

Plate-forme MIRAGE : Méso-Informatique Répartie pour des Applications en Géophysique et Environnement

Rapport d'activité 2003

1 Acteurs du projet et objectifs scientifiques

La modélisation numérique est un outil de recherche privilégié dans la plupart des domaines de la géophysique et des sciences de l'environnement. C'est donc de façon "naturelle", à partir de nombreuses collaborations pré-existantes, qu'a été créé en 1999 le projet MIRAGE. Celui-ci regroupe des chercheurs de 5 laboratoires grenoblois travaillant sur le thème de la modélisation numérique du climat et de l'environnement :

LECE	Laboratoire Ecosystèmes et Changements Environnementaux (EA3112)
LEGI	Laboratoire des Ecoulements Géophysiques et Industriels (UMR5519)
LGGE	Laboratoire de Glaciologie et Géophysique de l'Environnement (FRE2192)
LMC	Laboratoire de Modélisation et Calcul (UMR5523)
LTHE	Laboratoire d'étude des Transferts en Hydrologie et Environnement (UMR5564)

Les objectifs scientifiques et techniques du projet sont :

- Se doter localement d'une plate-forme de calcul commune, complémentaire du service fourni par les centres nationaux, et qui permette de gagner en efficacité pour le développement et la mise au point des codes de calcul, les travaux de pré- et post-traitement, et la réalisation des simulations de taille moyenne.
- Favoriser le partage de codes de calcul et d'outils informatiques et numériques
- Favoriser le dialogue et le partage de compétences scientifiques et techniques entre chercheurs

L'unité thématique du projet et sa taille "humaine" sont à cet égard des facteurs prépondérants.

2 Aspects matériels et financiers

2.1 Configuration matérielle

2.1.1 Configuration actuelle

La plate-forme de calcul MIRAGE est constituée actuellement de 7 stations de travail bi-processeurs Compaq DS-20 et d'une station quadri-processeurs ES-45.

La configuration des DS-20, qui ont bientôt 4 ans, est la suivante: processeurs EV6 (500Mhz), 2.5 Go de RAM, 5 Go de swap et environ 30 Go de mémoire disque disponible par machine.

L'ES-45 a été acquise en 2003. Elle est basée sur des processeurs EV68 (1250Mhz), et comporte 4 Go de RAM, et 90 Go de disque.

Ces machines sont réparties sur 2 sites (contre 3 précédemment): 4 DS-20 au LEGI-LTHE, 3 DS-20 et l'ES-45 au LMC. Sur chaque site, les machines sont reliées entre elles

par une liaison constructeur “Memory Channel” à très haut débit. Les sites sont reliés entre eux par le réseau du campus universitaire (liaison Gigabit ethernet).

La plate-forme dispose de plus de 3 baies-disques sécurisées (RAID-5) d’une centaine de Go chacune. Elle est par ailleurs équipée des outils logiciels standards de calcul scientifique.

2.1.2 Evolutions de l’année 2003

L’évolution matérielle majeure en 2003 a été l’achat du quadri-processeur ES-45, grâce à un financement de l’INRIA. Ceci a permis d’augmenter la capacité de calcul d’environ 60%. Par ailleurs, nous avons profité de la réorganisation due à l’arrivée de cette machine pour rapatrier au LMC la machine DS-20 qui était auparavant au LGGE, afin de la mettre avec sa baie RAID en mini-cluster avec l’ES-45. Ceci a permis au passage de supprimer le problème de la liaison à relativement faible débit (100 Mo/s) avec le LGGE. Les machines ne sont donc plus réparties maintenant que sur 2 sites.

2.1.3 Evolutions à venir

Une jouvence complète de la plate-forme est prévue fin 2004-début 2005, dans le cadre du contrat de plan état-région (228 Keuros prévus). A cette époque, les bi-processeurs DS-20 seront devenues quasiment obsolètes, tandis que l’ES-45 pourra être maintenue encore sans doute un à deux ans.

