



Compte-rendu des Journées CLIMERI-France 3-4 février 2022

Les journées CLIMERI-France se sont tenues en visio-conférence sur 4 demi-journées les 3 et 4 février 2022. Chacune des 4 sessions a réuni entre 70 et 80 personnes. Nous vous proposons ici un résumé synthétique des principaux points discutés.

Session 1. Points d'informations générales CLIMERI-France, outils et projets liés

La session d'introduction a permis d'informer les participants sur l'infrastructure de recherche (IR) CLIMERI-France et sur plusieurs projets en cours en lien avec l'IR :

- le projet financé Gaia Data qui a démarré fin 2021 pour 8 ans,
- le projet de PEPR TRACCS en cours d'évaluation dont le périmètre est plus large que celui de CLIMERI-France mais qui a un lien fort avec l'IR,
- une synthèse des activités du GDR DEPHY sur les enjeux des développements de la physique des modèles,
- et sur deux nouveaux projets européens ESM2025 et NextGEMS qui ont été sélectionnés lors de l'appel visant la préparation de la prochaine génération de modèles du système Terre.

Cette session a également permis d'informer sur les activités de l'IR liées à la plateforme d'analyse multi-modèles, les logiciels communs et le groupe de travail CLIMERI-tech.

Les présentations ont suscité des questions et discussions qui font ressortir les points principaux suivants :

- Un besoin de bien définir le contour de CLIMERI-France, en particulier vis-à-vis de la modélisation régionale (tout comme vis-à-vis des services labellisés, cf. la session 2).
- Si le projet TRACCS est retenu, le besoin de préciser les articulations avec l'IR mais aussi comment seront gérées les ressources humaines supplémentaires (CDD longs, CDI).
- Concernant la plateforme d'analyse et les liens avec le projet GAIA-DATA :
 - La communauté est très intéressée par le développement de nouveaux outils facilitant l'accès et l'analyse des données des modèles et d'observations. La solution PANGEO arrive à l'IPSL (Jupyter Notebook + catalogues intake + Dask, etc.) et permettra de répondre à la demande avec un catalogue de plus de 2 Po de données. GAIA-DATA apportera un soutien technique notamment pour le renouvellement du stockage à l'IDRIS et la mise en réseau direct du CNRM-CERFACS avec ESPRI. Le groupe ESPRI prévoit des formations et une documentation est en cours de construction pour accompagner les utilisateurs à utiliser ces nouveaux services.

- La communauté souhaiterait également savoir comment les développements qui seront faits pourront servir au-delà des analyses multi-modèles CMIP, en particulier l'analyse des simulations stockées sur les centres de calcul du TGCC et de Météo-France. L'infrastructure sous-jacente avec les centres HPC (TGCC et IDRIS) va être renforcée via les projets "GAIA-DATA" et "FITS" afin d'envisager i) l'accès aux données hors-CMIP6 depuis la plateforme CLIMERI (c.a.d. CICALAD+ClimServ) et ii) l'accès aux données depuis les centres eux-mêmes.
- La citation des données est devenue quasi obligatoire pour publier. L'accession à un DOI ou PID est nécessaire. ESPRI est habilitée à poser un DOI ou un PID sur un jeu de données, y compris pour des simulations hors CMIP6. Actuellement au cas par cas, la procédure vise à être automatisée via le projet GAIA-DATA avec le soutien de dépôts nationaux de données.
- Concernant le groupe CLIMERI-Tech, les participants ont soulevé des questions concernant la pérennité des outils développés par d'autres partenaires européens, par exemple les outils de contrôle de qualité du DKRZ (cadre IS-ENES).
- La présentation du GDR Dephy a permis d'informer sur les développements en cours et a suscité des questions autour du tuning (voir session 2) et les paramétrisations des nuages.
- La présentation des nouveaux projets ESM2025 et NextGEMS a soulevé beaucoup de débats, repris en partie dans les discussions de la session 4, principalement autour de l'intégration avec les SHS (ESM2025) et sur la place des modèles globaux kilométriques (NextGEMS).

