

radiatifs sauf sur les mois d'hiver où l'effet radiatif des nuages est peu biaisé dans le solaire et plutôt surestimé dans l'infrarouge. Cela conduit également à l'amélioration du bilan radiatif net à la surface tous les étés (Figure V.16 (a)), avec certains étés où des simulations du PPE⁴² atteignent les valeurs de référence.

L'augmentation du contenu intégré en eau liquide dans le PPE⁴² introduit en revanche un biais qui n'était pas ou peu présent dans la configuration de référence du modèle (Figure V.13 (c)). On retrouve également que les couvertures nuageuses totales ont été contraintes par l'exercice du tuning, avec des nuages moyens presque systématiquement plus couvrant et des nuages haut légèrement moins couvrant (Figure V.14). Contrairement à ce qui était vu sur l'année 2018, certaines périodes (d'été ou d'hiver) ont plus de couverture des nuages bas dans le PPE⁴², et d'autres en ont moins (Figure V.14 (b)). Finalement, les métriques de cette expérience de tuning restreignent bien les plages de paramètres des paramétrisations des nuages bas avec un effet sur leur couverture nuageuse.

La configuration de référence du modèle simulait systématiquement trop peu de précipitations (en moyenne sur 25 mailles), même au printemps-été, excepté au mois de juin 2018. La métrique de précipitations a donc été choisie sur un mois exceptionnellement pluvieux. Les différences entre les deux PPE sont plutôt petites.

Le PPE⁴² a tout de même tendance à faire un peu plus de précipitations convectives et un peu moins de précipitations grande échelle au printemps et en été (Figure V.15). L'intensité des orages points de grilles est également diminuée dans le PPE⁴², comme pour ceux très intenses d'avril 2016 ou d'avril 2020. L'écart type spatial des précipitations était surestimé sur toute la période dans la simulation de référence. Les simulations du PPE⁴² ont des écart-type spatiaux des précipitations plus faible sur presque toute la période que celles du PPE⁴¹ (Figure V.15 (b)).

L'ordre des trois meilleures simulations est le même sur toute la période 2016-2020 pour l'effet radiatif solaire des nuages, le LWP et le IWP, et la couverture des nuages haut. Ce n'est en revanche pas le cas pour l'effet radiatif des nuages dans l'infrarouge et la couverture des nuages bas, qui ne sont pas directement ciblés dans les expériences, ainsi que pour les précipitations régionales, qu'elles soient convectives ou grande échelle, ainsi que sur leur variabilité spatiale.

Ces résultats confirment que la simulation Best CREc fait décroître la variabilité des pluies et la ramène à un niveau nettement plus comparable aux observations. Il serait intéressant de voir le lien avec les paramètres contrôlant le déclenchement de la convection profonde. *[Eventuellement à vérifier d'ici la soutenance]*

Pour conclure, l'évaluation menée sur la période 2016-2020 montre que les métriques choisies contraignent de manière similaire l'année 2018 et la période 2016-2020. Cela valide la robustesse de ces métriques et illustre la pertinence de choisir des métriques orientées processus comme ici.

V.3 Place dans la chaîne de tuning actuelle

Dans la perspective d'intégrer les métriques développées ici dans la chaîne de tuning actuelle de LMDZ, nous évaluons l'effet du pré-conditionnement 1D sur les simulations au SIRTA, puis l'effet des métriques 1D+SIRTA sur la simulation du climat global.

V.3.1 Effet du pré-conditionnement 1D sur les simulations au SIRTA

Cette évaluation est faite en comparant un ensemble PPE¹ de trente simulations tirées au hasard dans l'hypercube initial non restreint de l'expérience de tuning, avec trente simulations tirées au hasard dans le PPE⁴¹, ayant expérimenté 40 vagues de réduction de l'espace des paramètres libres avec les métriques 1D uniquement.

