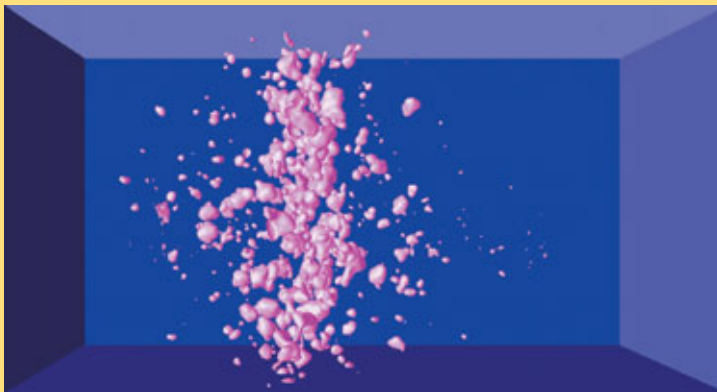




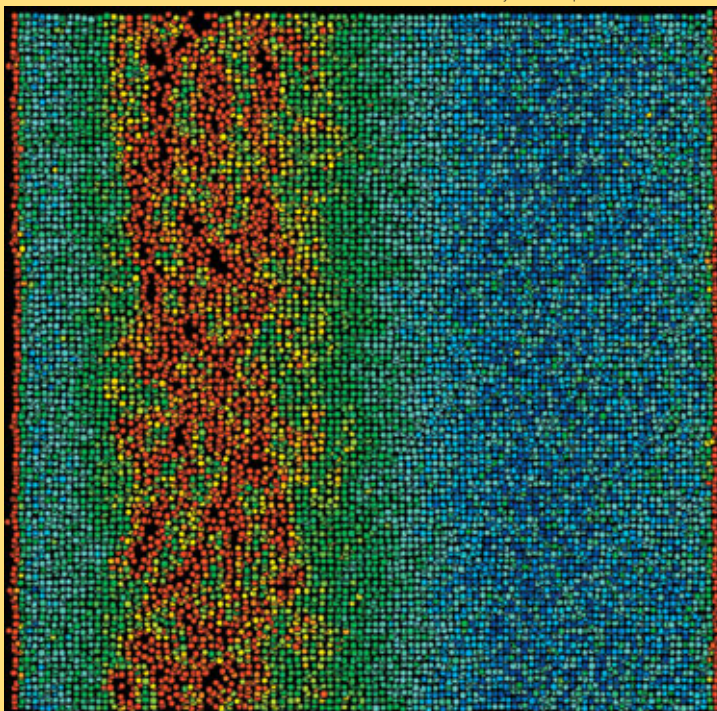
I. SIMULER POUR COMPRENDRE

Plus encore que dans le développement technologique, où elle a fait la preuve de son efficacité, la simulation occupe un rôle particulier dans la recherche scientifique, puisque c'est précisément à elle de traduire les effets des phénomènes fondamentaux, bien que certains puissent être insuffisamment connus ou même inconnus. Les quelques sujets de ce chapitre montrent comment la simulation aide à comprendre les phénomènes, mais aussi qu'elle n'est qu'un maillon de la compréhension, entre théorie, modélisation et expérimentation, cette dernière restant le "juge de paix". Ceci est par exemple spectaculaire dans le domaine des matériaux, où elle jette un pont entre la théorie des phénomènes à l'échelle atomique et la constatation des effets macroscopiques. En climatologie, la modélisation de l'ensemble constitué par l'atmosphère et les océans est indispensable pour prédire le sort de la Planète. Les nanotechnologies ont besoin de la modélisation des surfaces, interfaces et autres nanostructures. Les biotechnologies, au sens large, ne peuvent plus se passer de la modélisation moléculaire. S'agissant des déchets nucléaires, la simulation s'avère indispensable pour prévoir leur comportement à très long terme une fois entreposés dans un milieu géologique et pour optimiser ainsi la conception des colis, des conteneurs et du site.

Dans un autre ordre d'idées, la modélisation "comportementale" permet de s'assurer que des expérimentations ciblées et réduites en nombre peuvent être conduites sans risquer de "faire l'impasse" sur des paramètres déterminants.



CEA – analyse réalisée par J. Bontaz-Carion et al.



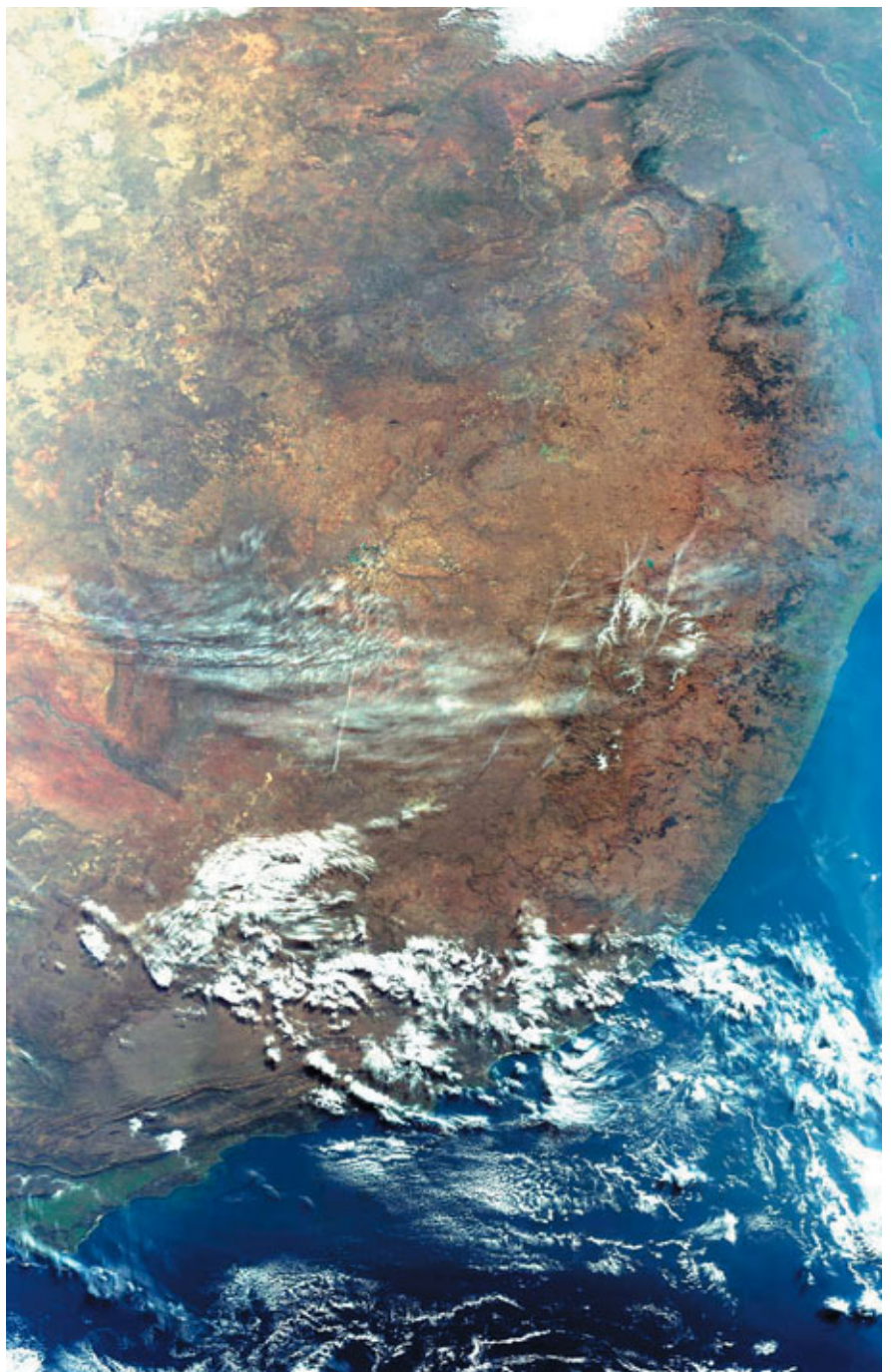
CEA – calcul réalisé par B. Magne

Comparaison de la microtomographie d'un échantillon de cuivre endommagé par choc (en haut) et du même phénomène d'impact simulé par dynamique moléculaire (en bas). Pour l'échantillon analysé, les dimensions du parallélépipède sont de $1\,200\ \mu\text{m} \times 1\,200\ \mu\text{m} \times 2\,400\ \mu\text{m}$. Celles du système simulé sont de $20\ \text{nm} \times 20\ \text{nm} \times 3,6\ \text{nm}$ pour 100 000 atomes à densité normale. Une tranche de 20 000 atomes épaisse de 0,36 nm est représentée ici. Le temps d'expérience est de 7 picosecondes. La distribution en taille des pores de l'échantillon analysé suit une loi d'échelle comparable à celle observée en dynamique moléculaire.

LA MODÉLISATION DU CLIMAT

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. Ces modèles numériques, connus aussi sous le vocable de "modèles de circulation générale", permettent de simuler l'évolution temporelle des caractéristiques tridimensionnelles de l'atmosphère et de l'océan, en tenant compte de leurs interactions avec les surfaces continentales et glacées (calottes de glace et glace de mer).

La région orientale de l'Afrique du Sud observée par l'instrument Meris du satellite Envisat en juillet 2002. Une telle image, où apparaissent à la fois l'océan, les terres et les nuages, de grandes villes côtières, la neige sur les montagnes du Lesotho et les sables du désert du Kalahari (en haut à gauche) illustrent la complexité du système climatique.