Session 2. Services labellisés, évaluation et tuning

- Composantes : Besoin de clarifier le rôle de CLIMERI par rapport aux services labellisés. Certains pensent que CLIMERI doit représenter ces services et en porter les besoins et les demandes (par exemple défendre la science, le besoin de CDI, etc.) "plus haut" au niveau des tutelles et du ministère car CLIMERI serait vu comme en étant l'interlocuteur. En pratique, CLIMERI ne joue pas ce rôle actuellement, mais relaie les aspects communs à l'ensemble des modèles climatiques et qui concernent l'infrastructure de modélisation du climat.
- Post-traitement : Le groupe de travail auto-constitué depuis environ 1 an a effectué un bon inventaire des outils permettant de réaliser atlas et post-traitement et progressé sur les différentes questions liées à leur développement, leur partage et les différents types d'utilisation. Depuis la réunion, une page dédiée au travail du groupe a été ouverte sur le site CLIMERI <https://climeri-france.fr/ressources/gt-atlas-post-traitements-simulations-climatiques/>
- Tuning : Les présentations ont suscité un vif intérêt et de nombreuses questions pour élargir à l'ensemble des composantes et des modèles couplés. Un relais CLIMERI sur ces questions semble pertinent, en lien avec la stratégie à adopter pour la réduction des biais des modèles. Approfondir ces questions demande aussi des accès à des modèles de type LES ayant les caractéristiques adéquates pour explorer les processus physiques ciblés qui ne sont pas nécessairement disponibles. Des questions similaires se posent certainement pour les autres milieux que l'atmosphère, ainsi que pour les couplages. Une coordination CLIMERI pourrait permettre d'avoir une dynamique

nationale cohérente sur ces sujets avec l'ensemble de la communauté concernée. On peut noter que la communauté française a une position assez avancée sur cette thématique émergente.

- Organiser un atelier dédié sur les incertitudes structurelles vs incertitudes paramétriques est vu comme un moyen de progresser, et constituerait une première étape par rapport aux objectifs énoncés dans le PEPR TRACCS sur ce sujet.

Session 3 Les enjeux du HPC et la préparation à l'exascale

Les journées CLIMERI-France ont été l'occasion d'aborder la question des enjeux du calcul de haute performance et l'arrivée des architectures accélérées pour la modélisation du climat. Les présentations ont été l'occasion de définir le contexte national de préparation à l'exascale ainsi que le contexte international. La session a également permis de discuter la problématique des nouvelles architectures plus hétérogènes et de faire un premier tour des réflexions en cours pour les deux modèles de climat IPSL et CNRM-Cerfacs ainsi que pour NEMO.

Le sujet a suscité un vif intérêt et a également mis en avant des inquiétudes de la communauté. Plusieurs points ressortent :

- Un besoin de clarifier les différents concepts concernant l'évolution des architectures de calcul. Un webinaire pourrait permettre de les approfondir.
- Un point d'information sur les stratégies et avancées des autres modèles de climat pour passer à l'exascale pourrait également être utile. Le workshop communautaire organisé par IS-ENES3 du 9 au 11 mai 2022 à Barcelone (et en hybride) pourra fournir des informations qui pourraient être ensuite relayées par un webinaire CLIMERI.
- Si la communauté a bien conscience des gains d'efficacité et de consommation électrique associés à l'utilisation de GPU, elle attire l'attention sur la nécessité de garder une partie non négligeable de CPU au sein des moyens de calcul nationaux, à la fois en raison de l'incertitude encore grande sur la capacité finale de nos modèles à utiliser efficacement les GPU et également du temps nécessaire pour le passage à ces architectures pour nos modèles couplant plusieurs codes. Il est important que ce message remonte de notre communauté.
- Si l'augmentation de résolution est mise en avant quand on parle de l'exascale, il serait important de faire le point sur les différentes opportunités que ces machines pourraient offrir, en particulier vis-à-vis de grands ensembles de simulations. Une opportunité quant à la quantification des incertitudes des modèles (voir session 2) est sûrement à saisir ici.
- L'analyse faite dans le cadre du projet exascale-France met en avant la nécessité d'augmenter les ressources humaines, à la fois sous forme de CDD longs et de CDI. Cependant, la communauté soulève la grande difficulté de recruter des experts de haut niveau dans le domaine du calcul haute performance. Climeri-France peut avoir un rôle à jouer ici (voir session 1).