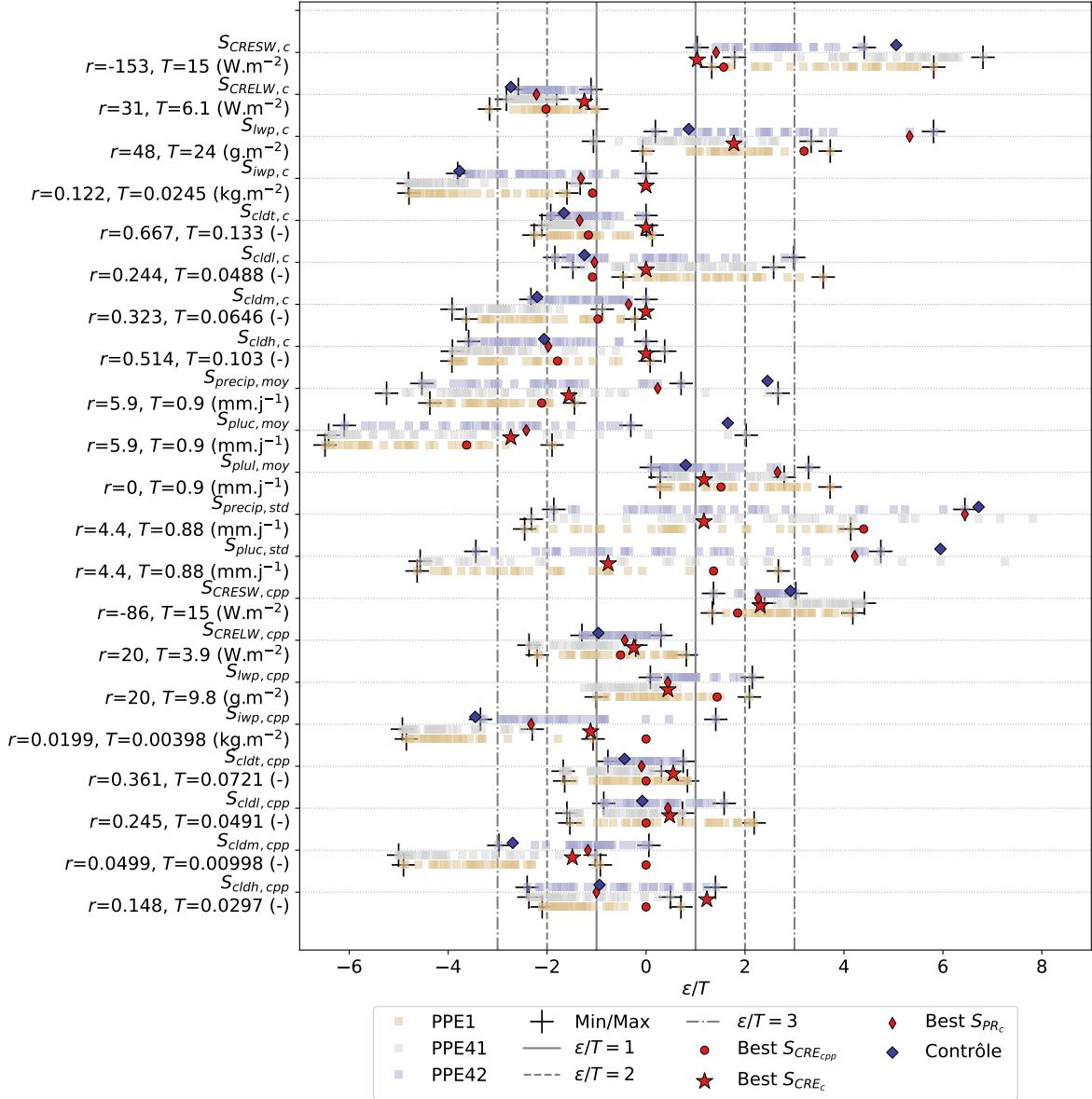


FIGURE V.17 – Même figure que Figure V.8 pour trente membres quelconques du PPE¹ en jaune, du PPE⁴¹ en gris et du PPE⁴² en bleu. Les scores sur la période convective du 22 mai au 12 juin 2018 sont notés _c, et ceux sur la période de convection peu profonde du 13 au 30 juin 2018 sont notés _{cpp}. Les variables tracées sont liées à la simulation des nuages et des précipitations : effets radiatifs des nuages dans le solaire (*CRESW*) et l'infrarouge (*CRELW*), quantité d'eau condensée intégrée sur la verticale pour l'eau liquide (*lwp*) et la glace (*iwp*), couvertures nuageuses pour les nuages bas (*cldl*), moyen (*cldm*) et haut (*cldh*), ainsi que la couverture nuageuse totale (*cldt*). Enfin, sont tracés les scores des précipitations totales en moyennes régionales (*precip, moy*) et leur écart type spatial (*precip, std*), ainsi que les scores des précipitations régionales grandes échelles (*plul, moy*) et convectives (*pluc, moy*) avec leur variabilité spatiale (*pluc, std*). Les croix noires marquent les valeurs minimales et maximales des scores des ensembles. Les scores de la simulation de contrôle sont indiqués avec un marqueur bleu, et ceux des trois simulations sélectionnées “Best S_{CRE} ”, “Best S_{PR} ” et “Best S_{CRE} ” sont indiqués avec des marqueurs rouges. Pour les couvertures nuageuses pour lesquelles nous ne disposons pas d'observations, celles-ci sont remplacées par valeurs des simulations “Best S_{CRE} ”, “Best S_{PR} ” pour les périodes de convection peu profonde et profonde, respectivement.

