



HAL
open science

Attribution et reconstruction du rôle de la variabilité interne et des forçages externes sur le climat passé récent et du dernier millénaire

Beyrem Jebri

► To cite this version:

Beyrem Jebri. Attribution et reconstruction du rôle de la variabilité interne et des forçages externes sur le climat passé récent et du dernier millénaire. Milieux et Changements globaux. Sorbonne Université, 2020. Français. NNT : 2020SORUS162 . tel-03383697

HAL Id: tel-03383697

<https://theses.hal.science/tel-03383697>

Submitted on 18 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT
DE SORBONNE UNIVERSITÉ

Spécialité : Sciences de l'Environnement

École doctorale : « Sciences de l'Environnement d'Île-de-France »

réalisée au

Laboratoire d'Océanographie et du Climat
Expérimentations et Approches Numériques

présentée par

Beyrem JEBRI

pour obtenir le grade de :

DOCTEUR DE SORBONNE UNIVERSITÉ

Attribution et reconstruction du rôle de la variabilité interne et des forçages externes sur le climat passé récent et du dernier millénaire

soutenue le 10/06/2020

devant le jury composé de :

Pr. Francis CODRON	LOCEAN Sorbonne U., Paris	<i>Président</i>
Dr. Christophe CASSOU	CERFACS CNRS, Toulouse	<i>Rapporteur</i>
Pr. Hugues GOOSSE	UC de Louvain, Louvain-la-Neuve de Belgique	<i>Rapporteur</i>
Dr. Sophie CRAVATTE	LEGOS IRD, Toulouse	<i>Examinatrice</i>
Dr. Aurélien RIBES	CNRM, Toulouse	<i>Examineur</i>
Dr. Myriam KHODRI	LOCEAN IRD, Paris	<i>Directrice de thèse</i>
Pr. Sylvie THIRIA	LOCEAN Sorbonne U., Paris	<i>Directrice de thèse</i>

Remerciements

Ces quelques lignes sont l'occasion de rappeler qu'une thèse est loin d'être la simple rédaction d'un manuscrit et qu'il s'agit d'un parcours long, très riche en enseignements et de surtout beaucoup de soutiens.

Je tiens tout d'abord à remercier mes encadrantes Myriam Khodri et Sylvie Thiria. Ayant rencontré dans un premier temps Sylvie je la remercie de m'avoir soutenu et encouragé dans ma volonté de faire une thèse. Je la remercie également pour sa disponibilité tout au long de cette thèse, ainsi que pour sa confiance, sa pédagogie et sa bienveillance à mon égard.

Je tiens à remercier tout particulièrement Myriam pour son aide durant ces nombreuses années. Venant d'une formation en mathématiques/informatique/statistique, le climat et sa variabilité étaient des notions très vagues à mon arrivée à LOCEAN. Je la remercie pour sa patience et sa bienveillance à mon égard ainsi que pour tout le temps qu'elle m'a consacré tout au long de cette thèse.

Je remercie Hugues Goosse et Christophe Cassou pour avoir accepté d'être rapporteur de ma thèse et pour leurs remarques très constructives sur mon manuscrit. Je remercie également Francis Codron, Aurélien Ribes et Sophie Cravatte d'avoir bien voulu faire partie des membres de mon jury et ce malgré des circonstances un peu particulières.

Je remercie les nombreux scientifiques avec lesquels j'ai eu l'occasion d'échanger et qui m'ont fait profiter de leurs connaissances notamment Guillaume Gastineau, Vincent Echevin, Jérôme Vialard, Nicolas Lebas, Emilie Dassié, Anastase, Julien Brajard et Michel Crepon.

Merci également à toute l'équipe de thésard et post doc Alex, Emilie, Chris, Djallel, Khalil, Aniss, Vio, Amélie, Victor, Julián, Camille, Leo et bien d'autres qui ont fait de cette expérience et de ce labo un endroit si exceptionnel et sans qui ces nombreuses années de thèses n'auraient pas été si sympa.

Un grand merci à toi, Asma, pour ton réconfort, ton sourire et ton soutien quotidien durant ces longues années. Merci pour ta patience durant cette thèse que tu as également vécue jour après jour.

J'ai également une petite pensée pour ma première fille Wided arrivée et partie trop tôt et pour sa petite sœur Firdaws qui ne cesse de grandir à toute vitesse.

Pour finir un très grand merci à mes amis, mes proches et en particulier à mes parents et à mon frère pour leurs soutiens, et sans qui je ne serais pas celui que je suis aujourd'hui.

Résumé

Dans le contexte du changement climatique émergeant actuellement en réponse au forçage d'une concentration de gaz à effet de serre en forte croissance, une difficulté est de savoir détecter toute anomalie climatique significative, puis lui « attribuer » des causes possibles liées, soit aux forçages naturels (activité solaire, éruptions volcaniques), soit à la variabilité interne du climat, soit à l'impact des forçages anthropiques. Cette question est d'autant plus complexe que l'activité anthropique peut s'exprimer aussi plus régionalement en modulant certains processus caractéristiques d'une région donnée. Ceci peut s'illustrer par la notion de « temps d'émergence », c'est à dire de savoir, suivant les régions, la variable ou système climatique considérés (température, précipitation, upwelling de bord Est, ...), quand, et avec quelle incertitude, un signal climatique dû au forçage anthropique va dépasser le « bruit » de la variabilité climatique interne et naturelle, ou bien si cela est déjà effectif. Il est possible de répondre en partie à ces questions en utilisant des observations instrumentales, des archives naturelles ou proxys résolvant l'échelle pluriannuelle et des modèles climatiques. La combinaison de ces approches permet d'étendre la fenêtre d'observation au-delà de la période instrumentale pour une meilleure caractérisation de la variabilité naturelle.

Cette thèse a pour objectif de développer une expertise sur les processus physiques par lesquels les forçages externes (volcans, CO₂ et formation du trou d'ozone stratosphérique) peuvent interférer avec la variabilité interne à l'échelle globale et régionale ainsi que le développement de techniques d'assimilation de données paléoclimatiques dans le modèle de l'IPSL permettant de mettre à profit les données PAGES-2K récemment publiées. Nous avons dans un premier temps cherché à comprendre les tendances au refroidissement observées dans la zone d'upwelling au large des côtes du Nord Chili depuis 1980 et du Pérou depuis 1950. Dans un contexte de réchauffement global, les résultats montrent que l'expansion de la cellule de Hadley dans l'hémisphère Sud en réponse aux forçages anthropiques (CO₂ et formation du trou d'ozone stratosphérique) est un résultat robuste dans nos grands ensembles de simulations du modèle IPSL-CM5A, les observations et Réanalyses depuis les années 80s et est responsable de l'intensification de l'upwelling. A l'échelle régionale celle-ci se traduit par un renforcement de l'anticyclone subtropical du Pacifique Sud et des alizées qui s'étendent depuis la côte Chilienne vers le Pacifique équatorial. L'analyse du bilan de chaleur de la couche de mélange océanique montre qu'en réponse à l'expansion des alizés l'approfondissement de la couche de mélange et le pompage d'Ekman s'oppose au signal régional du réchauffement climatique dans le Pacifique Sud-Est et l'annule près de la côte Chilienne.

Bien que cette première étude nous ait permis d'isoler et de comprendre certains phénomènes observés, il m'est apparu clair que la fenêtre d'observation instrumentale restreinte est un des freins majeurs dans la compréhension de la dynamique naturelle de notre climat. La deuxième partie de ma thèse a consisté à mettre en place une méthode

d'assimilation de données dans le modèle de l'IPSL permettant la reconstruction de la variabilité naturelle passée en contexte pré industriel. Les méthodes d'assimilations de données dites séquentielles comme les filtres de Kalman d'ensemble et de filtres particulaires sont très difficiles à mettre en œuvre dans un modèle de circulation général couplé océan-atmosphère (OAGCM) et sont très coûteuses en temps de calcul. Elles offrent néanmoins l'avantage de constituer une puissante méthode de validation et d'identification des mécanismes clés à la base des changements observés. Ce type de méthode est confronté à un problème de dégénérescence associée à la résolution d'un problème de grande dimension avec un nombre de particules restreint. Pour lever ce verrou j'ai développé un filtre particulaire dans le modèle de l'IPSL ("SIR") utilisant un émulateur statistique du modèle de l'IPSL (LIM) servant comme modèle d'intégration. Cette nouvelle méthode baptisée SIR-LIM est proposée et testée « en modèle parfait » pour reconstruire la variabilité préindustrielle des températures de surface dans un contexte sous observé analogue à celui du dernier millénaire. La validation de cette méthode à travers plusieurs ensembles d'expériences permet aujourd'hui d'envisager la reconstruction de la variabilité climatique des siècles passés par l'assimilation d'observations et de proxys dans un OAGCM tout en préservant la cohérence physique le long de la simulation.

Table des matières

Remerciements	iii
Résumé	v
Table des matières	ix
Abréviations	xi
I Introduction	1
1 Introduction Générale	3
1.1 Motivations : Un changement climatique sans précédent ?	4
1.2 Variabilité naturelle ou signature du forçage anthropique ?	5
1.3 Objectifs et plan de la thèse	6
2 Le système climatique et sa variabilité	9
2.1 Introduction	10
2.2 Présentation du système climatique	10
2.2.1 Circulation atmosphérique moyenne	11
Circulation méridienne	12
Circulation zonale	14
2.2.2 Circulation océanique	15
La circulation océanique grande échelle	15
Les upwellings de Bord-Est	16
2.3 La variabilité climatique	19
2.3.1 Variabilité dans l’Océan Pacifique	20
2.3.2 Variabilité décennale dans l’Océan Atlantique	24
2.4 Les forçages externes	25
2.4.1 Les forçages naturels	26
2.4.2 Les forçages anthropiques	27
2.5 Les modèles climatiques	28
2.5.1 Présentation	28
2.5.2 Les modèles de circulation générale	28
2.5.3 Description et évaluation succincte du modèle IPSL-CM5A-LR	30
2.5.4 Des modèles de circulation générale aux modèles statistiques	32
2.5.5 Utilisation	33

2.5.6	Exercice d'inter-comparaison des modèles CMIP	35
2.6	Convergence entre modèles et observations	39
II	Étude de la variabilité climatique sur la période historique	43
3	Tendance et variabilité climatique dans l'océan Pacifique au 20^{ème} siècle	45
3.1	Préambule	46
3.2	Évaluation de la variabilité interne aux échelles de temps décennales	46
3.3	Rôle des forçages externes sur la variabilité et tendances observées et simulées	50
3.3.1	Impacts des forçages externes sur l'ENSO et l'IPO	50
3.3.2	Impacts des forçages externes sur les tendances à long terme	52
3.3.3	Tendances robustes sur la période historique : variabilité interne ou signature des forçages anthropiques ?	55
3.4	Article : Contributions of internal variability and external forcing to the recent trends in Southeastern Pacific and Peru-Chile upwelling system	58
3.5	Expressions océaniques de l'expansion de la cellule de Hadley	82
III	Reconstruction de la variabilité climatique des 2000 dernières années	85
4	Modélisation et méthodes d'assimilation de données paléoclimatiques	87
	Notations	89
4.1	Préambule	90
4.2	PAGES2k Consortium	90
4.3	L'assimilation de données en contexte paléoclimatique	91
4.3.1	Approches et méthodes d'assimilations de données	91
4.3.2	Approches séquentielles et variationnelles	93
4.3.3	La règle de Bayes	94
4.3.4	Filtrage Bayésien	94
4.4	Les filtres particuliers	95
4.4.1	Le filtre particuliers séquentiel pondéré	95
4.4.2	Le filtre particuliers avec ré-échantillonnage	97
4.4.3	Méthode de ré-échantillonnage	98
4.4.4	Taille effective de l'échantillon N_{eff}	99
4.4.5	En pratique	100
4.5	Le LIM	100
4.5.1	Présentation	100
4.5.2	Définition du modèle LIM	101
4.5.3	Le LIM couplé	102
4.5.4	Le LIM en tant que modèle d'intégration	103
4.6	Indices de performances	104

5	Reconstruction de la variabilité des températures de surface en contexte paléoclimatique avec le SIR-LIM	107
	Notations	109
5.1	Préambule	110
5.2	Protocole expérimental	110
5.3	Le SIR-LIM	111
5.4	Expériences et validations	114
5.5	Calcul des EOFs	116
5.6	Construction et Validation du LIM	119
5.7	Le LIM en tant qu'émulateur des modèles climatiques	123
5.8	Reconstructions spatiales de la variabilité des températures de surface . .	129
5.9	Reconstructions des principaux modes de variabilité pour le modèle IPSL-CM5A-LR	133
5.10	Sensibilité aux paramètres d'assimilation	138
	5.10.1 Nombre de particules et nombre d'EOFs	138
	5.10.2 Estimation de la matrice d'erreur C	140
6	Conclusions et Perspectives	143
6.1	Synthèse des résultats et Discussion	143
6.2	Perspectives	147
	6.2.1 Améliorations du SIR-LIM	147
	6.2.2 Reconstructions paléoclimatiques avec des GCMs	149
	Bibliography	153
	Annexe A: Tableaux et Figures supplémentaires	179
	A.1 Chapitre 3	179
	A.2 Chapitre 6	181
	Annexe B: SomVar	187

Abréviations

AMV	A tlantic M ulti-decadal V ariability (Variabilité multi-décennale de l'Atlantique)
AMOC	A tlantic M eridional O verturing C irculation (Circulation méridienne de retournement Atlantique)
CMIP	C oupled M odel I ntercomparaison P rotocol (Projet d'intercomparaison de modèles couplés)
ENSO	E l Niño S outherne O scillation (El Niño - Oscillation australe)
EOF	E mpirical O rthogonal F unction (Analyse en composantes principales)
GHG	G reenhouse G as (Gaz à effet de serre)
GCM	G eneral C irculation M odel (Modèle de circulation générale)
IPCC	I ntergovernmental P anel on C limate C hange (Groupe d'experts intergouvernemental sur l'évolution du Climat)
IPO	I nter-decadal P acific O scillation (Oscillation inter-décennale du Pacifique)
ITCZ	I nter-Tropical C onvergence Z one (Zone de convergence Inter-Tropicale)
PC	P rincipal C omponent (Composante principale)
SPCZ	S outh P acific C onvergence Z one (Zone de convergence du Pacifique Sud)
SST	S ea S urface T emperature (Température de surface de la mer)
SSH	S ea S urface H eight (Hauteur de la surface de la mer)

Partie I

Introduction

Chapitre 1

Introduction Générale

Sommaire

1.1 Motivations : Un changement climatique sans précédent ?	4
1.2 Variabilité naturelle ou signature du forçage anthropique ?	5
1.3 Objectifs et plan de la thèse	6

1.1 Motivations : Un changement climatique sans précédent ?

De nos jours, nous savons que la Terre a été soumise au fil des millénaires à des alternances de périodes chaudes et froides. L'impact de ces variations touchant tous les aspects de notre société, les conséquences économiques, sociétales ou biologiques de ces variations ne sont aujourd'hui plus à démontrer (McMichael, 2000; DeMenocal, 2001; Haug et al., 2003; Buckley et al., 2010; Büntgen et al., 2011). Bien que les aléas du climat du type vagues de froid, sécheresses, inondations, etc. aient longtemps été considérés aux yeux des hommes comme imprévisibles, il devient évident par l'observation de la succession d'événements extrêmes ces dernières années (records de températures, cumuls de précipitations induisant des crues rapides, sécheresses exceptionnelles dans certaines régions) que le climat de la Terre se modifie à l'échelle planétaire.

Dans son 5^{ème} rapport publié en 2013 (IPCC, 2013), le Groupe d'experts Intergouvernemental sur l'Évolution du Climat (GIEC) conclut que le changement climatique global observé depuis le début de l'air industriel se fait à un rythme rapide et avec une amplitude sans précédent. En 2019, le taux atmosphérique en dioxyde de carbone (CO₂) a atteint plus de 400 parties par million¹ (ppm) (Figure. 1.1). En 1850, avant la révolution industrielle, ce taux s'élevait à seulement 284 ppm, et la température moyenne du globe était 1°C plus froide qu'aujourd'hui. Il s'agit d'une première dans l'histoire de l'humanité, la barre symbolique des 400 ppm ayant été dépassée pour la dernière fois il y a trois millions d'années, durant le Pliocène, la sixième époque du tertiaire qui s'étend de 5 à 2 millions d'années. Les températures étaient alors 3 à 4°C plus élevées, des arbres poussaient en Antarctique et le niveau de la mer était plus haut de 15 mètres (Willeit et al., 2019). Les activités anthropiques sont donc aujourd'hui un facteur de forçage du système climatique à part entière.

Dans ce contexte du changement climatique émergeant actuellement en réponse au forçage d'une concentration de gaz à effet de serre en forte croissance, une difficulté est de savoir détecter toute anomalie climatique significative, puis lui « attribuer » des causes possibles liées, soit aux forçages naturels (activité solaire, éruptions volcaniques), soit à la variabilité interne du climat, soit à l'impact des forçages anthropiques. Il est probable que la zone intertropicale sera avec les zones polaires, le lieu où cette émergence se fera le plus rapidement (Christensen et al., 2013).

Cette question est d'autant plus complexe que l'activité anthropique peut s'exprimer aussi plus régionalement en modulant certains processus caractéristiques d'une région donnée. Ceci peut s'illustrer par la notion de « temps d'émergence », c'est à dire de savoir, suivant les régions et la variable climatique considérées (température, précipitation, ...), dans combien d'années, et avec quelle incertitude, un signal climatique dû au forçage anthropique va dépasser le « bruit » de la variabilité climatique interne et naturelle, ou bien si cela est déjà effectif. Pour cela il est nécessaire d'être capable d'observer (pour détecter), puis de comprendre (pour attribuer), pour enfin prévoir l'évolution de l'empreinte du

¹Nombre de molécules du gaz à effet de serre considéré par million de molécules d'air.

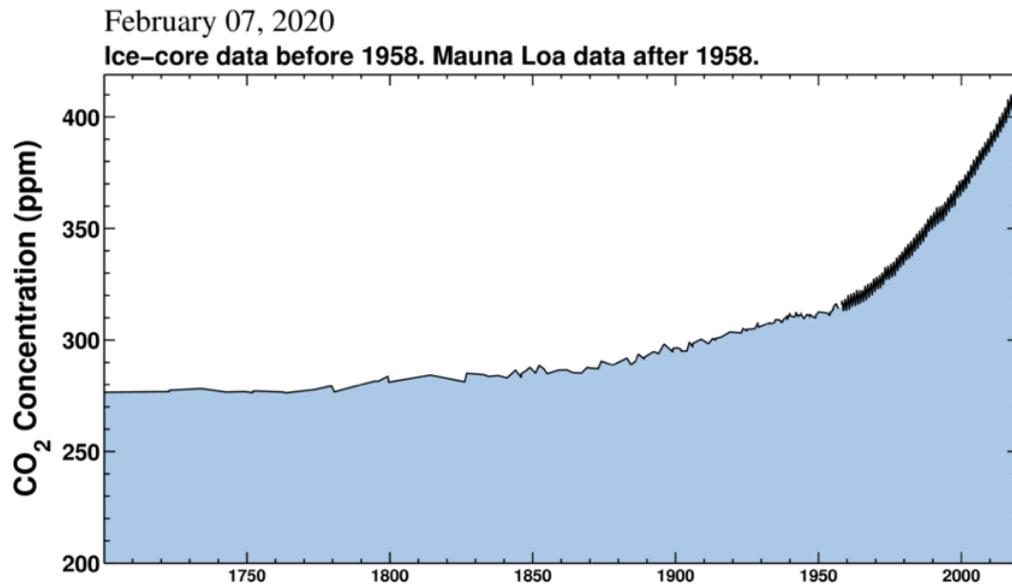


Figure 1.1: Changements de la concentration en CO₂ atmosphérique déduite de l'analyse des carottes de glace de Low Dome en Antarctique de 1700 à 1958 (Macfarling Meure et al., 2006; Joos and Spahni, 2008) et mesurée à l'Observatoire du Mauna Loa (Keeling and Keeling, 2017), Hawaï de 1958 au 7 Février 2020.

changement climatique sur l'environnement et les sociétés.

1.2 Variabilité naturelle ou signature du forçage anthropique ?

Pour pouvoir assurer le suivi de l'évolution de notre climat en transition, et plus généralement notre environnement, et pouvoir interpréter les changements observés et en attribuer la ou les causes probables, il est tout d'abord nécessaire de pouvoir s'appuyer sur des réseaux de données d'observations et des observatoires pérennes et pluridisciplinaires. La combinaison d'observations instrumentales aux variables climatiques déduites d'archives naturelles et résolvant l'échelle annuelle à pluriannuelle, telles que les coraux en domaine marin ou les spéléothèmes en domaine continental collectées à l'échelle du globe, permet d'étendre la fenêtre d'observation au-delà de la période instrumentale et d'obtenir une meilleure caractérisation de la variabilité naturelle. Ces informations témoignent d'une variabilité importante du climat, aux échelles de temps interannuelle, décennale et séculaire, et permettent de caractériser le temps de retour d'événements météorologiques rares (fortes tempêtes, vagues de chaleur ou de froid). Cette période permet donc de caractériser la variabilité « naturelle » du climat, c'est-à-dire indépendante de l'action de l'homme sur le bilan radiatif de la Terre et pose la question des conséquences de cette variabilité du climat sur les sociétés humaines et leur adaptation. Savoir comment et pourquoi le climat a évolué au cours des dernières décennies à siècles est essentiel pour contextualiser les tendances récentes observées et détecter l'émergence de l'influence anthropique.

La phase de compréhension du fonctionnement du système climatique et environnemental Terre peut passer, d'une part par l'utilisation d'outils statistiques appliqués à l'analyse de ces données d'observations (instrumentales et proxy), d'autre part par des approches diagnostiques mettant en jeu certains concepts ou théories, et plus largement par des outils de modélisation représentant la complexité des processus et mécanismes physiques en jeu dans le système Terre, tout en appréhendant les différentes échelles spatiales et temporelles. Le meilleur moyen de prévoir le climat étant de comprendre sa dynamique, les climatologues s'affairent donc à comprendre et modéliser les lois qui gouvernent notre climat. Les modèles numériques les plus complexes sont basés sur les équations de la dynamique des fluides et sont utilisés pour prédire et comprendre l'impact des changements climatiques récents sur notre climat. Ces modèles de climat sont très élaborés, mais comportent encore de forts biais et incertitudes (Dufresne et al., 2013; Wang et al., 2014). Des travaux de validation à l'aide d'observations adaptées sont donc indispensables pour évaluer par la suite le niveau de confiance et d'incertitude des projections climatiques futures fournis par ces modèles.

La compréhension et l'anticipation de la variabilité dite forcée, et la meilleure modélisation et compréhension de la variabilité dite « interne » font partie des enjeux majeurs des sciences du climat. Pour optimiser cette confrontation et faciliter l'exploitation scientifique des observations et proxy climatiques, des méthodes d'assimilation de données existent et sont mises en œuvre à l'aide de modèle de climat. Les méthodes d'assimilation de données utilisant des modèles de circulation générale couplé océan-atmosphère sont très difficiles à mettre en œuvre surtout dans la cadre d'approches ensemblistes très coûteuses en temps de calcul. Elles offrent en revanche l'avantage de constituer une puissante méthode de validation des modèles et permet l'identification des mécanismes clefs à la base des changements climatiques des derniers siècles à décennies.

1.3 Objectifs et plan de la thèse

Cette thèse a pour objectif de développer une expertise sur les processus physiques par lesquels les forçages externes naturels (volcans, activité solaire) et anthropiques (CO₂ et formation du trou d'ozone stratosphérique) peuvent interférer avec la variabilité interne à l'échelle globale et régionale ainsi que le développement de techniques d'assimilation de données paléoclimatiques permettant de mettre à profit les données PAGES-2K récemment publiées (PAGES2k Consortium, 2017). J'ai pour cela développé durant ma thèse différents types d'approches combinant observations, modèles climatiques et modèles statistiques afin d'aborder les questions suivantes:

1. Est-il possible de distinguer, l'action des forçages naturels de celle des forçages anthropiques, liés par exemple à l'évolution de la concentration des gaz à effet de serre dans l'atmosphères ?
2. Dans quelles mesures les modèles climatiques ou système Terre sont-ils des outils fiables pour reproduire et comprendre les processus contrôlant les principaux modes

de variabilité interne et la réponse aux forçages externes ? Quelles en sont les incertitudes ? Peut-on les réduire ?

3. De quelle manière peut-on combiner les informations provenant des modèles et des observations (ou proxys) pour reconstruire le climat passé ? Quelles en sont les incertitudes ? Peut-on les réduire ?

Il est possible de répondre en partie à ces questions en utilisant des observations directes du climat récent (depuis 1850 environ) et de l'utilisation de modèles climatiques. Néanmoins la reconstruction de variables climatiques sur de plus longues échelles de temps est nécessaire pour mettre en perspective les caractéristiques climatiques de la période récente. Aujourd'hui, grâce à l'augmentation récente des capacités de calcul numérique il est possible d'effectuer des grands ensembles de simulations avec des modèles globaux sur plusieurs centaines d'années. Ces modèles peuvent être validés notamment en confrontant les résultats des simulations climatiques avec des observations ou reconstructions. Ils fournissent notamment des estimations de la réponse du système climatique à l'évolution des différents forçages connus. Ces modèles sont également régulièrement utilisés pour identifier les processus et la signature spatiale de la réponse climatique à certains de ces forçages.

Présenté à travers les différents chapitres, le cheminement scientifique suivi durant ma thèse est le suivant. Après avoir réalisé une première étude sur le rôle de la variabilité forcée et la contribution de la variabilité interne sur les changements climatiques observés au 20^{ème} siècle dans l'océan Pacifique, il m'est apparu clair que la fenêtre d'observation instrumentale restreinte constitue l'un des freins majeurs dans la compréhension de la dynamique naturelle de notre climat aux échelles de temps décennales à séculaires. Bien que cette première étude nous ait permis d'isoler et de comprendre certains phénomènes observés, il m'est apparu primordial d'aller plus loin via la mise en place d'approches permettant la reconstruction de la variabilité passée en contexte pré industriel et contraint par les observations. Ce cheminement scientifique explique la structure de ce manuscrit qui s'organise en un premier chapitre d'introduction puis en 3 grandes parties:

- La première partie constituée des **chapitres 1 et 2**, introduit le contexte de notre travail, les enjeux, les outils disponibles ainsi que les concepts de base nécessaires à la compréhension des chapitres suivants. Dans ces chapitres, j'introduis les éléments généraux de la thèse. J'aborde la notion de système climatique en détaillant ses composantes et les principaux cycles de mouvements et d'énergie actifs en son sein. Dans un contexte de réchauffement climatique, je définis les principaux modes de variabilité dans les océans Atlantique, Pacifique, les systèmes d'upwelling du Bord-Est et les limites dans la représentation de ce type de système par les modèles de climat. Enfin, je présente l'état de l'art de la modélisation climatique, notamment en décrivant et en évaluant brièvement le modèle IPSL-CM5A-LR (Dufresne et al., 2013) utilisé dans cette thèse.

- La deuxième partie se compose du **chapitres 3**, et porte sur la détection et l'attribution des changements climatiques récents et tendances observés dans l'océan Pacifique. Dans un premier temps, nous décrivons les données et méthodes statistiques utilisées. Nous effectuons une étude préliminaire qualitative des réponses forcées observées en fonction de celles simulées pour la période historique depuis le début de l'ère industrielle par les modèles climatiques de classe CMIP5 (Taylor et al., 2012) et faisons un état de l'art des mécanismes proposés. Nous discutons des incertitudes présentes dans les observations et de la pertinence de l'utilisation des modèles climatiques pour estimer et comprendre le rôle des différentes sources de variabilité. A travers une étude s'inscrivant dans la continuité du chapitre et faisant l'objet d'une publication dans la revue scientifique *Journal of Climate*, nous évaluerons à l'aide du modèle IPSL-CM5A la contribution de la variabilité interne et des différents forçages externes sur les tendances observées dans le Pacifique Sud-Est et de l'upwelling situé le long des côtes du Pérou-Chili et enquêterons sur les processus à l'origine de ces tendances.
- La troisième partie se compose quant à elle des **chapitres 4 et 5**, et se focalise sur la reconstruction par assimilation de données de la variabilité climatique en contexte préindustriel et sous observé analogue à celui du climat du dernier millénaire. Une méthode de reconstruction baptisée *SIR-LIM* est proposée et testée pour reconstruire la variabilité naturelle des températures de surface. Dans le **chapitre 4** nous effectuons un bref état de l'art des méthodes d'assimilation de données dans les modèles de climat. Nous donnons les concepts théoriques des méthodes utilisées par la suite et décrivons les jeux de données utilisés. Dans le **chapitre 5**, nous décrivons la méthode d'assimilation proposée et développée « en modèle parfait »² ainsi qu'un ensemble d'expériences permettant de la valider.
- Enfin le **chapitre 6** reprend les conclusions et les perspectives liées à mon travail de thèse.