Les choix techniques pour cette jouvence seront discutés dans l’année qui vient. On s’oriente en tout cas très certainement vers une gestion des machines sur un site unique. Rappelons en effet que un souhait fort exprimé par le COMI-CNRS et le MENRT lors de la genèse de Mirage en 1998-1999 était que nous expérimentions une architecture de calcul physiquement répartie sur plusieurs sites. Ceci complique toutefois la gestion quotidienne des machines et la fiabilité de l’ensemble. Il nous semble qu’un tel choix n’est plus d’actualité aujourd’hui pour une machine de la taille de Mirage, l’enjeu du calcul sur grille étant plutôt de faire communiquer entre elles différentes plate-formes.

2.2 Rappel de la chronologie

La chronologie de mise en place de ces moyens a été la suivante :

- 1999 Montage du dossier, demandes de financement, choix du matériel
- 01/2000 Arrivée des machines
- 05/2000 Finalisation d’une première configuration de travail satisfaisante
- 09/2000 Recrutement de Laurence Viry, IR UJF en calcul scientifique, et responsable technique de la plate-forme.
- 01/2001 Ajout d’un routeur sur le site LEGI-LTHE (amélioration de la stabilité et du débit du réseau)
- 05/2001 Passage à la version 5.1 de *True64 Unix*, qui apporte une gestion “cluster” des machines sur chaque site et une meilleure fiabilité du système.
- 09/2001 Installation de LSF, logiciel gestionnaire de batchs.
- 02/2003 Installation de l’ES45 et reconfiguration en deux sites

2.3 Budget

Le budget initial d'acquisition et installation du matériel s'élevait à 235 Keuros, doté par le CNRS-COMI, l'UJF/MENRT, l'IMAG, l'INPG, l'INSU, et une participation des laboratoires.

Le budget de fonctionnement s'élève environ à 12-15 Keuros par an, doté environ pour moitié par les organismes (30 Keuros UJF sur 1999-2001, 15 Keuros INRIA sur 2001-2003, 15 Keuros CPER en 2003), l'autre moitié étant de l'auto-financement via des contrats de recherche.

L'upgrade préparé mi-2001 (achat de disques et upgrade système) et finalisé début 2003 (achat de l'ES-45) a été réalisé grâce à une dotation de l'INRIA (75 Keuros, répartis sur 2001, 2002 et 2003).

3 Utilisation et formation

3.1 Utilisation des ressources

D'après les chiffres fournis par LSF (Load Sharing Facility), le taux d'utilisation des machines est bon, les temps d'attente pour la soumission des jobs restent peu élevés pour les jobs séquentiels, un peu plus importants pour les jobs parallèles car pour des raisons de performance on impose l'exclusivité du processeur pour chaque processus. Depuis avril 2003, on a pu observer que l'utilisation de l'ES45 (4 processeurs) en terme de temps CPU consommé est multipliée par 3 par rapport au temps moyen consommé par chacun des DS20.

Les machines essentiellement utilisées en séquentiel dans un premier temps sont de plus en plus utilisées par des applications parallèles utilisant un modèle:

- à mémoire distribuée
 - sur le cluster de 4 DS20 et la bibliothèque constructeur MPI bénéficiant des performances du réseau Memory Channel constructeur (les spécificités MPI utilisées sont limitées au standard MPI-1)
 - la bibliothèque LAM/MPI et le réseau ethernet lorsque l'application fait appel à des spécificités de la norme MPI-2 comme la gestion dynamique des processus dans des applications de couplage de codes par exemple, les E/S parallèles ou l'utilisation du parallélisme inter-cluster. L'utilisation de la version MPICH a été abandonnée car elle n'apporte par rapport à ces deux bibliothèques ni les performances ni le standard MPI-2
- à mémoire partagée avec les directives de compilation OpenMP. Ce type d'application est favorisé par l'arrivée en début d'année de l'ES45 quadri-processeurs (ce modèle de programmation parallèle étant simple à implémenter et efficace lorsque l'application est adaptée).

3.2 Complémentarité avec les centres nationaux

L'usage de la plate-forme Mirage est bien conforme à nos souhaits initiaux. Les ressources sont essentiellement utilisées pour des tâches expérimentales de mise au point

d’algorithmes et de méthodes numériques, et de réalisation de simulations relativement légères. Par contre, tous les travaux de production nécessitant des moyens lourds sont systématiquement réalisés sur les calculateurs nationaux.