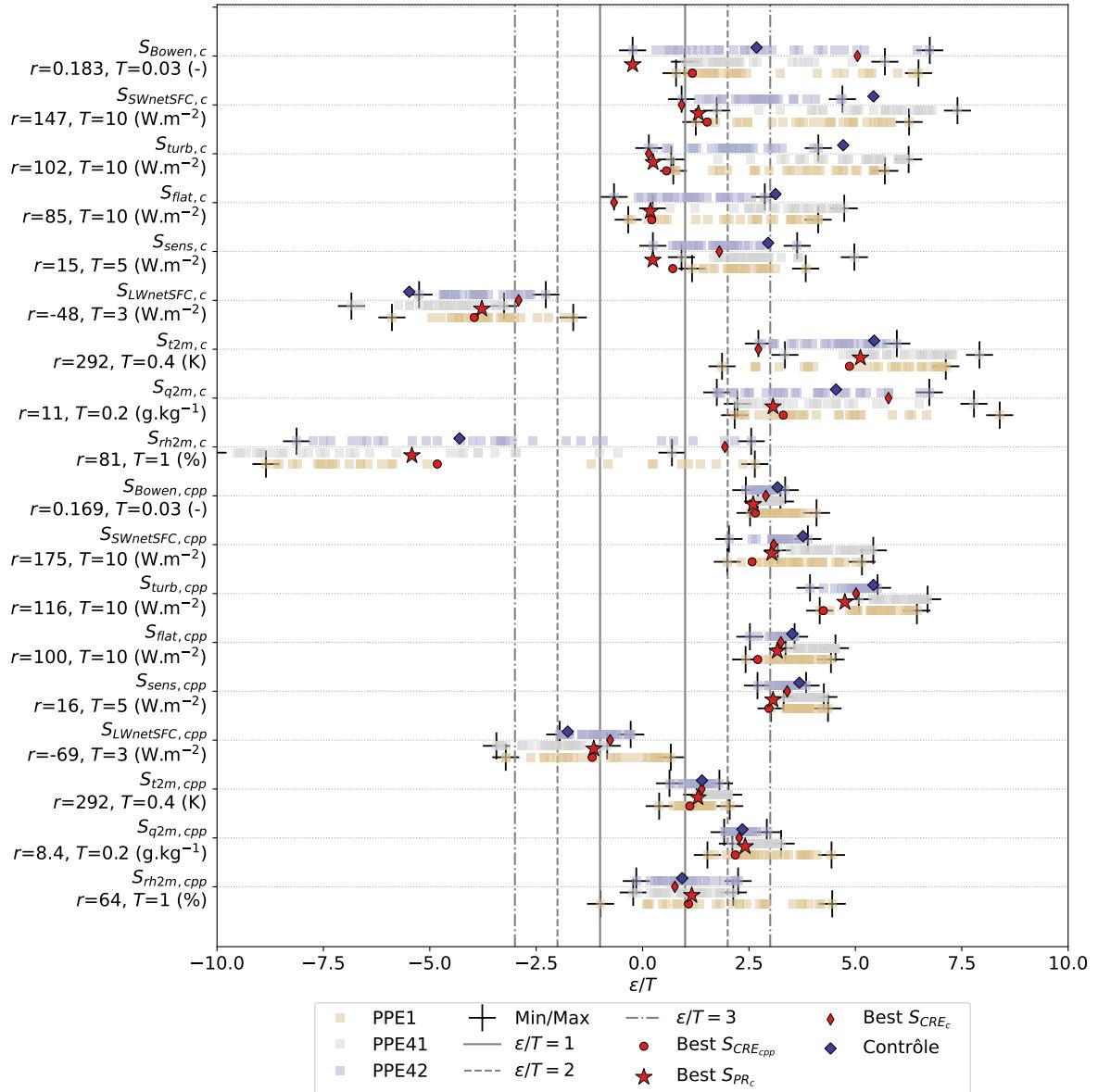


FIGURE V.18 – Même figure que Figure V.17 pour les métriques liées au bilan d'énergie à la surface : rapport de Bowen (*Bowen*), rayonnement solaire net (*SWnetSFC*), somme des flux turbulents (*turb*), flux de chaleur latente (*flat*) et sensible (*sens*), rayonnement infrarouge net à la surface (*LWnetSFC*), température atmosphérique à deux mètres (*t2m*), humidité spécifique à deux mètres (*q2m*) et humidité relative à deux mètres (*rh2m*). Toutes les valeurs de référence utilisées ici pour le calcul des scores sont issues des observations du SIRTA. Les valeurs des tolérances à l'erreur sont ici choisies comme des valeur absolues, toujours égales pour les métriques des deux périodes d'une même variable.

Le pré-conditionnement par des métriques 1D portant sur des cas de convection peu profonde a finalement un effet assez modéré sur la simulation des nuages de la période de convection peu profonde du 12 au 30 juin 2018 au SIRTA (métriques indicées *cpp* en bas de la Figure V.17). Ce pré-conditionnement entraîne en moyenne une légère diminution de la couverture des nuages bas ($S_{cldl,cpp}$) et du LWP ($S_{LWP,cpp}$) entraînant une légère diminution de l'effet radiatif solaire et infrarouge des nuages (valeurs moins négatives de $S_{CRESW,cpp}$ et moins positives de $S_{CRELW,cpp}$). La sous-estimation moyenne de l'effet radiatif solaire des nuages, cible prioritaire du tuning au SIRTA pour cette période, est renforcée par le pré-conditionnement 1D, sans doute parce les

vecteurs de paramètres conduisant à la simulation excessive de nuages bas sur les cas unicolumnes sont éliminés. Cela n'empêche pas que les 30 simulations de la vague 42, après l'ajout des contraintes au SIRTA, sont toutes au moins aussi bonnes que la simulation de contrôle sur cette métrique.

La sous-estimation de l'effet radiatif des nuages dans le solaire ($S_{CRESW,cpp}$), plus forte en moyenne sur le PPE⁴¹ que sur le PPE¹, entraîne une augmentation du rayonnement solaire net à la surface ($S_{SWnetSFC,cpp}$, Figure V.18) et de la somme des flux turbulents ($S_{turb,cpp}$) dont l'augmentation est essentiellement due à l'augmentation du flux de chaleur latente ($S_{flat,cpp}$) ; ce qui résulte en une augmentation de la surestimation de la température à deux mètres.