²Expérience qui consiste à utiliser les résultats d'un modèle comme représentatif du processus réel

Chapitre 2

Le système climatique et sa variabilité

Sommaire

2.1	Introduction	10
2.2	Présentation du système climatique	10
2.2.1	Circulation atmosphérique moyenne	11
	Circulation méridienne	12
	Circulation zonale	14
2.2.2	Circulation océanique	15
	La circulation océanique grande échelle	15
	Les upwellings de Bord-Est	16
2.3	La variabilité climatique	19
2.3.1	Variabilité dans l’Océan Pacifique	20
2.3.2	Variabilité décennale dans l’Océan Atlantique	24
2.4	Les forçages externes	25
2.4.1	Les forçages naturels	26
2.4.2	Les forçages anthropiques	27
2.5	Les modèles climatiques	28
2.5.1	Présentation	28
2.5.2	Les modèles de circulation générale	28
2.5.3	Description et évaluation succincte du modèle IPSL-CM5A-LR	30
2.5.4	Des modèles de circulation générale aux modèles statistiques	32
2.5.5	Utilisation	33
2.5.6	Exercice d’inter-comparaison des modèles CMIP	35
2.6	Convergence entre modèles et observations	39

2.1 Introduction

Ce chapitre a pour objectif de décrire le système climatique et d'introduire les concepts et notions clés nécessaires à la compréhension des chapitres suivants. L'étude de la variabilité climatique à l'échelle décennale fait l'objet de cette thèse et nécessite donc une description des facteurs à son origine ainsi qu'une présentation des modèles et observations utilisés pour l'étudier.

Pour permettre une vision d'ensemble je décrirai dans un premier temps la notion de système climatique et présenterai les différentes enveloppes fluides qui le composent. Je décrirai brièvement les interactions en son sein et la manière dont la circulation atmosphérique s'organise dans sa composante latitudinale et longitudinale pour redistribuer l'énergie du Soleil à l'échelle globale. Je présenterai également la manière dont la circulation océanique s'organise en surface et en profondeur. Dans le cadre d'une étude spécifique sur l'upwelling du Pérou Chili dans le chapitre 3, je donnerai ici également une description plus détaillée de la physique des systèmes d'upwellings de bord-Est.

Je décrirai dans un second temps les différents éléments internes ou externes pouvant agir sur le climat. Je détaillerai ces notions en présentant les principaux modes variabilités opérants aux échelles de temps décennales et décrirai les principaux forçages externes selon leur origine naturelle ou anthropique.

Enfin je présenterai les outils et sources de connaissances dont on dispose pour étudier le climat passé et prévoir les changements futurs. Je présenterai pour cela la manière dont le système climatique est modélisé dans les modèles les plus complexes dits de circulation générale comme celui de l'Institut Pierre Simon Laplace (IPSL) que je placerai dans un contexte plus large, par rapport aux évolutions des modèles de climat au cours des dernières décennies. Je présenterai également une seconde variété de modèles basés sur des approches statistiques et discuterai leur utilisation dans les sciences du climat. Pour finir je présenterai les principales informations dont on dispose pour décrire les variations du climat au dernier millénaire et les méthodes d'assimilation de données qui permettent de le reconstruire en combinant les observations, archives naturelles et modèles climatiques.

2.2 Présentation du système climatique

Le système climatique tel qu'on se le représente est un système ouvert qui en permanence reçoit par rayonnement de l'énergie du soleil et en transmet vers l'espace. Il se compose de l'ensemble atmosphère, hydrosphère (eau sous forme liquide), cryosphère (eau sous forme solide), lithosphère (enveloppe rigide de la surface de la Terre), et biosphère (ensemble des organismes vivants) et évolue au fil du temps sous l'effet de processus internes et de contraintes externes appelées forçages externes. Les interactions entre ces composantes pouvant se faire via des échanges de quantité de mouvement, d'énergie et/ou de matière, toutes variations d'une de ces composantes peut induire un changement sur le climat à

l'échelle globale (Fig 2.1).

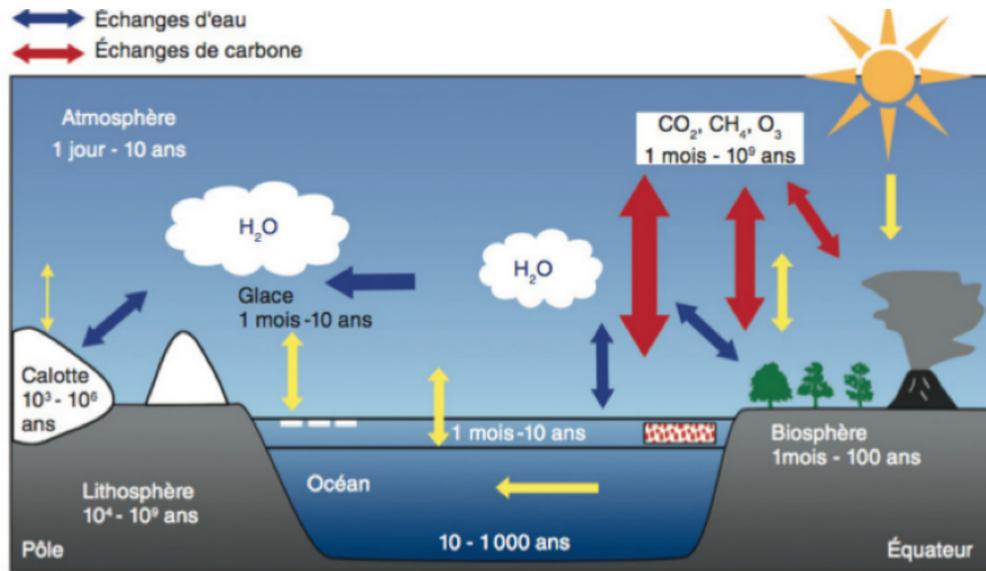


Figure 2.1: Schéma du système climatique présentant ses différentes composantes: atmosphère, océan, cryosphère, biosphère et lithosphère, leurs constantes de temps et leurs interactions en termes d'échanges d'énergie, d'eau et de carbone (source :Jeandel and Mosseri, 2011).

La principale source d'énergie du système climatique étant le rayonnement solaire, celui-ci est absorbé, transformé et en grande partie réémis sous forme de rayonnement infrarouge vers l'espace. Les régions tropicales reçoivent en moyenne annuelle d'avantage d'énergie solaire qu'elles n'en réémettent par rayonnement infrarouge (Fig. 2.2). Cette différence entre le rayonnement solaire absorbé et le rayonnement infrarouge réémis engendre un excédent d'énergie au niveau de l'équateur que l'on trouve également en été aux moyennes latitudes. Dans les régions polaires et en hiver aux moyennes latitudes l'inverse se produit avec un rayonnement infrarouge émis supérieur au rayonnement solaire absorbé causant ainsi un déficit d'énergie radiative. L'océan joue alors le rôle de réservoir thermique aux moyennes latitudes redistribuant l'excédent d'énergie reçu en été durant les périodes déficitaires. Le déséquilibre causé par la répartition inégale du rayonnement solaire sur le globe induit également un transport d'énergie méridien des zones excédentaires vers les zones déficitaires. Ce transport d'énergie s'effectue principalement via les deux composantes fluides du système climatique qui sont: l'atmosphère et l'océan, le transport réalisé via ces deux fluides étant du même ordre de grandeur (Fig. 2.2).

2.2.1 Circulation atmosphérique moyenne

L'inégale répartition du rayonnement solaire sur le globe est le principal moteur de la circulation atmosphérique à l'échelle planétaire. Cette circulation est influencée par la rotation de la Terre via la force de Coriolis qui s'annule près de l'équateur et augmente avec la latitude. La circulation atmosphérique dépend également de la distribution des océans et de la topographie des continents.

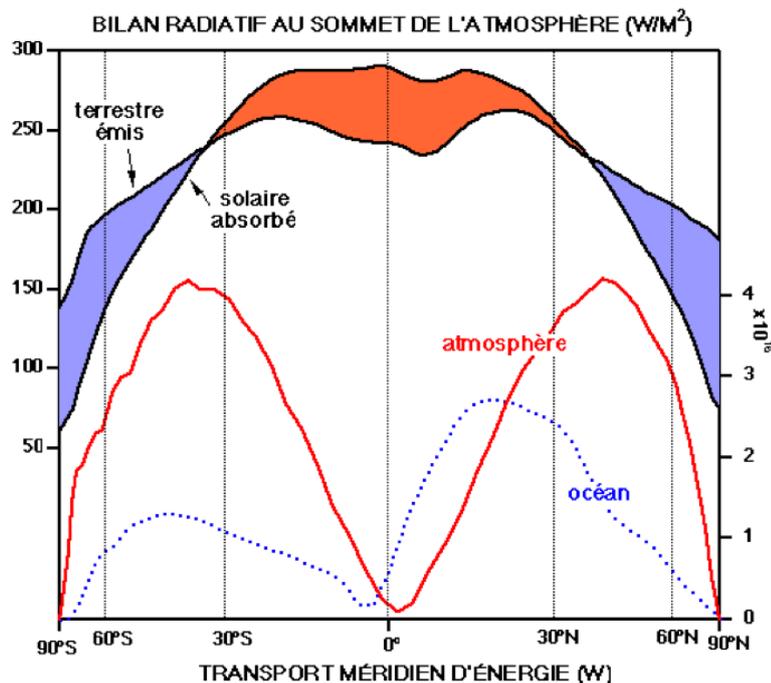


Figure 2.2: Estimation du bilan radiatif annuel au sommet de l'atmosphère, en moyenne zonale (échelle à gauche: énergie solaire absorbée, énergie terrestre émise), ainsi que du transport méridien de chaleur par les circulations atmosphériques et océaniques, également en moyenne zonale annuelle (échelle de droite). Le bilan radiatif est positif entre 40°S et 40°N, et négatif aux plus hautes latitudes.

Circulation méridienne

Au niveau des tropiques le transport d'énergie atmosphérique de l'équateur vers les pôles s'effectue principalement via sa circulation méridienne. Cette circulation ne pouvant cependant pas s'étendre de l'équateur jusqu'au pôle à cause de contraintes physiques liées à la rotation de la Terre et à la conservation du moment cinétique, celle-ci s'organise sous forme de cellules appelées cellule de Hadley aux basses latitudes, cellules de Ferrel aux moyennes latitudes et cellules polaires aux hautes latitudes (Fig. 2.3). La cellule de Hadley se caractérise par des mouvements de masses d'airs se dirigeant en surface depuis les pôles vers les basses latitudes et en altitude depuis l'équateur vers les pôles. Elle génère donc deux cellules de part et d'autre de l'équateur permettant de redistribuer l'énergie accumulée dans les tropiques vers les plus hautes latitudes des deux hémisphères.

L'ascension de masses d'airs chaudes et humides légèrement au nord de l'équateur entraîne d'importantes précipitations localisant l'équateur météorologique appelé zone de convergence inter tropicale ou ITCZ (pour en anglais: Intertropical Convergence Zone) (Fig. 2.4). Pouvant s'élever jusqu'à la tropopause, une fois en altitude ces masses d'airs chaudes vont se diriger vers les pôles et redescendre aux moyennes latitudes. Ces zones de subsidences sont associées à des centres de hautes pressions subtropicaux et des régions arides subtropicales. A la surface dans la bande intertropicale les branches ascendantes et

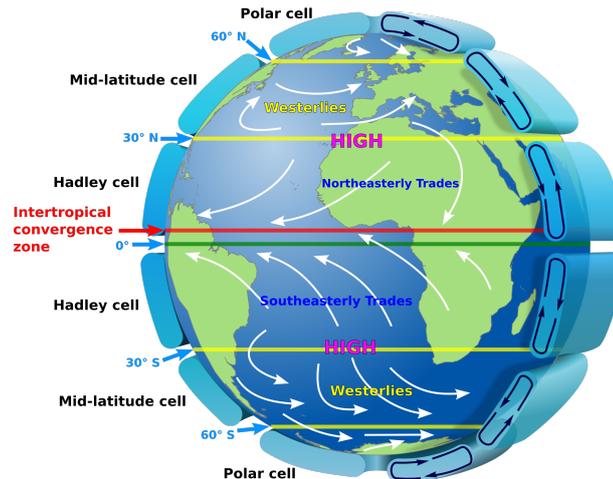


Figure 2.3: Représentation schématique de la circulation atmosphérique générale avec les différentes cellules de circulation et les vents de surface. (Source: NASA)

descendantes des cellules de Hadley sont reliées par des vents d'Est dirigés vers l'ITCZ appelés Alizés qui plus réguliers que les vents d'ouest des zones tempérées possèdent en revanche une vitesse moyenne plus faible. A cette circulation méridienne, est superposée une circulation zonale sous forme de cellule est-ouest liée à la distribution des continents et des océans.

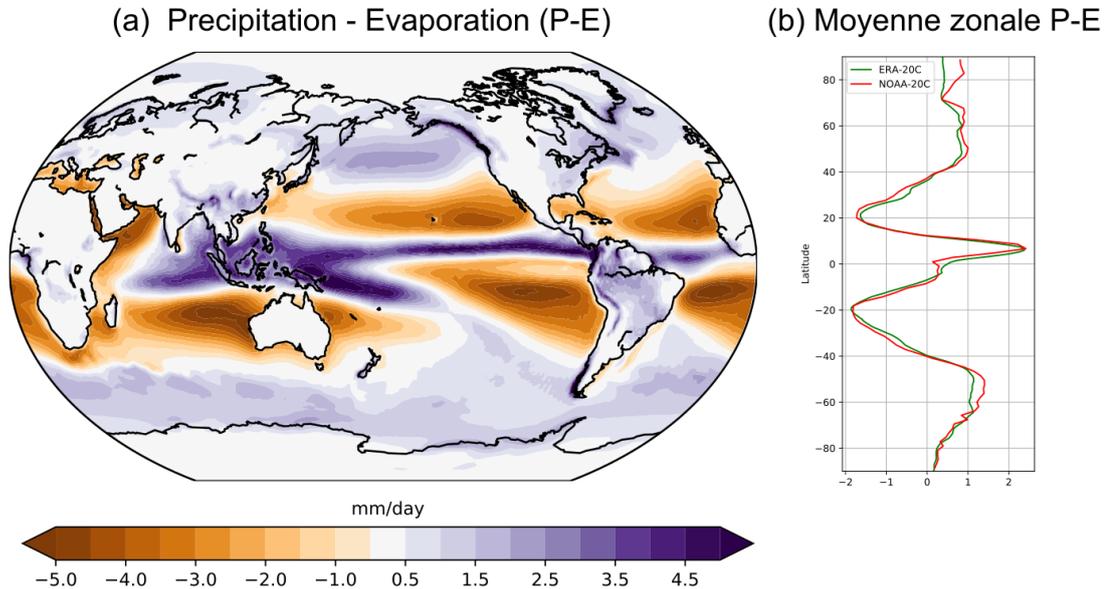


Figure 2.4: Climatologie pour la période 1900-1930 des (a) Précipitation moins Evaporation (P-E) pour la Réanalyse ERA-20C et (b) moyennes zonales de (P-E) pour ERA-20C (en vert) et NOAA-20C (en rouge). Les zones d'ascendances et de subsidences de la cellule de Hadley sont localisées par les régions où $P - E > 0$ et $P - E < 0$ respectivement.

Circulation zonale

La composante longitudinale de la circulation atmosphérique à l'équateur s'appelle la circulation de Walker. L'existence de cette circulation longitudinale s'explique par les fortes variations des températures tropicales dues aux vents de surface, à l'absence de force de Coriolis dans les tropiques et à la distribution spatiale des océans et continents. Cette circulation est illustrée dans la (Fig. 2.5), avec en condition moyenne des zones d'ascendances localisées au niveau du continent maritime, de l'Afrique et de l'Amazonie et des zones de subsidences localisées à l'ouest de l'océan Indien et à l'Est de l'océan Pacifique.

La branche la plus intense et persistante de cette circulation est caractérisée par les Alizés soufflant d'Est en Ouest et convergeant vers l'ITCZ. Dans les océans Pacifique et Atlantique, sous l'action des Alizés les eaux chaudes de surface sont poussées et confinées vers l'ouest du bassin tandis que ces eaux sont remplacées à l'est du bassin par des remontées d'eaux profondes localement plus froides. Dans les océans Atlantique et Pacifique, ces « langues » d'eaux froides de surface ou « cold tongue » s'étendent généralement vers l'ouest le long de l'équateur à partir des marges continentales. Elles sont constituées d'eau remontant des profondeurs de l'océan et sont entourées d'eau de surface plus chaude dans les deux hémisphères (Dijkstra and Neelin, 1995).

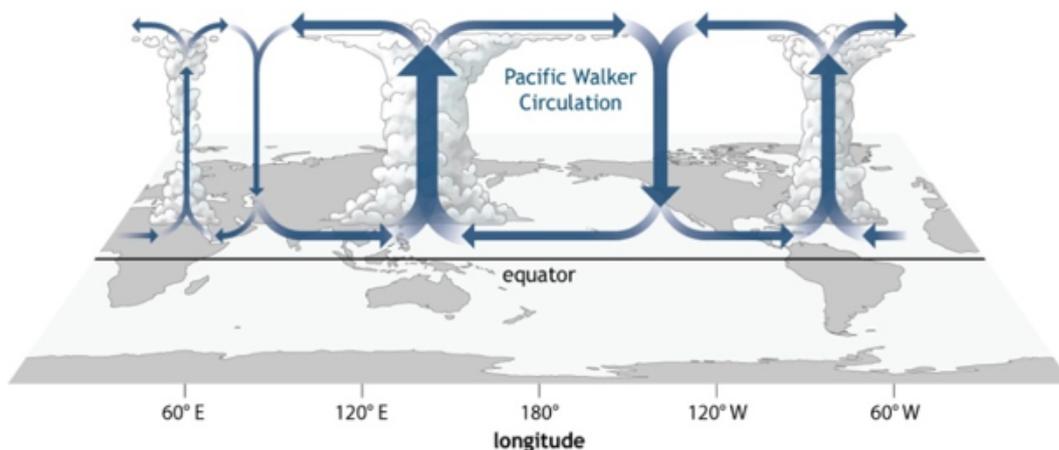


Figure 2.5: Illustration schématique de la circulation de Walker pour des conditions moyennes. Les zones d'ascendances sont associées à la formation de couches nuageuses et des précipitations plus importantes (Source : NOAA)

Dans le bassin Pacifique, le gradient de température Est-Ouest génère un gradient de pression avec un système de haute pression à l'est du bassin et de basse pression à l'ouest du bassin. L'élévation des masses d'air chaudes et humides au niveau du continent maritime engendre la formation de nuages et d'importantes précipitations au Sud-Est de l'Asie (Fig. 2.6).

On observe également la zone de convergence du Pacifique Sud (SPCZ, en anglais South Pacific Convergence Zone), une bande de convergence, de nébulosité et de précipitations de basses couches s'étendant depuis les eaux chaudes de la « Warm Pool » sur le continent maritime vers le sud-est, jusqu'à la Polynésie française et les Îles Cook

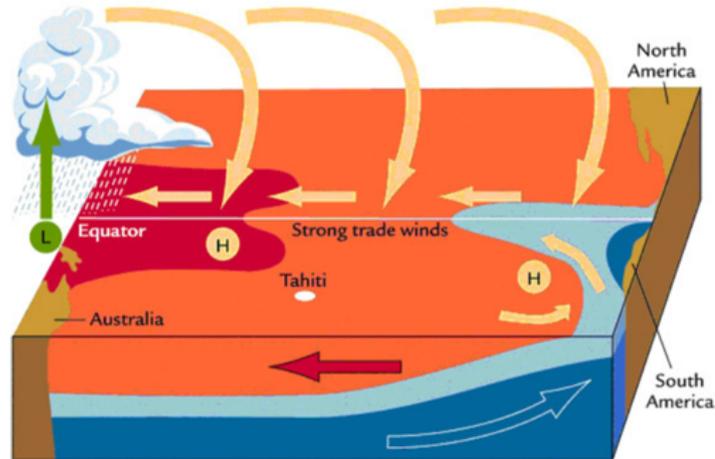


Figure 2.6: Conditions moyennes dans le Pacifique avec en couleur les températures allant du bleu pour froid au rouge pour chaud. Les flèches représentent le sens de la circulation océanique et atmosphérique et les sigles H et L les zones de hautes et basses pressions. (source : Ruddiman et al., 2001)

(160° W, 20°S). La SPCZ est la composante extra tropicale de l'ITCZ et peut affecter les précipitations sur les îles polynésiennes du sud-ouest de l'océan Pacifique.

2.2.2 Circulation océanique

La circulation océanique grande échelle

L'océan est une composante essentielle du système climatique. Il occupe les trois quarts de la surface terrestre et influence le climat global à travers le stockage et le transport de grande quantité de chaleur, de sel et de CO₂. La capacité thermique de l'océan étant bien supérieure à celle de l'atmosphère, elle agit comme un tampon capable de ralentir le réchauffement global en emmagasinant jusqu'à 93% de l'excédent thermique de la planète (Church and White, 2011; Levitus et al., 2012; Rhein et al., 2013). A travers sa circulation, la chaleur stockée dans certaines régions peut parcourir plusieurs milliers de kilomètres avant d'être relâchée dans l'atmosphère sous forme de chaleur latente (évaporation) ou sous forme de chaleur sensible¹ (Trenberth and Caron, 2001; Ganachaud and Wunsch, 2003).

La relation qui existe entre la salinité, la température et la densité d'une masse d'eau explique la circulation thermohaline. Alors que les eaux froides et salées tendent à s'enfoncer sous les eaux de surface, les eaux plus chaudes et moins salées et donc moins denses restent en surface. Les courants océaniques des couches superficielles sont pour leur part essentiellement forcés par les vents de surface et soumis à la déviation de leur trajectoire par la force de Coriolis (Fig. 2.6). On distingue les gyres subtropicaux anticycloniques et subpolaires cycloniques dans les bassins de l'Atlantique et du Pacifique. La branche ouest des gyres subtropicaux est constituée de courants chauds qui s'écoulent vers le Nord puis vers le Nord-Est, c'est le cas du Gulf Stream le long des côtes américaines

¹conduction à l'interface

dans l'Atlantique Nord et du Kuroshio le long des côtes du Japon. Ces courants chauds participent à la formation de perturbations synoptiques dans l'atmosphère, notamment en hiver, où les masses d'airs continentales froides et sèches qui viennent de l'Ouest par exemple entrent brutalement au contact de l'océan et se chargent d'humidité en même temps qu'elles vont se réchauffer.

Bien que l'influence des vents de surface sur la circulation océanique se limite essentiellement aux eaux au dessus de la thermocline² supérieure, lorsque les conditions s'y prêtent les courants de surface entraînés par les vents peuvent néanmoins induire des courants verticaux comme dans le cas des systèmes d'upwellings. Ce type de phénomène se produit principalement au niveau de l'équateur et le long des bords Est des bassins océaniques aux moyennes latitudes. Je donne dans la section suivante une description plus détaillée des processus à l'origine des phénomènes d'upwellings de bord Est.

Les upwellings de Bord-Est

Le vent qui souffle vers l'équateur le long des bords Est des bassins océaniques pousse les eaux côtières de surface vers le large. La dérive de ces eaux de surface vers le large est alors compensée par des remontées d'eaux profondes plus froides et généralement riches en nutriments à la côte. Le stress du vent sur la surface de l'océan et la rotation de la Terre étant à l'origine de ce phénomène, il s'agit en réalité d'un cas particulier de la théorie d'Ekman. Nous reprenons ci-dessous les équations décrivant le transport et pompage d'Ekman afin de comprendre la dynamique à l'origine de ces upwellings.

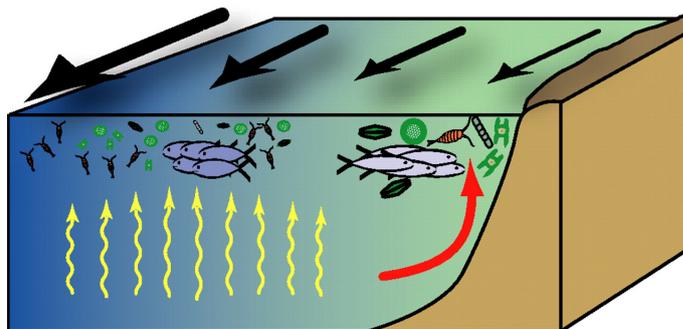


Figure 2.7: Représentation schématique du mécanisme d'upwelling côtier. Les vents parallèles à la côte (flèches noires) déplacent les masses d'eau superficielles vers le large qui induisent des remontées d'eau profonde au niveau de la côte (flèches rouges). L'intensification des vents au large induit également des remontées d'eau via le mécanisme de pompage d'Ekman (flèches jaunes) (Source: Rykaczewski and Checkley, 2008)

Transport d'Ekman:

Suite aux observations faites par le physicien Nansen lors d'une campagne dans l'océan Arctique sur la dérive des icebergs selon un certain angle à droite de la direction du vent, Ekman proposa en 1905 une théorie pour expliquer la déviation des masses d'eaux à la

²Limite entre les eaux profondes, froides, et les eaux de surface, plus chaudes

surface par rapport à la direction du vent (Ekman, 1905). Pour cela, Ekman considéra l'équation de la quantité de mouvement et l'état stationnaire d'un océan forcé par le vent:

$$\begin{cases} f v + A_z \frac{\partial^2 u}{\partial z^2} &= \frac{1}{\rho_0} \frac{\partial P}{\partial x} \\ -f u + A_z \frac{\partial^2 v}{\partial z^2} &= \frac{1}{\rho_0} \frac{\partial P}{\partial y} \end{cases} \quad (2.1)$$

où x, y, z sont respectivement dirigés vers l'Est, le Nord et le haut. u et v sont respectivement la composante zonale et méridienne de la vitesse du vent, f le paramètre de Coriolis, P la pression, ρ_0 la densité de l'eau constante et A_z le coefficient de viscosité turbulent vertical. En décomposant le champ de vitesse (u, v) en une composante géostrophique (u_G, v_G) et une composante turbulente (u_E, v_E) , on obtient les deux systèmes d'équations suivants:

- Les équations d'écoulements géostrophiques :

$$\begin{cases} f v_G &= \frac{1}{\rho_0} \frac{\partial P}{\partial x} \\ f u_G &= -\frac{1}{\rho_0} \frac{\partial P}{\partial y} \end{cases} \quad (2.2)$$

- Et les équations d'Ekman décrivant l'équilibre des forces entre la tension turbulente et la force de Coriolis :

$$\begin{cases} f v_E + A_z \frac{\partial^2 u_E}{\partial z^2} &= 0 \\ -f u_E + A_z \frac{\partial^2 v_E}{\partial z^2} &= 0 \end{cases} \quad (2.3)$$

En supposant que la tension verticale est uniquement due à l'influence du vent, que celui-ci est stationnaire, que la surface de l'océan est lisse et que la densité de l'eau est constante et égale ρ_0 on a: $\tau_x = \rho_0 A_z \frac{\delta u}{\delta z}$, $\tau_y = \rho_0 A_z \frac{\delta v}{\delta z}$ et l'équilibre entre la force de Coriolis et la tension de vents se traduit par les équations suivantes:

$$\begin{cases} f \rho_0 v_E + \frac{\partial \tau_x}{\partial z} &= 0 \\ -f \rho_0 u_E + \frac{\partial \tau_y}{\partial z} &= 0 \end{cases} \quad (2.4)$$

avec τ_x et τ_y les tensions de vents zonales et méridiennes. En intégrant ces équations de la surface où la tension turbulente est supposée égale à la tension de vent jusqu'à la profondeur $-H$ où la tension turbulente est nulle, on obtient ainsi le transport d'Ekman méridien $T_y = \int_{-H}^0 v_E dz$ et zonal $T_x = \int_{-H}^0 u_E dz$ sous la forme :

$$\begin{cases} T_x &= \frac{\tau_s^y}{\rho_0 f} \\ T_y &= \frac{\tau_s^x}{\rho_0 f} \end{cases} \quad (2.5)$$

où τ_s^x et τ_s^y sont les tensions de vents zonales et méridiennes à la surface. Comme on peut le voir la direction du transport d'Ekman est dirigée à 90° vers la droite dans l'Hémisphère Nord et à 90° vers la gauche dans l'Hémisphère Sud par rapport à la direction du vent.