L'avènement des **modèles** de circulation générale date des années soixante et leur complexification a été largement liée à l'évolution des **supercalculateurs** (encadré 1 et encadré B, *Les moyens informatiques de la simulation numérique hautes performances*). Que sont ces modèles ? Comment les utilise-t-on et quelles en sont les limitations ? Les lignes suivantes tentent de lever une partie du voile.

Les ingrédients d'un modèle de climat

Étudier le climat demande de prendre en compte la dynamique des fluides océan et atmosphère, les processus physico-chimiques des composés chimiques qu'ils contiennent et les interactions complexes avec la **biosphère** et la cryosphère⁽¹⁾ terrestre ou marine. La figure 1

(1) La cryosphère est la portion du système climatique qui comprend les masses de glace et les dépôts de neige du monde entier (inlandsis, plates-formes de glace, calottes glaciaires et glaciers, glaces de mer, couvert nival saisonnier, glaces de lac et de rivière, sols gelés saisonniers et pergélisol).



Une histoire récente

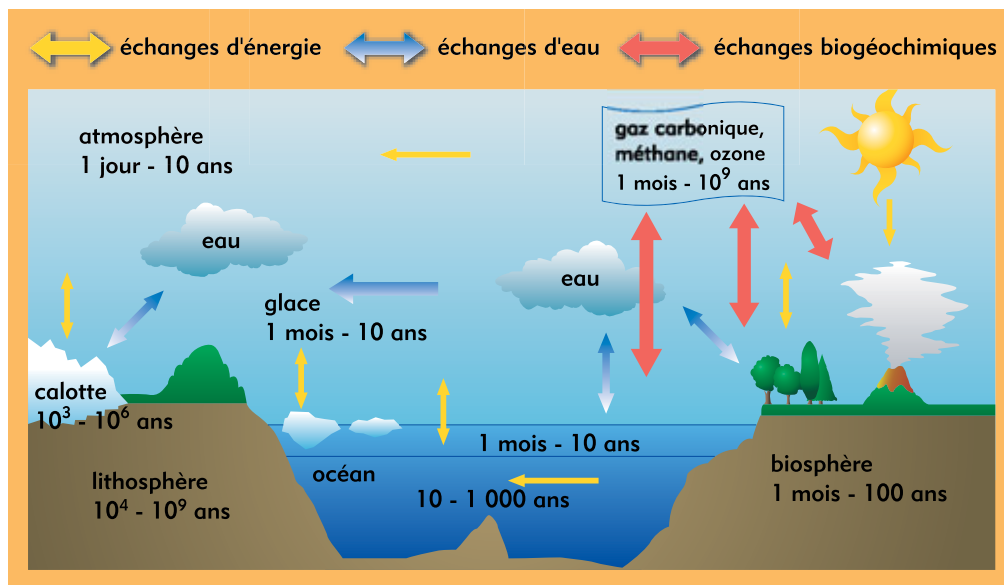
1

Les simulations climatiques, tout comme la prévision du temps, doivent leur essor au développement spectaculaire des ordinateurs au cours du xx^e siècle. C'est le traité de l'Anglais L. F. Richardson, *Weather Prediction by Numerical Process*, publié en 1922, qui le premier expose comment les tendances des variables atmosphériques (température, vents...) peuvent être calculées en différents points répartis dans l'espace, en substituant aux équations différentielles décrivant la circulation atmosphérique un jeu d'équations basées sur des différences algébriques. Mais 64 000 personnes auraient été nécessaires pour effectuer ces calculs ! Il fallut attendre 1950 pour que l'Américain J. G. Charney réalise la première prévision numérique avec un **modèle** d'atmosphère simplifié. Dans les années soixante, la prévision du temps devient petit à petit opérationnelle dans de nombreux pays. Démarrées dans les années soixante-dix avec des modèles d'atmosphère, les simulations climatiques n'ont depuis cessé de prendre de l'ampleur, considérant dans la dernière décennie non seulement l'atmosphère, mais aussi ses couplages avec les autres composantes du système climatique. Une activité importante autour du couplage océan-atmosphère s'est développée en France à partir du début des années quatre-vingt-dix, et le projet européen PRISM (*PRogram for Integrated earth-System Modelling*) (2002-2005) a pour objectif de faciliter l'interfaçage des différents modèles du système climatique disponibles en Europe.

montre que chaque composante du système climatique (atmosphère, océan, biosphère et cryosphère) a ses constantes de temps propres. L'atmosphère varie le plus rapidement et a peu de mémoire. L'océan profond au contraire peut emprisonner des perturbations pour des centaines d'années. Ces différentes constantes de temps permettent de comprendre la différence entre climatologie et météorologie. Prévoir le temps météorologique consiste à déterminer par exemple s'il y aura de la pluie dans les jours qui viennent. Comme l'océan varie moins vite que l'atmosphère, le prévisionniste a le droit de négliger ses variations sur la période de prévision. Étudier le climat, en revanche, nécessite des simulations plus longues (10 à 100 ans) pour obtenir des statistiques fiables. Il n'est plus question de reproduire la trajectoire du système, mais les statistiques de ces trajectoires. Dans ces conditions, les variations de

la circulation océanique et des autres réservoirs plus lents ne peuvent plus être négligées. Le système climatique à considérer dépend donc des processus et des échelles de temps étudiées. Depuis quelques années, le système couplé océan-atmosphère est devenu un élément central, ces deux fluides participant à la redistribution de l'excès d'énergie reçu dans les régions équatoriales vers les régions polaires, *via* les vents et les courants marins.

Un modèle de climat est avant tout un modèle physique, c'est-à-dire que les processus physiques à représenter sont mis sous forme d'équations mathématiques. Les principales composantes du système climatique font l'objet de sous-modèles permettant de calculer les processus internes à chacune et les couplages entre elles. Ainsi, le modèle d'atmosphère, qui n'est rien d'autre à la base qu'un modèle de prévision du temps, calcule le transport de



● ● ● ● ●
 Figure 1. Composantes du système climatique, d'après Sylvie Joussaume (Climat d'hier à demain, CNRS Éditions, Paris, 1993).

Les moyens informatiques de la simulation

Effectuer des **simulations numériques** plus précises impose de mettre en œuvre des **modèles** physiques et numériques eux-mêmes plus précis portant sur des descriptions plus fines des objets simulés (encadré A, *Qu'est-ce qu'une simulation numérique ?*). Tout ceci nécessite des progrès dans le domaine des logiciels de simulation mais aussi une augmentation importante de la capacité des équipements informatiques sur lesquels ces logiciels sont utilisés.

Processeurs scalaires et vectoriels

Au cœur de l'ordinateur, le processeur est l'unité de base qui, déroulant un programme, effectue les calculs. Il en existe deux grands types, les **processeurs scalaires** et les **processeurs vectoriels**. Les premiers exécutent des opérations portant sur des nombres élémentaires (scalaires), par exemple l'addition de deux nombres. Les seconds exécutent des opérations portant sur des ensembles de nombres (vecteurs), par exemple additionner deux à deux les nombres composant deux ensembles de 500 éléments. À ce titre, ils sont particulièrement adaptés à la simulation numérique : lors de l'exécution d'une opération de ce type, un processeur vectoriel peut fonctionner à une vitesse proche de sa performance maximale (crête). La même opération avec un processeur scalaire exige de nombreuses opérations indépendantes (opérations par composante des vecteurs) qui s'exécutent à une vitesse bien inférieure à sa vitesse crête. L'avantage principal des processeurs scalaires est leur prix : il s'agit de microprocesseurs généralistes dont les coûts de conception et de fabrication peuvent être amortis sur de larges marchés.

Forces et contraintes du parallélisme

Les processeurs récents permettent de hautes performances, d'une part en utilisant une fréquence de fonctionnement plus élevée, d'autre part en cherchant à exécuter en même temps

plusieurs opérations : c'est un premier niveau de **parallélisme**. L'accélération de la fréquence est limitée par l'évolution de la technologie micro-électronique, tandis que les dépendances entre instructions à exécuter par le processeur limitent le parallélisme possible. La mise en œuvre simultanée de plusieurs processeurs constitue un second niveau de parallélisme, qui permet d'obtenir des performances accrues à condition de disposer de programmes capables d'en tirer parti. Alors que le parallélisme au niveau des processeurs est automatique, celui *entre processeurs* dans un ordinateur parallèle est à la charge du programmeur, qui doit découper son programme en morceaux indépendants et prévoir entre eux les communications nécessaires. On procède souvent par un découpage du domaine sur lequel porte le calcul, chaque processeur étant chargé de simuler le comportement d'un domaine, et par l'établissement de communications régulières entre processeurs afin de garantir la cohérence d'ensemble du calcul. Pour obtenir un programme parallèle efficace, il faut s'assurer de l'équilibrage de charge entre processeurs et chercher à limiter le coût des communications.

Les différentes architectures

Les équipements informatiques ont différentes fonctions. À partir de son ordinateur de travail sur lequel il prépare ses calculs et en analyse les résultats, l'utilisateur accède à des moyens de calcul, de stockage, et de visualisation partagés, mais beaucoup plus puissants que les siens propres. L'ensemble de ces équipements sont reliés par des réseaux informatiques permettant de faire circuler les informations entre eux avec des débits compatibles avec le volume de données produites, pouvant atteindre 1 **téraoctet** (1 To = 10^{12} octets) de données pour une seule simulation.

Les grands équipements de calcul sont généralement appelés **supercalculateurs**. Ils atteignent aujourd'hui des puissances qui se chiffrent en **téraflops** (1 Tflops = 10^{12} opérations de calcul par seconde).