L'impact le plus clair du pré-conditionnement 1D est l'élimination des simulations avec un fort biais humide ($S_{q2m,cpp}$ et $S_{rh2m,cpp}$ Figure V.18). On sait en effet que l'humidité de surface est contrôlée au premier ordre par la convection de couche limite (Hourdin et al., 2020; Diallo et al., 2017b). Le pré-conditionnement élimine donc probablement des configurations insuffisamment convectives qui conduisent à une accumulation d'humidité dans la couche limite de surface. Une fois ces configurations éliminées, le rapport de Bowen devient fortement contraint dans les simulations au SIRTA ($S_{Bowen,cpp}$) dans lesquelles on a imposé en plus le facteur d'aridité.

Le pré-conditionnement 1D a un impact relativement faible en moyenne sur la simulation des nuages de la période de convection profonde (22 mai au 11 juin 2018), et ce dans le même sens que pour la période de convection peu profonde. Les nuages bas sont un peu moins couvrants en moyenne dans le PPE⁴¹ que dans le PPE¹ ($S_{cldl,c}$, Figure V.17) avec un contenu en eau plus faible ($S_{LWP,c}$), renforçant le manque d'effet (négatif) des nuages dans le solaire ($S_{CRESW,c}$).

De façon surprenante, les membres du PPE⁴¹ simulent des précipitations nettement plus diverses que celles du PPE¹ ($S_{precip,moy}$), avec quelques membres qui en simulent un peu moins, et surtout d'autres qui en simulent beaucoup plus. Ces fortes précipitations sont expliquées par les pluies convectives ($S_{pluc,moy}$). On constate simultanément une diminution en moyenne des pluies grande échelle ($S_{plul,moy}$) et une augmentation de la variabilité des précipitations ($S_{precip,std}$), cette dernière semblant donc liée à la paramétrisation de la convection. L'effet du pré-conditionnement 1D sur la simulation de la période de convection profonde peut s'expliquer par : (1) un effet direct de la convection peu profonde au travers du déclenchement et de la fermeture de la convection profonde, le pré-conditionnement 1D semblant cependant favoriser une convection peu profonde plus intense, comme le suggère la diminution de $S_{rh2m,c}$, favorisant donc le déclenchement de la convection profonde [ca aurait valu le coup de sortir ale et alp ...] et (2) un effet indirect, au travers d'une modification de l'état de l'atmosphère par la convection peu profonde.

En résumé :

- le pré-conditionnement 1D tend à sélectionner en moyenne des vecteurs de paramètres de convection peu profonde produisant moins d'effet radiatif des nuages, que ce soit dans les périodes de convection peu profonde que dans les périodes de convection profonde, ce qui semble renforcer les biais du modèle au SIRTA. Ce résultat peut paraître décevant dans la mesure où le tuning 1D ne suffit pas. En revanche, on voit clairement que la combinaison du tuning 1D et 3D permet de sélectionner à la vagues 42 des simulations dont les forçages radiatifs sont tous en meilleur accord avec les observations que la simulation de contrôle ;
- le pré-conditionnement 1D a un effet beaucoup plus marqué sur l'humidité en surface et le rapport de Bowen, confirmant l'importance de la convection de couche limite dans ce contrôle ;
- le pré-conditionnement 1D favorise la convection profonde pour des raisons qui restent à élucider.