Pompage d'Ekman:

Le stress du vent n'étant pas uniforme sur toute la surface de l'océan, les variations

spatiales du vent et du transport d'Ekman induisent des zones de convergences et de divergences des eaux de surface. Cette convergence ou divergence des eaux de surface est à l'origine du pompage d'Ekman noté W_{EK} . Celui-ci est défini comme la vitesse verticale à la base de la couche d'Ekman et s'obtient par l'équation de continuité appliquée au transports d'Ekman :

$$W_{EK} = \int_{-H}^0 \frac{\partial u_E}{\partial x} + \frac{\partial v_E}{\partial y} dz = \frac{1}{\rho_0 f} \left(\frac{\partial \tau_s^y}{\partial x} - \frac{\partial \tau_s^x}{\partial y} \right) \quad (2.6)$$

$$W_{EK} = \frac{rot(\tau_s)}{\rho_0 f} \cdot \mathbf{z} \quad (2.7)$$

où $-H$ représente la profondeur de la couche d'Ekman, $rot(\tau_s)$ le rotationnel de la tension de vent de surface et \mathbf{z} un vecteur unitaire dirigé vers le haut.

Saisonnalité et caractéristiques de l'upwelling de bord-est du Pérou-Chili:

Le système d'upwelling Humboldt Pérou-Chili (Strub, 1998) est l'une des régions océaniques les plus productives au monde en termes de phytoplancton (Chavez et al., 2008; Kämpf and Chapman, 2016), et représente l'une des principales pêcheries au monde avec une grande biodiversité (Cushing, 1971; Fréon et al., 2009). Comme dans d'autres systèmes d'upwelling de bord Est, l'action du stress du vent de surface pousse les eaux de surface au large et pompe des eaux plus profondes, froides et riches en nutriments (Sverdrup, 1937), déclenchant une productivité de surface élevée. Ce système d'upwellings au niveau du Pérou tend à se renforcer légèrement en hiver austral et diminuer en été tandis qu'aux plus hautes latitudes, celui du Chili est particulièrement intense au printemps et en été austral (Fig. 2.8b). Cette saisonnalité est grandement influencée par la migration saisonnière de l'ITCZ et plus particulièrement celle de l'anticyclone du Pacifique Sud Est qui tend à se déplacer vers le sud en été austral (Fig. 2.8a) et constitue le moteur principal de la circulation atmosphérique dans cette région (Muñoz and Garreaud, 2005; Garreaud and Falvey, 2009).

Avec la présence des Andes une importante chaîne montagneuse, et des sommets pouvant dépasser les 6000 mètres d'altitudes, la topographie de l'Amérique du Sud joue également un rôle important dans la structure des conditions atmosphériques dans cette région (Fig 2.8 et Figueroa and Moffat, 2000). La distance relativement faible entre la fosse océanique et la cordillère des Andes qui agit à la façon d'un mur naturel, force d'une part les alizés tropicaux et subtropicaux à longer la côte contribuant ainsi à la formation des vents côtiers à l'origine de l'upwelling du Pérou-Chili (Belmadani et al., 2014), et isole d'autre part le littoral des perturbations atmosphériques du continent et de la côte Atlantique (Xu et al., 2004; Sepulchre et al., 2009). Plusieurs questions se posent sur l'évolutions de ces systèmes upwellings dans un contexte de changement climatique (Aravena et al., 2014; Falvey and Garreaud, 2009; Gutiérrez et al., 2011). Nous aborderons plus en détail dans le chapitre 3 de ce manuscrit l'évolution du systèmes d'upwelling du Pérou-Chili en réponse au changement climatique récent et examinerons les liens

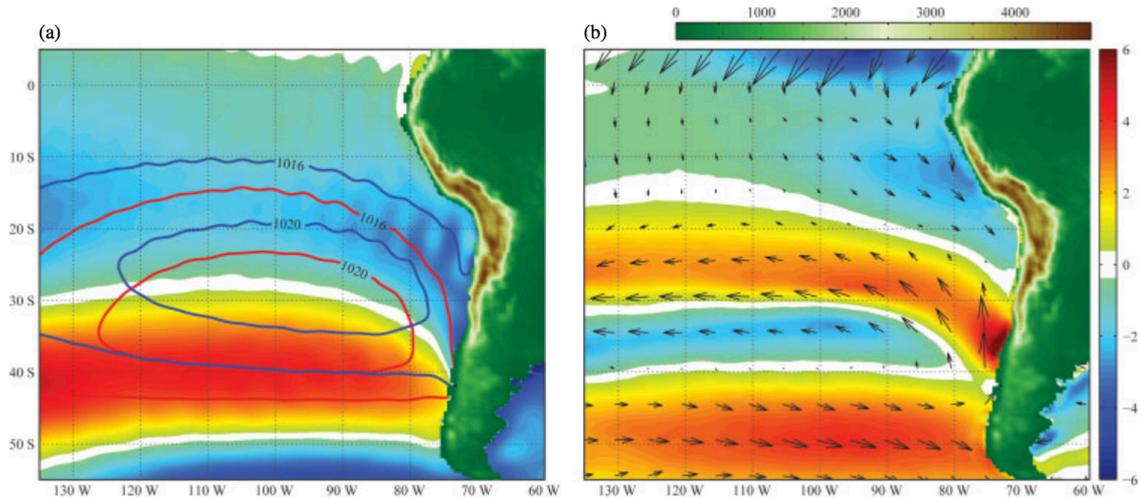


Figure 2.8: (a) Pression de surface océanique moyenne en hPa pour DJF (Décembre-Janvier-Février en contour rouge) et JJA (Juin-Juillet-Aout en contour bleu). L'amplitude saisonnière (DJF moins JJA) en hPa est indiquée en couleur. (b) Amplitude saisonnière (DJF moins JJA) des vents à 10m en $m.s^{-1}$ en couleur et les vecteurs associés. Données CFSR à 00 UTC (Source : [Rahn and Garreaud, 2014](#))

avec les tendances des températures de surface observées au large du Chili au cours des quatre dernières décennies ([Jebri et al., 2020](#)). Nous enquêterons également sur le rôle des forçages externes et les processus à l'origine de ces tendances.

2.3 La variabilité climatique

Ces circulations océaniques et atmosphériques décrites dans la section précédente caractérisent notre climat moyen défini comme la moyenne de toutes les situations météorologiques sur une période de plusieurs décennies. L'écart à ces situations moyennes caractérise la variabilité de notre climat. Cette variabilité se compose d'une part de la variabilité dite « interne », c'est à dire générée par les interactions spontanées entre ses différentes composantes, et d'autre part de la variabilité dite « forcée » induite par l'influence des variations des forçages externes sur le climat moyen et sa variabilité. La combinaison de ces deux types de variabilités et leur dynamique intrinsèque définissent la variabilité de notre climat. Je reviendrai plus en détail dans la section 2.4 sur la notion de forçages externes.

La variabilité interne du système climatique quant à elle s'organise selon des modes de variabilités préférentiels et sont généralement définis comme l'occurrence cyclique et/ou régulière d'anomalies de certains paramètres climatiques selon un même schéma spatial caractéristique. Divers indices sont utilisés pour caractériser et suivre l'évolution de ces modes de variabilité ceux-ci étant généralement défini à partir de moyennes de certaines variables dans certaines régions et/ou à partir d'une méthode statistique comme les analyses en composantes principales (ou EOFs pour en anglais Empirical Orthogonal Function) utilisées pour extraire les principaux modes spatio-temporels de variabilités en construisant par diagonalisation de la matrice de covariance, un ensemble d'axes

indépendants expliquant au mieux la variance des données (Navarra and Simoncini, 2010). L'analyse de cette variabilité à l'échelle globale est toutefois limitée par la couverture spatiale incomplète et par le nombre restreint de longues séries temporelles continues qui ne permettent pas de correctement estimer les caractéristiques des spectres de fréquences climatiques. L'analyse des séries annuelles de température indique que les variations à basse fréquence (longues échelles de temps, supérieures à la décennie) sont les plus importantes (Huybers and Curry, 2006). On parle alors de spectre de type « bruit rouge ». Une explication pour cette caractéristique spectrale des températures de surface peut s'appuyer sur les travaux pionniers de Frankignoul and Hasselmann, 1977 qui ont montré que le spectre de puissance au niveau de l'océan de surface est très souvent rouge. En raison de sa forte inertie thermique et dynamique, l'océan filtre en effet les variations turbulentes de l'atmosphère dont les fréquences sont beaucoup plus élevées celles-ci allant de la seconde à la semaine environ. Ainsi, les fluctuations atmosphériques les plus rapides n'ont pas le temps de mettre en mouvement ou de modifier la température de surface de l'océan, qui les intègre et varie beaucoup plus lentement.

Nous savons aujourd'hui que l'impact de ces modes de variabilités est important sur nos sociétés et que ceux-ci influencent de nombreuses régions autour du globe via des téléconnexions atmosphériques (Gill, 1980; Liu and Alexander, 2007). Afin d'avoir une meilleure vision des composantes mises en jeu, nous donnons ci-dessous une brève description de certains modes de variabilité du climat, de leur impact et des connaissances actuelles sur les processus associés.

2.3.1 Variabilité dans l'Océan Pacifique

L'océan Pacifique est le plus grand océan mondial et s'étend sur presque un tiers de la surface terrestre. Les circulations atmosphériques et océaniques y interagissent continuellement. L'analyse des séries de température dans cet océan, pouvant remonter sur plusieurs siècles met en évidence des réchauffements importants des eaux de surface équatoriales au large des côtes péruviennes tous les 2 à 7 ans (Jin et al., 2006), connus sous le nom d'El Niño. Ces épisodes chauds sont parfois suivis d'événements froids (La Niña). Durant les épisodes chauds de ce mode de variabilité, appelé oscillation australe ou ENSO (pour en anglais: El Niño Southern Oscillation), les alizés (vents de secteur Est soufflant sur la bande équatoriale) sont plus faibles qu'en temps normal. Des couplages océan-atmosphère permettent à ce type de situation de perdurer un an voire plus (Philander, 1989).

En condition normale, les alizés soufflent dans le Pacifique d'Est en Ouest et confinent à l'Ouest du bassin une masse d'eau chaude appelé « warm pool » (Fig. 2.6). A l'est du bassin, le long des côtes Péruvienne-Chilienne, les vents générés par la centre de haute pression subtropical induisent sous l'influence de la force de Coriolis un phénomène d'upwelling ramenant à la surface des eaux plus froides (Fig. 2.6). L'ENSO tendra à renforcer ce contraste Est-Ouest lors d'événement La Niña et au contraire l'amoindrir lors d'événement El Niño. Les phases positives de cette oscillation sont caractérisées par

un affaiblissement des alizées, une thermocline à l'Est anormalement profonde et des anomalies positives des températures de surface à l'Est du Pacifique équatorial. Afin de bien caractériser la fréquence et l'amplitude des différents types d'événements El Niño observés, de nombreux indices ont été définis dans les tropiques à partir des températures de surface de différentes régions du Pacifique équatorial (Fig. 2.9).

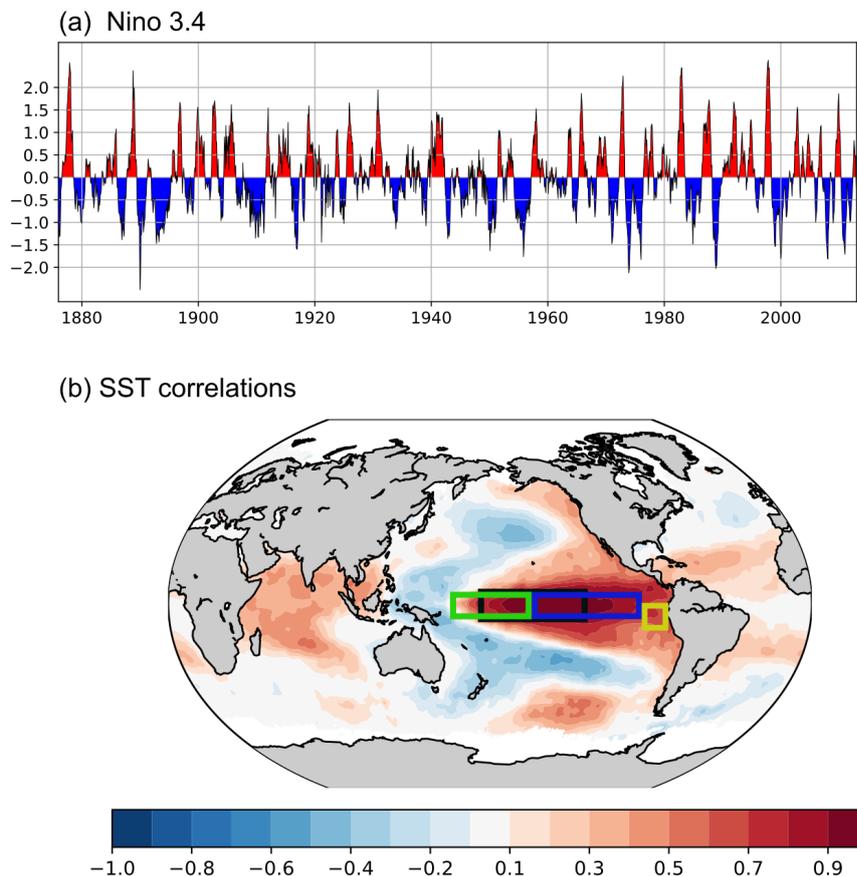


Figure 2.9: (a) Evolution temporelle de l'ENSO d'après l'indice Niño 3.4 calculé à partir des observations de température de surface océanique HadISST sur la période 1878 à 2018. (b) Corrélations entre les températures de surface océanique et l'indice d'ENSO. Les boîtes indiquent les régions Niño-3 (en bleu, 5°S-5°N et 150°W-90°W) pour les événements localisés à l'Est du bassin, Niño-3.4 (en noir, 5°S-5°N et 170°W-120°W) pour les événements centraux, Niño-4 (en vert, 5°S-5°N et 160°E-150°W) ceux piquant à l'ouest de la Date line (la ligne de changement de date internationale) et Niño1+2 (en jaune, 0°S-10°S, 90°W-80°W) pour les anomalies les plus à l'Est et correspondant à la région côtière d'Amérique du Sud.

Lors d'événements El Niño la localisation des branches ascendantes et descendantes de la circulation de Walker est modifiée (Fig. 2.10). La branche ascendante à l'Ouest de l'océan Pacifique tend à se déplacer vers le centre du bassin avec une subsidence accrue au niveau de l'Amérique centrale et l'Atlantique équatorial. La modification des vents de surface et les mouvements d'air associés à l'échelle du bassin Pacifique vont favoriser un approfondissement de la thermocline équatoriale et affaiblir la résurgence d'eau froide le long du Pérou et du Chili. Ces changements vont avoir un impact sur le climat global via des téléconnexions atmosphériques et ainsi favoriser par exemple des sécheresses en

Indonésie et dans le Nordeste Brésilien, de fortes précipitations et une diminution de la prise de pêche sur la côte Péruvienne.

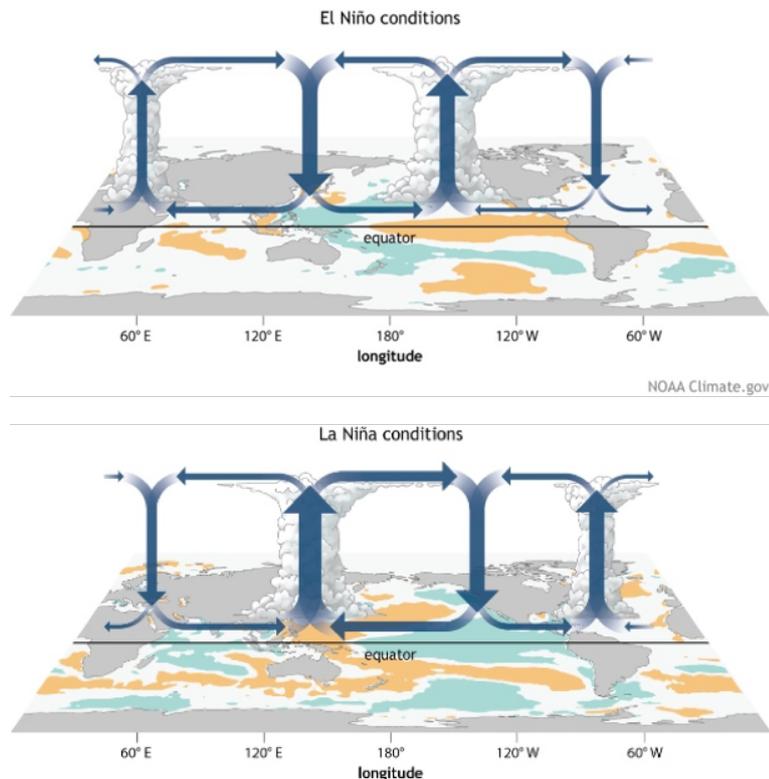


Figure 2.10: Illustration schématique de la circulation de Walker pour des conditions El Niño (en haut) et des conditions La Niña (en bas). (Source : NOAA)

De nombreuses théories ont été formulées sur les processus à l'origine de cette variabilité pseudo oscillatoire (Wang et al., 2012). Parmi les plus connues on trouve la théorie de l'oscillateur retardé (Suarez and Schopf, 1988; Graham and White, 1988; Battisti and Hirst, 1989), l'oscillateur du Pacifique ouest (Weisberg and Wang, 1997), l'oscillateur rechargé-déchargé (Jin, 1997a; Jin, 1997b), l'oscillateur advectif-réfléctif (Picaut et al., 1997) ou encore la théorie de la barrière de sel (Vialard and Delecluse, 1998) cette dernière mettant en avant le rôle actif de la salinité dans le déclenchement d'événements El Niño.

Aux échelles de temps décennales, on retrouve un phénomène similaire à l'ENSO en terme de signature spatiale, l'Oscillation Pacifique Interdécennale (en anglais Interdecadal Pacific Oscillation, IPO) (Power et al., 1999; Folland et al., 2002; Fig. 2.11). Principal mode de variabilité global opérant aux échelles de temps décennales, la structure spatiale des températures de surface dans l'océan Pacifique est similaire à l'ENSO dans la bande inter tropicale mais s'étale en revanche davantage vers les latitudes subtropicales et se distingue par des anomalies de températures et de vents plus marquées aux hautes latitudes (Fig. 2.11) (Tourre et al., 2001).

Les phases positives de l'IPO se traduisent par des anomalies de températures positives dans le Pacifique tropical Centre-Est et des anomalies négatives plutôt à l'ouest du Pacifique Nord et Sud (Fig. 2.11). Sans fréquence d'oscillation préférentielle unique bien définie, cette oscillation se trouve principalement dans une gamme de fréquence de

période 15-25 ans et 50-70 ans (Mantua and Hare, 2002; MacDonald and Case, 2005). Les indices proposés pour caractériser cette variabilité décennale Pacifique (voir Tableau 1 de Henley et al., 2015) se basent sur des analyses en composantes principales appliquées aux anomalies de températures de surface dans le Pacifique, ou à partir d'un indice tripolaire (en anglais, Tripole Index ou TPI; Henley et al., 2015) défini à partir des anomalies de températures de surface moyennes dans 3 régions distinctes. Une tendance négative particulièrement longue de l'IPO est observée depuis 1980 avec des phases positives durant les périodes 1915-1944 et 1978-1998 et une phase négative durant la période 1946-1977 (Fig 2.11a).

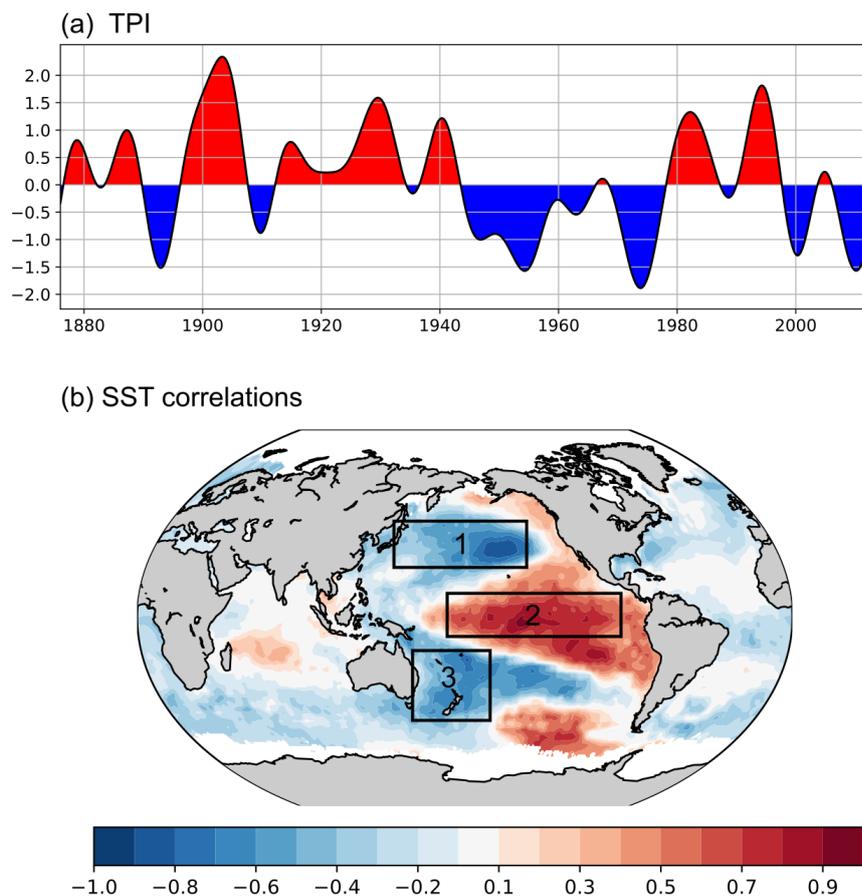


Figure 2.11: (a) Evolution annuelle de l'IPO calculée à partir des observations HadISST sur la période 1878 à 2018 d'après l'indice TPI tel que défini par (Henley et al., 2015). (b) Corrélations entre les températures de surface océaniques et l'indice TPI. Les 3 boîtes dans le Pacifique indiquent les 3 régions utilisées dans le calcul du TPI avec la region 1 entre 25°N–45°N et 140°E–145°W, la region 2 entre 10°S–10°N et 170°E–90°W et la region 3 entre 50°S–15°S et 150°E–160°W.

En plus de moduler l'impact de l'ENSO sur les précipitations et températures dans certaines régions (Meehl and Hu, 2006) comme en Australie (Power et al., 1999; Arblaster et al., 2002) ou en Amérique du Sud (Andreoli and Kayano, 2005), l'IPO influence le climat dans de nombreuses régions du globe (Dong and Dai, 2015) notamment au sud-ouest des Etats-Unis (Meehl and Hu, 2006; Dai, 2013), au sud de l'Afrique (Reason and Rouault, 2002), à l'ouest du Canada (Verdon-Kidd and Kiem, 2009; Whitfield et al., 2010), ou encore dans la région du Sahel (Mohino et al., 2011).

Depuis la mise en évidence du rôle de la phase négative de l'IPO dans le ralentissement du réchauffement global depuis le début des années 80s (en raison de la taille de l'Océan Pacifique et l'impact sur les températures globales via les téléconnexions atmosphériques) un intérêt croissant est porté à l'étude de ce mode de variabilité et son impact sur le contenu en chaleur de l'Océan (par exemple [Gastineau et al., 2019](#)) et les climats régionaux (par exemple [Falvey and Garreaud, 2009](#); [Gutiérrez et al., 2011](#) et [Vuille et al., 2015](#)). Je reviendrais sur ce point dans le chapitre 3 du manuscrit.

2.3.2 Variabilité décennale dans l'Océan Atlantique

À nos latitudes, dans la région de l'Atlantique Nord, un autre type d'organisation préférentielle aux échelles de temps décennales opère à travers l'Oscillation Multidécennale Atlantique ou AMO (en anglais, Atlantic Multidecadal Oscillation) ou AMV (en anglais, Atlantic Multidecadal Variability). Ce mode se caractérise par des anomalies de température de surface alternativement chaudes puis froides sur tous le bassin de l'Atlantique Nord avec une période d'oscillation comprise entre 60 et 80 ans (Fig. 2.12a) ([Schlesinger and Ramankutty, 1994](#); [Knight et al., 2005](#); [Delworth et al., 2007](#)). Son indice est souvent défini à partir de la moyenne des anomalies de températures de surface dans le bassin Atlantique Nord. Sur la période instrumentale, des phases négatives de l'AMO ont été observées pour les périodes 1900-1920, 1960-1995 et positives sur les périodes 1930-1960 et de 1995 à aujourd'hui ([Knight et al., 2005](#); [Alexander et al., 2014](#)). Tout comme pour l'IPO ([Dai et al., 2015](#)), les oscillations associées à l'AMO ont une influence sur les hydroclimats régionaux du pourtour Atlantique ([Schlesinger and Ramankutty, 1994](#); [Wu et al., 2011](#); [Muller et al., 2013](#); [Tung and Zhou, 2013](#)).

Les variations de l'AMO ont été associées à des variations de températures ([Semenov et al., 2010](#)) et de précipitations ([Enfield et al., 2001](#)) dans l'Hémisphère Nord. Par exemple, les phases positives de l'AMO sont associées à une augmentation des températures sur la plupart de l'Amérique du Nord et l'Europe et d'avantage de précipitations au sud de l'Amérique du Nord et à l'ouest de l'Europe ([Enfield et al., 2001](#); [McCabe et al., 2004](#); [Sutton and Hodson, 2005](#)). Il a également été montré que le gradient de pression inter hémisphérique causé par le gradient de température entre l'Atlantique Nord et l'Atlantique Sud durant les phases positives de l'AMO conduisait à un déplacement vers le nord de l'ITCZ (la zone intertropicale de convergence où d'importantes précipitations se produisent) ([Levine et al., 2018](#)), causant ainsi une augmentation des précipitations dans la région du Sahel et une diminution des précipitations au Nord-Est du Brésil ([Zhang and Delworth, 2006](#); [Alexander et al., 2014](#); [Villamayor et al., 2018](#)).