Session 4 Retour sur CMIP6 & CORDEX et la suite

La session 4 des journées CLIMERI-France a permis de donner la parole d'une part à un certain nombre d'acteurs français de la modélisation régionale et de l'exercice CORDEX, et d'autre part de la modélisation globale et de l'exercice CMIP6. La session s'est terminée sur une présentation des perspectives en cours d'élaboration pour CMIP7.

Les principaux points abordés lors des discussions sont les suivants :

- Quel apport de CLIMERI pour l'animation nationale pour CORDEX ? La communauté CORDEX s'accorde sur le besoin d'une mutualisation de la plateforme d'analyse multi-modèles ESPRI et d'outils techniques liés au workflow et aux analyses. Par ailleurs, le besoin de favorisation des échanges GCM-RCM au niveau national sur des questions transverses, des liens entre groupes locaux et niveau national et des échanges nationaux CORDEX sous forme d'une journée (ou de séminaires réguliers ?) sont également soulignés.
- La valeur ajoutée/retirée des modèles régionaux pour des simulations longues de projections climatiques futures est un sujet à explorer. De même, les alternatives aux grands ensembles sont à discuter avec en particulier l'opportunité d'explorer des approches possibles pour les communautés régionales et globales. Le sujet des composantes à intégrer dans les futurs systèmes couplés de climat régional, selon les régions, a également été abordé.
- La question du couplage temporel entre CORDEX et CMIP : Le découplage entre la réalisation des simulations CMIP et leur utilisation pour CORDEX est intentionnel (d'autres communautés, comme celle des impacts, sont encore plus décalées). Il permet d'avoir des retours sur les modèles forceurs avant leur sélection (bogues, analyses multi-modèles poussées). Il y avait aussi la volonté de finaliser de grands ensembles CORDEX-CMIP5 et leur analyse pour publier à temps pour l'AR6.
- Beaucoup de discussions ont eu lieu autour de la haute voire très haute résolution, à la fois sur le régional et le global. La communauté dans son ensemble se questionne sur les avantages, d'une part d'un point de vue purement scientifique mais aussi des limitations que cela impliquerait en termes d'évaluation des incertitudes aux échelles climatiques.
- Des présentations ont synthétisé les derniers résultats de CMIP autour de la sensibilité climatique des nouveaux modèles d'une part, et de l'analyse de l'origine de la variabilité basse fréquence des modèles français.
- Plusieurs discussions ont eu lieu autour du bilan et des perspectives de l'exercice CMIP. Par rapport à l'exercice CMIP6, la question s'est posée de l'intérêt des très nombreux MIPs et du choix de la participation des laboratoires (sur la base du volontariat côté français). La question de l'ECS et de la réduction de valeurs probables dans l'AR6 à partir d'émulateurs calibrés sur les observations et autres sources indépendantes a également fait réagir.
Pour la suite, la question de l'utilité des exercices de type CMIP est posée pour les études sur l'impact sociétal du changement climatique comme leur influence sur les décideurs. La réalisation d'un exercice CMIP7 pour ~2025 pour l'AR7 risque de

provoquer un conflit de calendrier par rapport aux changements de calculateurs. Certains voudraient aussi pouvoir continuer à exploiter les modèles utilisés pour CMIP6 avant de se lancer dans un nouvel exercice.

La question d'un découplage entre les calendriers CMIP et les rapports IPCC devrait se poser. Des interrogations ont également été soulevées quant au traitement des réponses individuelles et institutionnelles au sondage du WCRP en cours ("CMIP survey"). La question de l'ajout de composantes aux prochains modèles globaux se pose en particulier pour les calottes.

Complément au CR.

Au cours de la réunion, de nombreux échanges ont eu lieu et des inquiétudes ont été soulevées sur la notion de "digital twins" (ou "jumeaux numériques") et la modélisation à ultra-haute résolution comme solution pour améliorer les projections climatiques. Les discussions ont également bénéficié d'échanges informels et d'une session complémentaire organisée en fin de réunion. L'inquiétude est exacerbée par la place prise par la notion de jumeaux numériques dans les activités de modélisation du climat et de ses changements et la façon dont ils sont relayés dans l'article de [Bauer et al, NCC, 2021](#). A l'issue de la réunion un groupe s'est constitué et a envoyé une lettre à la direction de CLIMERI-France demandant de relayer leur positionnement sur les jumeaux numériques.