A cet égard, la plate-forme Mirage est “processeur-compatible” avec les machines Compaq de l’IDRIS (Vedia, 6 noeuds quadri-processeurs ES40) et du CCRT-CEA (Ixia, 800 processeurs), ce qui permet une bascule extrêmement aisée des codes de calcul fonctionnant sur Mirage avec ces deux machines.

3.3 Objectifs de développement

Dans le domaine de l’environnement et du climat, le couplage de modèles est un point de passage obligé pour aller vers la simulation de nombreux phénomènes complexes. Du point de vue technique, le couplage de codes d’origines différentes (équipes de développement, structures de données, architectures matérielles, langages...) pose assez souvent des problèmes lourds et complexes d’adaptation.

Un premier travail de réflexion et d’expérimentation (CORBA, C++, JAVA...) a été réalisé sur la définition d’interface de couplage apportant l’interopérabilité, la coordination et la réutilisation des codes en utilisant les mécanismes des langages objets comme l’encapsulation, le polymorphisme et l’héritage.

3.4 Formation

Deux interventions de formation:

- dans le cadre de la formation doctorale, un module sur la décomposition de domaine pour la résolution d’équations aux dérivées partielles par E. Blayo et L. Viry
- Organisation par L. Viry avec l’aide de F. Roch et F. Berthoud d’une formation “Introduction au calcul scientifique” dans le cadre des Universités d’automne financées par le ministère de la recherche.

4 Production scientifique

Comme nous l’avons indiqué précédemment, les thématiques scientifiques abordées dans le cadre de ce projet sont centrées sur la modélisation numérique de l’environnement et du climat. Plus précisément, les principaux thèmes de recherche développés par les utilisateurs de la plate-forme Mirage sont essentiellement les suivants :

- Modélisation des calottes polaires
- Couplage entre glace et atmosphère
- Modélisation des circulations océaniques
- Assimilation de données en océanographie physique et biologique
- Modélisation de l’interface sol-végétation-atmosphère
- Modélisation hydrologique couplée à la modélisation météorologique de méso-échelle
- Modélisation de la chimie atmosphérique à l’échelle d’une vallée alpine

Plusieurs collaborations de recherche existent entre les chercheurs impliqués dans Mirage, sur des thèmes complémentaires ou transverses, comme par exemple le couplage de modèles ou le raffinement de maillage, et qui impliquent des co-directions de thèses.

L'objectif n'est pas ici de fournir un rapport d'activité scientifique complet sur tous ces thèmes. Nous avons plutôt choisi ici de donner quelques illustrations, pour la réalisation desquelles l'outil Mirage s'est révélé particulièrement adapté.

4.1 Pollution des vallées alpines

Intervenants: J.P. Chollet, G. Brulfert, E. Chaxel, C. Chemel, M.A. Kerbirou, C. Staquet (LEGI)

L'évolution de la pollution de l'atmosphère est devenue une préoccupation grandissante de la population et des pouvoirs publics. Les phénomènes mis en jeu sont complexes avec des interactions entre la dynamique de l'atmosphère qui assure le transport et la dispersion des polluants et la chimie de la basse troposphère qui produit différents polluants secondaires à partir des émissions primaires dues au trafic routier, aux activités domestiques et industrielles, ainsi qu'à la végétation. Le problème du trafic routier est particulièrement sensible pour les vallées alpines parcourues par de grands axes de transit international, les circulations atmosphériques y sont particulières avec les régimes locaux de vents de pente, vents de vallée et inversion. Ces phénomènes essentiellement d'origine thermique sont étroitement associés à l'ensoleillement et à son évolution en cours de journée.

L'incendie du tunnel du Mont Blanc a entraîné, pendant plusieurs mois, une réduction drastique du trafic routier dans la vallée de Chamonix avec un report vers la vallée de la Maurienne. Le projet POVA (Pollution des Vallées Alpines) compare les niveaux de pollution avant et après réouverture du tunnel dans les deux vallées. Le développement d'une modélisation numérique est associé aux campagnes de mesure sur site de manière à: (1) compléter des mesures nécessairement très ponctuelles pour l'analyse des phénomènes mis en jeu, (2) étendre l'étude au-delà des situations atmosphériques et de pollution rencontrées pendant les semaines de campagne, (3) produire des scénarios types pour l'évaluation de l'impact de politiques locales d'abattement, réduction de trafic poids lourd par exemple, (4) mettre en place les premiers éléments d'un modèle opérationnel de prévision.