En raison de sa période d'oscillation longue et le manque d'observations les processus à l'origine de cette variabilité restent mal compris ([Delworth and Greatbatch, 2000](#); [Knight et al., 2005](#); [Medhaug and Furevik, 2011](#); [Zhang et al., 2013](#); [Clement et al., 2015](#); [Zhang et al., 2016](#)). Bien que la théorie la plus acceptée soit celle d'un mode de variabilité associé à la circulation océanique de retournement dans l'Océan Atlantique ou AMOC (pour en anglais, Atlantic Meridional Overturning Circulation; [Zhang and Wang, 2013](#); [Marini and Frankignoul, 2014](#); [Alexander et al., 2014](#); [Clement et al., 2015](#)), d'autres théories suggèrent

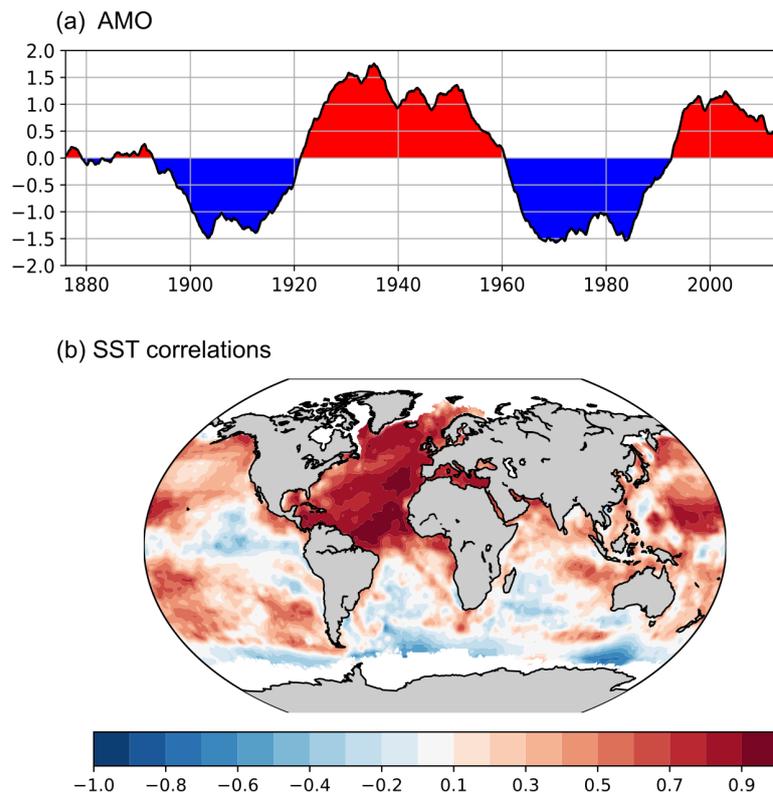


Figure 2.12: (a) Indice de l'AMO calculé comme la moyenne des températures de surface océaniques HadISST dans l'Atlantique Nord (0°N - 60°N ; 75°W et 7.5°W) sur la période 1878 à 2018. (b) Corrélations entre les températures de surface océaniques et l'indice d'AMO.

un lien avec l'Oscillation Nord Atlantique ou NAO (pour en anglais, North Atlantic Oscillation) (Gastineau and Frankignoul, 2012), ou encore une influence significative des forçages volcanique, solaire ou anthropique (Rotstayn and Lohmann, 2002; Otterå et al., 2010; Terray, 2012; Booth et al., 2012; Watanabe and Tatebe, 2019).

Parmi les questions scientifiques associées à l'influence de l'AMO sur le climat, de plus en plus d'études s'intéressent à l'impact de cette variabilité sur les autres bassins (Sutton and Hodson, 2005). En raison des changements de pression inter bassins intervenant lors des variations de l'AMO, il est probable que ce mode de variabilité ait un impact notamment sur l'ENSO ou l'IPO dans l'Océan Pacifique via des ponts atmosphériques altérant la circulation de Walker (McGregor et al., 2014; Gong et al., 2019).

2.4 Les forçages externes

Le terme de forçage externe se réfère à l'ensemble des facteurs extérieurs au système climatique pouvant modifier son bilan radiatif. Toute perturbation du bilan radiatif entraîne une période d'ajustement plus ou moins longue qui conduit le système climatique vers un nouvel état d'équilibre. L'impact régional ou global tout comme la durée et l'amplitude est spécifique à chaque forçage (Fig. 2.13).

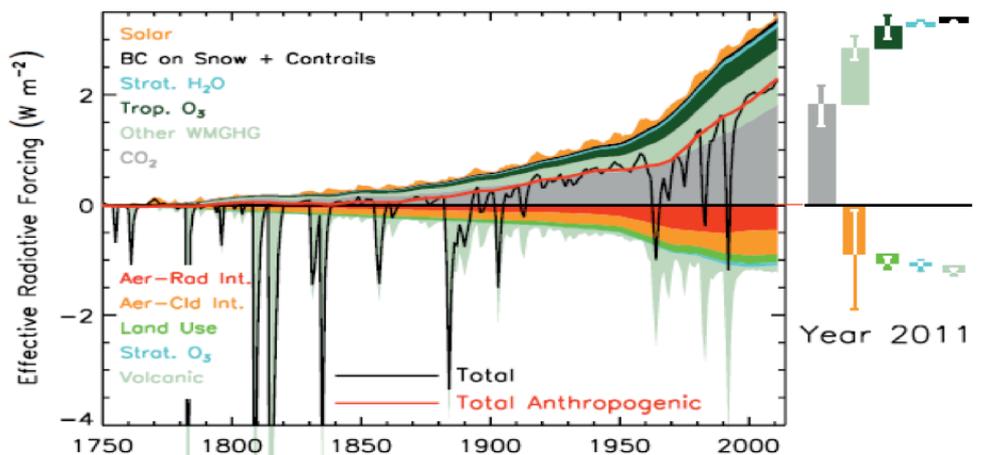


Figure 2.13: Évolution temporelle des mécanismes de forçage anthropiques et naturels (Source: Figure 8.18 de Myhre et al., 2013)

2.4.1 Les forçages naturels

Parmi les forçages naturels qui ont joué un rôle prépondérant avant le 19^{ème} siècle et l'ère industrielle, on trouve le forçage solaire lié à l'activité du soleil, le forçage astronomique lié aux caractéristiques orbitales de la Terre par rapport au Soleil et le forçage volcanique lié aux éruptions stratosphériques.

Alors que les variations très lentes des paramètres orbitaux de la Terre modifient la quantité de radiation solaire reçue et sa répartition, le rayonnement émis par le Soleil n'est quant à lui pas tout à fait constant (Lean, 1991; Hathaway, 2015). Il est en effet admis que les radiations émises par le Soleil fluctuent selon certains cycles notamment de 11 ans, 22 ans (appelé cycle de Hale) (Hale et al., 1919), 85 ans (appelé cycle de Gleissberg) (Gleissberg, 1939) et de 210 ans (appelé cycle de Suess) (Suess, 1980). En plus d'agir sur le bilan radiatif global, ces modulations graduelles de l'ordre de 0.1% du rayonnement solaire semblent affecter de façon spécifique les différentes latitudes et les différentes couches de l'atmosphère (Beer et al., 2000; Haigh, 2007; Pulkkinen, 2007) ce qui permet de déterminer l'empreinte spatiale de ces variations solaires sur les températures de surface (White et al., 1997; Tung and Camp, 2008). En outre, l'impact du forçage astronomique sur le climat des 1000 dernières années se limite aux hautes latitudes notamment en été dans l'Hémisphère Nord au-dessus de 65°N (Kaufman et al., 2009; Servonnat et al., 2010).

En opposition à ces forçages graduels et cycliques, on trouve le forçage volcanique qui s'exprime sous forme brutale et ponctuelle. Lors de grandes éruptions volcaniques des quantités considérables de gaz riches en dioxyde de soufre peuvent être projetées dans la stratosphère. En raison des faibles concentrations en vapeur d'eau et des mouvements verticaux limités dans la stratosphère, ces gaz peuvent être transformés par réactions chimiques en aérosols sulfatés capables de rester en suspension plusieurs mois voir années dans l'atmosphère avant de retomber au niveau des pôles (Trepte and Hitchman, 1992; Trepte et al., 1993; Holton et al., 1995). En raison de leur distribution et taille ces particules vont alors perturber le bilan radiatif terrestre en reflétant une partie du

rayonnement solaire (principalement dans l'ultra-violet et le visible) et absorbant une partie du rayonnement infrarouge. Bien qu'il soit de courte durée (2 ans ce qui correspond à la durée de vie des aérosols dans l'atmosphère) le forçage volcanique est de grande amplitude (Fig. 2.13) et peut donc influencer (notamment via la longue mémoire de l'océan) le climat et sa variabilité sur plusieurs années à décennies.

2.4.2 Les forçages anthropiques

L'influence de l'Homme sur son environnement n'a cessé de croître au cours du dernier siècle et son impact sur le climat n'est aujourd'hui plus à démontrer. Le terme d'ère Anthropocène a dès lors été proposé pour définir cette période où des changements environnementaux globaux sont associés aux activités humaines (Zalasiewicz et al., 2015; Crutzen, 2016). Parmi ces forçages on trouve l'utilisation des sols qui par la mise en culture, la déforestation et plus généralement la modification des surfaces végétales a influencé à l'échelle régionale les températures de surface et taux d'humidité des sols mais également de façon relativement faible la composition atmosphérique en intervenant dans le cycle global du carbone (Pielke Sr et al., 2002; Forster et al., 2007b). L'influence de ce forçage reste néanmoins relativement faible et ce n'est qu'après la révolution industrielle avec la destruction de la couche d'ozone, les émissions des gaz à effets de serre et d'aérosols que l'impact de l'activité humaine sur le climat est devenu flagrante.

L'influence des aérosols d'origine anthropique est complexe et dépend de leurs compositions et distributions en taille. Les aérosols sulfatés peuvent par exemple refléter vers l'espace le rayonnement solaire reçue à la surface, absorber une partie du rayonnement terrestre émis ou encore servir de noyaux de nucléation pour la formation des nuages par divers processus radiatifs et microphysiques complexes (Haywood and Boucher, 2000; Penner et al., 2001). Les gaz à effet de serre quant à eux ont tendance à laisser pénétrer le rayonnement solaire (situé principalement dans le visible) et à absorber les rayons infrarouges. La Terre émettant, comme tout corps noir, de la chaleur par rayonnement infrarouge vers l'espace, les gaz à effets de serre tendent à retenir les rayonnements infrarouges émis par le sol tout en laissant pénétrer une grosse partie du rayonnement solaire. Présent dans l'atmosphère sous forme de vapeur d'eau et de gouttelette d'eau, l'H₂O constitue à lui seul 72% de l'effet de serre et le CO₂ explique la majorité des 28% restant. L'émission de gaz à effet de serre par la combustion d'énergie fossile et en particulier de CO₂ a induit une augmentation du forçage radiatif de l'ordre de 2.4 W/m² depuis le début du 20^{ème} siècle (Fig. 2.13; IPCC, 2018). Les deux principales conséquences directes étant l'augmentation des températures (Bindoff et al., 2007) et l'acidification des océans (Doney et al., 2009). Les 4^{ème} et 5^{ème} rapport du GIEC (IPCC, 2007; IPCC, 2014) ont admis avec un fort degré de certitude que le réchauffement global observé est bien lié aux émissions de gaz à effet de serre issues de l'activité humaine.

2.5 Les modèles climatiques

2.5.1 Présentation

Les modèles climatiques sont des modèles numériques et mathématiques qui tentent d'approcher de manière plus ou moins détaillée les processus qui régissent le climat. Ne pouvant reproduire parfaitement la complexité des processus existants, une représentation simplifiée est faite dans ces modèles. Les approximations et simplifications effectuées sont réalisées selon les processus étudiés et les échelles spatiales et/ou temporelles considérées. Ces approches donnent lieu à une gamme très variée de modèles climatiques plus ou moins complexes depuis les modèles de bilan d'énergie (ou EBM pour en anglais Energy Balance Model) utilisés pour étudier de manière détaillée les processus radiatifs jusqu'aux modèles de circulation générale (ou GCM pour en anglais Général Circulation Model) qui caractérisent de manière tridimensionnelle les différentes composantes du système climatiques. Tandis que les modèles les plus simples permettent une bonne approximation du système au premier ordre, les plus complexes tentent d'approcher au mieux la réalité. Ce large éventail de modèle a permis de grandes avancées dans la compréhension de notre système climatique et des différents processus qui le gouvernent.

2.5.2 Les modèles de circulation générale

Les modèles de circulation générale se basent sur les lois fondamentales de la physique, telles que la conservation de la masse, de l'énergie et de la quantité de mouvement pour mimer les interactions et processus dynamiques qui gouvernent notre climat (Edwards, 2011). Lors de l'implémentation de ces modèles sur ordinateurs, de nombreuses approximations sont faites sur les équations issues des lois fondamentales de la physique pour pouvoir être discrétisées à la fois temporellement et spatialement sur un maillage 3D (Fig. 2.14). Cette représentation simplifiée de la réalité, permet la résolution à chaque point de grilles des équations fondamentales à l'origine des interactions spatiales et de la variabilité temporelle du système. De nos jours ces modèles représentent l'essentiel des modèles climatiques utilisés pour estimer les changements climatiques futurs. Ils sont constitués d'un ensemble de modules représentant chacune des composantes du système climatique. Chaque composante ayant sa propre physique, grille et résolution, un coupleur est généralement utilisé comme interface entre les différents modules.

Bien que très utile, il reste néanmoins difficile d'estimer la fiabilité de ces modèles en raison du manque d'observations disponibles et de la complexité du système qu'ils représentent. Chaque modèle reproduisant plus ou moins bien divers aspects du système climatique, il s'agit d'estimer dans quelle mesure un modèle reproduit les observations et permet de répondre à une question posée. L'une des principales limites de ces modèles provient de la résolution trop grossière utilisée pour représenter explicitement certains processus à fine échelle. Pour compenser cette limitation, une paramétrisation est utilisée

pour représenter ces processus par le modèle comme par exemple la formation des nuages dont l'échelle spatiale caractéristique est inférieure à la maille des modèles.

Cette étape de paramétrisation appelée « tuning » permet de se rapprocher le plus possible des observations et est généralement effectuée paramètre par paramètre, puis module par module avant d'être réalisée pour le modèle complet (Dufresne et al., 2013). La non linéarité et la sensibilité très importante du modèle à certains de ces paramètres révèle la très grande complexité de cette phase de calibration et l'enjeu majeur lié à cette étape de paramétrage. Celle-ci est d'autant plus critique que le manque de données instrumentales disponibles ne permet pas la mise en place de procédure de calibration puis validation rigoureuse de cette série de paramètres sur des jeux de données distincts.

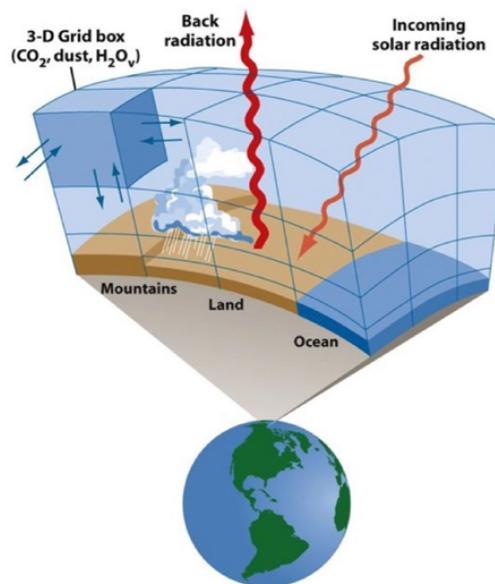


Figure 2-19
Earth's Climate: Past and Future, Second Edition
© 2008 W. H. Freeman and Company

Figure 2.14: (a) Représentation schématique de la discrétisation spatiale effectuée lors de l'implémentation des modèles couplés. (source : Ruddiman et al., 2001)

En pratique, malgré les énormes progrès réalisés ces dernières décennies en terme de puissance de calcul disponible, les ressources informatiques et la puissance des supercalculateurs restent souvent un des grands facteurs limitant lors de la réalisation de simulations avec ce type de modèle. Le temps de calcul augmentant de façon exponentielle avec la résolution du modèle, il est généralement nécessaire de trouver un compromis entre le nombre d'heure de calcul requis, la durée de la ou les simulations, sa résolution et le nombre de simulations à réaliser. Par exemple dans le cadre de simulations sur plusieurs milliers d'années la solution consiste à dégrader la résolution spatiale du modèle.

2.5.3 Description et évaluation succincte du modèle IPSL-CM5A-LR

Le modèle IPSL-CM5 de l'Institut Pierre Simon Laplace a été développé afin de permettre différentes configurations avec différents degrés de complexités pour différentes résolutions et composantes (Figure 1 de [Dufresne et al., 2013](#)). Nous avons utilisé la version IPSL-CM5A-LR (où l'acronyme LR signifie low-resolution) du modèle pour la majorité des expériences réalisées dans le cadre de l'exercice CMIP5 ([Dufresne et al., 2013](#)).

Cette configuration comporte 3 composantes, le modèle d'atmosphère LMDZ ([Hourdin et al., 2013a](#)) couplé à un modèle de surface continentale ORCHIDEE ([Krinner et al., 2005](#)), le modèle d'océan NEMO ([Madec, 2008](#)) qui inclue un modèle de glace de mer LIM2 ([Fichefet and Maqueda, 1997](#)) et de biogéochimie marine PISCES ([Aumont and Bopp, 2006](#)), et enfin un coupleur OASIS ([Valcke, 2013](#)) (Fig. 2.15).

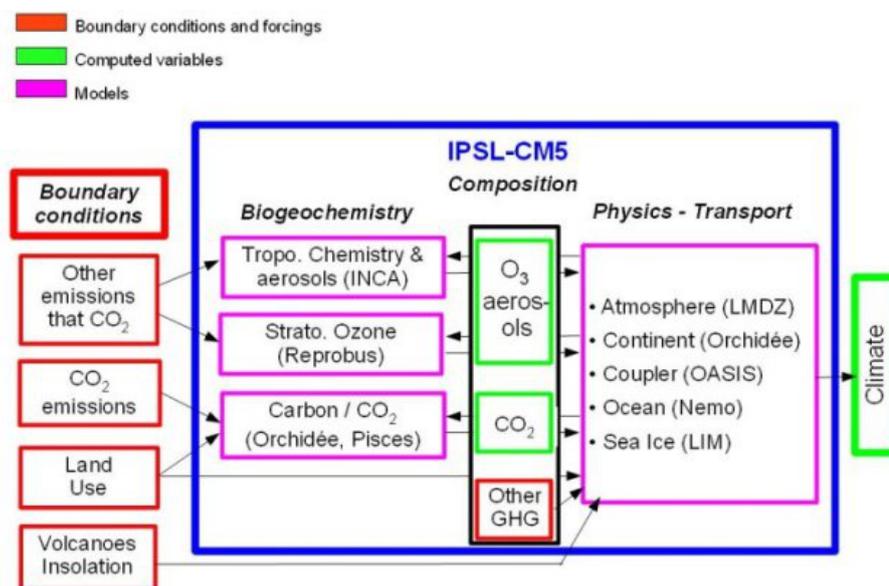


Figure 2.15: Représentation des différents modules du modèle de l'IPSL-CM5. (Source: [Dufresne et al., 2013](#))

La version 5 du modèle atmosphérique LMDZ du Laboratoire de Météorologie Dynamique (avec "Z" pour "zoom") a une résolution de 2.5° pour les longitudes et de 1.25° pour les latitudes. Le nombre de niveaux verticaux atmosphériques est passé de 19 pour la version 4, utilisé lors de l'exercice CMIP3, à 39 pour la version 5. En revanche, les paramétrisations de la version LMDZ5A sont restées quant à elles très similaires à celles de la version LMDZ4 ([Hourdin et al., 2006](#)).

Le modèle LMDZ5A est couplé avec le modèle de surface terrestre ORCHIDEE (pour ORganizing Carbon and Hydrology In Dynamic EcosystEms). Ce module simule le cycle du carbone terrestre, les échanges hydrologiques et énergétiques entre le sol, la végétation et l'atmosphère ainsi que la composition et la distribution de la végétation ([Krinner et al., 2005](#)).

Pour la partie océanique, la version 3.2 du module NEMO (pour Nucleus for European Modeling of the Ocean) est utilisée ([Madec, 2008](#)). Le module NEMO se compose d'un modèle thermodynamique-dynamique de la glace de mer à deux niveaux LIM2

(pour Louvain-la-Neuve Sea Ice Model, Version 2) (Fichefet and Maqueda, 1997; Fichefet and Maqueda, 1999), du module OPA pour la dynamique océanique et du module de biogéochimie marine PISCES (pour Pelagic Interaction Scheme for Carbon and Ecosystem Studies) (Aumont and Bopp, 2006). La configuration de la grille océanique utilisée présente une résolution spatiale moyenne de 2° avec un raffinement de l'ordre de 0.5° pour les latitudes proches de l'équateur. La partie océanique s'étend sur 31 niveaux verticaux d'épaisseurs décroissantes allant de 10m près de la surface à 500m près du fond.

Chaque composante ayant sa propre physique, grille et résolution, le coupleur OASIS3 est utilisé comme interface entre la partie océanique et atmosphérique du modèle (Fig. 2.15). Son rôle est d'interpoler et synchroniser les échanges et de s'assurer de la conservation locale de l'énergie et de l'eau entre les différents modules (Valcke, 2013).

Le modèle IPSL-CM5A-LR possède comme tout modèle de nombreux biais qu'il est nécessaire de prendre en compte lors de son utilisation. En plus d'un biais froid dans la moyenne des températures globales de l'ordre de 1.4°C (Dufresne et al., 2013), il présente le biais classique de nombreux modèles (Dai, 2006) au niveau de la « cold tongue » (Fig. 2.16) et une double ITCZ (Fig. 2.17). Le modèle montre également un biais froid important dans l'Atlantique Nord (Fig. 2.16 et Hourdin et al., 2013b) sans doute associé à une sous-estimation de l'AMOC (Fig. 20 dans Dufresne et al., 2013) par rapport aux valeurs observées (Kanzow et al., 2010).

Le modèle reproduit correctement la variabilité observée de l'ENSO avec un premier pic spectral entre 3 et 3.5 ans et un second au-delà de 4 ans (Fig 11a de Dufresne et al., 2013; Bellenger et al., 2014). La saisonnalité simulée ne correspond en revanche pas à celle observée avec des pics d'ENSO entre Mai-Juin plutôt que Novembre-Janvier et des minima en Octobre-Novembre plutôt qu'en Avril (Bellenger et al., 2014). La variabilité observée dans le Pacifique Nord est bien simulée par le modèle (Fleming and Anchukaitis, 2016) et le lien entre ENSO et IPO semble bien reproduit (Nidheesh et al., 2017). Comme de nombreux modèles, celui-ci tend en revanche à produire lors d'événement El Niño des anomalies de températures s'étendant trop loin à l'Ouest du Pacifique (Fig 10 de Dufresne et al., 2013) et à sous-estimer la variabilité associée à l'IPO dans le Pacifique Sud (Henley et al., 2017).

Le modèle IPSL-CM5A-LR fait partie des modèles les plus sensibles avec une réponse de 2°K par doublement de CO_2 (Andrews et al., 2012; Forster et al., 2013) alors que les modèles CMIP5 possèdent une réponse moyenne par doublement de CO_2 de l'ordre de $1.82^\circ\text{K} \pm 0.63^\circ\text{K}$. Cette grande sensibilité du modèle a été attribuée à la boucle de rétroaction positive des nuages et en particulier des nuages à basses altitudes qui ont la particularité de refroidir la surface terrestre en reflétant les rayonnements solaires (Brient and Bony, 2013). Cette rétroaction positive des nuages bas sur l'albédo terrestre en réponse aux forçages anthropiques constitue la principale source d'incertitude à l'origine des différences de sensibilité entre les modèles CMIP5 (Dufresne and Bony, 2008). En terme de distribution régionale le réchauffement simulé par le modèle IPSL-CM5A-LR est plus important sur les continents que les océans, sur les continents de l'Hémisphère Nord que ceux de l'Hémisphère Sud et plus faible dans l'Atlantique Nord et dans l'océan

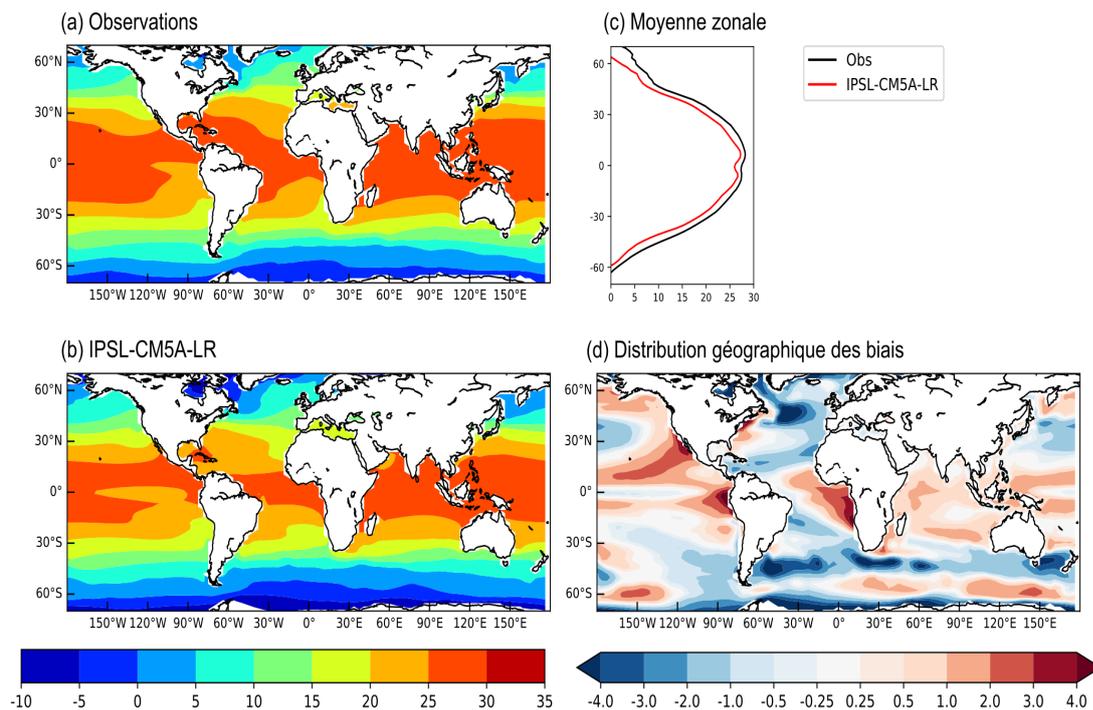


Figure 2.16: Moyennes annuelles des températures de surface océaniques entre 70°S et 70°N sur la période 1980-2005, pour (a) le jeu d’observations HadISST (Rayner et al., 2003) et (b) simulées par le modèle IPSL-CM5A-LR. (c) Moyennes zonales des températures de surface océaniques observées dans HadISST (en noir) et simulées par le modèle IPSL-CM5A-LR (en rouge). (d) Distribution géographique des biais de température de surface océanique entre 70°S et 70°N estimés comme la différence entre les moyennes annuelles simulées et les observations après soustraction dans les valeurs simulées du biais froid moyen de -2°C entre 70°S et 70°N.

Austral (Hourdin et al., 2013b).

2.5.4 Des modèles de circulation générale aux modèles statistiques

Comme nous le verrons plus en détail dans le chapitre 4, les GCMs ne constituent pas l’unique outils à la disposition des chercheurs pour étudier le climat et sa dynamique. En parallèle du développement et de l’améliorations de ces modèles on distingue le développement et l’apparition de modèles dit statistiques. Tandis que les GCMs se basent sur les lois fondamentales de la physique, les modèles statistiques cherchent à déterminer la dynamique d’un système à partir de l’observation de certaines variables d’états. Dopés par l’augmentation du nombre d’observations disponibles ainsi que l’augmentation de la puissance de calcul informatique, ces modèles pourraient à terme devenir de plus en plus complexes et performants. Comme exemple de modèle statistique nous citerons le LIM (pour en anglais, Linear Inverse Model) qui a montré son utilité aussi bien dans la prédiction et l’étude des variables océaniques (Penland and Sardeshmukh, 1995; Penland and Matrosova, 1998; Johnson et al., 2000) que dans celles de variables atmosphériques

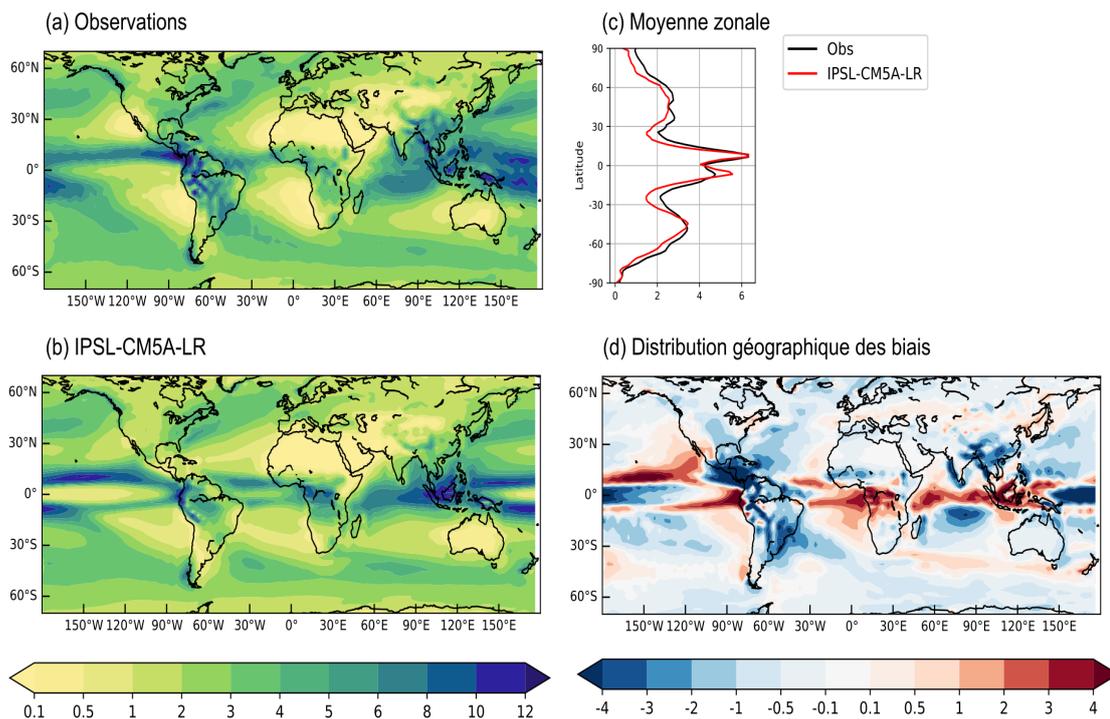


Figure 2.17: (a-b) Moyennes annuelles des précipitations entre 70°S et 70°N sur la période 1980-2005, pour (a) les observations GPCP (Adler et al., 2003) et (b) le modèle IPSL-CM5A-LR. (c) Moyenne zonale des précipitations dans GPCP (en noir) et simulées par le modèle (en rouge). (d) Distribution géographique des biais de précipitations entre 70°S et 70°N estimés comme la différence entre les moyennes annuelles des précipitations dans le modèle et les observations.