Nous reportons ci-dessous les points de positionnement extraits de cette lettre qui complètent les éléments du CR des journées :

- La place prise par les Digital Twins tels que présentés par exemple dans l'article de Bauer et al., NCC, 2021, inquiète une partie de la communauté CLIMERI et le commentaire de Bauer, Stevens et Hazeleger qui en fait la promotion dans Nature Climate Change pose un véritable problème d'éthique scientifique ; ceux qui ont fait un travail d'analyse scientifique du texte y voient, soit une ignorance des enjeux de la modélisation climatique, soit une forme de malhonnêteté intellectuelle.
- Les modèles globaux à résolution kilométrique (au coeur de ces soi-disant "jumeaux numériques") peuvent permettre d'éclairer certaines questions de physique du climat, mais ils ne pourront pas dans un terme envisageable permettre de réaliser des simulations globales de plusieurs milliers d'années, absolument nécessaires pour : i) initialiser les composantes lentes, ii) ajuster les paramètres, iii) explorer les incertitudes de modélisation. Aucun des trois points ne disparaîtra avec la montée en résolution.
- Les modèles à physique paramétrée ne doivent pas "co-exister" sur un pied d'égalité avec de soi-disant "jumeaux numériques" dans un éco-système de modélisation. Le rôle central des modèles à physique paramétrée dans l'anticipation du changement climatique et dans notre pensée même du climat doit être réaffirmé. Le développement physique de nos modèles se base depuis longtemps sur une cascade de configurations de modèles (LES, 1D, WTG, zoomés, guidés, aquaplanètes). Nul ne conteste que des modèles globaux kilométriques permettant de réaliser des simulations de quelques mois ou quelques années à terme entrent dans cette

hiérarchie d'outils, même si les développeurs de paramétrisations atmosphériques regrettent que l'accent ne soit pas mis davantage sur des simulations LES (à 100 m de résolution ou moins) sur des grands domaines. En revanche, cette utilisation d'une hiérarchie de modèles ne doit pas fragiliser le rôle très particulier et incontournable des modèles à physique paramétrée (globaux ou régionaux). L'idée même que l'incertitude paramétrique disparaîtrait avec la montée en résolution est d'ailleurs très probablement fausse.

- Les mots de "Digital Twins" ou "jumeaux numériques" doivent être proscrits de notre vocabulaire quand on parle de modélisation climatique. Ils ont une existence bien définie en ingénierie). Nous sommes face à un acte de communication non scientifique. Nous ne devons pas essayer de nous y raccrocher, soit en disant que nos modèles globaux le seraient déjà (ce que certains ont revendiqué à un moment), soit en se disant qu'on ne va en retenir que les aspects d'interface avec les questions de société, dans un esprit de "service climatique". Le "comment" dans Nature Climate Change en lui-même est un excellent plaidoyer pour éliminer ces termes de nos discussions.
- La communauté a bien conscience des gains d'efficacité/de consommation électrique potentiels liés au portage sur GPU de nos codes, et le travail de portage est maintenant bien engagé. Mais ces portages représentent un défi majeur, en partie spécifique à chaque code. Les conditions d'attractivité de la recherche très dégradées et les surcharges de tâches administratives résultant de la complexification des modes de financements, des réductions de personnels support/soutien dans les laboratoires (REF -11% de personnels au CNRS depuis 2010) ... créent un contexte globalement défavorable. La mise à disposition de personnels à fort niveau d'expertise sur des statuts CDD nous semble illusoire. En l'absence d'une politique de recrutement ambitieuse, le portage sur GPU se fera sur des échelles de temps incompatibles avec celles envisagées pour l'arrivée des nouvelles machines exascale (2024). De ce fait, les développeurs attendent de leurs relais dans ces instances que tout soit fait pour qu'une large part de CPU traditionnels soit maintenue sur les gros calculateurs pour une phase transitoire dont il est pour l'instant difficile de déterminer la durée.

Signataires de la lettre et porteurs des points listés ci-dessus par ordre alphabétique :

V. Balaji, Xavier Capet, Frédérique Cheruy, Maëlle Coulon-Decorzens, Fleur Couvreur, Julie Deshayes, Frédéric Hourdin, Claire Levy, Jean-Baptiste Madeleine, Sébastien Masson, Juliette Mignot, Catherine Rio, Romain Roehrig, Roland Seferian, Etienne Vignon, Najda Villefranque