L'absence de rétro-action significative de la chimie sur la dynamique conduit à une stratégie d'utilisation découplée des outils de modélisation de l'atmosphère locale (le code ARPS), et de l'outil de modélisation de la photochimie de la basse troposphère (le code TAPOM). Les champs de vitesse, pression, température, humidité sont périodiquement transmis du premier au deuxième maillon de cette chaîne de calcul. Même dans les conditions les plus propices au développement de la pollution chimique, les circulations de vent locales aux vallées dépendent de l'atmosphère aux méso échelles régionales. Un forçage est donc exercé par des champs météo à travers une imbrication en un ou deux niveaux.

Les calculs sont relativement lourds puisque les temps caractéristiques associés à la dynamique sont aussi petits que quelques secondes. Les maillages sont typiquement de

83x93 noeuds en horizontal et 30 niveaux verticaux. Pour une photochimie réaliste, 77 concentrations d'espèces chimiques interagissent dans 191 étapes de réaction. Le module de dynamique atmosphérique est parallélisé à travers les procédures MPI. Les réponses du modèle sont fortement conditionnées par la qualité des données. Celles-ci, fournies ou établies en collaboration avec Air APS, portent sur la topographie, la nature de la canopée, l'occupation du sol et les cadastres d'émission.

La figure 1 compare la concentration d'ozone calculée par le modèle à la valeur mesurée durant une journée de la campagne POVA d'août 2000. Les niveaux sont correctement reproduits ainsi que l'évolution avec l'heure de la journée. La figure 2 illustre un exemple d'analyse des modes de production de lozone grece à un index construit comme le rapport des concentrations dozone et doxydes dazote O_3/NO_x . La différence de concentration dozone (en ordonnée) est calculée par différence entre simulation avec toutes les émissions et simulation avec une réduction arbitraire de 50% de COV ou de NO_x . On constate que c'est la diminution des COV qui conduit à la réduction de la concentration d'ozone. On pourrait en déduire, au moins dans la situation de cette journée, une stratégie d'abattement des émissions.

Références :

G. Brulfert, C. Chemel, J.P. Chollet: Numerical simulation of air quality in Chamonix valley, impact of road traffic 16/6-18/6 2003, Avignon. Actes INRETS n92- vol 2, pp 39-44. ISSN 769-0266, ISBN 2-85782-588-9. 12th International Scientific Symposium Transport and Air Pollution (INRETS, TUG, NCAR).

G. Brulfert, J.P. Chollet, C. Chemel, and M.A. Kerbirou: Numerical simulation of photochemical episodes in alpine valleys. 26th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application. 26/5-30/5 2003, Istanbul, Turkey.

G. Brulfert, E. Chaxel, C. Chemel and J.P. Chollet: Numerical simulation of air quality in Chamonix valley, use of different chemistry indicators. 14th IUAPPA International Conference Air quality Assessment and policy at local, regional and global scales. Dubrovnik, Croatia, 6/10-10/10 2003.

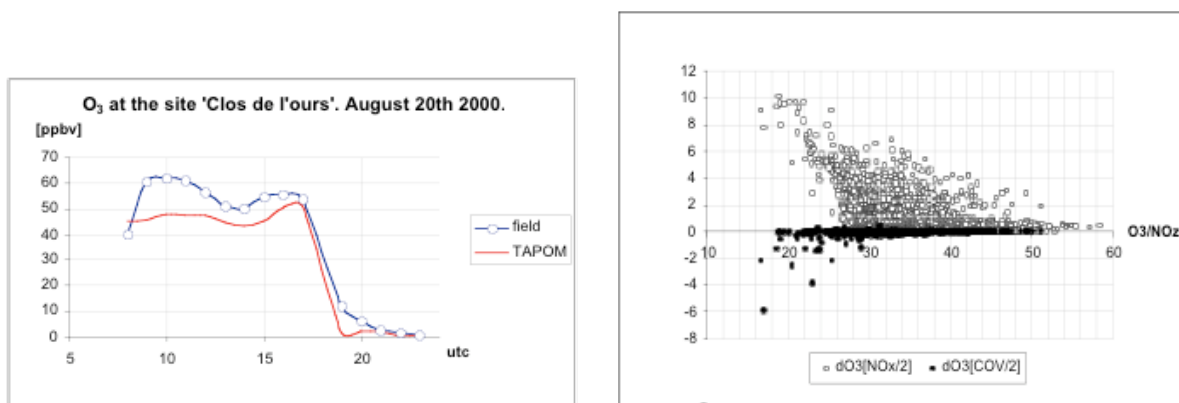


FIG. 1 – A droite : simulation d'une journée d'été, évolution de la concentration d'ozone. A gauche : O_3/NO_x et différence de concentration d'ozone avec réduction des COV ou des NO_x .