(Winkler et al., 2001). Le LIM s'est également avéré utile dans l'étude des interactions air-mer sur la variabilité tropicale (Newman et al., 2009) et dans la séparation de la variabilité forcée et interne (Frankignoul et al., 2017).

2.5.5 Utilisation

Comme énoncé par le statisticien George Box « Tous les modèles sont faux mais certains sont utiles », et c'est en grande partie grâce aux progrès réalisés dans la modélisation physique de notre climat que notre compréhension des processus à l'origine de sa variabilité a considérablement progressée. Les GCMs en particulier ont permis, malgré leurs nombreuses imperfections, d'énormes progrès dans la compréhension de notre climat passé et présent afin d'appréhender les changements futurs. Newman, 2013 a pour sa part, montré que dans le cadre des prédictions décennales des températures de surface, les prédictions obtenues avec un LIM aux horizons 2-5 ans et 6-9 ans sont aussi bonnes (voir meilleurs dans certains cas) que celles réalisées avec des GCMs durant l'exercice de prédiction décennale CMIP5 (Taylor et al., 2012). Une description plus détaillée du LIM que j'ai mis au point dans le cadre de cette thèse et les étapes de sa construction sont données dans le chapitre 4, ce LIM étant utilisé dans le chapitre 5 comme un émulateur des GCMs.

La réalisation d'un éventail très large d'approches et de protocoles expérimentaux permettent d'aborder de très nombreuses questions scientifiques. Chaque question scientifique nécessitant la mise en place d'un protocole particulier, nous nous intéresserons par la suite à l'étude de la variabilité interne liée aux interactions spontanées au sein du système climatique et à l'étude de la variabilité forcée.

Etude de la variabilité interne:

Si l'on souhaite étudier la variabilité interne du climat, c'est-à-dire les interactions spontanées du système en s'affranchissant de la variabilité forcée, une des façons d'y parvenir est de réaliser des simulations très longues dites de contrôles. Durant ces simulations un forçage externe constant représentatif des conditions moyennes préindustrielles est appliqué. Ces simulations permettent d'explorer la gamme de variabilité du système climatique et peuvent servir de référence pour estimer la contribution de la variabilité forcée dans d'autres simulations (par exemple [DelSole et al., 2011](#)). De par l'initialisation imparfaite de certains champs du modèle comme l'océan profond, une période transitoire est couramment observée au début de ces simulations. Il est donc nécessaire que ces simulations soient suffisamment longues afin de s'assurer que la dynamique observée ne soit pas liée à l'initialisation imparfaite de la simulation. Une fois cet état d'équilibre atteint, ces simulations peuvent également être utilisées pour initialiser d'autres simulations et ainsi éviter le choc dynamique au début de l'expérience.

Etude de la variabilité forcée:

Comme introduit dans le chapitre précédant l'un des enjeux majeurs des sciences du climat est de distinguer le rôle et la contribution de chacun des acteurs du système climatique. La détection d'une anomalie climatique vise à identifier la présence d'un changement significatif de l'état moyen du système climatique. L'attribution consiste quant à elle à identifier la source de cette anomalie, qu'elle soit liée à des processus purement internes au système climatique ou à l'influence d'un ou plusieurs forçages externes. Une première approche de détection un peu naïve a été de considérer que la variabilité lente (souvent représentée par des tendances sur plusieurs décennies) correspond à la variabilité forcée et que les modes plus rapides correspondent à la variabilité interne. D'autres approches un peu plus avancées consistent à distinguer à l'aide de divers approches numériques le rôle des différents acteurs en fonction de la fréquence et des patterns spatiaux ([Schneider and Held, 2001](#); [Thompson et al., 2009](#); [Hegerl and Zwiers, 2011](#); [Frankignoul et al., 2017](#)). Mais dans le cadre d'un système aussi complexe et avec si peu d'observations disponibles, l'efficacité de ces approches restent souvent discutables et leur utilisation dans le cadre de l'attribution de ces changements à un forçage spécifique très compliquée.

Les méthodes de détection/attribution les plus couramment utilisées de nos jours consistent à se placer dans le monde des modèles climatiques afin d'isoler certaines composantes forcées et de s'affranchir de la variabilité interne ([Hegerl and Zwiers, 2011](#); [Bindoff et al., 2013](#)). Ces approches visent à estimer la contribution des divers forçages

externes dans les changements observés en comparant les observations disponibles avec les résultats obtenus dans le cadre de simulations numériques. Un ensemble de simulations soumises aux mêmes forçages, mais initialisées à partir de situations différentes est pour cela réalisé. La moyenne d'ensemble de ces simulations est alors associée à la réponse forcée tandis que l'enveloppe décrite par les membres autour de cette moyenne exprime la contribution de la variabilité interne. Cette expérience pouvant être reproduite plusieurs fois avec des jeux de forçages spécifiques, il est possible d'étudier la réponse individuelle ou commune aux différents forçages. En général, l'attribution d'un changement observé dans la réalité à un forçage en particulier s'effectue lorsque des changements similaires sont retrouvés dans les observations et dans la moyenne d'ensemble des simulations incluant le forçage en question. Ces expériences ont permis des avancées majeures dans la compréhension du rôle des différents forçages externes dans les changements climatiques récents (Hegerl and Zwiers, 2011). Il est en revanche important de noter que la somme des réponses obtenues dans différentes expériences pour différents forçages, n'est en rien représentative de la réponse conjointe des forçages en question. Toutefois il reste intéressant de comparer la somme de ces réponses individuelles à la réponse conjointe afin d'estimer la part additive de la réponse dynamique totale. S'apparentant à des méthodes de régressions linéaires multiples, des méthodes statistiques dites de *fingerprint* (Hasselmann, 1997) peuvent également être mises en œuvre pour reconstruire de manière optimale (au sens de la vraisemblance) les observations à partir des réponses forcées obtenus dans les différents jeux de simulations. Avec ces approches la variabilité observée est alors reconstruite à partir d'une combinaison pondérée des effets associés aux différents forçages externes et du bruit de la variabilité interne. Ces méthodes sophistiquées et leurs variantes (e.g Ribes et al., 2009) permettent d'estimer la contribution significative ou non des différents forçages externes sur les changements observés via des tests statistiques de nullité sur les coefficients de régression associés aux différents forçages.

2.5.6 Exercice d'inter-comparaison des modèles CMIP

Le programme international d'inter-comparaison des modèles climatiques CMIP (pour en anglais Coupled Model Intercomparison Project) débuté en 1997 a permis la coordination des grands groupes de modélisations dans le cadre d'exercice d'inter-comparaison de modèle. Le principe repose sur l'élaboration par des grands groupes de modélisations de jeux de simulations standards selon un protocole précis. Ces simulations coordonnées permettent une meilleure estimation des différences entre modèles et sont révélatrices des incertitudes présentes dans les projections futures. Le protocole en question n'a cessé d'évoluer durant les phases successives du projet CMIP en fonction des questions définies comme prioritaires par la communauté scientifique. Parmi les simulations effectuées dans le cadre de la phase 5, on trouve les simulations du dernier millénaire, de la période historique, les prévisions décennales et les projections climatiques pour le 21^{ème} siècle (Taylor et al., 2012). Les résultats obtenus dans le cadre de cet exercice sont pris en compte lors de l'évaluation des connaissances sur le climat par le GIEC (Groupe d'experts inter-gouvernemental sur l'évolution du climat) et lors des discussions inter-gouvernementales

sur les politiques de limitation des émissions à gaz à effets de serre.

Simulations du dernier millénaire et de la période historique :

Dans le cas de simulations du dernier millénaire de l'an 850 à 1850, les forçages naturels liés à l'activité solaire et volcanique sont prépondérant par rapport aux autres forçages. Pour ces simulations les principaux forçages pris en compte sont donc l'activité solaire qui module d'environ 0.1% la quantité d'énergie solaire reçu par la Terre à travers des cycles de 11 ans, 22 ans (cycle de Hale), 85 ans (cycle de Gleissberg) et de 210 ans (cycle Suess), les variations du forçage orbitale et les éruptions volcaniques qui possèdent la signature interannuelle la plus forte mais dont l'action est limitée dans le temps.

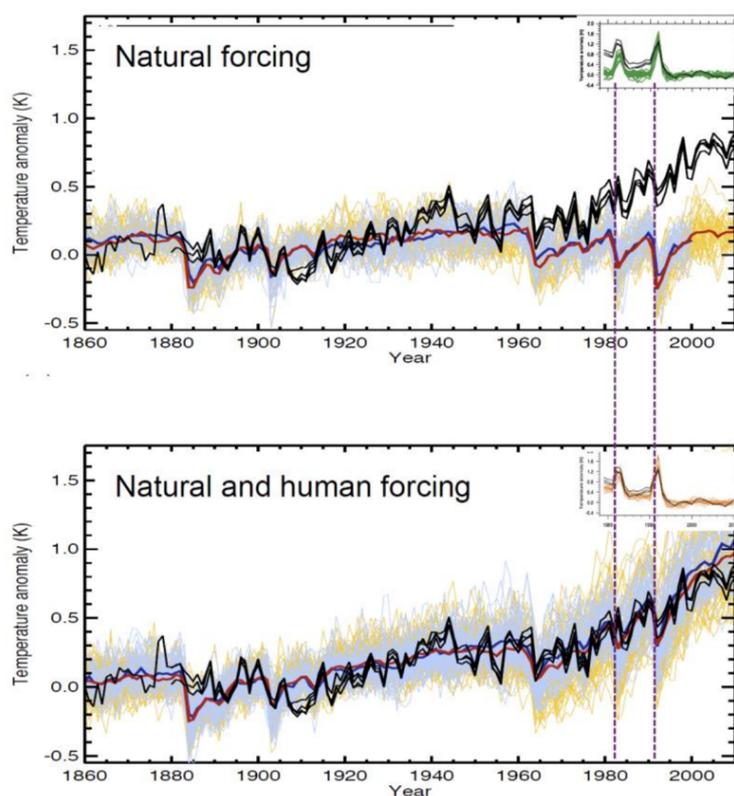


Figure 2.18: Comparaison entre les anomalies de températures globales de surface (°C) observées (en noir) et simulées par plusieurs modèles climatiques couplés globaux de 1860 à 2010. Dans le panel du haut, seuls les forçages naturels ont été considérés tandis que dans le panel du bas, les forçages anthropiques et naturels sont pris en compte (Source : IPCC, WP1, Chapitre 10, [Bindoff et al., 2013](#)).

Les simulations historiques démarrent généralement en 1850 à partir de simulations de contrôle. Pour capturer et reproduire les tendances observées depuis 1850 un jeu de forçages complet est nécessaire et inclus les gaz à effet de serre (CO₂, CH₄, N₂O, ozone, ...), l'évolution des changements d'usages des sols, les variations de l'irradiance solaire, les principales éruptions volcaniques, généralement les aérosols, la formation du trou d'ozone stratosphérique, etc. Comme le montre la figure 2.18, l'évolution de la température de surface en moyenne globale est correctement simulée par l'ensemble des modèles CMIP5. Dans le cas d'expériences de détections/attributions décrites dans la section précédente,

certains de ces forçages sont volontairement omis afin de révéler leur contribution sur les tendances observées (Fig. 2.18).

Les principaux résultats de ce type d'approche décrits dans le dernier rapport du GIEC (2013) montrent que le signal de réchauffement en fin de 20^{ème} siècle, ne peut être simulé par aucun modèle lorsque l'augmentation des concentrations des GES n'est pas prise en compte (Fig. 2.18, en haut). Ainsi, malgré la diversité des modèles, il s'agit là d'une indication (parmi d'autres) suggérant que les émissions des GES d'origine anthropique ont très probablement contribué à l'augmentation de la température globale d'environ 0,7°C depuis 1900.

Une perspective à plus long terme apportée par le climat du dernier millénaire montre le rôle important joué par les forçages externes dans la variabilité climatique à l'échelle hémisphérique, même à l'époque préindustrielle. La compilation de l'ensemble des données couvrant le dernier millénaire révèle que l'Hémisphère Nord a connu des conditions climatiques relativement chaudes jusqu'au 13^{ème} siècle environ, « l'optimum médiéval », période suivie d'un « petit âge glaciaire » (du milieu du 16^{ème} siècle au milieu du 19^{ème} siècle environ). Bien que la variabilité interne, avec sa capacité à déplacer la chaleur au sein du système climatique, soit importante à l'échelle hémisphérique, il est très peu probable que les températures reconstruites depuis l'an 850 puissent être expliquées uniquement par la variabilité interne.

Les simulations de modèles climatiques qui n'incluent que les forçages naturels peuvent expliquer une partie de la variabilité inter-décennale pré-industrielle de la température depuis le début du millénaire à l'échelle hémisphérique, dominée pour l'essentiel par le forçage volcanique (Schurer et al., 2014). Cependant, de telles simulations n'expliquent pas un réchauffement plus récent depuis 1950 sans l'inclusion d'augmentations anthropiques des concentrations de gaz à effet de serre. Le réchauffement en moyenne hémisphérique depuis 1950 est bien en dehors de la fourchette de tendances de longueur similaires estimées en variabilité résiduelle à partir des reconstructions du dernier millénaire.

Projections climatiques:

Les projections faites pour la fin du 21^{ème} siècle visent à estimer la réponse du climat à l'augmentation de la concentration des gaz à effet de serre dans l'atmosphère. Un ensemble de scénarios possibles appelés RCP (pour Representative Concentration Pathway) dans le cadre de CMIP5 (Van Vuuren et al., 2011) et SSP (Shared Socioeconomic Pathways) pour CMIP6 (O'Neill et al., 2016), ont été défini principalement en fonction des évolutions possibles des émissions de gaz à effet de serre. Les projections réalisées sur les températures de surface par un ensemble de modèles pour les différents scénarios CMIP5 sont illustrées dans la Fig. 2.19. Alors que le scénario rcp8.5 prévoit une augmentation significative des émissions de gaz à effets de serre avec un forçage radiatif allant jusqu'à 8.5 W/m² d'ici 2100, le scénario rcp2.6 suppose une diminution rapide des émissions de gaz à effets de serre après 2015 (Riahi et al., 2007). Il est important de noter que ces

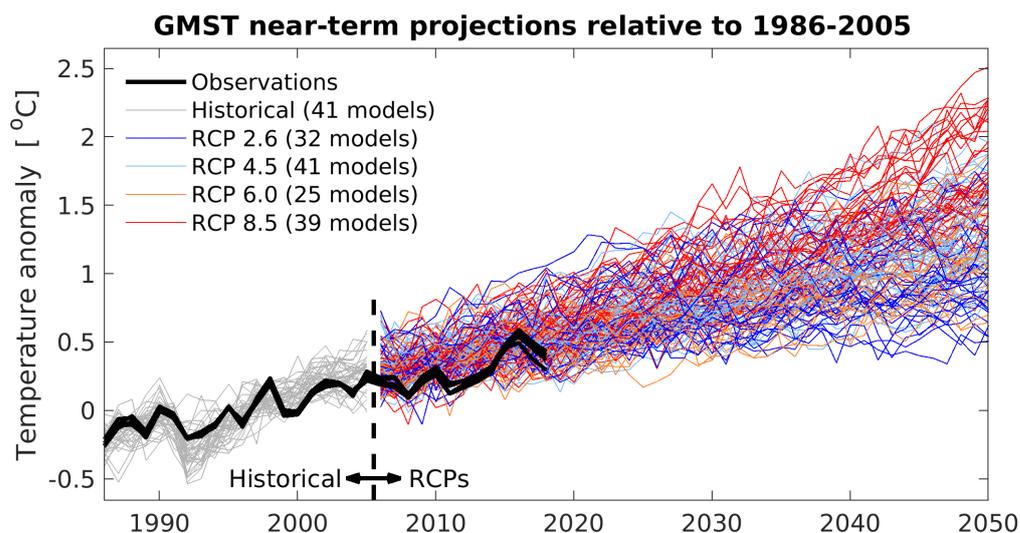


Figure 2.19: Version mise à jour de la Figure 11.25a du rapport de l’IPCC AR5, montrant les observations et les projections futures des modèles CMIP5 par rapport à la période de référence 1986-2005. Les courbes noires représentent les jeux d’observations (HadCRUT4.6, NASA GISTEMP, NOAA GlobalTemp, BEST)

scénarios ne concernent que l’évolution des émissions ou concentrations de gaz à effet de serre et d’aérosols dues aux activités humaines et que ceux-ci ne prennent donc pas en compte l’évolution des autres types de forçages externes. Aujourd’hui les scénarios SSPs dans le cadre de CMIP6 prennent en compte d’autres paramètres comme par exemple l’évolution de l’usage des sols selon divers scénarios économiques (O’Neill et al., 2016).

Ce type d’exercice est réalisé avec les modèles en date par les grands groupes de modélisations pour évaluer l’impact des émissions futures des gaz à effet de serre et autres forçages anthropiques sur le climat. Le très récent rapport spécial du GIEC (SR1.5) sorti en 2018 indique que la teneur en CO₂ dans l’atmosphère aujourd’hui de 400 ppm a déjà modifié le climat et que l’évaluation de scénarios climatiques à bas niveau d’émissions limitant le réchauffement global futur sous la barre des 2°C suggère des impacts régionaux catastrophiques d’ici 2050.

Prévisions décennales:

La prévision décennale est une science jeune, et alors que dans les projections futures seule la réponse aux forçages anthropiques est considérée, dans le cadre de prévisions décennales toutes les sources de variabilités confondues doivent être prises en compte. Les modèles ne sont alors pas initialisés à partir de simulations de contrôle, mais de l’état observé du climat. Bien que de nombreux éléments tels que les éruptions volcaniques soient imprévisibles, il est possible à partir de l’état actuel du climat et de scénarios disponibles sur les forçages externes (naturels et anthropiques) de se prêter à ce type d’expériences (Meehl et al., 2009b; Taylor et al., 2012). Ces exercices effectués à partir de simulations initialisées ou non par les observations, permettent d’évaluer la capacité des modèles à reproduire des périodes passées et de prévoir des évolutions probables du climat futur à court terme, entre 2016 et 2035 pour CMIP5 (Oldenborgh et al., 2012; Kim

et al., 2012; Kirtman et al., 2013; Newman, 2013). Utilisé comme moyen de comparaison des modèles durant l'exercice CMIP5, les résultats indiquent que l'initialisation à partir d'observation est la principale source de prévisibilité de la température en moyenne annuelle pour les premières années, en particulier dans l'Atlantique Nord, les régions du Pacifique Sud et l'océan Indien tropical. Au-delà des premières années, la prévisibilité obtenue pour les moyennes annuelles et pluriannuelles de température et de précipitations est principalement due aux forçages radiatifs externes. A l'heure actuelle les scores obtenus pour les prédictions des anomalies de température de surface par des modèles linéaires simplifiés tels que le LIM sont comparables aux prédictions obtenues avec les modèles CMIP5 (Hawkins and Sutton, 2009; Newman, 2013).

2.6 Convergence entre modèles et observations

Pour optimiser et faciliter l'exploitation scientifique des observations, les méthodes d'assimilation de données dans les GCMs, bien que très difficile à mettre en œuvre et très coûteuses en temps de calcul, offrent l'avantage de constituer une puissante méthode de validation des modèles et d'identification des mécanismes clés à la base des changements climatiques observés ainsi que de leur prévisibilité. Le principe de ces méthodes est de reconstruire le climat et sa variabilité en combinant de façon optimale les informations provenant à la fois des modèles et des observations tout en tenant compte de leurs incertitudes (Talagrand, 1997; Fig. 2.20). Je reviendrais dans le Chapitre 4 sur les principales méthodes d'assimilation de données utilisées à l'heure actuelle et leurs bases théoriques. La qualité de la reconstruction obtenue dépend de la méthode appliquée, des biais des modèles utilisés mais également des observations disponibles. Le type d'observations disponibles, leur quantité et leur qualité dépendent néanmoins de la période considérée ce qui conditionne le type de méthodes d'assimilation pouvant être mis en œuvre (Fig. 2.20).

Limites et sources d'incertitudes sur la période historique:

Depuis 1900 environ, on dispose sur les continents d'une variété d'observations instrumentales (thermomètres, pluviomètres, etc.). Jusqu'en 1930, les observations des températures de surface océaniques se limitaient principalement aux zones de navigations laissant certaines zones très mal ou pas observées comme par exemple dans le Pacifique Sud (Folland and Parker, 1995; McKittrick, 2010). Ce n'est qu'à partir de 1978 et l'utilisation de données satellites que la couverture spatiale sur les continents et les océans, est devenue quasi complète. Pour étudier des champs pas ou peu observés, il est courant d'utiliser ce que l'on appelle des réanalyses combinant la physique décrite par les GCMs avec les observations disponibles. Depuis les premières réanalyses proposées par (Trenberth and Olson, 1988) et (Bengtsson and Shukla, 1988), d'énormes progrès ont été réalisés pour mettre en place des méthodes d'assimilation de données intégrant des modèles climatiques de plus en plus réalistes et performants afin d'obtenir des reconstructions spatiales cohérentes et homogènes. Les méthodes d'assimilation de données utilisées pour les réanalyses du 20^{ème} siècle sont généralement des méthodes d'assimilation variationnelles du type

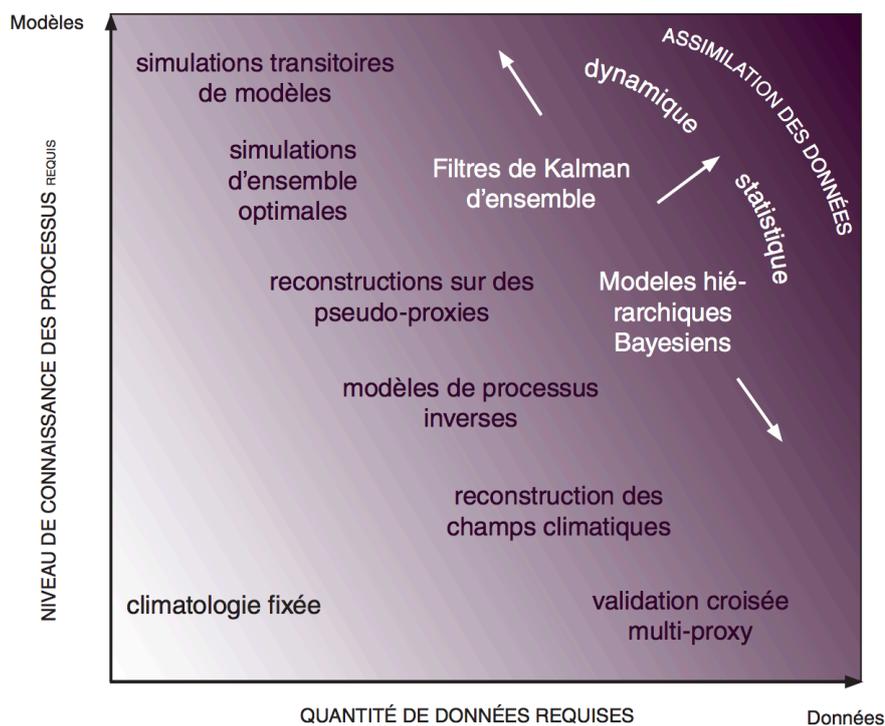


Figure 2.20: Illustration schématique de l'association modèles et observations en fonction de la connaissance des processus physiques qui gouvernent notre climat en ordonnées et de la quantité de données requise en abscisse (source: [Hughes et al., 2010](#))

3DVAR et 4DVAR comme par exemple pour NCEP/NCAR et ERA40 ([Kistler et al., 2001](#); [Uppala et al., 2005](#)) ou bien des méthodes du type filtre de Kalman d'ensemble comme par exemple pour 20CR ([Compo et al., 2011](#)). L'évolution des techniques de mesures, l'introduction de nouvelles sources de données, la qualité (discutable avant 1950) et la faible couverture spatiale au début de 20^{ème} siècle constituent une combinaison de facteurs sources de discontinuités et d'incohérences entre les jeux de réanalyses, ces incertitudes pouvant être source de tendances artificielles ([Kent et al., 2007](#); [Thompson et al., 2008](#); [Vecchi et al., 2008](#); [Woodruff et al., 2008](#)).

Limites et sources d'incertitudes en paléoclimat:

A la différence des observations instrumentales utilisées dans le cadre de réanalyses, les proxys utilisés dans le cadre de reconstructions paléoclimatiques sont plus rares, bien moins précises et possèdent des échelles de temps caractéristiques de l'ordre de l'année. Bien que plus compliqués à utiliser, la disponibilité croissante de ce type d'observations a donné lieu à de nombreuses études s'intéressant aux possibilités offertes par ces observations et à la manière d'intégrer celles-ci dans des simulations millénaires via l'utilisation de méthodes d'assimilation de données (par exemple [Goosse et al., 2010](#)). Les méthodes d'assimilation de données variationnelles du type 3DVAR et 4DVAR généralement appliquées dans le cadre de réanalyses du 20^{ème} siècle ne pouvant convenir dans ces circonstances ([Bhend et al., 2012](#); [Fang and Li, 2016](#)), des approches du type filtres de Kalman d'ensemble et filtres particuliers sont généralement utilisées en paléoclimat (par

exemple [Liu et al., 2017](#)). Malgré les avancées importantes réalisées ces dernières décennies ([Hakim et al., 2013](#)), ces méthodes restent confrontées dans le cadre de reconstructions de champs spatiaux à certaines limitations liées à l'utilisation de modèles climatiques coûteux sur de longue période et à la réalisation d'un nombre de simulations insuffisant pour pallier au fameux problème de dimension (ou en anglais « curse of dimensionality »). Les méthodes existantes ont donc dû être adaptées et des solutions pratiques développées. Ces adaptations ont par exemple donné lieu à des approches dites « offline » plus simples à implémenter mais ne garantissant pas nécessairement la cohérence physique des reconstructions obtenues dans le temps ([Matsikaris et al., 2015](#)). Une description plus complète de ces adaptations, des approches utilisées et des difficultés rencontrés lors de leurs applications est donnée dans le chapitre 4.

Partie II

Étude de la variabilité climatique sur la période historique

Chapitre 3

Tendance et variabilité climatique dans l'océan Pacifique au 20^{ème} siècle

Sommaire

3.1	Préambule	46
3.2	Évaluation de la variabilité interne aux échelles de temps décennales .	46
3.3	Rôle des forçages externes sur la variabilité et tendances observées et simulées	50
3.3.1	Impacts des forçages externes sur l'ENSO et l'IPO	50
3.3.2	Impacts des forçages externes sur les tendances à long terme . . .	52
3.3.3	Tendances robustes sur la période historique : variabilité interne ou signature des forçages anthropiques ?	55
3.4	Article : Contributions of internal variability and external forcing to the recent trends in Southeastern Pacific and Peru-Chile upwelling system	58
3.5	Expressions océaniques de l'expansion de la cellule de Hadley	82

3.1 Préambule

L’objectif de ce chapitre est d’étudier la variabilité décennale et les tendances au 20^{ème} siècle dans l’Océan Pacifique. Une grande partie de ce que l’on sait est basée sur la comparaison des observations aux résultats des modèles climatiques et notre compréhension physique des phénomènes robustes identifiés. Les forçages externes n’étant pas les uniques sources de la variabilité climatique au 20^{ème} siècle, nous évaluerons et discuterons dans un premier temps la capacité des modèles climatiques à reproduire la variabilité décennale interne à l’échelle du bassin Pacifique.