Il existe aujourd'hui trois grands types de supercalculateurs : les supercalculateurs vectoriels, les grappes de mini-ordinateurs à mémoire partagée et les grappes de PC (l'ordinateur que chacun possède chez soi). Le choix entre ces architectures dépend largement des applications et de l'utilisation visées. Les supercalculateurs vectoriels disposent de processeurs très performants mais dont il est difficile d'augmenter la puissance en ajoutant des processeurs. Les grappes de PC sont peu coûteuses mais mal adaptées à des environnements où de nombreux utilisateurs font beaucoup de calculs très gourmands en puissance machine, en mémoire et en entrées-sorties.

Ce sont ces considérations qui ont en particulier conduit la Direction des applications militaires (DAM) du CEA à choisir pour son programme simulation (voir *Le programme Simulation : la garantie des armes sans essais nucléaires*) les architectures de type grappe de mini-ordinateurs à mémoire partagée, encore appelées **clusters de SMP** (Symmetric Multi-Processor). Un tel système utilise comme brique de base un mini-ordinateur com-

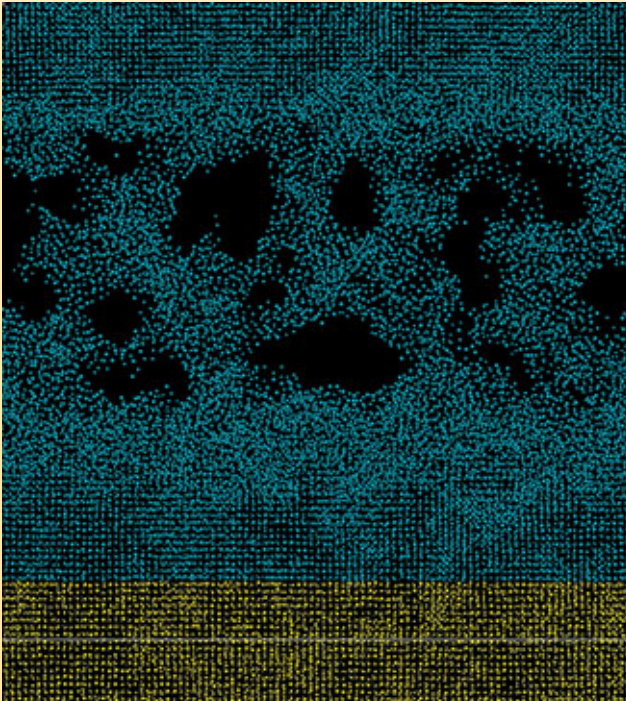


Installée en décembre 2001 au CEA (centre DAM-Ile de France) et conçue par Compaq (devenue depuis HP), la machine Tera a pour élément de base un mini-ordinateur à 4 processeurs Alpha à 1 GHz partageant une mémoire de 4 Go et fournissant une puissance totale de 8 Gflops. Ces éléments de base sont interconnectés par un réseau rapide conçu par la société Quadrics. Une opération de synchronisation sur l'ensemble des 2 560 processeurs s'effectue en moins de 25 microsecondes. Le système de fichiers global offre un espace de stockage de 50 téraoctets pour les entrées-sorties avec une bande passante agrégée de 7,5 Go/s.



numérique hautes performances

B



Les calculateurs parallèles sont adaptés aux méthodes numériques basées sur des maillages (encadré A, **Qu'est-ce qu'une simulation numérique ?**) mais aussi au traitement de calculs ab initio comme cette simulation par dynamique moléculaire de l'endommagement par choc de deux plaques de cuivre à 1 km/s (voir La simulation des matériaux). Le système considéré est constitué de 100 000 atomes de cuivre représentant un parallélépipède de section carrée (0,02 μm de côté) à densité normale. Les atomes interagissent suivant un potentiel EAM (embedded atom potential) pendant 4,6 picosecondes. Le calcul, effectué sur 18 processeurs du supercalculateur Tera de Bruyères-le-Châtel à l'aide du logiciel Stamp développé au CEA, a représenté une dizaine de minutes de temps "utilisateur" (calcul réalisé par B. Magne). Des tests impliquant jusqu'à 64 millions d'atomes ont été réalisés, mobilisant 256 processeurs pendant une centaine d'heures.

portant plusieurs microprocesseurs qui partagent une mémoire commune (figure). Ces mini-ordinateurs étant largement diffusés dans des domaines variés allant de la banque au serveur web

en passant par les bureaux d'études, ils offrent un excellent rapport performance/prix. Ces "briques" de base (encore appelées *nœuds*) sont reliées entre elles par un réseau d'interconnexion hautes performances : la puissance cumulée de plusieurs centaines de ces "briques" peut atteindre plusieurs téraflops. On parle alors d'**ordinateur massivement parallèle**.

Cette puissance peut être disponible pour une seule application parallèle utilisant toutes les ressources du supercalculateur mais aussi pour de multiples applications indépendantes, parallèles ou non, utilisant chacune une partie des ressources.

Si la caractéristique mise en avant pour décrire un supercalculateur est en général sa puissance de calcul, il ne faut pas négliger l'aspect entrées-sorties. Ces machines capables d'effectuer des simulations de grande taille doivent disposer de systèmes de disques avec des capacités et des performances adaptées. Dans les *clusters* de SMP, chaque mini-ordinateur dispose d'un espace disque local. Il n'est néanmoins pas judicieux d'utiliser celui-ci pour les fichiers utilisateurs, ce qui obligerait l'utilisateur à explicitement déplacer ses données entre les différentes phases de ses calculs. Pour cette raison, il est important de disposer d'un espace disque accessible par l'ensemble des mini-ordinateurs du supercalculateur. Cet espace est en général constitué de batteries de disques reliées à des nœuds dont la fonction principale est de les gérer. Comme pour le calcul, c'est le parallélisme pour les entrées-sorties qui permet d'offrir des performances élevées. Il faut, pour ce faire, disposer de systèmes de fichiers globaux parallèles permettant un accès rapide et sans contraintes à l'espace disque partagé.

Offrant des puissances de calcul considérables, les *clusters* de SMP posent néanmoins plusieurs défis. Parmi les plus importants, outre la programmation de logiciels de simulation capables de tirer parti du grand nombre de processeurs, il faut mettre au point des systèmes d'exploitation et les logiciels associés compatibles avec de telles configurations et tolérants vis-à-vis des pannes.

François Robin

Direction des applications militaires
CEA centre DAM-Ile de France

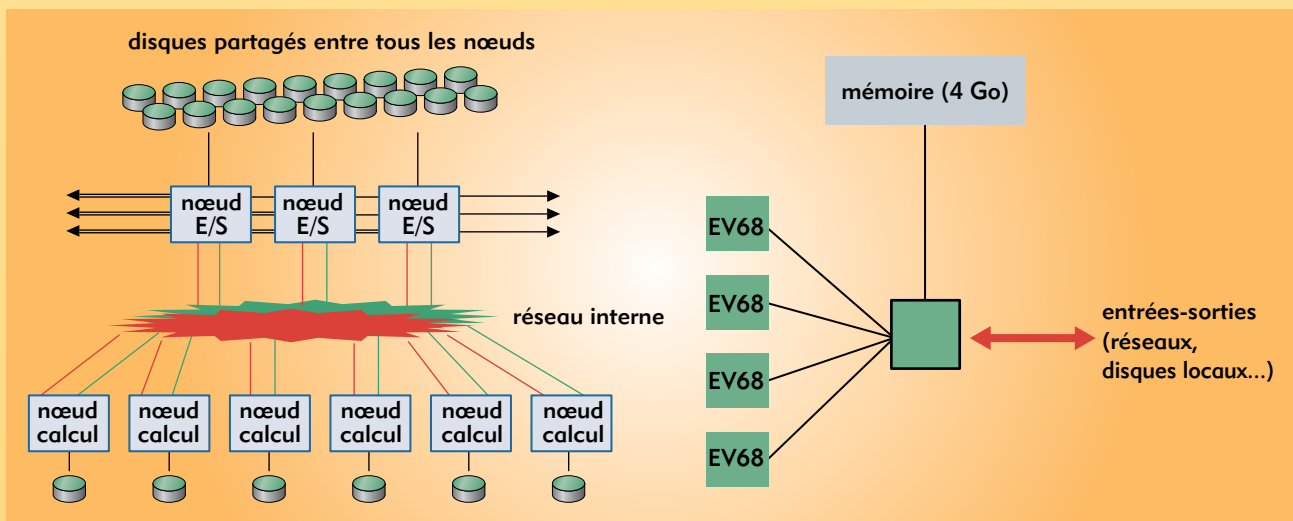


Figure. Architecture d'une machine du type "cluster de SMP". À gauche, l'architecture générale (E/S = entrée/sortie), à droite celle d'un nœud avec quatre processeurs Alpha EV68 cadencés à 1 GHz.

Les nuages constituent un des éléments les plus difficiles à mettre en équations dans les modélisations climatiques.