Le programme POVA est soutenu par la région Rhone Alpes et par l'ADEME dans le cadre du programme PRIMEQUAL.

4.2 Modélisation de l'écoulement d'eau dans des sols non saturés fortement hétérogènes

Intervenants : J. Lewandowska, A. Szymkiewicz, M. Vauclin (LTHE)

Ce travail, sujet de thèse d'A. Szymkiewicz, concerne les sols caractérisés par la présence de deux régions distinctes aux paramètres hydriques (conductivité) très contrastés (sols à double porosité). A cause de cette structure particulière, il est nécessaire d'utiliser des méthodes de modélisation mathématique d'écoulement adaptées, afin de prendre en compte l'influence des hétérogénéités locales sur le comportement macroscopique.

Le modèle mathématique d'écoulement à l'échelle locale est l'équation de Richards, formulée par rapport à la pression capillaire. C'est une équation de type parabolique, fortement non linéaire, car les coefficients de capacité et de conductivité dépendent de la pression capillaire. Dans le cas considéré, la difficulté de la modélisation numérique (non-linéarité) est encore augmentée à cause de la structure hétérogène et du contraste local de paramètres. Il est donc souhaitable (voire nécessaire) de remplacer la description mathématique à l'échelle locale par une description à l'échelle macroscopique - milieu "homogénéisé" avec des paramètres effectifs.

L'objectif principal est le développement de modèles d'écoulement à l'échelle macroscopique par méthode d'homogénéisation asymptotique. Cette méthode permet d'obtenir les modèles macroscopiques et les paramètres effectifs à partir de la description du phénomène physique à l'échelle locale, sans hypothèses auxiliaires sur la forme finale du modèle.

Dans le cadre de ce travail plusieurs calculs numériques sont effectués sur Mirage. Ils concernent principalement les deux problèmes suivants:

– *Résolution d'une équation macroscopique obtenue par homogénéisation*

Dans le cas d'un milieu bien perméable avec des inclusions faiblement perméables, le modèle macroscopique développé a une forme d'équation intégro-différentielle. Un modèle numérique a été développé et implanté dans le code DPOR-1D. Il permet la résolution de l'écoulement macroscopique 1D avec des inclusions 2D ou 3D. A cause de la résolution d'un grand nombre de problèmes locaux 2D/3D, le temps de calcul devient important (2-12 h sur Mirage en séquentiel).

– *Résolution d'équation de Richards à l'échelle locale*

La validation des modèles obtenus par homogénéisation est effectuée par comparaison avec une solution numérique à l'échelle fine (3D/2D), où la structure hétérogène du milieu est explicitement représentée (codes commerciaux SWMS-2D et SWMS-3D). Ce type de calcul ne peut être fait que sur les machines de grande puissance (temps de calcul typique: 4-10 jours en séquentiel).

A titre d'exemple, sont présentés ci-dessous les résultats de calcul pour un problème d'infiltration 1D dans une colonne du sol avec des inclusions peu perméables par rapport à la matrice (Fig. 1). Ils sont obtenus avec des modèles différents: modèle à simple porosité (où l'influence des inclusions est négligée), modèle obtenu par

homogénéisation, solution à l'échelle fine (solution de référence). La distribution de la pression capillaire obtenue par homogénéisation est comparée avec des résultats de la solution à l'échelle fine (Fig. 2). Le flux d'eau et l'évolution de la pression capillaire à la sortie de la colonne sont présentés en Fig. 3 et 4. La solution par homogénéisation est proche de la solution à l'échelle fine. Nous voyons bien que l'écoulement est ralenti par rapport au modèle à simple porosité (à cause de l'échange d'eau avec les inclusions). Cet effet est connu comme "l'effet de traîne".