Démêler le rôle relatif des forçages externes sur le climat moyen, les modes décennaux et les tendances dans l’Océan Pacifique représente un défi important en raison des échelles de temps similaires associées à ces deux types de variabilités, de la fenêtre d’observation relativement courte et des biais des modèles. Nous verrons dans un second temps, que des études cohérentes de la littérature récente, prenant en compte l’ensemble de ces facteurs et sources d’incertitudes permettent d’estimer la contribution de certains forçages dans la variabilité récente.

Des grands ensembles de simulations climatiques peuvent aussi être réalisés avec les forçages externes pris collectivement ou individuellement, permettant ainsi d’attribuer certains des phénomènes observés à des sources spécifiques. Une étude dédiée s’appuyant sur cette approche avec le modèle IPSL-CM5A-LR est présentée sous la forme d’une publication à la fin de ce chapitre. Cette étude vise à comprendre les tendances en SSTs dans le Pacifique Sud-Est depuis la fin des années 70s, le signal le plus robuste à l’échelle du bassin dans les modèles et les observations. Une discussion clos ce chapitre et permet de mettre en perspective les résultats obtenus.

3.2 Évaluation de la variabilité interne aux échelles de temps décennales

Bien que peu d’observations instrumentales suffisamment longues et homogènes soient disponibles, nous tentons dans cette section d’évaluer la capacité des modèles ayant participé au 5^{ème} exercice CMIP à caractériser correctement la variabilité décennale interne observée au 20^{ème} siècle dans l’Océan Pacifique. Cette étape est essentielle pour comprendre le rôle de la variabilité interne sur la période récente et anticiper les évolutions futures (Cassou et al., 2018; Kushnir et al., 2019). Les observations, réanalyses et simulations historiques des modèles CMIP5 que j’ai utilisé dans cette évaluation sont listés en Annexe (Tableaux A.1, A.2, A.3). Les résolutions spatiales sont très variables en particulier pour la grille atmosphérique. Au delà de la question de la résolution spatiale qui permet de discrétiser l’ensemble du globe de façon plus ou moins fine, la complexité des processus représentés varie d’un modèle à l’autre et explique les différences de comportement. Pour s’affranchir des variabilités interannuelles et tendances liées aux forçages externes, un filtre de Hanning de 13 ans (passe bas) et de 40 ans (passe haut) sont appliqués aux anomalies (par rapport à la climatologie de la période 1900-1930) des SSTs du Pacifique

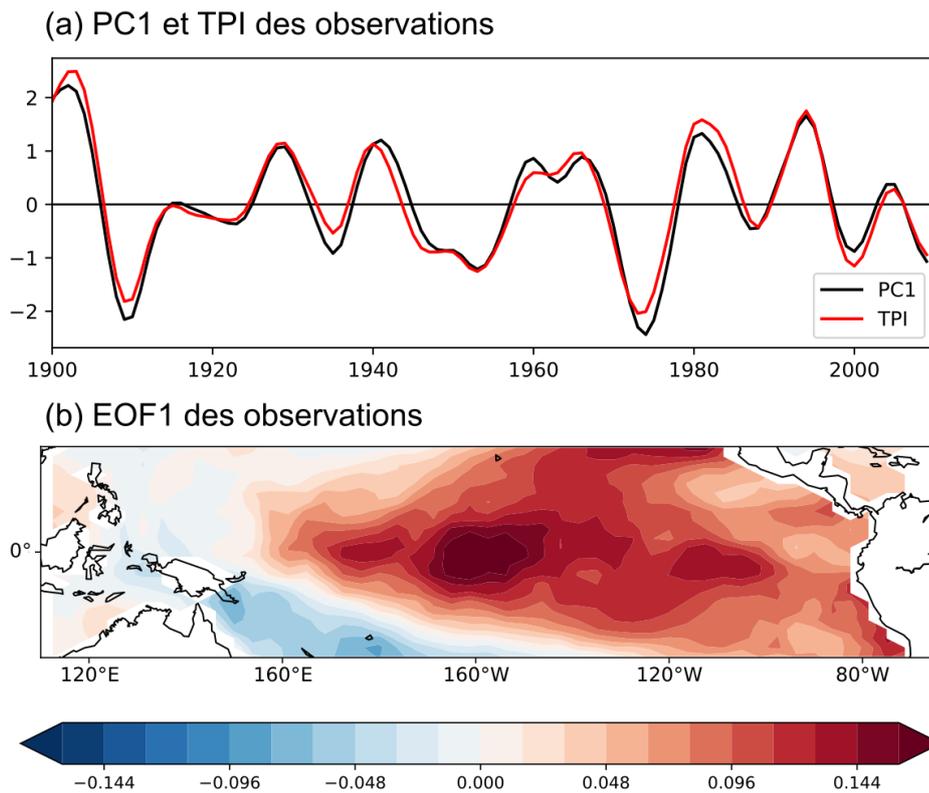


Figure 3.1: (a) Séries temporelles sur la période 1900-2009 des indices standardisés de la PC1 (en noir) et du TPI (en rouge); (b) EOF associée à la PC1 calculée à partir des observations HadISSTv1.1 filtrées (fenêtre 13-40 ans).

entre 30°S et 30°N. Une analyse en composante principale permet ensuite d'identifier le principal mode interne de la variabilité opérant aux échelles de temps décennales. La même analyse appliquée aux SSTs observées et simulées par chaque modèle CMIP5 sur la période commune 1900-2009 permet d'identifier la signature spatiale typique de l'IPO (Fig. 3.1). La PC1 associée est très corrélée ($r=0.97$) à l'indice tripolaire ou TPI (Henley et al., 2015) sur la période 1900-2009.

La figure 3.2 nous montre que chacun des 25 modèles CMIP5 simule sa propre variabilité décennale dans les tropiques avec des anomalies de SSTs plus ou moins marquées dans le Pacifique équatorial et qui s'étendent plus ou moins loin vers l'Ouest.

A l'échelle du Pacifique, la projection de l'indice TPI sur les SSTs observées montrent des anomalies tropicales positives qui s'étendent aux extra-tropiques au niveau des bords Est, avec une inversion du signe à l'Ouest de 160°E dans les tropiques et aux extra tropiques Sud-Ouest et Nord-Ouest (Fig. 3.3a). Les changements associés de la hauteur de la surface de la mer (ou SSH pour en anglais Sea Surface Height) et les vents de surface permettent également de caractériser les anomalies typiques liées à l'IPO (Deser et al., 2010). Les résultats obtenus illustrés sur les Figure 3.3b-c par la moyenne d'ensemble et ceux d'un seul modèle (IPSL-CM5A-LR) CMIP5, indiquent à première vue une assez bonne correspondance avec les observations. Toutefois, malgré les progrès réalisés dans la reproduction de l'IPO par les modèles de classe CMIP (Lyu et al., 2016), la majorité d'entre

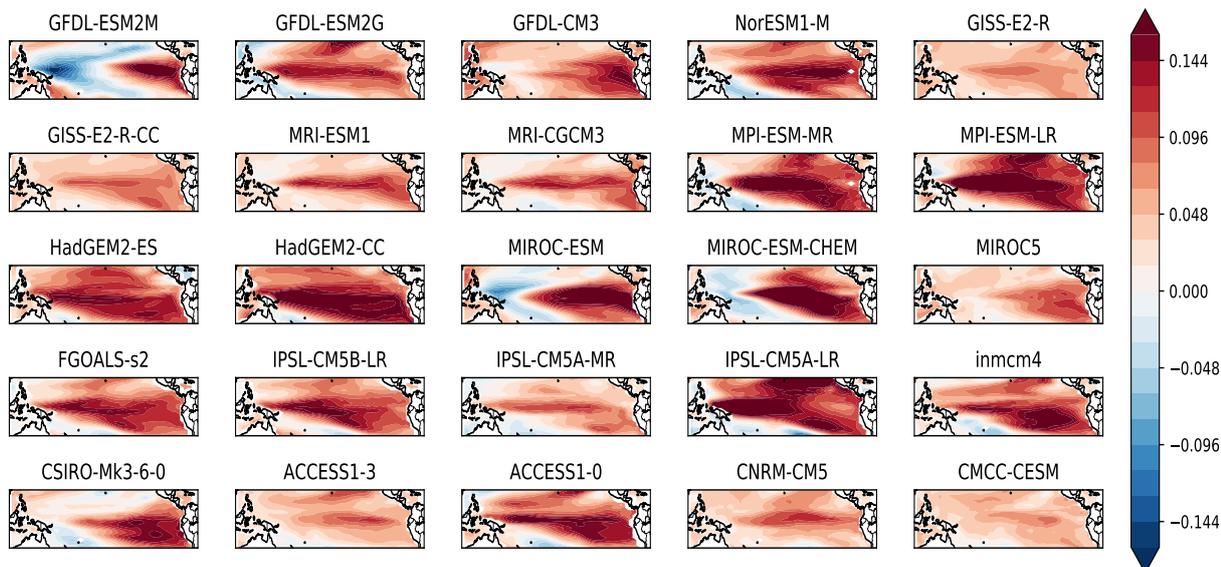


Figure 3.2: Première EOF des SSTs filtrées (fenêtre 13-40 ans) des modèles CMIP5 dans l'Océan Pacifique entre 30°S et 30°N sur la période 1900-2009.

eux continuent à simuler, en relation avec le biais de la « cold tongue », des anomalies s'étendant trop loin à l'Ouest dans la zone tropicale. L'amplitude des anomalies de SST, des vents de surface et de la SSH dans le Pacifique tropical-Est et le Pacifique Nord est sous-estimée tandis que dans le Pacifique Sud, les anomalies sont très mal reproduites.

Une majorité de modèles sous-estime également la variance associée à la variabilité décennale Pacifique (Fig. 3.4). Environ 68% ont une variance inférieure à celle des observations sur la période 1900-2009. Ce pourcentage s'élève à 76% pour la période 1950-2009. L'analyse de la variance associée aux 2 régions extratropicales et la région tropicale utilisées pour définir l'indice TPI montre que ce biais des modèles est encore plus marqué dans l'Hémisphère Sud où 96% des modèles CMIP5 sous-estiment la variance associée à la composante extratropicale sud de l'IPO pour les périodes 1900-2009 et 1950-2009 (Fig. 3.4).

L'analyse fréquentielle par ondelette telle que décrite dans (Torrence and Compo, 1998; Liu et al., 2007) appliquée aux indices TPI calculés sur différentes fenêtres de variabilité décennale, montre que la moyenne et la médiane des modèles sous-estiment la variance globalement et en particulier ne parviennent pas à reproduire le pic d'amplitude entre 12.5-13 ans présent dans les observations (Fig. 3.5). Les modèles tendent plutôt à simuler des pics d'amplitudes à 18.5 ans dans la fenêtre 13-40 ans et d'environ 9-10 ans dans les fenêtres 8-20 et 8-40. Le modèle IPSL-CM5A-LR simule pour sa part deux pics aux périodes 6.9 ans et 19.7 ans.

Les difficultés prononcées des modèles climatiques à reproduire le schéma spatial et le contenu spectral de la variabilité décennale, particulièrement marquée dans le Pacifique Sud, suggèrent un rôle important des interactions air-mer et des téléconnexions tropicales extra tropiques dans l'établissement du schéma inter-hémisphérique de l'IPO. Plusieurs théories ont été avancées et dans lesquelles l'atmosphère tropicale serait forcée par des

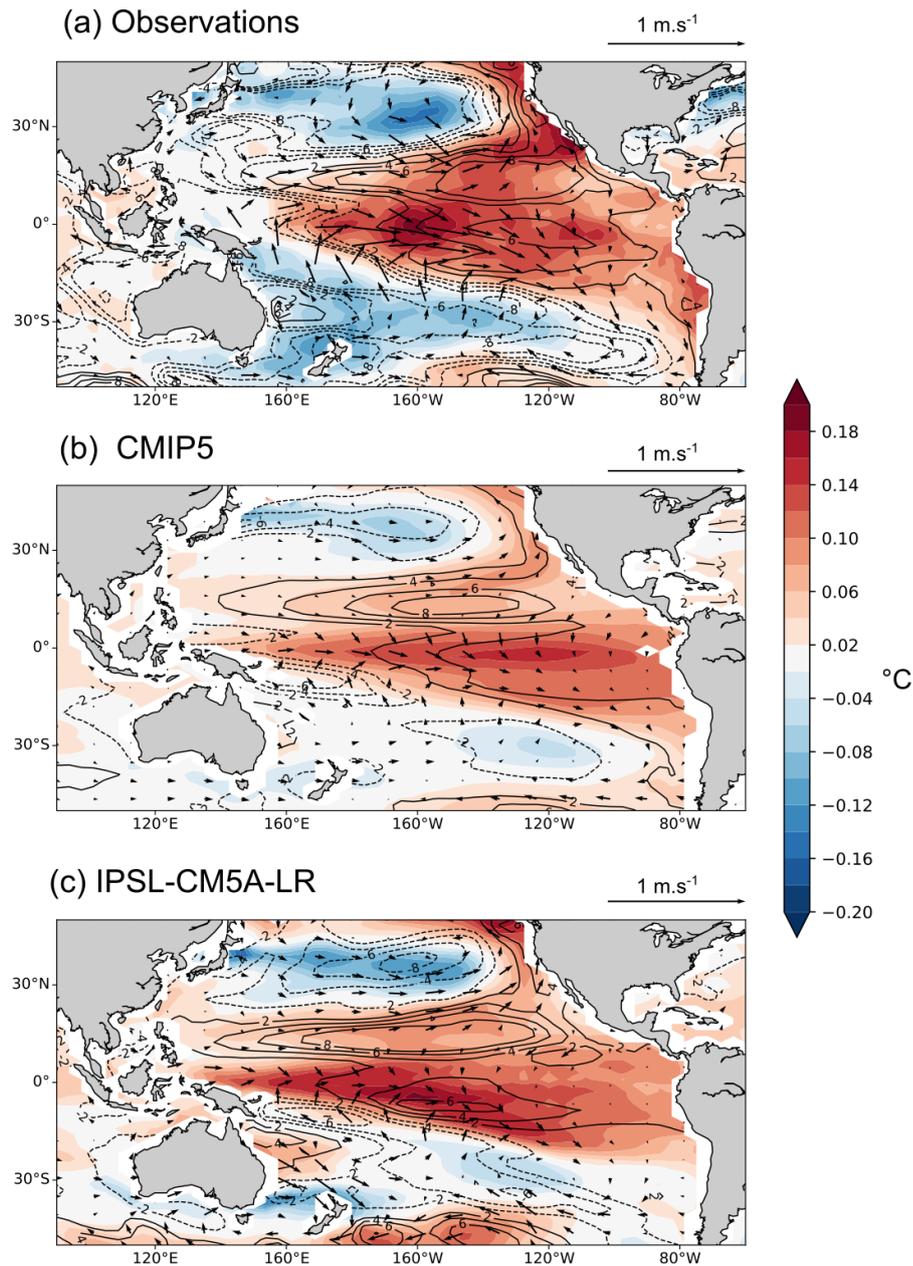


Figure 3.3: Régression de la SST (en couleur ; °C), des vents de surface (flèches ; m.s^{-1}) et de la SSH (contours noirs ; mm) sur l'indice du TPI standardisé sur la période 1900-2009 pour (a) les observations (HadISSTv1.1 pour la SST, Réanalyses NOAA-20thCv2c pour les vents de surface et ORA-20C pour la SSH), (b) la moyenne d'ensemble des modèles CMIP5 et (c) le model IPSL-CM5A-LR. Toutes les analyses sont effectuées sur des données filtrées (fenêtre 13-40 ans).

anomalies de SST générées par la variabilité atmosphérique des moyennes latitudes du Pacifique Nord opérant à différentes échelles de temps (Vimont et al., 2003; Farneti et al., 2014; Newman et al., 2016). Certaines études pointent également la contribution non négligeable des interactions avec le Pacifique Sud (Okumura, 2013; Zhang et al., 2014; Ding et al., 2015). Une autre hypothèse soutient également que l'IPO correspondrait en grande partie au résidu de l'asymétrie de l'ENSO intégrée par la couche de mélange océanique (Newman et al., 2003; Power and Colman, 2006).

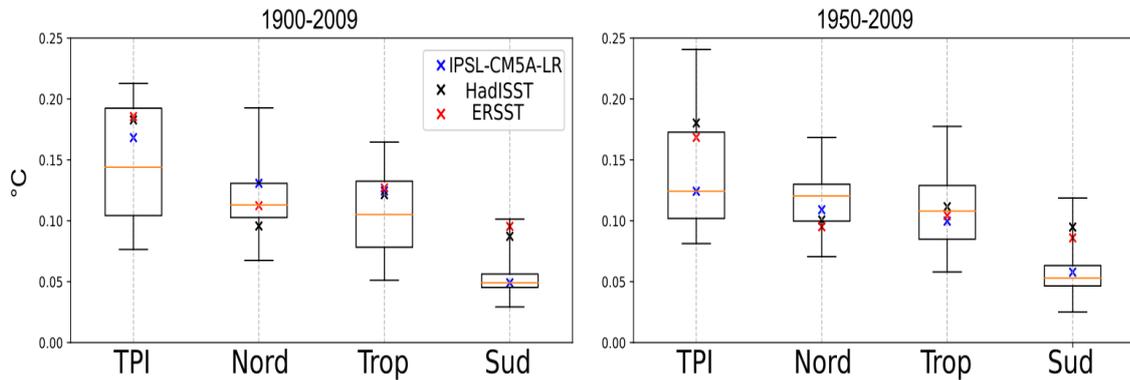


Figure 3.4: Box plots des déviations standards de l'indice TPI (Henley et al., 2015) calculées sur les périodes 1900-2009 (colonne de gauche) et 1950-2009 (colonne de droite) et les 3 régions utilisées dans la définition du TPI (boîtes du Pacifique Nord, Sud et équatoriale indiquées dans la Figure 2.11) sur les SST filtrées (fenêtre 13-40 ans). Les bornes supérieures et inférieures de chaque box plot indiquent le 1^{er} et 3^{ème} quartile, les extrémités des segments indiquent quant à eux les valeurs maximales et minimales et les lignes oranges représentent les médianes de l'ensemble des modèles CMIP5 considérés. Les croix noires indiquent les valeurs obtenues pour les observations HadISSTv1.1, les croix rouges les valeurs obtenues pour ERSSTv4 et les croix bleues les valeurs obtenues pour le modèle IPSL-CM5A-LR.

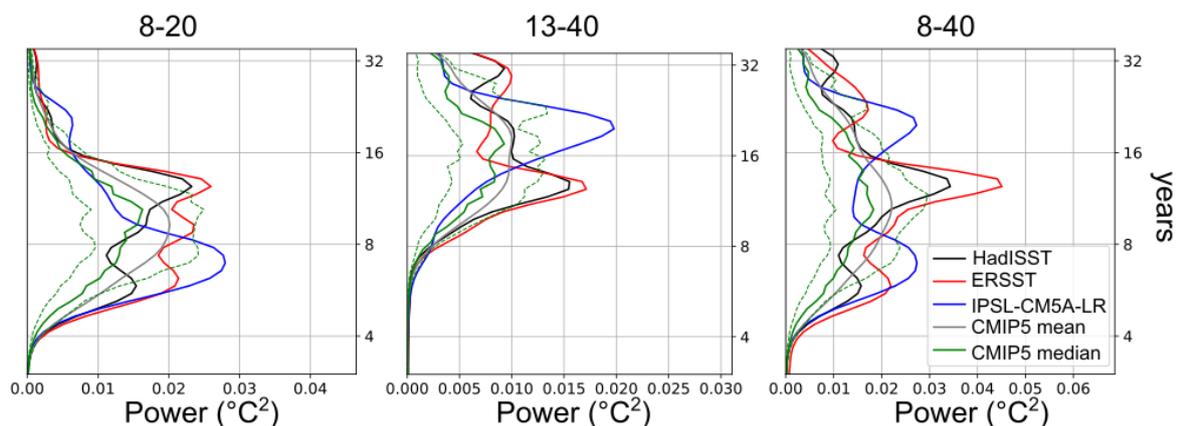


Figure 3.5: Spectre de puissance de l'analyse en ondelettes de l'indice TPI sur la période 1900-2009 calculé à partir des SSTs observées HadISSTv1 (courbe noire) et ERSSTv4 (courbe rouge) et simulées par le modèle IPSL-CM5A-LR (courbe bleue). La moyenne des spectres de puissances des modèles CMIP5 est indiquée en trait plein gris, la médiane est en trait plein vert et les 1^{er} et 3^{ème} quartiles sont en traits pointillés vert. Chaque colonne correspond à des fenêtres de filtres différentes appliquées aux anomalies de SSTs avant le calcul du TPI : fenêtre 8-20 ans dans la colonne de gauche, fenêtre 13-40 ans dans la colonne au centre et fenêtre 8-40 ans dans la colonne de droite.

3.3 Rôle des forçages externes sur la variabilité et tendances observées et simulées

3.3.1 Impacts des forçages externes sur l'ENSO et l'IPO

En plus de la variabilité résultant des interactions internes au climat (section précédente), le forçage radiatif externe, peut également moduler les cycles naturels des principaux modes de variabilité de l'océan Pacifique. Par exemple, parmi les forçages naturels, les variations décennales de l'irradiance solaire modulent de $\sim 0,1\%$ l'énergie solaire

reçue par la Terre avec une amplitude plus importante aux latitudes tropicales. Les mécanismes proposés liant les variations de l'irradiance solaire et le Pacifique tropical invoque des processus amplificateurs, dit de « *top down* » impliquant la chimie de l'ozone stratosphérique et son effet sur la propagation des ondes planétaires vers la surface, et « *bottom up* » impliquant le couplage air-mer dans l'Est du Pacifique tropical (Meehl et al., 2008; Meehl et al., 2009a; Lockwood et al., 2010). Toutefois, la détection de l'empreinte d'un forçage aussi faible ($\sim 0.25 \text{ W.m}^{-2}$ en surface) est difficile dans les observations et les modèles, car des fluctuations quasi décennales peuvent être intrinsèquement générées par le système couplé océan-atmosphère (Misios et al., 2016). Des recherches récentes ont également montrées que le forçage associé au volcanisme stratosphérique peut, en plus du refroidissement de surface global (Robock, 2000), influencer les processus de la variabilité interne dans le Pacifique tropical, principalement le cycle naturel d'ENSO (par exemple McGregor and Timmermann, 2011; Maher et al., 2014). Les observations instrumentales, les proxy d'ENSO et les modèles climatiques CMIP5 suggèrent en effet une probabilité accrue qu'un El Niño se produise dans les deux ans suivant une éruption stratosphérique (par exemple Khodri et al., 2017).

Les GES tels que la teneur en CO_2 atmosphérique ont augmenté régulièrement induisant un forçage radiatif positif relativement uniforme spatialement (GIEC AR5). Bien que certaines incertitudes persistent sur les variations d'intensité de l'ENSO en réponse aux forçages anthropiques (Vecchi and Wittenberg, 2010; Van Oldenborgh et al., 2005), de nombreuses études suggèrent néanmoins que le forçage anthropique tendrait à augmenter l'amplitude d'ENSO via un réchauffement plus rapide de l'Est du Pacifique équatorial par rapport au Pacifique Ouest (Xie et al., 2010; Power et al., 2013; Cai et al., 2015; Power et al., 2017; Liguori and Di Lorenzo, 2018; Zheng et al., 2018). Ces mêmes modèles suggèrent une augmentation de la variabilité décennale des précipitations dans certaines parties du Pacifique tropical (Pendergrass et al., 2017; Power et al., 2017). Certaines études mettent également en avant la composante non linéaire de cette réponse avec une intensification/affaiblissement des variations d'intensité de l'ENSO qui serait liée aux différents niveaux de concentrations de CO_2 dans l'atmosphère (Cherchi et al., 2008).

Contrairement aux concentrations de GES, les aérosols troposphériques anthropiques agissent généralement pour refroidir la température de surface globale (Forster et al., 2007a; Evan et al., 2009) et présentent de grandes variations temporelles et spatiales (Boucher et al., 2013; Undorf et al., 2018). Malgré les différences d'uniformité des forçages anthropiques des GES et des aérosols troposphériques, l'analyse des résultats des modèles CMIP5 suggèrent une signature spatiale très semblable sur la SST et les précipitations (Xie et al., 2013). Une étude récente a également montré que le déplacement des sources d'émissions des aérosols anthropiques depuis l'Amérique du Nord et l'Europe vers la Chine au début des années 80s, se projette sur une tendance des SSTs dans le Pacifique comparable à la transition récente observée de l'IPO vers une phase négative (Smith et al., 2016). Takahashi and Watanabe, 2016 ont suggéré qu'une partie de cette transition et l'accélération récente des alizés dans le Pacifique équatorial pourraient être dues aux

aérosols sulfatés, y compris de sources volcaniques. Néanmoins la complexité des processus et des interactions aérosols-nuages qui en découlent sont sources d'incertitudes significatives entre les différents modèles existants et questionnent fortement le rôle potentiel des aérosols anthropiques dans les tendances récentes (Boucher et al., 2013; Oudar et al., 2018).

3.3.2 Impacts des forçages externes sur les tendances à long terme

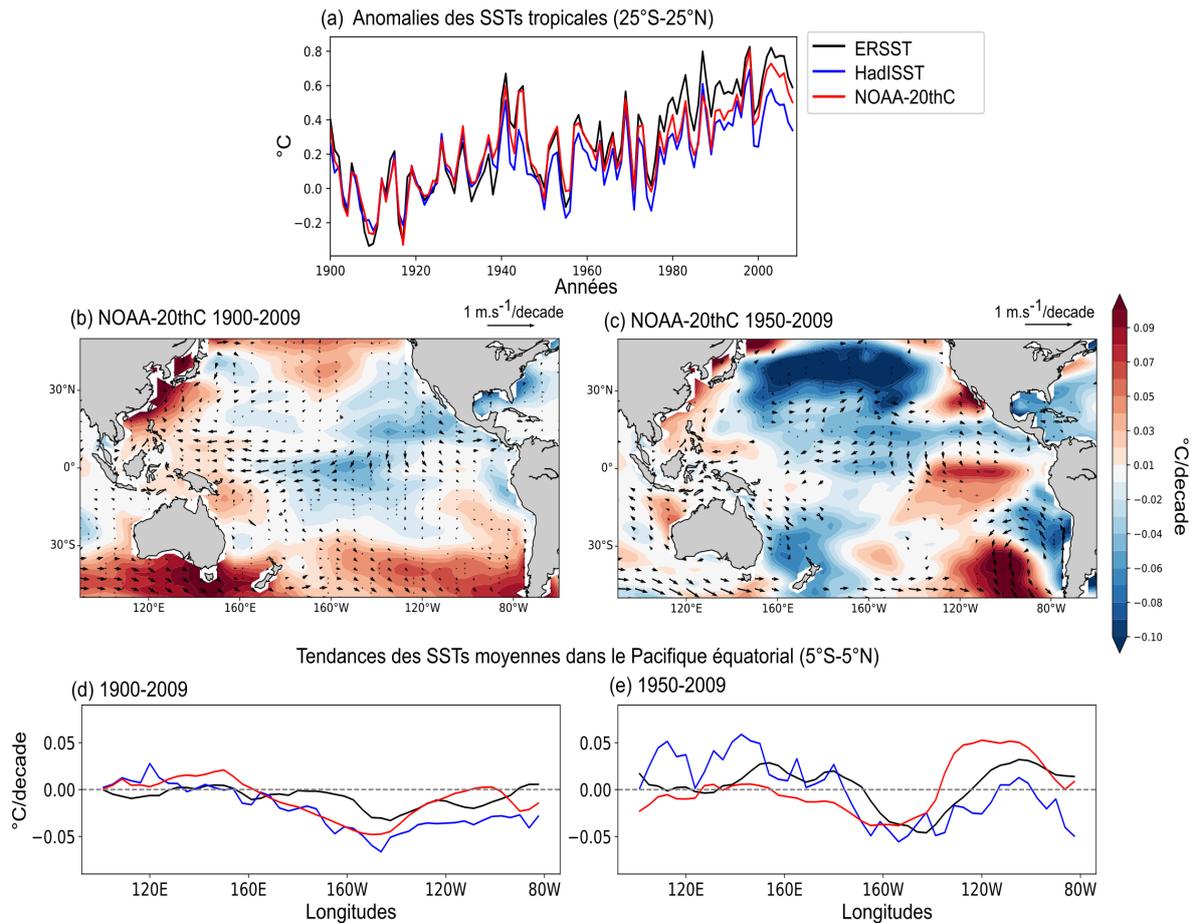


Figure 3.6: Tendances des conditions de surface observées dans l'Océan Pacifique durant la période 1900-2009 par rapport à la période de référence 1900-1930. (a) Séries temporelles des anomalies annuelles de SST tropicales (25°S-25°N) pour les observations ERSSTv4 (courbe noire), HadISSTv1.1 (courbe bleue) et NOAA-20thC (courbe rouge). (b,c) Tendances spatiales durant la période 1900-2009 (colonne de gauche) et 1950-2009 (colonne de droite) des anomalies de vents à 10m ($m.s^{-1}$; vecteurs) et des SSTs relatives par rapport à la moyenne tropicale ($^{\circ}C$; en couleur) pour les réanalyses NOAA-20thC. (d-e) Tendances zonales des SSTs relatives dans le Pacifique équatorial (entre 5°S et 5°N) pour les périodes (d) 1900-2009 et (e) 1950-2009.