PhotoDisc

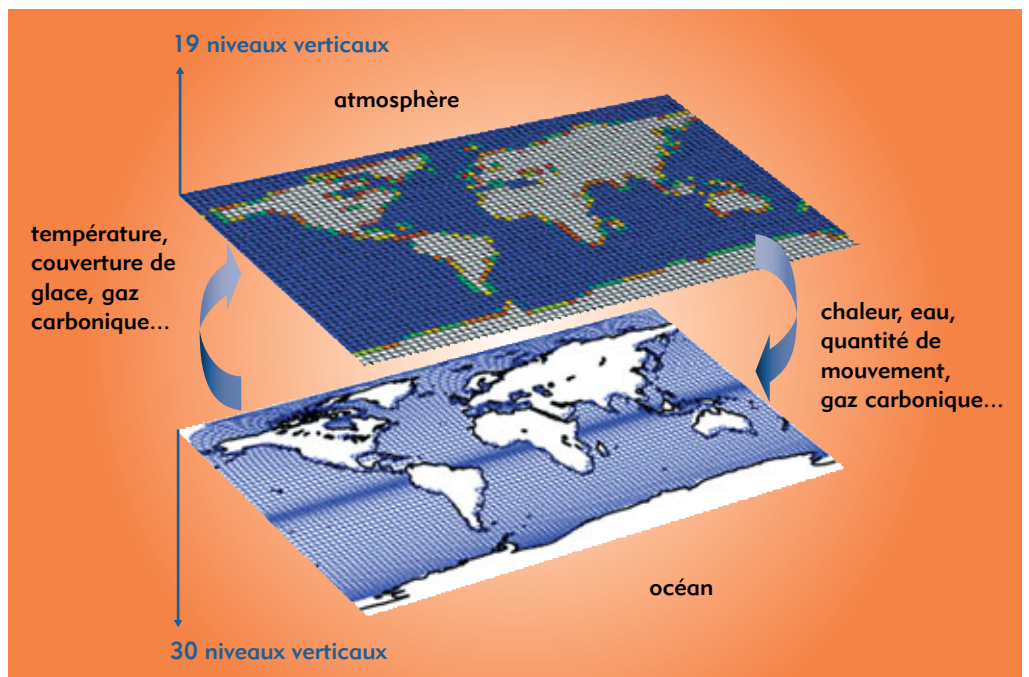
chaleur et d'humidité par les vents, les échanges de quantité de mouvement, de chaleur et d'eau entre l'atmosphère et les surfaces océaniques et terrestres. S'y ajoutent la condensation d'humidité dans les nuages et les précipitations, l'absorption et la diffusion du rayonnement solaire incident, ainsi que l'émission et l'absorption des radiations infrarouges par les nuages, les gaz atmosphériques et les surfaces océaniques et terrestres. Les différents facteurs susceptibles de modifier les processus de base sont également introduits. La glace de mer, la neige et le couvert végétal, par exemple, modulent la quantité de rayonnement solaire absorbé par la surface, par leur capacité à réfléchir le rayonnement solaire incident vers l'espace. Le modèle d'océan tient compte des interactions avec l'atmosphère *via* les vents, les flux de chaleur

et les radiations solaires et infrarouges. Les mouvements **convectifs** dans la colonne d'eau sont déterminés par les variations de température et de salinité.

L'art de mettre le monde dans un ordinateur

Le terme "modèle" de climat peut être trompeur, car après le modèle physique intervient le modèle numérique. Mettre la Terre dans un ordinateur nécessite de résoudre les équations sur une grille tridimensionnelle à l'échelle du globe (figure 2). Les **méthodes numériques** mises en œuvre pour le climat doivent conserver des quantités intégrales comme l'énergie ou la masse d'eau totale. Aux **résolutions** généralement utilisées pour le climat, les équations sont résolues toutes

Figure 2. Grilles des composantes océanique et atmosphérique du modèle couplé de l'IPSL (Institut Pierre Simon Laplace) dans une version dite basse résolution. Au cours d'une intégration, le couplage des deux modèles est effectué une fois par jour simulé, par l'intermédiaire d'un coupleur qui gère la communication entre les modèles et les interpolations entre les grilles. Pour assurer la conservation d'énergie entre les grilles, les mailles d'atmosphère peuvent contenir différents types de surfaces (terre, océan, glace de mer et glace continentale) et le trait de côte est obtenu par projection sur la grille atmosphérique du masque terre/mer du modèle d'océan. Les couleurs de la grille atmosphérique indiquent les pourcentages de terre de chaque maille. La grille océanique est plus raffinée en latitude vers l'équateur pour mieux représenter les phénomènes tropicaux.





Comment valider les modèles et estimer leur précision ?

2

La validité des résultats dépend de plusieurs facteurs. Le premier repose sur les processus physiques reproduits dans le **modèle**. Ce seul élément ne suffit pas et il faut mettre en œuvre toute une batterie de tests pour vérifier que le climat simulé est en accord avec les phénomènes observés. Les nombreux aspects chaotiques du système climatique (leur évolution est très sensible à de petites perturbations dans les conditions initiales et les conditions limites) obligent à recourir à des ensembles de simulations, ou à se baser sur les résultats de plusieurs modèles. C'est ainsi que de nombreux projets de comparaisons de résultats de modèles sur le climat actuel ou les climats du passé ont montré que les modèles représentaient de façon satisfaisante les principaux aspects du climat et de ses variations, mais qu'aucun modèle n'était actuellement capable de reproduire correctement toutes les facettes du climat. Une des principales sources d'incertitude est liée à la représentation des nuages, qui conditionne en grande partie la réponse énergétique des modèles à une perturbation. Les simulations des climats passés offrent aussi la possibilité d'évaluer si les modèles sont capables de reproduire des climats différents du climat actuel. C'est l'un des objectifs du projet international PMIP (*Paleoclimate Modeling Inter-comparison Project*) coordonné par Sylvie Joussaume au LSCE. Les résultats obtenus pour l'holocène moyen (il y a



CNRS Photothèque - Nicole Petit-Maire

Ces dépôts lacustres retrouvés au Sahara témoignent pour cette région d'un passé plus humide qu'aujourd'hui.

6 000 ans) ont montré, par exemple, que le système couplé océan-atmosphère-végétation devait être considéré dans son ensemble pour reproduire les conditions plus humides révélées par les dépôts lacustres et les pollens dans des régions aujourd'hui sèches (figure).

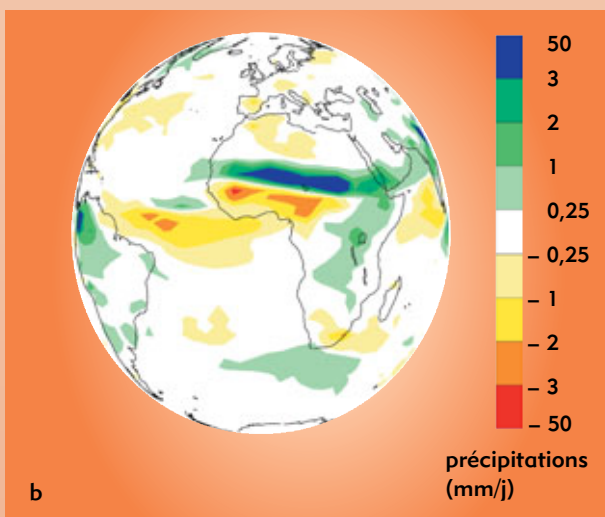
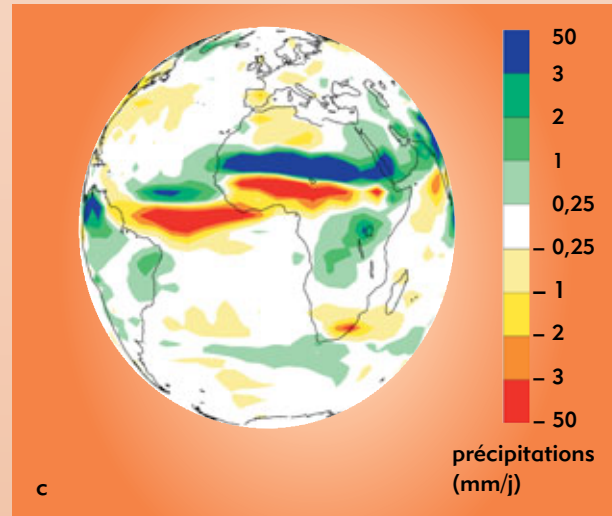
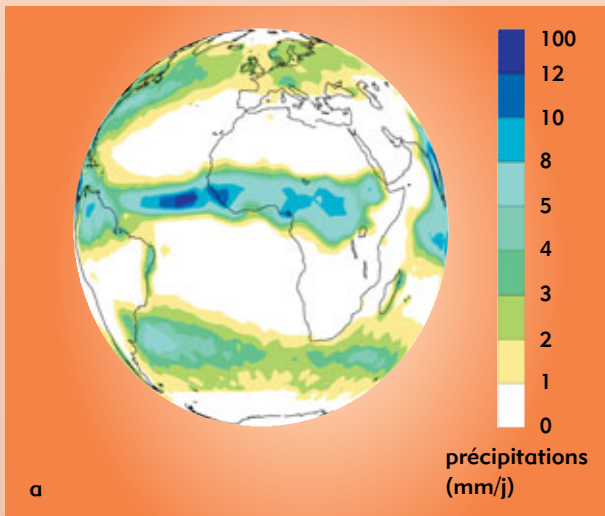
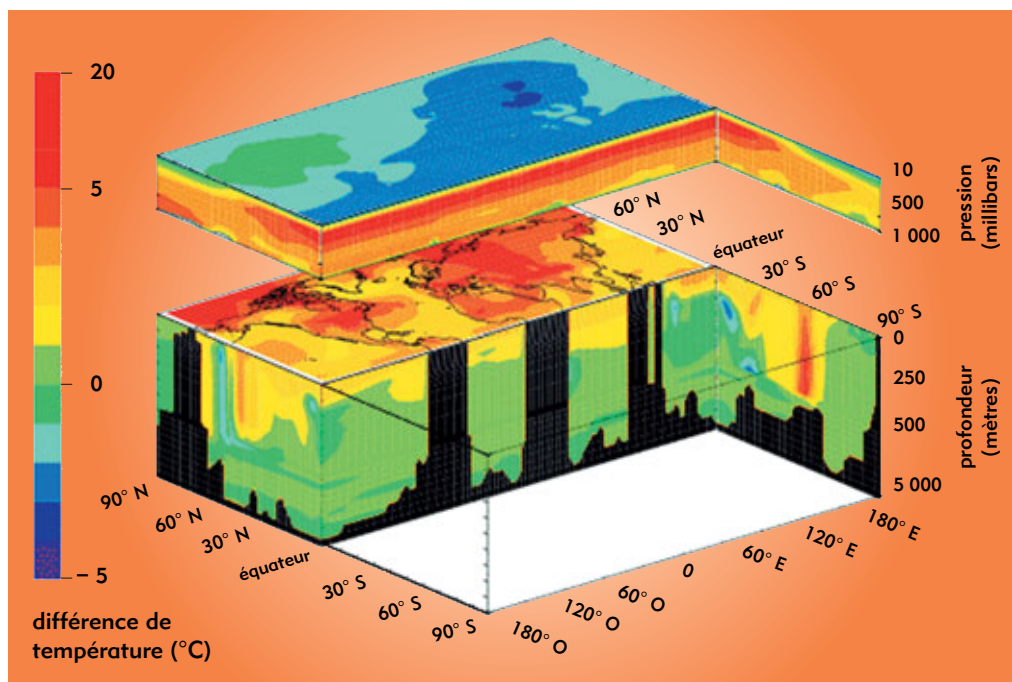