Références :

Lewandowska J., A. Szymkiewicz, K. Burzyski, and M. Vauclin, 2003: Modeling of unsaturated water flow in double-porosity soils by the homogenization approach. A paraître dans *Advances in Water Resources*.

Lewandowska J., A. Szymkiewicz and J.-L. Auriault, 2003: Modeling of unsaturated water flow in soils with highly conductive inclusions by homogenization. Soumis à *Advances in Water Resources*.

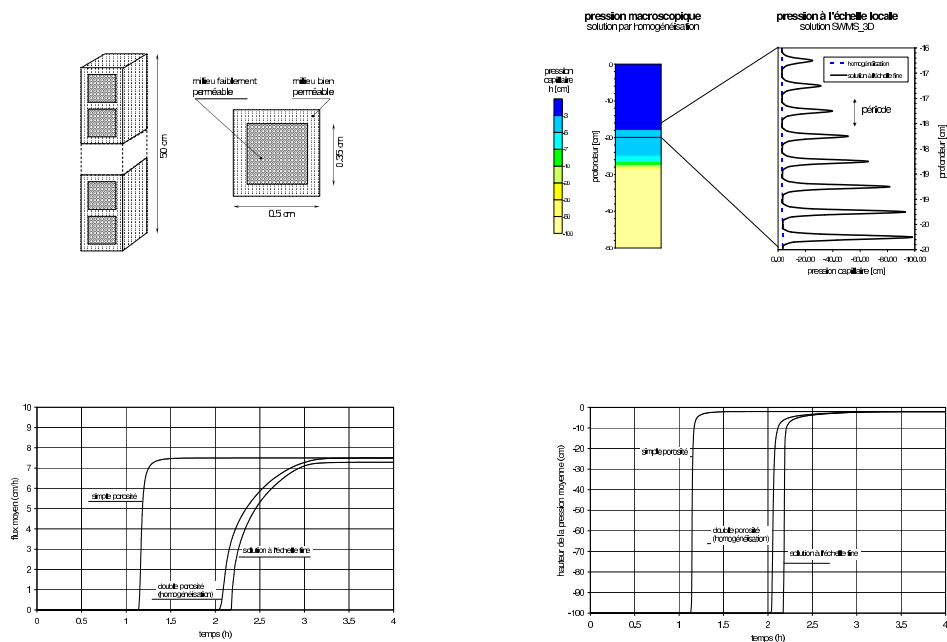


FIG. 2 – En haut : géométrie du milieu et distribution de la pression capillaire dans la colonne après 1h d'infiltration. En bas : flux moyen et pression moyenne à la sortie de la colonne.

4.3 Composition isotopique des calottes de glace

Intervenants: N. Lhomme, C. Ritz (LGGE)

Au cours des deux derniers millions d'années, le climat de la Terre a été marqué par une série de cycles oscillant entre de courtes périodes aux températures similaires au présent, et des longues périodes glaciaires où de vastes calottes de glaces recouvraient le nord de l'Europe et de l'Amérique. Afin de mieux comprendre les mécanismes responsables de ces larges bouleversements climatiques, des campagnes de mesures ont été entreprises au Groenland et en Antarctique pour étudier le comportement de ces masses de glace et extraire de la glace des informations sur notre passé. Des forages ont été effectués sur ces deux grandes calottes polaires pour en extraire des carottes de glace et mesurer la composition chimique des couches de glace, car la proportion d'isotopes dans la glace dépend d'abord des températures qui régnaient au-dessus des régions où la neige est tombée. Des modèles simples permettent d'estimer l'âge de la glace et la température à la surface de la calotte. Cependant, de larges incertitudes quant à la durée et à l'amplitude des événements enregistrés demeurent. Par exemple, l'écoulement tri-dimensionnel d'une calotte de glace et sa réaction complexe à des changements de température déforme les couches (modifiant l'âge apparent) et influe sur l'altitude de surface (modifiant la température apparente, et aussi la quantité de neige qui tombe et donc l'épaisseur des couches). L'altitude, l'étendue et le volume d'une calotte de glace a aussi une influence sur le comportement de l'atmosphère et la composition et la température des océans, qui en retour conditionne la composition de l'eau dans les nuages. Autant de complexité que de simples modèles de comptage des couches de glace ne peuvent incorporer, mais incertitudes cruciales lorsqu'il s'agit de prédire à quel point le Groenland et l'Antarctique ont fondu dans le passé et dans quelle mesure ils peuvent demain contribuer à la montée du niveau des mers.