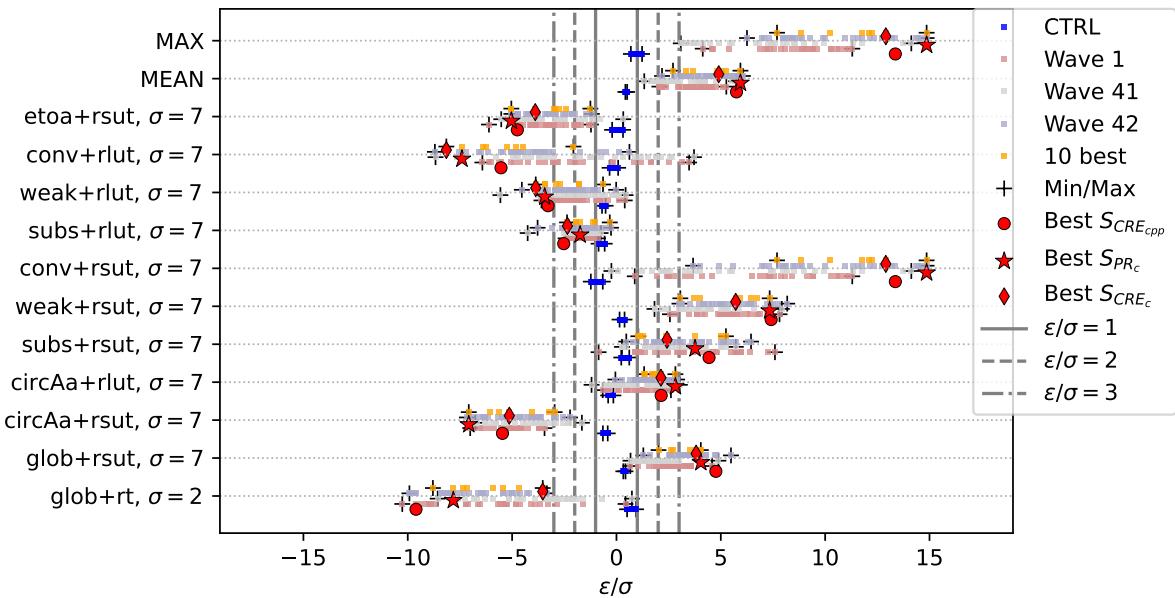


FIGURE V.19 – Scores des métriques appliquées aux simulations 3D globales, utilisées classiquement pour le tuning de LMDZ. Les métriques sont calculées sur la seconde année d'une simulation de deux ans réalisée sur des SST climatiques. Les tolérances appliquées pour calculer les scores sont de 2 W m^{-2} pour le bilan énergétique global (`glob+rt`) et de 7 W m^{-2} pour les autres métriques. Les scores sont calculés par rapport aux observations CERES-EBAF (Loeb et al., 2009).

V.3.2 Evaluation sur la simulation du climat global

Si cela avait été possible, la compatibilité des métriques 1D et SIRTA avec les métriques globales de LMDZ aurait été évaluée en rajoutant celles-ci dans les vagues suivantes. Puisque le NROY⁴² est très petit, il n'est pas possible de sélectionner suffisamment de bons vecteurs de paramètres dans le NROY⁴² pour itérer sur une nouvelle vague 43 (par abus de langage on dira que le NROY⁴² est vide).

Pour anticiper un futur exercice de tuning et contourner ce problème, trente vecteurs de paramètres des vagues 1, 41 et 42 sont sélectionnés aléatoirement, ainsi que les vecteurs de paramètres des dix meilleures simulations, pour réaliser des simulations globales de deux ans avec la configuration standard de LMDZ. Les métriques radiatives définies Section III.4.2 sont calculées sur la deuxième année de simulation, la première étant considéré comme un spin-up.

Il s'avère que les simulations ainsi sélectionnées sont loin d'être optimales du point de vue du climat global. La métrique `conv+rsut`, rayonnement solaire montant au sommet de l'atmosphère dans les régions convectives, est particulièrement mal simulée par les ensembles de la vague 42 et les dix meilleures simulations au SIRTA. Ce qui semblait un avantage important pour les simulations au SIRTA, à savoir le renforcement de l'effet radiatif des nuages dans les situations de convection profonde, se traduit dans les tropiques par une forte surestimation du rayonnement réfléchi. Cet effet est visible à l'échelle du globe avec un `glob+rsut` lui aussi sous-estimé, produisant un déficit de bilan global `glob+rt` d'une dizaine de W m^{-2} .

En revanche, si on regarde les métriques associées aux zones de subsidence forte (`subs`) ou faible (`weak`, typiquement les zones de cumulus d'alizé), on constate que les meilleures simulations de la vague 42 restent proches des meilleures simulations de la vague 41. On peut donc penser que pour ces métriques, le tuning 3D pourra permettre de sélectionner des valeurs raisonnables des métriques.

Ceci paraît plus incertain pour les métriques convectives, pour lesquelles on semble avoir

une contradiction forte entre les zones de paramètres privilégiées par le tuning tropical et par le tuning au SIRTA. Plusieurs explications sont possibles. Il se peut qu'il y ait un manque de physique pour expliquer un comportement différencié de la convection au SIRTA et dans les tropiques. Il est possible également que le tuning au SIRTA ait privilégié un renforcement de l'effet radiatif des nuages associés à la convection pour compenser un autre problème. Une autre possibilité est que le réglage des paramètres convectifs soit dépendant de la configuration, soit via la résolution qui est plus fine en mode zoomé (on peut s'attendre à ce qu'il faille des paramétrisations moins actives à des résolutions plus fines où une partie du transport va être réalisé par la dynamique du modèle), soit via l'environnement qui est contrôlé en mode guidé (si les paramétrisations sont réglées pour répondre correctement étant donné un bon profil en entrée en guidé, elles risquent de produire de mauvais profils en sortie si les profils en entrée ne sont pas bons en standard). Autant de questions ouvertes par ce travail qui soulignent le potentiel des approches utilisant les PPEs sous contrainte.

Une autre question ouverte par ce travail est celle de savoir dans quelle mesure il pourrait être souhaitable ou non de privilégier un tuning d'une version régionale (zoomée ou à aire limitée) au dépend de l'ajustement global si l'objectif est d'étudier le climat régional uniquement.

Si en revanche on souhaite introduire les métriques du SIRTA dans un processus de tuning global, il faut certainement introduire beaucoup plus tôt des métriques globales pour éviter de “perdre du temps” (de calcul et d'ajustement) dans des zones de l'espace des paramètres non pertinentes.

De là sont nées des discussions, au sein de l'équipe LMDZ et dans le cadre d'un atelier de travail avec Saloua Balhane, Gerhard Krinner et Aude Champouillon, pour voir dans quelle mesure il serait possible de contraindre les métriques globales sur la base de configurations plus légères, avec l'idée à la fois de réduire la durée de la simulation (certaines métriques globales s'ajustent en moins d'un an) et la taille de la grille horizontale.