Figure. En été, les continents de l'hémisphère nord se réchauffent plus fortement que l'océan, qui a une plus grande capacité thermique. L'air chaud est moins dense, ce qui se caractérise par le développement d'une basse pression sur le continent, dont le minimum se situe au pied des reliefs himalayens et du plateau tibétain et s'étend jusqu'au Sahara. Les vents sont ainsi aspirés de l'océan vers le continent. Ils se chargent en eau au cours de leur trajet sur les océans tropicaux chauds et se déversent sous forme de pluies torrentielles sur le continent. Ce sont les pluies de mousson, comme ici en Afrique en juillet-août (a). Il y a 6 000 ans, la Terre recevait plus de rayonnement solaire en été dans l'hémisphère nord, ce qui renforçait le gradient terre, océan et la circulation de mousson. Les vents de mousson convergeaient plus au nord en Afrique, rendant humides des régions aujourd'hui sèches. Les modèles d'atmosphère soumis au rayonnement solaire d'il y a 6 000 ans montrent tous ce mécanisme (b). Ces simulations négligent les changements de circulation océanique et de végétation. Les simulations réalisées avec le modèle couplé océan-atmosphère-végétation de l'IPSL (Institut Pierre Simon Laplace) montrent que les rétroactions liées à l'océan et à la végétation amplifient la réponse de la circulation de mousson (c). Les résultats de cette simulation sont en meilleur accord avec les données disponibles.

Figure 3. Représentation tridimensionnelle du changement de température simulé avec le modèle de climat de l'IPSL (Institut Pierre Simon Laplace) pour un climat perturbé par le gaz carbonique (CO₂) anthropique au bout de 240 ans. La référence correspond à un climat préindustriel, le taux de CO₂ étant fixé à celui des années 1860 (180 ppm). La partie supérieure de la figure se rapporte à l'atmosphère et la partie inférieure à l'océan. Les isolignes sont portées tous les 1 °C (ligne 0 omise) jusqu'à 5 °C, puis tous les 5 °C.



L. Fairhead (CNRS/LMD)

les demi-heures au cours d'une intégration du modèle. De nombreux processus, comme ceux qui font intervenir les nuages, sont actifs à une échelle bien inférieure à celles de la **maille** ou du pas de temps du modèle, et ne peuvent donc pas être modélisés de façon explicite. Leurs effets sur les échelles représentées sont considérés en les modélisant à partir de relations physiques avec les variables de grande échelle. Cette technique s'appelle **paramétrisation**.

Ensuite vient le modèle informatique, ou plus prosaïquement le "**code**". Le Fortran est aujourd'hui largement employé par héritage historique et parce qu'il est généralement plus performant que d'autres langages sur les **machines vectorielles** sur lesquelles ces types de modèles sont exploités. Le choix de la résolution et de la complexité du modèle résulte aussi d'un compromis entre le "réalisme recherché" et le coût informatique. Il faut calculer plus vite, beaucoup plus vite, que le temps qui passe. Pour les études des climats passés menées au Laboratoire des sciences du climat et de l'environnement (CEA-CNRS/LSCE), un modèle couplé océan-atmosphère, dont la composante atmosphérique a une résolution de 72 points en longitude, 45 points en latitude et 19 niveaux verticaux, est mis en œuvre. Avec cette résolution, il faut 200 heures de calcul sur les calculateurs vectoriels les plus puissants pour produire 100 ans de climat simulé. Ce temps de calcul est multiplié par 10 pour augmenter la résolution spatiale d'un facteur 2 dans les trois directions. De telles simulations sont réalisables sur station de travail, mais il faut compter un facteur dix sur le temps de restitution.

Au code de simulation lui-même s'ajoute une suite d'outils destinés à gérer les simulations de plusieurs centaines ou milliers

d'heures d'ordinateur. Des moyens de stockage puissants et des outils d'analyse permettent d'exploiter l'information de ces machines à fabriquer des chiffres. L'analyse graphique des résultats et les analyses statistiques, allant de simples moyennes à des méthodes rendant possible la caractérisation des situations climatiques et leurs occurrences temporelles, sont couramment employées en climatologie (encadré 2).

Les images du climat de demain

Les modèles de climat ont de multiples applications, comme les études des processus climatiques, des fluctuations interannuelles à pluridécennales du climat ou des changements climatiques à venir ou passés. Pour étudier les modifications du climat induites par les activités humaines, la façon de procéder consiste tout d'abord à réaliser une simulation représentative d'un climat non perturbé par l'homme. Ensuite, une simulation du climat perturbé est effectuée en modifiant par exemple la concentration en gaz carbonique (CO₂) dans l'atmosphère, en suivant un scénario proposé par les économistes.

La figure 3 montre le réchauffement obtenu pour un tel scénario en 2100 dans une simulation réalisée avec le modèle de climat de l'Institut Pierre Simon Laplace (IPSL, Paris). Par rapport à la simulation du climat préindustriel, les continents de l'hémisphère nord marquent un réchauffement important. Aux hautes latitudes, la couverture de neige a diminué. Le sol nu réfléchit moins le rayonnement solaire incident que la neige, ce qui favorise le chauffage de la surface continentale. L'océan, en raison de sa plus grande capacité calorifique que le sol, subit un

réchauffement moindre. Le réchauffement n'a pas atteint l'océan profond. La stabilité verticale de l'océan est augmentée, ce qui réduit les échanges verticaux et modifie le rôle climatique de l'océan. L'effet de serre augmenté isole la très haute atmosphère qui refroidit. Le devenir du cycle de l'eau associé à un tel réchauffement climatique est encore mal connu. Le cycle hydrologique fait intervenir des processus de petite échelle qui sont difficiles à représenter dans les modèles et la palette des résultats de différents modèles est bien plus large que celle obtenue pour le changement de température. C'est à partir de ce type de simulations, en combinant des approches variées faisant intervenir les simulations du climat actuel, du climat futur et des climats passés que petit à petit les rouages des multiples mécanismes survenant dans la machine climatique se dégagent. La modélisation du climat est actuellement un domaine en pleine expansion. Il est maintenant possible de coupler les mécanismes climatiques avec les processus biogéochimiques. Les climatologues cherchent à ne plus simplement étudier l'impact des modifications des cycles biogéochimiques sur le climat, mais aussi l'impact du climat sur ces cycles, ce qui va très certainement dans les prochaines années modifier leur façon de considérer le monde et leur poser encore de nombreuses questions sur les méthodes adaptées pour introduire cette physique dans les calculateurs. ●

Pascale Braconnot et Olivier Marti
Laboratoire des sciences du climat
et de l'environnement
CEA-CNRS
Direction des sciences de la matière
Saclay

Qu'est-ce qu'une simulation numérique ?

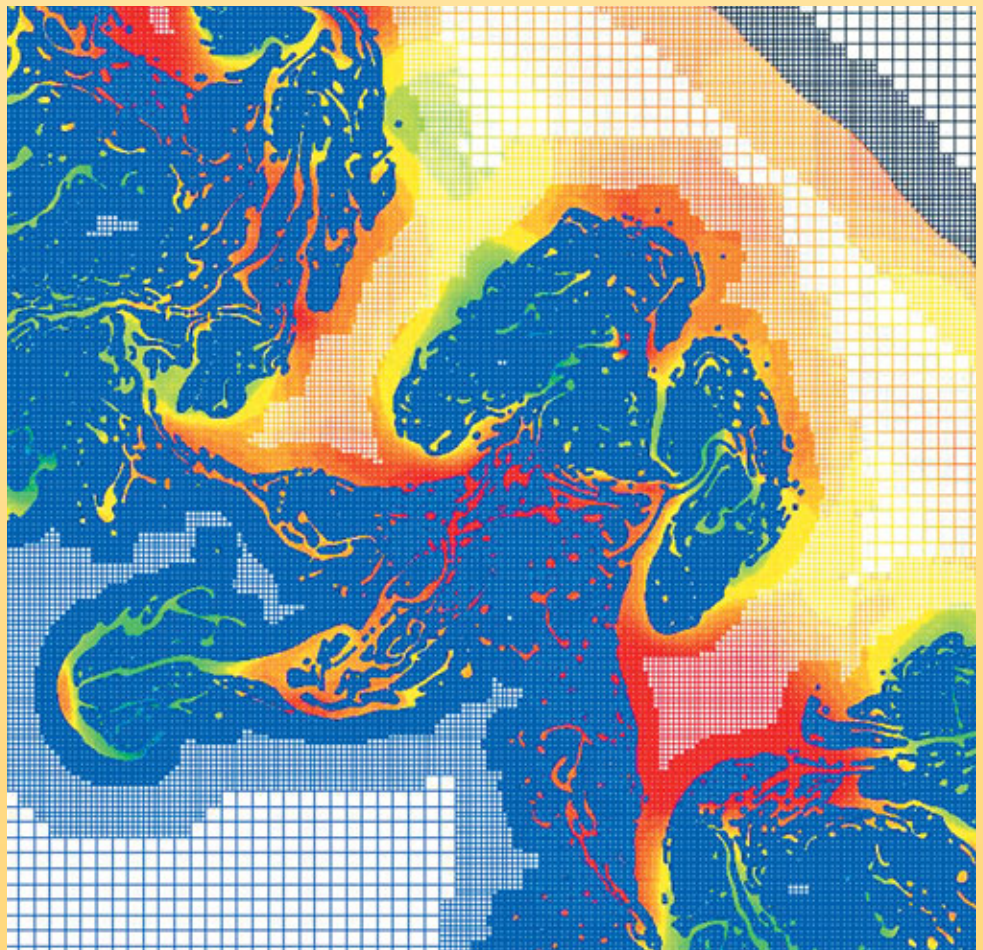
La **simulation numérique** consiste à reproduire par le calcul le fonctionnement d'un système, préalablement décrit par un ensemble de **modèles**. Elle s'appuie sur des méthodes mathématiques et informatiques spécifiques. Les principales étapes de la réalisation d'une étude par simulation numérique sont communes à de nombreux secteurs de la recherche et de l'industrie, en particulier le nucléaire, l'aérospatial ou l'automobile.