Pour mieux comprendre ces enregistrements climatiques, nous avons développé au Laboratoire de Glaciologie et de Géophysique de l'Environnement des modèles qui expriment la complexité des calottes de glace et qui permettent de simuler ces fameuses carottes de glace. Ces modèles s'appuient sur la résolution tri-dimensionnelle d'équations dynamiques et thermodynamiques qui régissent l'évolution de vastes calottes de glace. La résolution de ces équations est très intensive en terme de calculs, d'autant plus que notre problématique nécessite de suivre l'évolution de ces masses de glace pour des périodes couvrant des centaines de milliers d'années. Une seule simulation pouvait durer deux à trois semaines il y a trois ans à peine. Grâce notamment à la puissance de calcul du réseau de machines de Mirage, il est aujourd'hui possible de se plonger dans la complexité de ces mécanismes climatiques et de tenter de répondre de manière plus documentée à la problématique scientifique des changements de climat passés et futurs.

La figure ci-jointe montre notre estimation de la contribution de l'eau de fonte du Groenland à la montée du niveau des mers simulé par notre modèle de calotte de glace lors de la dernière période climatique similaire au présent, pour les scénarios climatiques les plus probables. Suivant le cas de figure envisagé, le Groenland augmente le niveau des mers de 0,80 à 4,40m, pour un même enregistrement climatique simulé. Les derniers développements de nos modèles qui simulent à la fois le Groenland et l'Antarctique et toutes les carottes de glace qui y ont été prélevées tendent à montrer que le Groenland

serait le principal responsable de la montée du niveau marin enregistré lors de la dernière période inter-glaciaire.

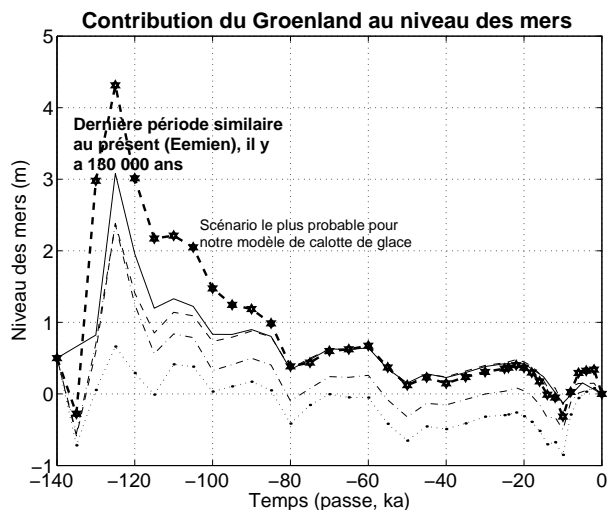


FIG. 3 – Contribution de l’eau de fonte du Groenland à la montée du niveau des mers au cours du dernier cycle glaciaire simulé par notre modèle de calotte

4.4 Assimilation de données de rang réduit en océanographie

Intervenants : E. Blayo, C. Robert (LMC)

Les méthodes d’assimilation de données sont des techniques mathématiques issues des théories du contrôle optimal et du filtrage stochastique. Leur but est de fournir une estimation de l’état d’un système en mixant différentes sources d’information de natures différentes (typiquement un modèle numérique du système et des données d’observation). Ces méthodes sont très utilisées notamment en météorologie et en océanographie, car une connaissance aussi bonne que possible de la condition initiale est absolument nécessaire en vue de la prévision. Or l’état présent de l’atmosphère et de l’océan n’étant que très partiellement observé, il doit être reconstitué au mieux par ces techniques en assimilant les mesures disponibles dans les modèles numériques.

Ces méthodes sont extrêmement coûteuses en place mémoire et temps calcul, et accroissent typiquement le coût d’un modèle numérique d’un à deux ordres de grandeur. De nombreuses recherches sont donc en cours actuellement pour développer de nouvelles méthodes moins coûteuses.