L'analyse des tendances relatives des SSTs observées dans tout le Pacifique sur la période 1900-2009 (Fig. 3.6b) par rapport au réchauffement moyen de la zone tropicale (Fig. 3.6a) révèle un refroidissement relatif au centre du Pacifique équatorial robuste dans les observations HadISST1.1, ERSSTv4 et NOAA-20Cv2c (Fig. 3.6d). Ces analyses sont en accord avec les travaux pionniers de Vecchi et al., 2008. Sur la période mieux observée après 1950, les tendances semblent être à l'inverse indiquer un réchauffement

relatif dans l'Est du Pacifique équatorial (Fig. 3.6c). La divergence persistante entre les jeux d'observations dans cette région sur la période 1950-2009 (Fig 3.6e) complique néanmoins la détection de tendances claires en SSTs (Deser et al., 2010). Une intensification des alizés dans le Pacifique tropical cohérente avec les tendances de la SST relative de NOAA-20C, est également présente dans d'autres réanalyses comme ERA-20C (Li et al., 2019). Ces observations suggèrent un renforcement de la circulation de Walker bien que la fenêtre d'observation 1950-2009 soit trop courte pour s'affranchir efficacement de l'influence des modes internes décennaux comme l'IPO (Tableau 3.1).

Table 3.1: Tendances de l'indice de l'IPO observé (HadiSSTv1) en unité de déviation standard par décennie pour différentes fenêtres de temps.

	1900-2009	1950-2009	1980-2009
IPO	-0.1	0.21	-0.65

Comme pour les observations nous avons analysé des conditions de surface dans le Pacifique pour l'ensemble des 25 simulations CMIP5 soumises aux mêmes forçages externes naturels et anthropiques de la période historique (Taylor et al., 2012). Chaque simulation échantillonne des phases aléatoires de la variabilité climatique interne ce qui garantit et assure la non co-variabilité de celle-ci entre les différents membres. Estimé à partir d'un nombre important de membre, la variabilité de la moyenne d'ensemble, si celle-ci est significative, peut alors être considérée comme la réponse associée aux forçages externes. En outre l'enveloppe des réponses individuelles est représentative à la fois de la contribution de la variabilité interne et de l'incertitude associée à la réponse aux forçages.

L'analyse des tendances obtenues pour les 25 modèles CMIP5 et celles du modèle de l'IPSL-CM5-LR sur la période 1900-2009 est présentée dans la Figure 3.7. L'augmentation des SSTs dans le Pacifique équatorial en réponse au flux radiatif associé majoritairement aux GES n'est pas uniforme. Il n'est pour autant pas en accord avec le schéma spatial des tendances observées qui indiquent un refroidissement équatorial (Fig. 3.6).

Les modèles CMIP5 indiquent un réchauffement accru au centre et à l'Est du Pacifique équatorial, probablement en raison d'un refroidissement lié à l'évaporation plus faible en réponse à l'affaiblissement de la WC et des alizés à l'équateur (Xie et al., 2010). L'augmentation plus significative des précipitation dans les projection CMIP5 à l'ouest qu'au centre et à l'Est du Pacifique (Held and Soden, 2006; Sohn et al., 2019) semble toutefois contradictoire avec l'affaiblissement simulé des cellules de Walker. Une des raisons possibles avancée par (Sohn et al., 2019) serait lié à une intensification régionale de la circulation de Hadley. Ces variations zonales de la circulation méridionale pourraient expliquer l'augmentation des précipitations plus importante à l'Ouest du Pacifique et certainement contribuer à l'affaiblissement du gradient de SST zonal simulé par les modèles CMIP5 dans le Pacifique tropicale (Fig. 3.7b).

Le modèle de l'IPSL-CM5A-LR à la différence de la moyenne multi-modèles simule un renforcement du gradient de température zonale avec un réchauffement plus important

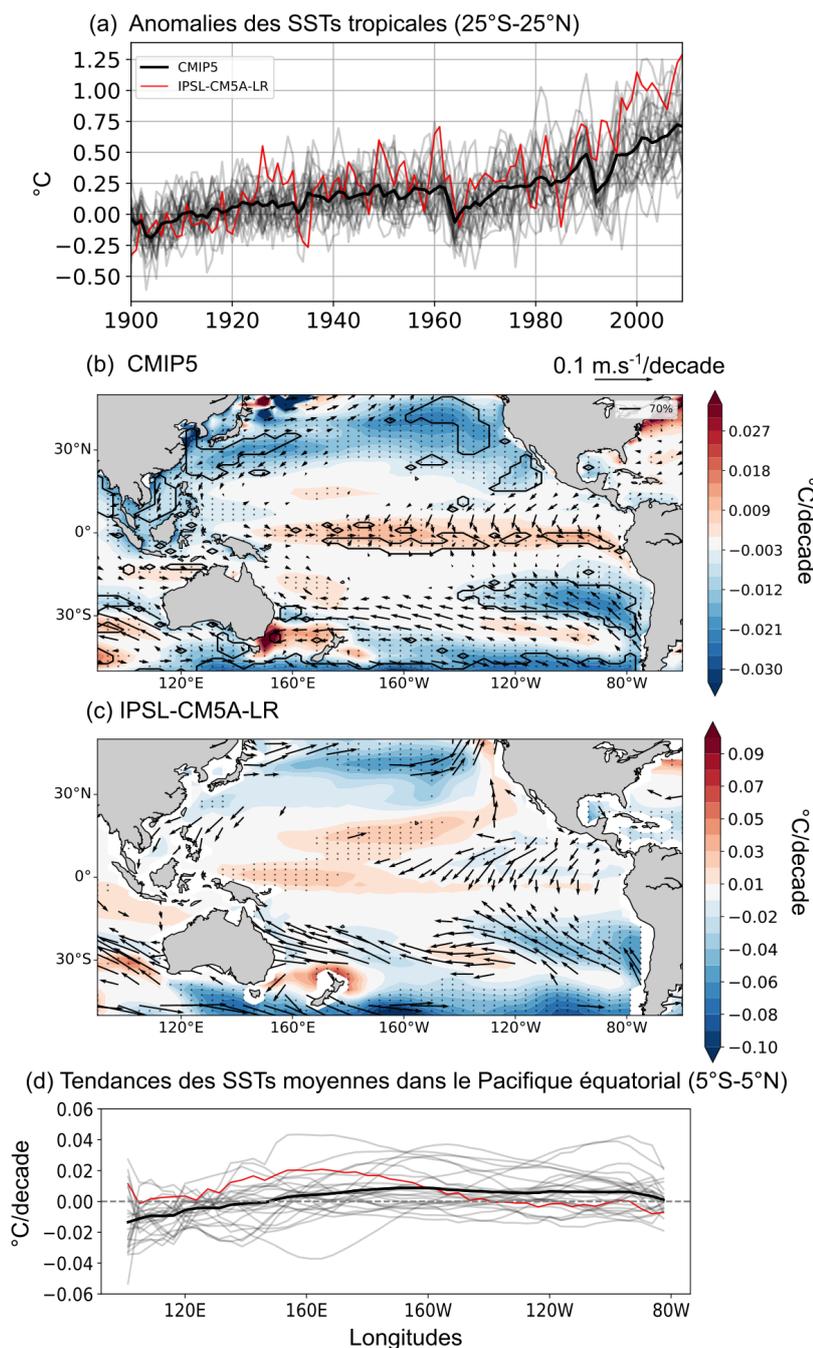


Figure 3.7: Tendances des conditions de surface simulées par les modèles CMIP5 dans l'Océan Pacifique sur la période 1900-2009 par rapport à la période de référence 1900-1930. (a) Séries temporelles des anomalies annuelles de SST tropicales (25°S-25°N) pour chaque modèle CMIP5 (courbes grises), la moyenne multi-modèle (courbe noire en gras) et le modèle IPSL-CM5A-LR (courbe rouge). (b-d) Tendances spatiales des anomalies de vents à 10m (m.s^{-1} ; vecteurs) et des SSTs relatives par rapport à la moyenne tropicale ($^{\circ}\text{C}$; en couleur) pour (b) la moyenne multi-modèles et (c) le modèle IPSL-CM5A-LR. Les zones pointillées désignent les régions où la tendance de la moyenne d'ensemble est significative à 95% et les contours sur le panneau (b) indiquent les régions où plus de 70% des modèles CMIP5 ont des tendances de même signe. (d) Tendances zonales des SSTs relatives dans le Pacifique équatorial (entre 5°S et 5°N) pour chaque modèle CMIP5 (courbes grises), la moyenne multi-modèle (courbe noire en gras) et le modèle IPSL-CM5A-LR (courbe rouge).

dans l'Ouest du Pacifique équatorial. Ce modèle fait également partie des quelques modèles projetant un renforcement de la circulation de Walker cohérent avec l'augmentation du gradient zonal de SST (Fig. 3.7c). Ces différences illustrent le niveau d'incertitude dans les tendances à long terme simulées dans le Pacifique équatorial par les modèles de classe CMIP5. La plupart des modèles climatiques souffrent en effet de nombreux biais dans la reproduction des caractéristiques climatiques moyennes dans cette région. Parmi les biais persistants et les plus importants dans le Pacifique équatorial, on trouve le problème de la double ITCZ (Dufresne et al., 2013) et celui de la « cold tongue » excessive (Dufresne et al., 2013). Une relation directe a été observée entre les biais moyens des modèles et les différences dans leurs projections futures (Li and Xie, 2014). Ces biais limitent l'utilité des modèles climatiques pour la compréhension des phénomènes observés dans le Pacifique équatorial et l'anticipation des changements à venir (Ying et al., 2019).

Dans les moyennes latitudes les tendances des SSTs relatives observées après les années 50s sont caractérisées par un refroidissement relatif important dans le Pacifique Nord, le Pacifique Sud-Ouest (aux alentours de la Nouvelle-Zélande) et le Pacifique Sud-Est le long des côtes du Pérou-Chili (Fig. 3.6c). La moyenne multi-modèles montre également un refroidissement relatif dans le Pacifique Nord et le Pacifique Sud-Est le long des côtes d'Amérique du Sud (Fig. 3.7b). La réponse associée des alizés sur la période 1900-2009 est également très similaire dans la moyenne CMIP5 et le modèle IPSL-CM5A-LR, avec une intensification de part et d'autre de l'équateur. Ces résultats suggèrent que contrairement à la réponse de la WC, l'expansion des cellules de Hadley dont les branches inférieures convergent à l'équateur, est une caractéristique robuste se produisant en réponse aux forçages anthropiques (Staten et al., 2018). Plusieurs ensembles de données d'observation et réanalyses montrent en effet de manière cohérente avec les modèles climatiques que la marge sud de la cellule Hadley de l'Hémisphère Sud s'est déplacée vers le pôle à un rythme d'environ $0,5^\circ$ de latitude par décennie depuis 1979 en réponse à l'appauvrissement de la couche d'ozone et la concentration croissante des GES (par exemple Grise et al., 2019). Le signal n'est en revanche pas robuste pour la cellule Hadley de l'Hémisphère Nord. Bien que l'expansion simulée dans l'Hémisphère Sud soit inférieure à celle observée, celle-ci peut expliquer au moins en partie l'intensification des alizés au sud de l'équateur. Elle pourrait aussi contribuer au refroidissement relatif observé sur la période 1950-2009 dans le Pacifique Sud-Est et le long des côtes Pérou-Chili (Fig. 3.6c).

3.3.3 Tendances robustes sur la période historique : variabilité interne ou signature des forçages anthropiques ?

Bien qu'une majorité de modèles climatiques suggèrent une réduction future de la circulation de Walker et un réchauffement équatorial accru, le degré de confiance dans ces résultats est relativement faible pour plusieurs raisons. Premièrement, la forte variabilité climatique interne ne permet pas de voir clairement ces deux signaux dans les observations (Fig. 3.8a). Alors que les premières études observationnelles semblaient confirmer la tendance à l'affaiblissement de la WC et des alizés à l'équateur (Vecchi et al., 2006; Power and Kociuba, 2011; Tokinaga et al., 2012), le fort renforcement observé depuis le début des

années 1980, qui est sous-estimé par de nombreux modèles (England et al., 2014; Kociuba and Power, 2015; Luo et al., 2018) cache probablement dans les observations une telle tendance au réchauffement équatorial due aux GES.

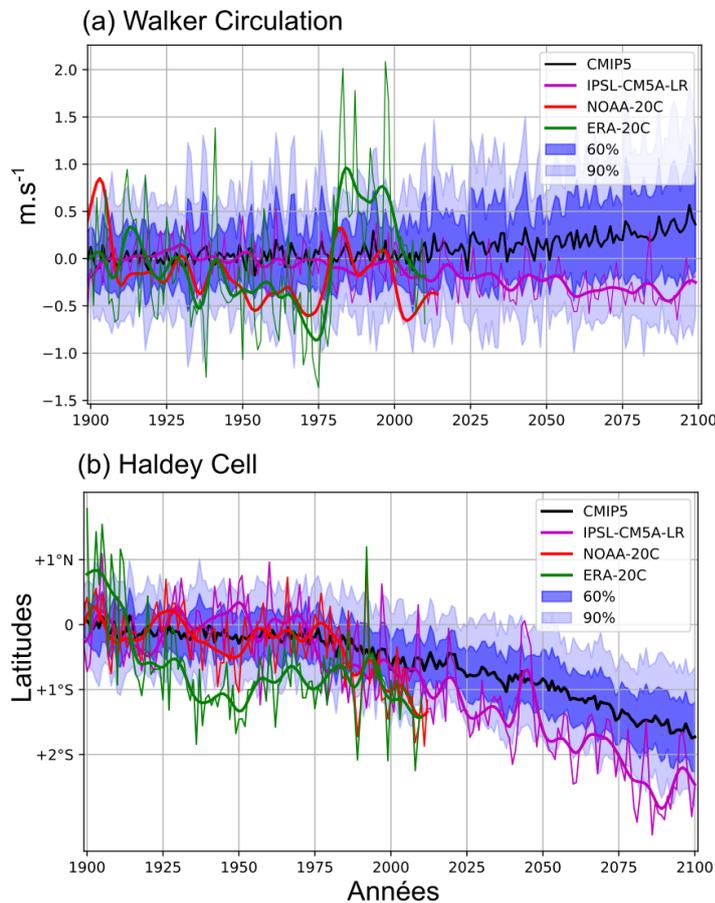


Figure 3.8: Evolutions temporelles des circulations de Walker et de Hadley sur la période 1900-2100. (a) Anomalies des vents zonaux à 10m ($m.s^{-1}$) du Pacifique équatorial ($5^{\circ}N-5^{\circ}S$). Les valeurs positives signifient un affaiblissement de la circulation de Walker et les valeurs négatives un renforcement. (b) Evolution de la position latitudinale (degrés de latitudes sud, en ordonnée) de la limite sud de la cellule Hadley diagnostiquée par la position latitudinale où les précipitations moins l'évaporation en moyenne zonale est égale à 0 ($P-E=0$). Les simulations CMIP5 de la période historique sont utilisées entre 1900 et 2009 et les scénarios RCP8.5 (Van Vuuren et al., 2011) entre 2009 et 2100. Toutes les valeurs correspondent à des anomalies par rapport à la période de référence 1900-1930. Les courbes noires en gras indiquent la moyenne multi-modèle CMIP5 et les enveloppes bleues foncées et bleues claires, l'intervalle de confiance à 60% et 90% selon une distribution t de student. Les résultats pour le modèle IPSL-CM5A-LR sont indiqués par la courbe violette. Les courbes rouges et vertes correspondent aux valeurs des Réanalyses NOAA-20C et ERA-20C respectivement.

Deuxièmement, alors que la majorité des modèles simule une tendance à l'affaiblissement de la WC, certains d'entre eux montrent une WC renforcée sous l'influence du réchauffement climatique avec des mécanismes suggérés tout à fait plausibles (par exemple Solomon and Newman, 2012; Luo et al., 2012; L'Heureux et al., 2013; Plesca et al., 2018). Troisièmement, il existe également un lien statistique solide entre les biais actuels communs aux modèles tels qu'un Pacifique équatorial trop froid et sec et la réponse projetée

des SSTs (Li and Xie, 2014), ce qui réduit la confiance des projections pour le Pacifique tropical. Il est donc actuellement difficile de déterminer s'il y a une diminution à long terme de l'intensité de la WC, masquée qui plus est par une variabilité climatique interne de grande amplitude associée à l'IPO, elle-même sous-estimée par l'ensemble des modèles climatiques (par exemple Chung et al., 2019). Alternativement, l'effet du forçage externe peut être reproduit de manière erronée par les modèles ce qui pourrait aussi expliquer une partie de l'inadéquation entre les changements décennaux récents observés et modélisés.

L'expansion des branches Sud de la cellule de Hadley observée et simulée par l'ensemble des modèles CMIP5 (Fig. 3.8b) fait en revanche partie des réponses cohérentes qui ont été associées aux émissions de gaz à effets de serre et à la destruction de la couche d'Ozone stratosphérique (Grise et al., 2019). Nous avons donc une confiance relativement forte dans l'expansion projetée de la cellule Hadley en réponse au forçage anthropique. Une émergence aussi claire de l'empreinte anthropique sur la tendance HC dans l'Hémisphère Sud pourrait être due à l'amplitude plus faible de la variabilité interne par rapport à l'influence combinée des GES et de l'appauvrissement de la couche d'ozone (Fig. 3.8b). Ces résultats ont une forte implication en ce qui concerne l'expression océanique de l'expansion de la cellule de Hadley et les tendances en SSTs dans l'océan Pacifique Sud. En effet, l'augmentation des vents d'Est dans le Pacifique Sud-Est et la tendance relative au refroidissement dans cette région par rapport au reste de l'océan Pacifique, pourrait être attribuée à l'influence de l'élargissement des cellules de Hadley sur la dynamique océanique du Pacifique Sud. Une telle tendance pourrait aussi influencer l'intensité de l'upwelling côtier le long du Pérou-Chili (Belmadani et al., 2014). Le rôle de la variabilité interne et la tendance négative de l'IPO depuis 1980 pourrait néanmoins aussi expliquer les tendances observées dans cette région. Pour répondre à toutes ces questions une étude détaillée est présentée sous forme d'un article dans la section suivante.

3.4 Article : Contributions of internal variability and external forcing to the recent trends in Southeastern Pacific and Peru-Chile upwelling system

Contributions of Internal Variability and External Forcing to the Recent Trends in the Southeastern Pacific and Peru–Chile Upwelling System

BEYREM JEBRI, MYRIAM KHODRI, VINCENT ECHEVIN, GUILLAUME GASTINEAU, SYLVIE THIRIA, JÉRÔME VIALARD, AND NICOLAS LEBAS

LOCEAN-IPSL, Sorbonne-Université, CNRS/IRD/UPMC/MNHN, Paris, France

(Manuscript received 25 April 2019, in final form 30 December 2019)

ABSTRACT

In a warming world context, sea surface temperature (SST) off central-south Peru, northern Chile, and farther offshore increases at a slower rate than the global average since several decades (i.e., cools, relative to the global average). This tendency is synchronous with an interdecadal Pacific oscillation (IPO) negative trend since ~1980, which has a cooling signature in the southeastern Pacific. Here, we use a large ensemble of historical coupled model simulations to investigate the relative roles of internal variability (and in particular the IPO) and external forcing in driving this relative regional cooling, and the associated mechanisms. The ensemble mean reproduces the relative cooling, in response to an externally forced southerly wind anomaly, which strengthens the upwelling off Chile in recent decades. This southerly wind anomaly results from the poleward expansion of the Southern Hemisphere Hadley cell. Attribution experiments reveal that this poleward expansion and the resulting enhanced upwelling mostly occur in response to increasing greenhouse gases and stratospheric ozone depletion since ~1980. An oceanic heat budget confirms that the wind-forced upwelling enhancement dominates the relative cooling near the coast. In contrast, a wind-forced deepening of the mixed layer drives the offshore cooling. While internal variability contributes to the spread of tendencies, the ensemble-mean relative cooling in the southeastern Pacific is consistent with observations and occurs irrespectively of the IPO phase, hence, indicating the preeminent role of external forcing.

1. Introduction

The Humboldt Peru–Chile upwelling system (PCUS; Strub et al. 1998) is one of the most productive oceanic regions globally in terms of phytoplankton (Chavez et al. 2008; Kämpf and Chapman 2016), and represents one of the leading fisheries in the world, thriving with high biological diversity (Cushing 1971; Fréon et al. 2009). As in other eastern boundary upwelling systems, alongshore equatorward wind stress and the associated wind stress curl drive coastal upwelling. The action of the surface wind stress pushes surface waters offshore and pumps up deeper, cold and nutrient-rich waters (Sverdrup 1938), triggering the high surface productivity. Ecosystems and fisheries are already under strong pressure from climate change globally, with projections pointing to an overall oceanic warming, acidification,

deoxygenation, and reduced primary productivity (e.g., Bopp et al. 2013, Beaugrand et al. 2019). Due to the strong socioeconomic relevance of the PCUS (Ibarra et al. 2000; Pérez et al. 2010), it is critical to further investigate and quantify the potential impacts of climate change in this region.

Previous works documenting the PCUS long-term upwelling trends have mostly relied on wind stress based upwelling indices over recent decades (e.g., Aravena et al. 2014) or estimates from phase 5 of the Coupled Model Intercomparison Project (CMIP5) global circulation models to infer wind-driven upwelling intensity variations into the twenty-first century (e.g., Rykaczewski et al. 2015; Wang et al. 2015). Analyses from general circulation models have suggested mixed results and even contradictory trends for the Humboldt system (Oyarzún and Brierley 2019). On the other hand, sea surface temperatures (SST) cooling trends (Figs. 1a,b) have been observed off Chile since the 1980s (Falvey and Garreaud 2009) and off central-south Peru since the 1950s (Gutiérrez et al. 2011). These coastal trends have been interpreted as an indirect evidence of wind-driven upwelling intensification.

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Beyrem Jebri, beyrem.jebri@locean-ipsl.upmc.fr

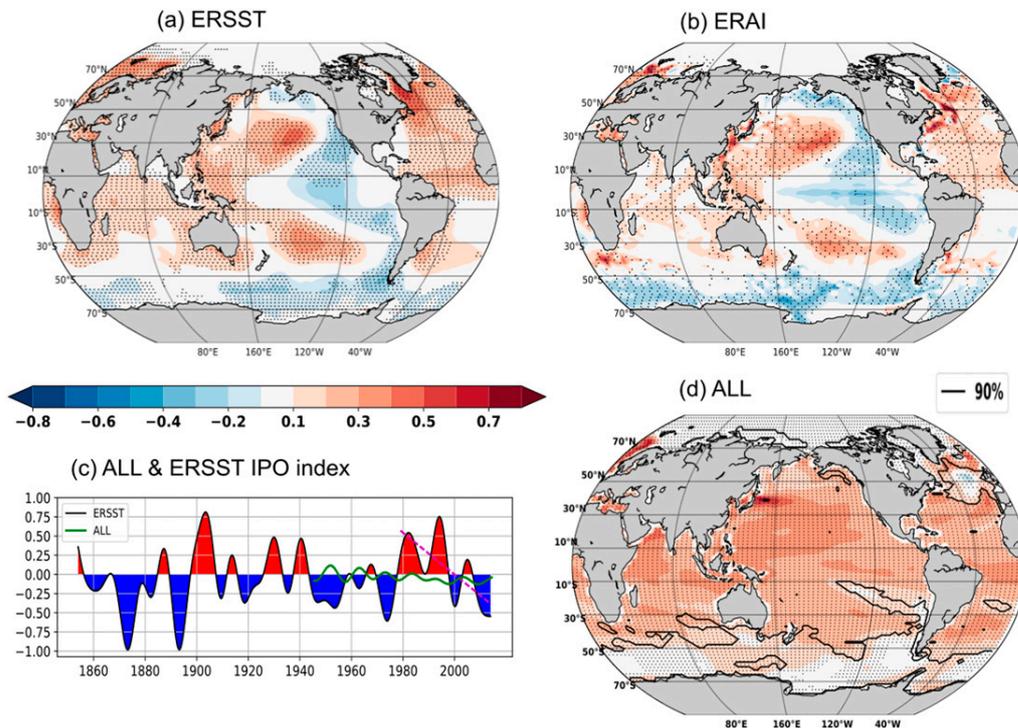


FIG. 1. Sea surface temperature trend in austral summer (December–January–February) for the period 1980–2014, in $^{\circ}\text{C decade}^{-1}$ for (a) ERSSTv4, (b) ERA-Interim, and (d) ALL ensemble mean. Stippling denotes regions where the observed or ensemble mean trend is significant at the 90% confidence level. The black contour lines in (d) indicate where at least 90% of the ALL ensemble members have a positive sea surface temperature trend. (c) IPO TPI index from ERSSTv4 dataset. The green curve shows the ALL ensemble-mean IPO index. The purple dashed line corresponds to the 1980–2014 IPO index linear trend from ERSSTv4.

Low-frequency SST changes in this region can also be associated with internal climate variability (Montecinos et al. 2003; Vargas et al. 2007). While El Niño Southern Oscillation (ENSO; e.g., McPhaden et al. 2006) is the leading mode of interannual climate variability in the Pacific Ocean, its decadal fluctuations imprint a pan-Pacific pattern of coherent SST decadal variability (Newman et al. 2016) through atmospheric teleconnections (e.g., Alexander et al. 2002). This pattern is known as the interdecadal Pacific oscillation (IPO), and is the leading mode of internal, decadal to multidecadal variability in the Pacific Ocean (Folland et al. 1999; Power et al. 1999). The IPO SST pattern is broadly similar to that of ENSO, but less equatorially confined in the eastern Pacific and more prominent in the extratropical latitudes (Minobe 1997; Minobe 1999; Deser et al. 2004; Deser et al. 2010). The IPO has experienced a trend from a positive (i.e., El Niño-like) to a negative (La Niña-like) phase over the 1980–2014 period (Fig. 1c), associated with an intensification of Pacific trade winds and with a shoaling of the thermocline in the Eastern Equatorial Pacific (e.g., Balmaseda et al. 2013; de Boissésou et al. 2014; England

et al. 2014; Kociuba and Power 2015; Choi et al. 2016; Gastineau et al. 2018). Poleward propagating coastal waves and a shoaling of the nearshore thermocline may also play a key role in modulating SST at decadal time scales off Chile (e.g., Montecinos et al. 2007). Such an IPO phase transition during the 1980–2014 period is indeed associated with a central and eastern equatorial Pacific cooling that extends toward midlatitudes and in particular into the southeastern Pacific (hereafter, SEP) and the broad PCUS region (Figs. 1a,b). Thus, the IPO might have contributed to the recent cooling trend in this region.