En chaque point de l'"objet" considéré, plusieurs grandeurs physiques (vitesse, température...) décrivent l'état et l'évolution du système étudié. Celles-ci ne sont pas indépendantes, mais reliées et régies par des **équations**, généralement **aux dérivées partielles**. Ces équations constituent la traduction mathématique des lois de la physique qui modélisent le comportement de l'objet. Simuler l'état de ce dernier, c'est déterminer – idéalement en tout point – les valeurs numériques de ses **paramètres**. Comme il y a un nombre infini de points, donc une infinité de valeurs à calculer, cet objectif est inaccessible (sauf dans des cas bien particuliers où l'on peut résoudre les équations de départ à l'aide de formules analytiques). Une approximation naturelle consiste donc à ne considérer qu'un nombre fini de points. Les valeurs des paramètres à calculer sont ainsi en nombre fini et les opérations nécessaires deviennent abordables grâce à l'ordinateur. Le nombre effectif de points traités dépendra bien sûr de la puissance de celui-ci : plus il sera élevé, meilleure sera finalement la description de l'objet. À la base du calcul des paramètres comme à la base de la simulation numérique, il y a donc la réduction de l'infini au fini, la **discrétisation**.

Comment opère-t-on précisément à partir des équations mathématiques du modèle ? Deux méthodes sont très souvent utilisées, respectivement représentatives des méthodes de **calcul déterministe**, qui résolvent les équations régissant les phénomènes étudiés après avoir discrétisé les variables, et des méthodes de **calcul statistique** ou **probabiliste**.

Le principe de la première, connue sous le nom de **méthode des volumes finis**, est antérieur à l'usage des ordinateurs. Chaque point de l'objet est assimilé simplement à un petit volume élémentaire (un cube par exemple), d'où le nom de *volume fini*. Un plasma, par exemple, est ainsi vu comme un ensemble ou un réseau de volumes contigus qui, par analogie avec la trame d'un tissu, sera dénommé **maillage**. Les paramètres de l'état de l'objet sont maintenant définis dans chaque maille du maillage. Pour chacune d'elles, en reformulant les équations mathématiques du modèle par des moyennes volumiques, il sera alors possible de construire des *relations algébriques* entre les paramètres de la maille et ceux de ses voisins. Au total, il y aura autant de relations que de paramètres inconnus et ce sera à l'ordinateur de résoudre le *système* de relations obtenu. Il faudra pour cela recourir aux techniques de l'**analyse numérique** et programmer des **algorithmes** spécifiques.

L'accroissement de la puissance des ordinateurs a permis d'augmenter la finesse de discrétisation, permettant de passer de quelques dizaines de mailles dans les années soixante à plusieurs dizaines de milliers dans les années quatre-vingt, à des millions dans les années quatre-vingt-dix et jusqu'à la dizaine de milliards de mailles aujourd'hui (machine Tera de la Direction



Exemple d'image d'une simulation 2D d'instabilités réalisée avec le supercalculateur Tera du CEA. Le calcul a fait appel au maillage adaptatif, qui se fait plus fin dans les zones où les phénomènes sont les plus complexes.

des applications militaires du CEA), chiffre qui devrait décupler à la fin de la décennie.

Un raffinement du maillage, le **remaillage adaptatif**, consiste à ajuster la taille des mailles en fonction des circonstances, par exemple en les rendant plus petites et plus serrées aux interfaces entre deux milieux, là où les phénomènes physiques sont les plus complexes, ou là où les variations sont les plus importantes.

La méthode des volumes finis s'applique dans des contextes physiques et mathématiques très variés. Elle autorise toute forme de maille (cube, hexaèdre, tétraèdre...) et le maillage peut être modifié durant le calcul, en fonction de critères géométriques ou physiques. Enfin, elle est aisée à mettre en œuvre dans le contexte des **ordinateurs parallèles** (encadré B, **Les moyens informatiques de la simulation numérique hautes performances**), le maillage pouvant en effet faire l'objet d'un découpage pour des calculs sur ce type de machines (exemple figure B, p. 13).

Appartiennent à la même famille la **méthode des différences finies**, cas particulier de la **méthode des volumes finis** où les côtés des mailles sont orthogonaux, et la **méthode aux éléments finis**, qui peut juxtaposer divers types de mailles. La deuxième grande méthode, dite de **Monte-Carlo**, est particulièrement adaptée pour simuler le *transport de particules*, par exemple des neutrons ou des photons d'un **plasma** (voir *Les simulations en physique des particules*). Un tel transport est en fait caractérisé par une succession d'étapes lors desquelles chaque particule peut subir différents événements (diffusion, absorption, émission...) possibles *a priori*. Les probabilités élémentaires de chacun de ces événements sont connues individuellement pour chaque particule.

Il est alors naturel d'assimiler un point du plasma à une particule. Un ensemble de particules, en nombre fini, va constituer un échantillon représentatif de l'infinité de particules du plasma, comme lors d'un sondage statistique. D'étape en étape, l'évolution de l'échantillon sera déterminée grâce à des tirages au hasard (d'où le nom de la méthode). L'efficacité de cette méthode, mise en œuvre à Los Alamos dès les années 1940, dépend bien sûr de la qualité statistique des tirages au hasard.

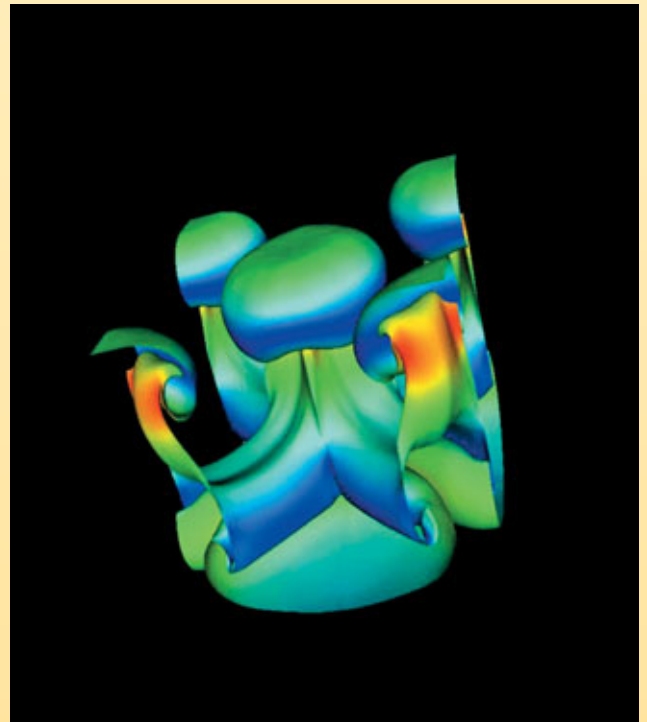
Comment se déroule une simulation numérique ?

Il est souvent question d'*expérience numérique* pour souligner l'analogie entre la pratique d'une simulation numérique et la conduite d'une expérience de physique.

Brièvement, cette dernière utilise un dispositif expérimental, configuré selon des conditions initiales (de température, de pression...) et des paramètres de contrôle (durée de l'expérience, des mesures...). Durant l'expérience, le dispositif produit des points de mesures qui sont enregistrés. Ces enregistrements sont ensuite analysés et interprétés.

Dans une simulation numérique, le dispositif expérimental consiste en un ensemble de programmes informatiques exécutés sur des ordinateurs. Les **codes** ou **logiciels de calcul** sont la traduction, à travers des algorithmes numériques, des formulations mathématiques des modèles physiques étudiés. En amont et en aval du calcul, les *logiciels d'environnement* effectuent la gestion de plusieurs opérations complexes de préparation des calculs et de leur dépouillement.

Les données initiales de la simulation comporteront d'abord la délimitation du domaine de calcul à partir d'une représentation approchée des formes géométriques (produite par le dessin et la CAO, conception assistée par ordinateur), suivie de la discrétisation de ce



Simulation 3D réalisée à l'aide du supercalculateur Tera installé fin 2001 au centre CEA/DAM Île-de-France à Bruyères-le-Châtel (Essonne).