C’est le cas notamment au LMC et au LEGI, où l’on s’intéresse à la mise au point de méthodes d’assimilation de rang réduit. L’idée de base consiste à ne corriger les prévisions du modèle que dans un nombre restreint de directions privilégiées, choisies à l’avance en fonction de la dynamique du système. Ainsi, on utilisera typiquement un sous-espace de correction de dimension 20 ou 30, alors que la variable d’état du modèle est de dimension $10^6 - 10^7$. La pertinence du choix de ces directions est donc un aspect essentiel pour espérer une bonne efficacité de la méthode.

Nos travaux récents mettent en évidence l'intérêt de deux familles de vecteurs : les EOFs (issus d'une analyse statistique de trajectoires précédentes du modèle) et les *bred modes* (vecteurs d'amplification maximale). Nous avons montré qu'une approche variationnelle de rang réduit peut permettre de réduire significativement le temps de calcul par rapport à une méthode classique opérant dans tout l'espace d'état, tout en améliorant le processus d'identification grâce à l'information supplémentaire contenue dans la base de correction. Ce résultat, obtenu dans un premier dans une configuration simplifiée dite d'expériences jumelles, a été étendue récemment au cas beaucoup plus difficile d'expériences avec données réelles.

Références :

Blayo E., S. Durbiano, P. Vidard and F.-X. Le Dimet, 2003: Reduced order strategies for variational data assimilation in oceanic models. Dans *Data Assimilation for Geophysical Flows*, Springer-Verlag, à paraître.

Durbiano S., E. Blayo, J. Verron, J. Blum, F.-X. Le Dimet and C. Robert, 2003: A reduced order strategy for 4D-Var data assimilation. Soumis à *Quart. J. Roy. Meteorol. Soc.*.

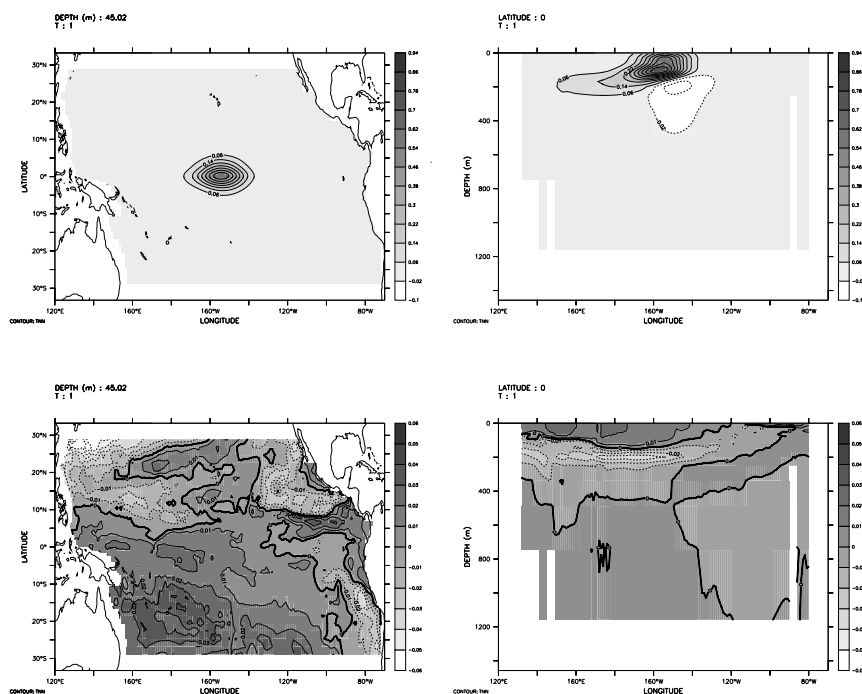


FIG. 4 – Un exemple de l'apport d'une base réduite : correction optimale sur le champ de température lors de l'assimilation d'une seule mesure ponctuelle dans un modèle réaliste de l'océan Pacifique équatorial. On montre ici la structure de cette correction (à gauche : coupe horizontale à 50m de profondeur; à droite : coupe verticale au niveau de l'équateur), pour l'approche usuelle avec l'espace d'état complet (en haut), et pour l'approche de rang réduit (en bas). L'information nouvelle apportée par la base réduite enrichit de façon très significative le contenu physique de la correction optimale.