
- la complexification du système pour tenir compte non seulement de l'impact du changement de climat sur les cycles biogéochimiques, mais aussi du couplage entre les cycles et le climat,
- la réalisation d'ensembles de simulations pour tenir compte des incertitudes des modèles et de l'aspect chaotique du système.,
- la réalisations de simulations longues pour obtenir une bonne caractérisation de la variabilité climatique à différentes échelles de temps,
- la possibilité d'étudier pouvoir étudier les interactions d'échelles et les non-linéarités du système,
- de simulations régionales à haute résolution pour pouvoir traiter correctement les problèmes d'impact et faire le lien avec les mesures ponctuelles.