Il existe pour cela des méthodes de nombres aléatoires, bien adaptées au traitement par un ordinateur.

Les méthodes des volumes finis et de Monte-Carlo ont suscité et suscitent de nombreuses études mathématiques. Ces études s'attachent notamment à préciser la convergence de ces méthodes, c'est-à-dire comment la précision de l'approximation varie avec le nombre de mailles ou de particules. Cette question est naturelle lors de la confrontation des résultats de la simulation numérique à ceux de l'expérience.

domaine de calcul sur un maillage, ainsi que les valeurs des paramètres physiques sur ce maillage et les paramètres de contrôle du bon déroulement des programmes... Toutes ces données (produites et gérées par les logiciels d'environnement) seront saisies et vérifiées par les codes. Les résultats des calculs proprement dits, c'est-à-dire les valeurs numériques des paramètres physiques, seront sauvegardés au fur et à mesure. En fait, un protocole spécifique structurera les informations produites par l'ordinateur afin de constituer une base de données numériques.

Un protocole complet organise l'échange informatique des informations requises (dimensions notamment) suivant des formats prédéfinis : *modeleur*⁽¹⁾, *mailleur*⁽²⁾, *découpeur de maillage*, *code*

- (1) Le *modeleur* est un outil qui permet la création et la manipulation de points, courbes et surfaces en vue par exemple de la création d'un maillage.
- (2) Les formes géométriques d'un maillage sont décrites par des ensembles de points reliés par des courbes et des surfaces (de Bézier par exemple) qui en représentent les frontières.



de calculs, logiciel de visualisation et d'analyse. Les études de *sensibilité* des résultats (au maillage et aux modèles) font partie des "expériences" numériques.

À l'issue des calculs (résolution numérique des équations décrivant les phénomènes physiques qui se déroulent dans chaque maille), l'analyse des résultats par des spécialistes reposera sur l'exploitation de la base de données numériques. Elle comportera plusieurs étapes : extraction sélective des données (selon le paramètre physique recherché) et visualisation, extraction et transfert des données pour calculer et visualiser des diagnostics.

Le parallèle entre la conduite d'un cas de calcul, d'une expérience numérique et la conduite d'une expérience physique ne s'arrête pas là : les résultats numériques seront comparés aux résultats expérimentaux. Cette analyse comparative, effectuée sur la base de critères quantitatifs standardisés, fera appel et à l'expérience

et à l'art de l'ingénieur, du physicien, du mathématicien. Elle débouchera sur de nouvelles améliorations des modèles physiques et des programmes informatiques de simulation.

Bruno Scheurer

Direction des applications militaires
CEA centre DAM-Ile de France

Frédéric Ducros et Ulrich Bieder

Direction de l'énergie nucléaire
CEA centre de Grenoble

L'exemple d'un calcul de thermohydraulique

La mise en œuvre d'un protocole de simulation numérique peut être illustrée par les travaux réalisés par l'équipe de développement du logiciel de calcul **thermohydraulique** Trio U. Ces travaux se sont déroulés dans le cadre d'une étude faite en collaboration avec l'Institut de radioprotection et de sûreté nucléaire (IRSN). L'objectif était d'obtenir des données très précises pour fournir à l'ingénieur les sollicitations en température à la paroi des composants d'un réacteur à eau sous pression dans le cas d'un accident grave impliquant une circulation naturelle turbulente de gaz chauds. Cette étude requiert la modélisation simultanée d'effets "système" à grande échelle et de phénomènes **turbulents** à petite échelle (encadré F, **Modélisation et simulation des écoulements turbulents**).

Elle débute par la définition du modèle de calcul global (figure A), suivie de la réalisation du modèle CAO et du maillage correspondant avec des logiciels du commerce (figure B). Les maillages de plus de cinq millions de mailles exigent l'utilisation de puissantes stations graphiques. Dans cet exemple, le maillage d'un générateur de vapeur (figures C et D) a été découpé pour répartir les calculs sur huit processeurs d'un ordinateur parallèle du CEA : chaque couleur symbolise une zone affectée à un processeur particulier. Les calculs, dont les conditions aux limites sont données par un calcul "système" (Icare-Cathare), produisent des résultats qu'il appartient aux spécialistes d'interpréter. En l'occurrence, les visualisations sur des stations graphiques des valeurs instantanées des champs de vitesse montrent l'impact d'un panache chaud sur la plaque tubulaire du

générateur de vapeur (coupe dans le champ de vitesses à gauche de la figure E) et la température instantanée dans la boîte à eau (à droite).

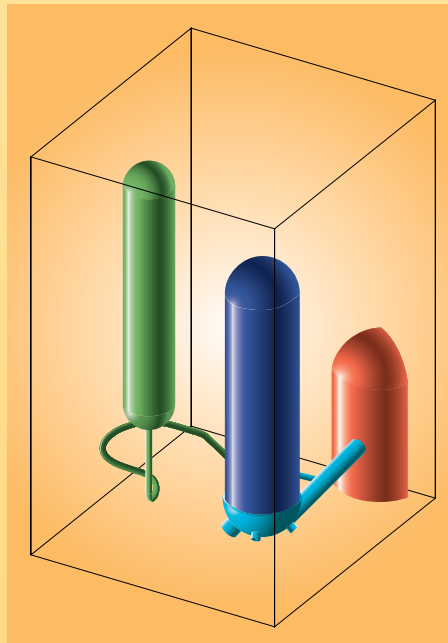


Figure A. Domaine de calcul global incluant une partie de la cuve réacteur (rouge), la conduite de sortie (branche chaude en bleu clair), le générateur de vapeur (bleu foncé) et le pressuriseur (vert).

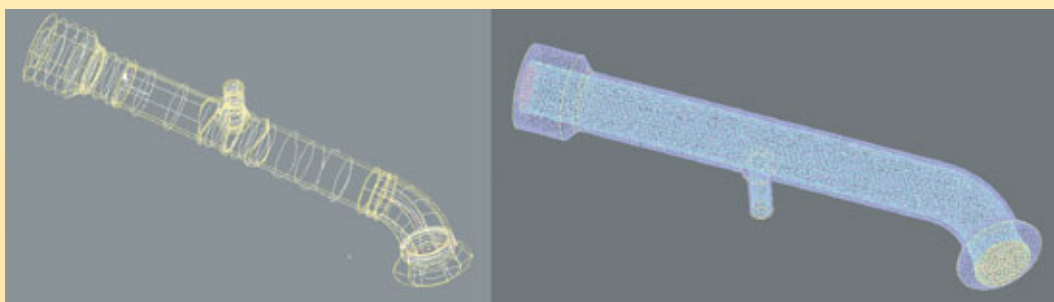
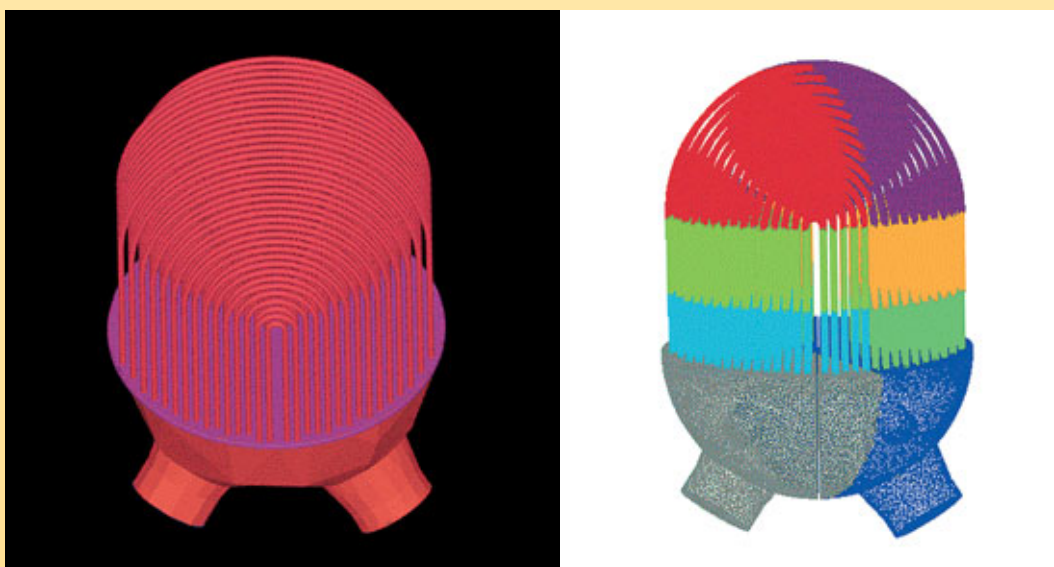


Figure B. Modèle CAO de la branche chaude en sortie de la cuve réacteur (à gauche) et son maillage non structuré (à droite).



Figures C et D.

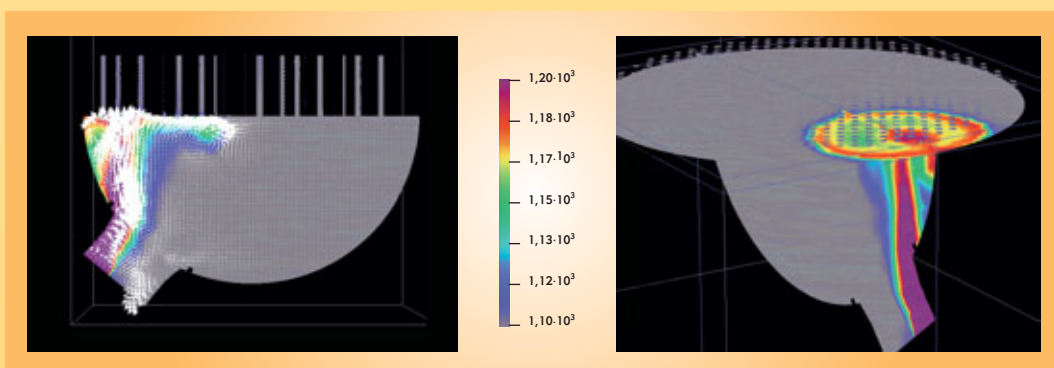


Figure E.