Suivant une proposition de Gerhard Krinner, nous nous sommes focalisés sur le mois d'octobre, qui est à la fois proche de l'équinoxe, donc avec une certaine symétrie du rayonnement en latitude, et pas encore trop de neige dans l'hémisphère nord.

En termes de résolution nous avons utilisé, en plus de la résolution standard LR (144 points en longitude et 143 en latitude), la configuration 5LR utilisée pour CMIP6 (96 points en longitude et 95 en latitude) et une Very Low Resolution (VLR avec 64 points dans les deux directions).

La Figure V.20 montre une comparaison, pour trois métriques et en grille LR (la grille standard des simulations CMIP6 avec IPSL-CM6ALR), du calcul en moyenne mensuelle sur le mois d'octobre versus en moyenne annuelle (se reporter à la Figure III.15 pour la description des métriques en question). Pour évaluer la part des comportements provenant de la variabilité interne, nous montrons également les résultats pour 20 années simulées avec la configuration de contrôle. Sans surprise, la variabilité interne est plus forte quand on se concentre sur un mois d'octobre. La seconde ligne de la figure montre les mêmes ensembles de simulations du mois d'octobre réalisées sur deux calculateurs différents (sauf pour la simulation de contrôle où les 20 années n'ont été réalisées que sur l'un des calculateurs). Sur deux calculateurs différents, et sauf présence de bug dans le modèle ou les compilateurs, les simulations doivent avoir le même comportement climatique. Mais les différences dans les opérations faites à la précision de la machine introduisent une perturbation qui s'amplifie avec le temps. Au bout d'une quinzaine de jours les trajectoires des deux mêmes simulations vont ainsi totalement diverger. On voit que la variabilité interne de la simulation de contrôle (dispersion sur l'axe des ordonnées) est tout à fait comparable à celle obtenue en changeant de calculateur (dispersion des PPEs autour de la diagonale).

On constate cependant que pour ces trois métriques au moins, l'écart entre métrique annuelle

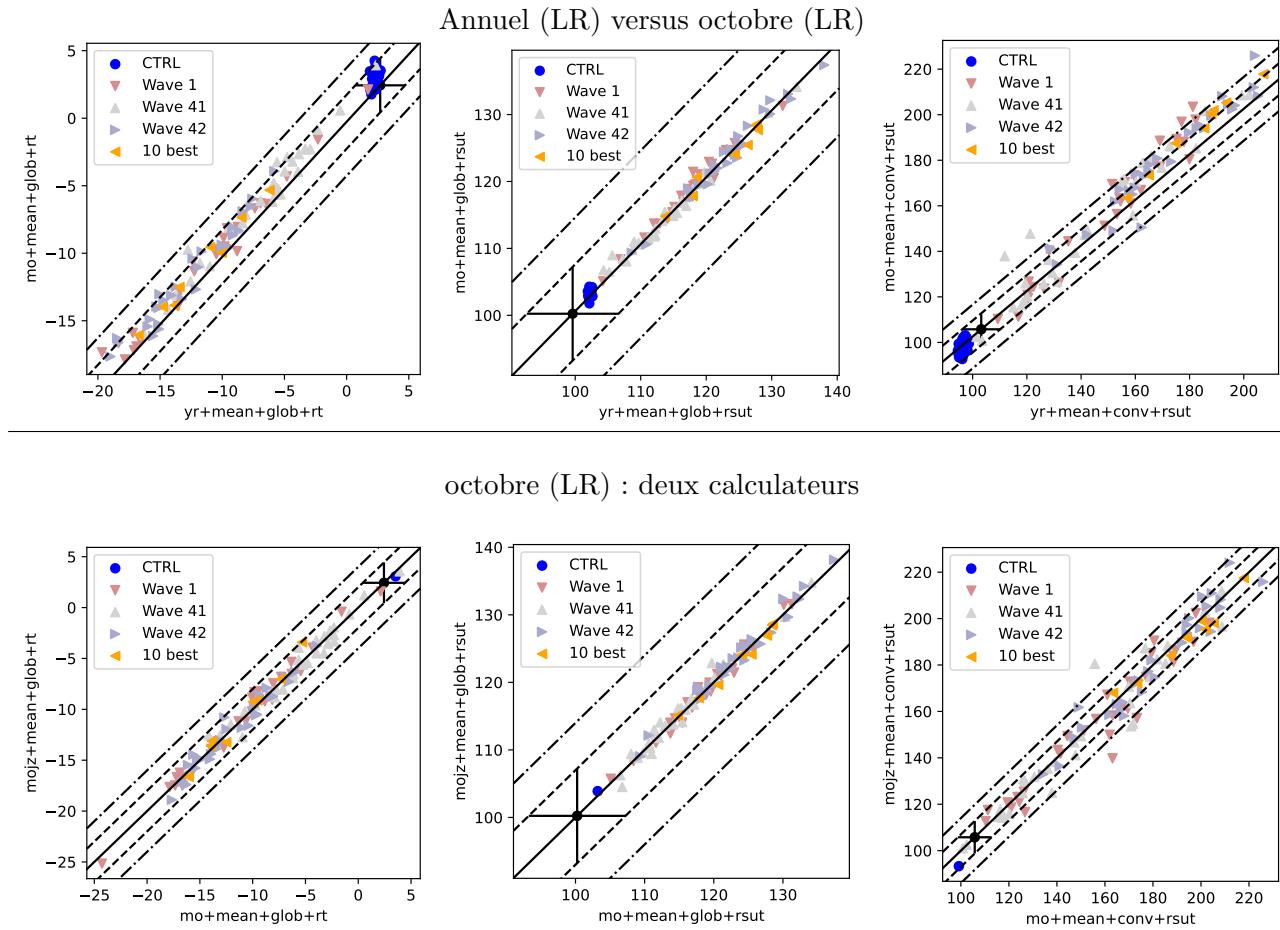


FIGURE V.20 – Diagramme de dispersion comparant trois métriques pour différentes configurations du modèle : colonne de gauche le bilan total au sommet (`glob+rt`), colonne du milieu le flux solaire montant global (`glob+rsut`), colonne de droite le flux solaire montant moyen dans les régions convectives sur les océans tropicaux (`conv+rsut`). La ligne du haut montre la comparaison du mois d'octobre (ordonnées) avec la moyenne annuelle (abscisse) pour chacune des métriques en configuration LMDZ6ALR. La ligne du bas compare les métriques obtenues pour les mêmes PPE pour le mois d'octobre, mais en réalisant les simulations sur deux calculateurs différents (le calculateur `jean-zay` de l'ideis et le calculateur `adastral` du Cines).

et mensuelle reste inférieur à deux fois la tolérance à l'erreur. Sur ces métriques en particulier, il n'y a pas de biais systématique entre les deux calculs (les points étant distribués autour de la diagonale) et il est donc envisageable, et à notre avis souhaitable, d'inclure des vagues avec cette configuration pour économiser du temps humain et de l'énergie dans le processus de tuning.

Si on regarde maintenant le changement de résolution Figure V.21 on constate que des différences supplémentaires viennent s'ajouter, avec une dispersion qui s'accroît et des biais (décalages systématiques par rapport à la diagonale). Notamment le flux solaire montant total au sommet de l'atmosphère est renforcé en configuration VLR, et c'est encore plus vrai dans la région de convection. La dégradation obtenue pour la configuration VLR semble en général trop importante pour utiliser cette configuration légère dans la chaîne de tuning.

Il apparaît clairement qu'un tuning sur le mois d'octobre en configuration LR et même 5LR devrait suffire à restreindre l'espace des paramètres dans une région compatible avec le tuning annuel de la configuration standard, sous réserve, suivant les métriques, d'accroître un peu la tolérance à l'erreur, ou de cibler une valeur légèrement différente de celle donnée par les