Modélisation et simulation des écoulements turbulents

La **turbulence**, ou l'agitation de l'écoulement dit turbulent, se développe dans la plupart des écoulements qui conditionnent notre environnement immédiat (rivières, océan, atmosphère). Elle se révèle être aussi un, sinon le, paramètre dimensionnant dans un bon nombre d'écoulements industriels (liés à la production ou la conversion d'énergie, à l'aérodynamique...). Il n'est donc pas étonnant que soient entrepris des efforts visant sa prédiction – fût-elle encore imprécise – surtout lorsqu'elle se trouve combinée à des phénomènes qui la compliquent : stratification, combustion, présence de plusieurs phases... C'est que, paradoxalement, même s'il est possible d'anticiper la nature turbulente d'un écoulement et même, d'un point de vue théorique, de dégager certaines caractéristiques communes et apparemment universelles aux écoulements turbulents⁽¹⁾, leur prédiction dans

des cas précis reste délicate. Celle-ci doit en effet prendre en compte l'importante gamme d'échelles spatiales et temporelles⁽²⁾ impliquées dans tout écoulement de ce type.

Les chercheurs ne sont pourtant pas démunis, aujourd'hui, pour aborder ce problème. En premier lieu, les équations qui régissent l'évolution spatio-temporelle des écoulements turbulents (équations de Navier-Stokes⁽³⁾) sont connues. Leur résolution complète, dans des cas très favorables, a conduit à des descriptions prédictives. Mais l'emploi systématique de cette méthode de résolution se heurte à deux difficultés rédhibitoires : d'une part, il nécessiterait la connaissance complète et simultanée de toutes les variables attachées à l'écoulement et des forçages s'exerçant sur lui⁽⁴⁾ et, d'autre part, il mobiliserait des moyens de calculs irréalistes pour encore des décennies.

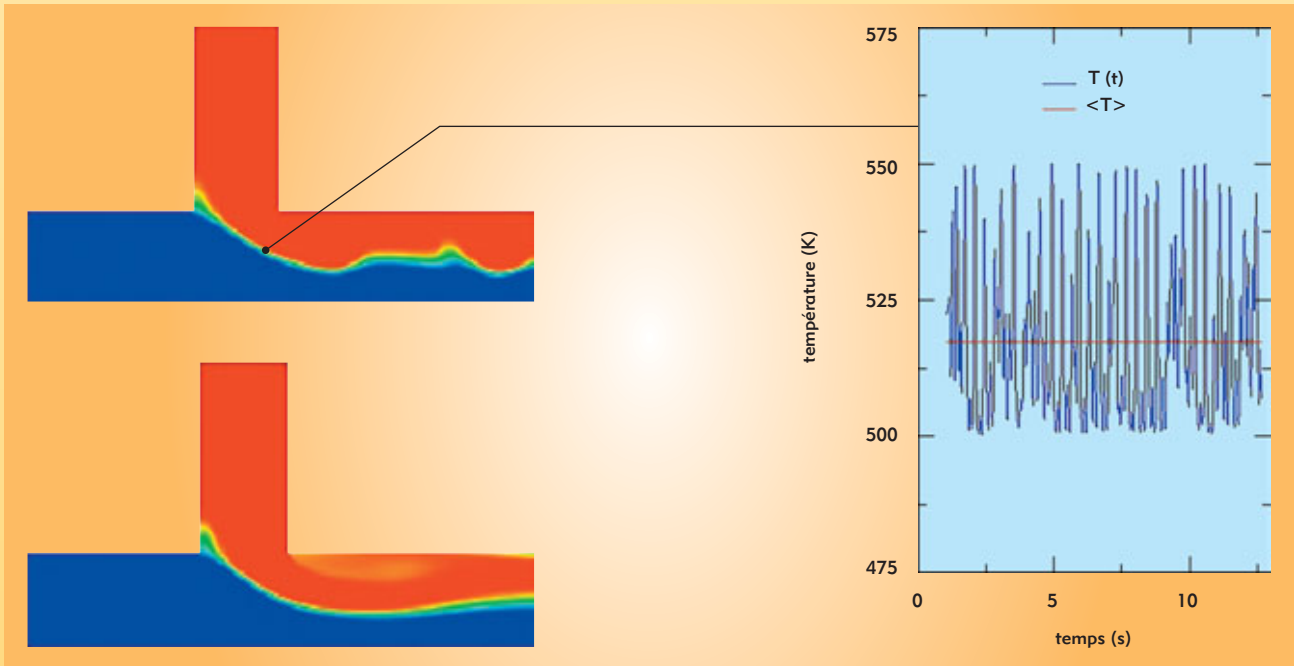


Figure. Champ de température instantané (haut) et moyenné (bas) dans une situation de mélange. La courbe donne l'historique de la température en un point : valeur instantanée fluctuante en bleu et moyenne en rouge (d'après la thèse d'Alexandre Chatelain [DEN/DTP/SMTH/LDTA]).

Il faut donc se résoudre, en s'appuyant sur le caractère fluctuant dû à l'agitation turbulente, à définir et utiliser des moyennes. Une des approches les plus répandues consiste à aborder le problème sous un angle statistique. Les moyennes d'ensemble de vitesse, de pression, de température... dont la distribution caractérise l'écoulement turbulent sont définies comme les variables principales de l'écoulement qu'on cherche à qualifier par rapport à ces moyennes. Ceci conduit à une décomposition du mouvement (dite de Reynolds) en champs moyen et fluctuant, ce dernier mesurant l'écart instantané et local entre chaque grandeur réelle et sa moyenne (figure). Ces fluctuations représentent la turbulence et couvrent une partie importante du spectre de Kolmogorov⁽¹⁾.

Cette opération réduit considérablement le nombre de degrés de liberté du problème et le rend « manipulable » informatiquement. Elle comporte aussi de nombreuses difficultés : il faut tout d'abord constater que, précisément en raison des non-linéarités des équations du mouvement, toute moyenne fait surgir des termes nouveaux et inconnus qu'il faut estimer. En fermant la porte à la description complète et déterministe du phénomène, on ouvre celle de la modélisation, c'est-à-dire à la représentation des effets de la turbulence sur les variables moyennes.

Beaucoup de progrès ont été accomplis depuis les premiers modèles (Prandtl, 1925). Les modélisations n'ont cessé d'évoluer vers plus de complexité, se basant sur le fait généralement vérifié que toute nouvelle extension permet de conserver les propriétés antérieurement acquises. Il faut aussi constater que, même

si de nombreux développements remettent en avant la nécessité de traiter les écoulements en respectant leur caractère *instationnaire*, les modélisations les plus populaires ont été développées dans le cadre des écoulements *stationnaires*, pour lesquels on n'accède donc qu'à une représentation de la moyenne temporelle de l'écoulement : dans le modèle mathématique final, les effets de la turbulence proviennent ainsi intégralement de la modélisation.

Il est également remarquable que, malgré de nombreux travaux, aucune modélisation n'est aujourd'hui capable de rendre compte de l'intégralité des phénomènes qui influencent la turbulence ou sont influencés par elle (transition, instationnarité, stratification, compression, etc.). Ce qui semble pour l'instant empêcher les modélisations statistiques de nourrir une ambition d'universalité.

Malgré ces limitations, la plupart des modélisations statistiques courantes sont maintenant disponibles dans les codes commerciaux et les outils des industriels. Il n'est pas possible de prétendre qu'elles permettent des calculs prédictifs dans toute situation. Leur précision est variable, offrant des résultats utiles pour l'ingénieur dans des situations maîtrisées et favorables (prédiction de la trainée avec une précision de 5 % à 10 % d'erreur [parfois mieux] sur certains profils), mais parfois faux dans des situations qui se révèlent, après coup, en dehors du champ de validité du modèle. Tout emploi maîtrisé d'une modélisation repose donc sur une qualification particulière au type d'écoulement à traiter. Des modélisations alternatives, répondant au besoin d'une plus grande précision sur des gammes d'échelles spatiales et temporelles plus étendues et donc basées sur un opérateur de "moyenne" d'une nature différente, sont actuellement en développement et représentent des voies nouvelles.

Le paysage des modélisations de la turbulence est aujourd'hui très complexe et l'unification des points de vue et des divers concepts de modélisation est une gageure. La tentation de l'universalité des modélisations reste donc hors de propos. Leur mise en œuvre réelle relève la plupart du temps de compromis généralement guidés par le savoir-faire de l'ingénieur.

(1) On peut faire référence à la répartition spectrale de l'énergie cinétique turbulente, connue comme le "spectre de Kolmogorov", qui illustre de manière très simple la hiérarchie des échelles, des grandes échelles porteuses d'énergie aux échelles de plus en plus petites et de moins en moins énergétiques.

(2) Cette étendue est le résultat des non-linéarités des équations du mouvement qui donne naissance à une gamme étendue d'échelles spatiales et temporelles. Cette gamme est une fonction croissante du nombre de Reynolds, Re , mesurant le rapport entre force d'inertie et force visqueuse.

(3) L'hypothèse selon laquelle la résolution complète des équations de Navier-Stokes permet la simulation de la turbulence est généralement admise, tout du moins dans la gamme des écoulements sans choc.

(4) Il s'agit d'un problème régi par des conditions initiales et aux limites.