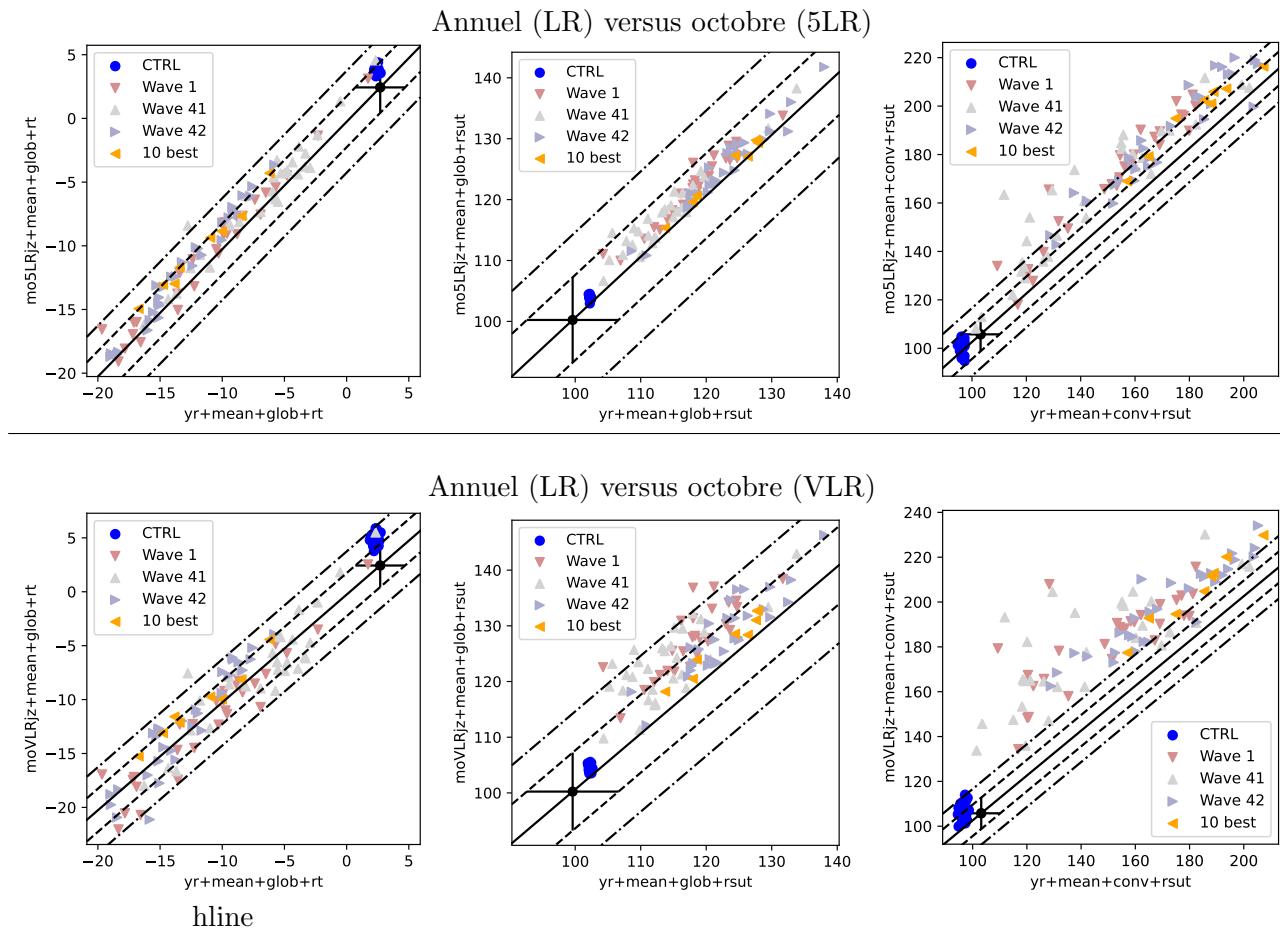


FIGURE V.21 – Comme la première ligne de la Figure V.20 excepté que les grilles utilisées pour les simulations mensuelles sont la grille 5LR (ligne du haut) et VLR (ligne du bas).

observations afin de corriger un biais de configuration.

V.4 Conclusion

Comme on a pu le voir, la mise en place de cette expérience de tuning au SIRTA a été semée d'embuches.

C'est d'abord la mise en place de la configuration, détaillée au chapitre précédent, et nous a finalement conduit à revoir à la baisse nos ambitions d'étudier le potentiel d'un tuning conjoint de deux modèles (ici LMDZ et ORCHIDEE) en ciblant explicitement leur couplage.

La conduite de l'expérience de tuning a ensuite butté sur un espace NROY vide du fait d'un mauvais choix de tolérance à l'erreur lié au fait d'avoir oublié le passage de 3 à 2 du seuil sur la sélection des invraisemblances. Cependant, la méthode consistant à régler des tolérances pour conserver au moins une simulation compatible avec les cibles s'est avérée très fructueuse dans des expériences menées depuis lors, et notamment dans celles du chapitre suivant.

Enfin, les résultats préliminaires semblent indiquer que les configurations permettant de se rapprocher des observations au SIRTA en renforçant le contenu en eau des nuages pendant la période de pluies convectives a tendance à surestimer ces nuages dans les tropiques, quand les mêmes vecteurs de paramètres sont utilisés avec la configuration climatique globale. Ces

résultats sont un peu trop préliminaires et se basent sur quelques simulations sélectionnées par le tuning au SIRTA. Mais si cette incompatibilité est confirmée, elle pourrait suggérer un besoin de développement physique permettant de contraster le comportement du modèle dans sa représentation des nuages convectifs tropicaux ou des moyennes latitudes.

Ces difficultés rencontrées et les erreurs commises font parties de la recherche scientifique et de la mise en pratique d'une nouvelle méthodologie et d'un nouvel outil. Elles sont porteuses d'enseignements et laissent entrevoir des perspectives positives que nous n'avons pu tester dans cette thèse faute de temps. Nous avons cependant pu expérimenter et proposer une méthodologie pour choisir les métriques et leurs tolérance à l'erreur à partir de PPEs, et avons montré la robustesse d'utiliser des métriques orientées processus au SIRTA. Nous avons beaucoup appris sur les informations qui pouvait être extraites ou non des PPEs et de l'analyse des meilleures simulations, et avons globalement tiré des leçons de cette expérience, et en particulier des erreurs faites, détaillées plus amplement au Chapitre VII. Cette expérience nous a également fait prendre conscience de la nécessité de porter nos efforts sur la conception de configurations légères de tuning global pour accompagner la mise en place de tuning plus ciblés, de façon à ne pas explorer inutilement des régions de l'espace des paramètres dans lesquelles les métriques radiatives de grande échelles sont trop mal simulées. L'utilisation d'un tuning en grille relativement grossière et sur le seul mois d'octobre semble prometteuse.

La sensibilité de la variabilités des précipitations aux paramètres du modèle laisse entrevoir un espoir de s'affranchir de cette variabilité et de pouvoir effectuer un tuning couplé intégrant des paramètres du modèles de surface, tout en ouvrant un certain nombre de questions. La sensibilité de la pluie à la résolution peut-elle être réglée par un tuning de certains paramètres ? Est-ce que le raffinement de la la grille lors du tuning pour limiter les erreurs de représentativité n'introduit pas autant de problèmes qu'il n'en règle ? Autant de questions ouvertes pour des études futures.

La difficulté à correctement simuler les effets radiatifs des nuages au SIRTA, et plus globalement à la surface des continents, l'arrivée dans le paysage du modèle ecRad qui lève une partie de ces limites et enfin le fait de disposer de références radiatives sur des scènes LES ont motivé en grande partie le chapitre suivant, portant sur les interactions nuages rayonnement et les compensations d'erreurs.