Nevertheless, the PCUS trends over recent decades could also result from external forcing. Under greenhouse gas forcing continental surface temperatures increase more than nearby SSTs because of the drier environment and larger lapse-rate over land (e.g., Joshi et al. 2008; Dommenges 2009). This lowers pressure over land (e.g., Bayr and Dommenges 2013) and potentially strengthens the land–sea sea level pressure gradient and the upwelling-favorable alongshore wind stress (Bakun et al. 2010). Indeed, CMIP5 model projections point to a

significant increase in the duration and intensity of upwelling-favorable winds in the poleward portions of the four major eastern boundary upwellings (California, Canary, Humboldt, Benguela), which highly correlates to the land–sea temperature gradient (Wang et al. 2015).

The alongshore winds intensification is, however, not accompanied by a clear reinforcement of the land low pressure system, neither in CMIP5 historical simulations nor in projections (Rykaczewski et al. 2015). Belmadani et al. (2014) suggest that the enhancement of upwelling-favorable winds over the PCUS are rather due to subtropical southerly wind anomalies linked to a poleward expansion of the Southern Hemisphere Hadley cell (HC). Several observational datasets and reanalyses indeed show that the Southern Hemisphere HC southern edge ($\sim 20^\circ\text{S}$ in the current climate) has shifted southward by about 0.3° – 0.5° of latitude per decade since 1979 (Johanson and Fu 2009; Hu et al. 2011, 2013; Choi et al. 2014; Staten et al. 2018; Grise et al. 2019), with the largest poleward expansion during austral summer (Chen et al. 2014). Climate model experiments suggest that the HC poleward expansion is forced by the ozone depletion from 1979 to the late 1990s (e.g., Son et al. 2010; McLandress et al. 2011; Hu et al. 2013; Min and Son 2013) and by the rising concentration of atmospheric greenhouse gases (Lu et al. 2007; Staten et al. 2012; Quan et al. 2014; Adam et al. 2014). However, the tendency toward more La Niñas in relation with the recent IPO trend may also contribute to the HC poleward shift (Nguyen et al. 2013).

To our knowledge, previous studies have not specifically examined (i) the respective role of anthropogenic forcing and internal climate variability in long-term trends in the PCUS, (ii) the specific role of the HC expansion on PCUS changes and its attribution to specific sources of external forcing, and (iii) the physical processes responsible for the weak SST warming in the southeastern Pacific (relative to the global mean) over the last decades. These are the questions addressed in this study. Here, we assess the 1940–2020 PCUS climate changes, using observations and a set of 20–30-member ensembles conducted with the IPSL-CM5A-LR (Dufresne et al. 2013) coupled general circulation model. The paper is organized as follows: in section 2, we describe the model, data and methods. Section 3 presents the globally averaged SST trends in the model, and then focuses on the PCUS region, for a validation of the ensemble-mean climatology and overview of the ensemble trends. In section 4, we explore the relative roles of externally forced versus internal variability. Section 5 describes the contribution of various external forcing (greenhouse gases, stratospheric ozone, volcanoes) on the Hadley cell expansion and PCUS.

In section 6, we investigate the mechanisms responsible for the relative SST cooling in the southeastern Pacific. Results are discussed in section 7 before concluding.

2. Methods and data

a. The coupled model

We used the CMIP5 version of the L'Institut Pierre-Simon Laplace (IPSL) atmosphere–ocean general circulation model, called IPSL-CM5A-LR (Dufresne et al. 2013). LR stands for low resolution, as the atmospheric grid resolution is 1.87° in latitude and 3.75° in longitude. It includes the LMDZ5A atmospheric component (Hourdin et al. 2013), with 39 hybrid sigma-pressure levels, including 18 in the stratosphere. LMDZ5A is coupled to the ORCHIDEE (Krinner et al. 2005) land surface component. The oceanic component uses Nucleus for European Modelling of the Ocean (NEMO), version 3.2 (Madec 2008), with 31 vertical levels (and resolution decreasing from 10 m near the surface to 500 m toward the bottom), a mean spatial horizontal resolution of about 2° (with a refinement to a 0.5° latitudinal resolution near the equator). The ocean model includes the LIM2 sea ice model (Fichefet and Maqueda 1997).

Compared to other CMIP5 models, IPSL-CM5A-LR has a relatively large transient climate response, reaching 2 K per CO_2 doubling. It thus ranks among the high-sensitivity climate models of CMIP5 (multimodel mean of 1.82 ± 0.63 K) (Andrews et al. 2012; Forster et al. 2013), due to a large positive low-level clouds feedback (Brient and Bony 2013). Previous works show that this model reproduces ENSO variability with two spectral peaks around 3–3.5 years and beyond 4 years, which is in good qualitative agreement with observations (Bellenger et al. 2014). IPSL-CM5A-LR is also realistic for other aspects of the climate variability. It simulates well the characteristics of observed North Pacific variability (Fleming and Anchukaitis 2016) and the links between ENSO and the IPO (Nidheesh et al. 2017). However, like many global climate models, IPSL-CM5A-LR simulates a double intertropical convergence zone (ITCZ) and the classical cold tongue bias in the Pacific (Dufresne et al. 2013). The observed ENSO seasonal phase locking is also not well captured (Bellenger et al. 2014). Last, IPSL-CM5A-LR also has a rather strong cold bias for the surface air temperature, in particular in the midlatitudes.

b. Experimental protocol

We performed four IPSL-CM5A-LR ensembles with historical external forcing for the 1940–2020 period (see Table 1). The first ensemble of 30 members (referred to as the ALL ensemble) uses the historical CMIP5

TABLE 1. Ensemble size and external forcing considered over the 1940–2020 period for the ensemble IPSL-CM5A-LR experiments used in this study (see section 2b).

Experiment name	Used external forcing			Model members
	Greenhouse gases	Ozone	Volcanic forcing	
ALL	×	×	×	30
GHG	×	—	—	20
OZ	—	×	—	20
VOLC	—	—	×	20

(Taylor et al. 2012) concentrations of greenhouse gases (GHG) from 1940 to 2005 and from RCP8.5 between 2006 and 2020 (Riahi et al. 2011). The standard CMIP5 ozone concentration used in IPSL-CM5A-LR was calculated from a separate atmospheric simulation with interactive atmospheric chemistry (Szopa et al. 2013). We also use the historical stratospheric aerosol forcing from the main volcanic eruptions by prescribing the aerosol optical depth, as given in Sato et al. (1993) (2012 updated version).

Three other ensembles use each forcing individually, namely either volcanic (referred to as VOLC), greenhouse gases (referred to as GHG), or ozone (referred to as OZ) forcing, while maintaining other forcing at their 1940 level. The VOLC, GHG, and OZ ensembles have 20 members each. Each ensemble member starts in January 1940 from randomly sampled initial conditions (31 December) chosen over 1920–60 in a set of IPSL-CM5A CMIP5 historical simulations (initiated in 1870) using all forcings (i.e., similar to ALL). Tropospheric aerosols forcing is maintained to the preindustrial level in all simulations.

Taking initial conditions that sample random phases of the internal climate variability guarantees that this variability is largely uncorrelated across members from the start of the experiments. The ensemble averaging of the large number of members hence damps the internal climate variability by a factor of \sqrt{n} in a n -member ensemble, so that it is reduced by a factor of ~ 5.5 and 4.5 in the 30- and 20-member ensembles, respectively. It allows effective isolation of the externally forced variability. This is demonstrated in Fig. 1c, which shows the ALL ensemble mean IPO index (green curve) whose variability is much smaller than the observed one, confirming that the large ensemble averaging successfully damps out internal variability.

c. Trend analysis and statistical testing

All anomalies are calculated by removing the 1941–61 monthly mean seasonal cycle to the time series. We then applied a Hanning filter with a 7-yr cutoff period to

focus only on the decadal variability. All the linear trends in the paper are computed from the low-pass-filtered time series, using least squares. On time series plots, the ensemble mean is plotted, along with an estimate of the intermember spread. Depending on the figure, this spread is either estimated from the 5th and 95th percentiles or from the 90% confidence interval, assuming Gaussian distributions and considering each member as independent. Trend maps are computed over 1980–2014 and regions where the trend (ensemble mean trend for simulations) is significantly different from zero at 90% confidence level are estimated with a Wald test using a t distribution for the test statistic. We also indicate regions where more than 60% and 90% of the members display anomalies of the same sign than the ensemble mean trend.

d. Upwelling and IPO indices

We considered four main variables to characterize the variations of the PCUS in each ensemble: meridional surface wind stress (τ_y), wind stress curl (ζ), SST, and vertical velocity at 100-m depth (w_{100}). Because the region over which the variability of these variables is homogenous are not necessarily square boxes, and may depend on the variable, we have defined specific regions for each variable according to their covariability. We defined regions in the broad southeast Pacific region (10°N–52°S, 113°–66°W) by clustering, for each variable, nearby grid points with similar temporal variations based on correlations of their 1940–2020 time series. To do so we combined the ClustVar (Vigneau and Qannari 2003) algorithms and the self-organizing map (SOM) (Kohonen 1998) topological ordering proprieties in order to aggregate variables together according to their covariability. We then extract spatial regions of coherent temporal variability from this clustering. To analyze the IPO, we use the SST tripole index (TPI; Henley et al. 2015), defined as the central equatorial Pacific SST (10°S–10°N, 170°E–90°W) minus the average of northwest (25°–45°N, 140°E–145°W) and southwest (15°–50°S, 150°E–160°W) Pacific SST, smoothed using a 13-yr Chebyshev low-pass filter.

e. Mixed layer heat budget

We compute a simplified mixed layer heat budget for each member of the ALL ensemble using monthly data, in order to quantify the respective contributions of surface heat flux and oceanic processes to the SST tendencies. The mixed layer heat budget is expressed as follows:

$$\Delta \text{SST}(t) = \int_{t_0}^t \left[\frac{Q_{\text{atm}}(t')}{\rho \times c_p \times h(t')} \right] dt' + \text{OCE}_{\text{dyn}}(t), \quad (1)$$

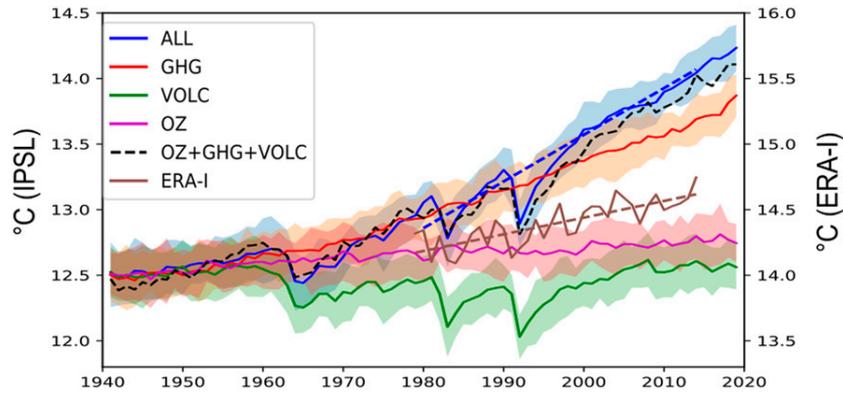


FIG. 2. Global annual mean 2-m temperature, in °C, for ALL (blue), GHG (red), VOLC (green), OZ (purple), and ERA-Interim (brown). The sum of corresponding anomalies from GHG, VOLC, and OZ is shown as a black dashed curve; note that the left y axis for the ensemble simulations is offset by 1.5°C as compared to the right y axis for ERA-I. Thick solid lines indicate the ensemble means; shading shows the spread corresponding to 90% of the members according to a t distribution. Dashed blue and brown lines show the 1980–2014 linear trends.

with

$$Q_{\text{atm}} = Q_{\text{lw}} + Q_{\text{flat}} + Q_{\text{sens}} + Q_{\text{sw}} - Q_{\text{swh}}, \quad (2)$$

where $\Delta\text{SST}(t)$ is the SST monthly variation between t_0 (taken as 1950) and t . The first term of the right-hand side in Eq. (1) represents the SST anomaly due to surface heat flux. The $Q_{\text{atm}}(t)$ represents the anomalous net surface heat flux into the oceanic mixed layer. The $h(t)$ is the thickness of the mixed layer defined as the depth at which the vertical eddy diffusivity coefficient falls below a given value (here taken equal to background value of $5 \text{ cm}^2 \text{ s}^{-1}$) and can be interpreted as a maximum penetration depth of the turbulent motion in the surface layer. Here C_p and ρ denote the seawater specific heat ($4000 \text{ J kg}^{-1} \text{ K}^{-1}$) and a reference density (1025 kg m^{-3}), respectively. We compute the first term on the right-hand side of the equation using monthly mean averages. The net surface heat flux is composed of Q_{sw} the shortwave radiative flux at the ocean surface, Q_{swh} the shortwave radiative flux transmitted through the bottom of the mixed layer, Q_{lw} the long-wave radiative flux, Q_{flat} the latent heat flux, and Q_{sens} the sensible heat flux. The net heat flux is defined positive downward. The second term on the right-hand side of Eq. (1) is the SST anomaly due to ocean dynamics $\text{OCE}_{\text{dyn}}(t)$ (including the local influence of upwelling) estimated as a residual. It corresponds to ocean processes such as the horizontal and vertical advection of temperature, horizontal and vertical turbulent mixing. Since the ocean dynamical term is estimated as a residue from the SST change and the atmospheric heat flux effect, it may also contain errors associated with neglecting the submonthly variations of surface

heat flux and mixed layer depth when computing the atmospheric forcing term.

f. Observations and reanalyses

Our study uses global gridded monthly analyses of NOAA Optimal Interpolation Sea Surface Temperature, version 2 (OISSTv2; Reynolds et al. 2002), which begins in December 1981 and ERSSTv4 (Huang et al. 2015), which starts in 1854. We also analyze the SST seasonal cycle and the wind stress patterns from SODA2.2.1 Ocean Reanalyses climatology (Carton and Giese 2008), representative of the 1941–61 period. Last, we use sea level pressure (SLP), surface wind stress, 2-m air temperature, 3D wind, and pressure fields (to compute the atmospheric meridional overturning streamfunction) from the European Centre for Medium-Range Weather Forecast (ECMWF) ERA-interim reanalysis (ERA-I) over 1979–2014 (Dee et al. 2011), twentieth-century NOAA–CIRES (NOAA-20C; Compo et al. 2011), and NCEP (Kalnay et al. 1996).

3. Model validation

a. Global warming signal and contributions from individual forcing sources

Figure 2 displays the global annual mean 2-m air temperature time series, in ALL, VOLC, GHG, OZ, and ERA-I. The ALL simulation, which includes all sources of forcing, significantly overestimates the observed global heating rate since ~ 1980 . This is mainly due to two factors. As noted in the introduction, the tropospheric aerosol change and the cooling effects associated to sulfate aerosols (e.g., Boer et al. 2000;

Allen and Ajoku 2016) are omitted in these simulations (see section 2b). The overestimated warming also originates from the large climate sensitivity of IPSL-CM5A-LR (Forster et al. 2013). Besides, the (GHG + VOLC + OZ) mean 2-m air temperature fits well that of ALL, suggesting linear additivity of the response to individual external forcing, in terms of global mean 2-m air temperature anomalies. GHG forcing is clearly the main driver of the 2-m air temperature increase and explains most (+0.75°C) of the 1980–2014 warming trend in ALL (+0.99°C), consistent with the behavior of other CMIP5 models (Knutson et al. 2017). Only a weak warming (+0.18°C) is simulated in OZ over the period. This probably reflects the transient adjustment of the climate system toward a warmer state. It could also be due to the small increase in tropospheric ozone acting as a greenhouse gas. VOLC illustrates the strong cooling contribution to ALL a few years following the three main historical volcanic eruptions, namely Agung in 1963, El Chichón in 1982, and Pinatubo in 1991. We further use the GHG, VOLC, and OZ experiments in section 5c to attribute changes in the PCUS to specific forcing sources.

b. Global warming trend pattern

The comparison of the ALL ensemble-mean and observed 1980–2014 SST trends reveals large differences (Figs. 1a,b,d). The two observational products display significant positive SST trends over the North Atlantic Ocean, most of the western Indian Ocean, the tropical and subtropical western Pacific, while weak cooling trends are observed over much of the eastern equatorial Pacific. The observed SST trends capture the negative IPO-like trend pattern (Figs. 1a,b; Deser et al. 2010), with coastal cooling off California and Chile–Peru (Falvey and Garreaud 2009; Gutiérrez et al. 2011) and over most of the equatorial Pacific. The ensemble averaging in ALL filters out internal variability such as that of the IPO, hence reveals the externally forced warming over the period. As a result, the ALL experiment (Fig. 1d) features a large-scale positive trend, largest in subtropical latitudes (Foster and Rahmstorf 2011). The lack of tropospheric aerosol forcing and overestimated climate sensitivity also probably contribute to the more preminent warming than in observations. Despite these differences, the ALL ensemble mean reveals a weak warming trend over much of the SEP, the Southern Ocean and the subpolar Atlantic, consistently with the SST patterns simulated by most models in response to natural and anthropogenic forcing of the twentieth century (e.g., Stouffer et al. 1989). In particular, no significant SST trend is simulated off the coasts of central and south Chile (30°–50°S). This weak surface warming off Chile can be

interpreted as an indirect evidence of the intensification of upwelling in the PCUS in response to the anthropogenic forcing, as suggested by Belmadani et al. (2014), Wang et al. (2015), and Rykaczewski et al. (2015).

c. Peru–Chile upwelling system

In this subsection, we validate the climatological structure of the PCUS in the ALL experiment. Coastal upwelling dynamics are very sensitive to the wind stress structure near the coast (Capet et al. 2004; Small et al. 2015; Renault et al. 2016). Alongshore wind stress at the coast drives the offshore Ekman transport that leads to coastal upwelling. The wind stress also weakens when approaching the coast due to the enhanced friction on land, generating negative wind stress curl and upward Ekman pumping (e.g., Belmadani et al. 2014). Despite its coarse resolution, the model reproduces the first-order seasonal characteristics of the PCUS, namely the intensification of meridional wind stress, associated coastal cooling and positive oceanic velocities off Chile (~20°–30°S) in boreal winter (DJF) and off south Peru (~15°–20°S) in boreal summer (JJA) (Fig. 3). The known “cold tongue” bias of coupled climate models (e.g., Li and Xie 2014) is visible in JJA, but does not extend into the PCUS.

d. Hadley circulation

We here diagnose the Hadley circulation in our model based on the meridional streamfunction. The meridional streamfunction in ALL (Figs. 4a,b, black contours) is consistent with results from previous model studies (Gastineau et al. 2008; Kim et al. 2017) and reanalyses (Mitas and Clement 2005). We define the Southern Hemisphere Hadley cell edge as the linearly interpolated latitude for which the 500-hPa meridional streamfunction (ψ_{500}) is equal to zero. There is an overall good agreement between the ALL and ERA-I ψ_{500} (Fig. 4c) despite a southern edge located too much equatorward (2.2° latitude in DJF and 0.9° in JJA) as in other coupled general circulation models. This discrepancy is linked to the positions of subtropical jets that are also positioned too close to the equator (Barnes and Polvani 2013; Arakelian and Codron 2012).

4. Long-term regional trends and external forcing contribution

a. Regional trends in the Peru–Chile upwelling system

We already underlined that the broad SEP region warms less than the global average in the previous section. To visualize that signal more clearly, and to get rid of the overestimated warming in our simulations, we now examine SST trends relative to the global average

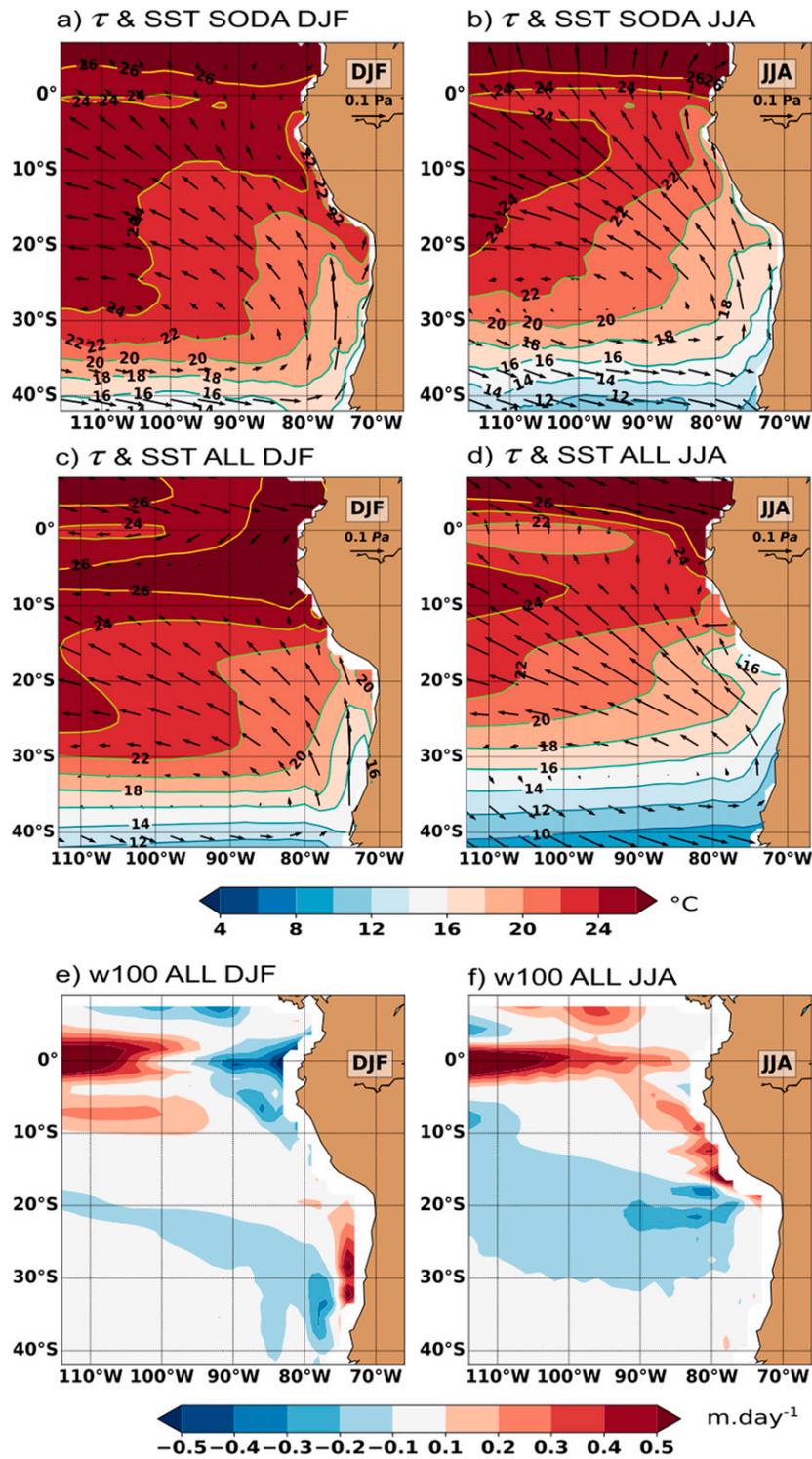


FIG. 3. Mean climatological sea surface temperature (SST, in °C) and surface wind stress (τ , in Pa) for the 1941–61 period for (a),(b) SODA2.2.4 and (c),(d) ALL ensemble mean, in December–January–February (DJF) and June–July–August (JJA), respectively. (e),(f) Mean climatological ocean vertical velocity at 100-m depth (W100, m day⁻¹) over the 1941–61 period for ALL ensemble mean in DJF and JJA. Red (blue) shading corresponds to upward (downward) velocity.

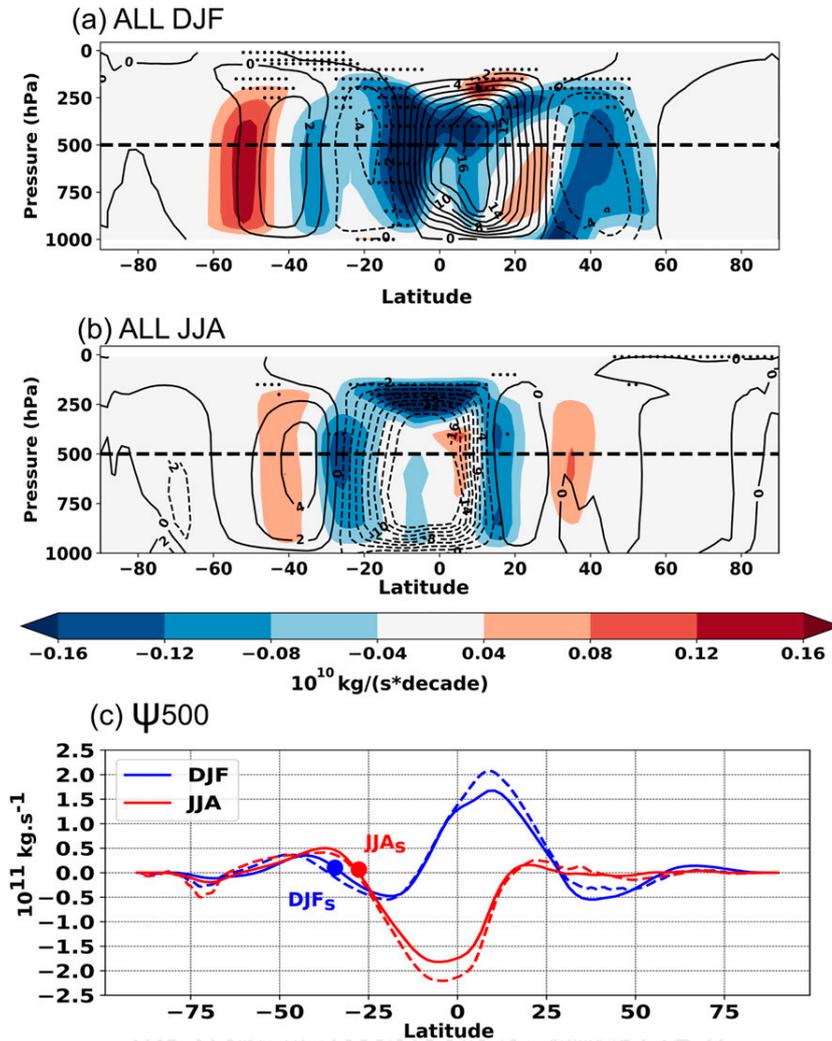


FIG. 4. Mean meridional streamfunction, in $10^{10} \text{ kg s}^{-1}$, positive clockwise, for (a) December–January–February (DJF) and (b) June–July–August (JJA) for the ALL ensemble mean. The black contour corresponds to the climatological values with contours every $2 \times 10^{10} \text{ kg s}^{-1}$. The color shading shows the trend in $10^{10} \text{ kg s}^{-1}$ per decade for the 1980–2014 period. Stippling denotes regions where at least 90% of the ensemble members have anomalies of the same sign. The horizontal dashed lines locate the 500-hPa level. (c),(d) Mean climatological meridional streamfunction at 500 hPa (ψ_{500}) for the 1941–61 period for ALL (thick solid lines) and ERA-I (dashed lines) in DJF (blue) and JJA (red). The dots locate the southernmost (DJFs, JJAs) boundaries of the Southern Hemisphere Hadley cell (see text for details).

(hereafter relative SST trends, similar to, e.g., Vecchi and Soden 2007). The ERA-I SST displays an extended coastal cooling in both seasons (Figs. 5a,b), with a maximum negative trend off Chile between 35° and 43°S in DJF, and off Peru between 10° and 27°S in JJA. In ALL, the cooling trend off the South American coast occurs between 20° and 35°S during both seasons (Figs. 5e,f). Although the cooling trend pattern in ALL broadly agrees with that of the observed negative SST trend, it is shifted northward by about 5° in DJF and southward by 5° in JJA.

Consistent and significant southerly (i.e., upwelling-favorable) wind stress trends develop off Chile in DJF and off north Chile and south Peru in JJA in ERA-I and ALL, albeit with a 4 times stronger amplitude in ERA-I. The amplitude difference is probably related to ensemble averaging, which damps the internal climate variability by a factor of \sqrt{n} in an n -member ensemble. Such positive wind stress trends are, however, located at the poleward boundaries of their climatological pattern, suggesting a southward extension of the southerly wind

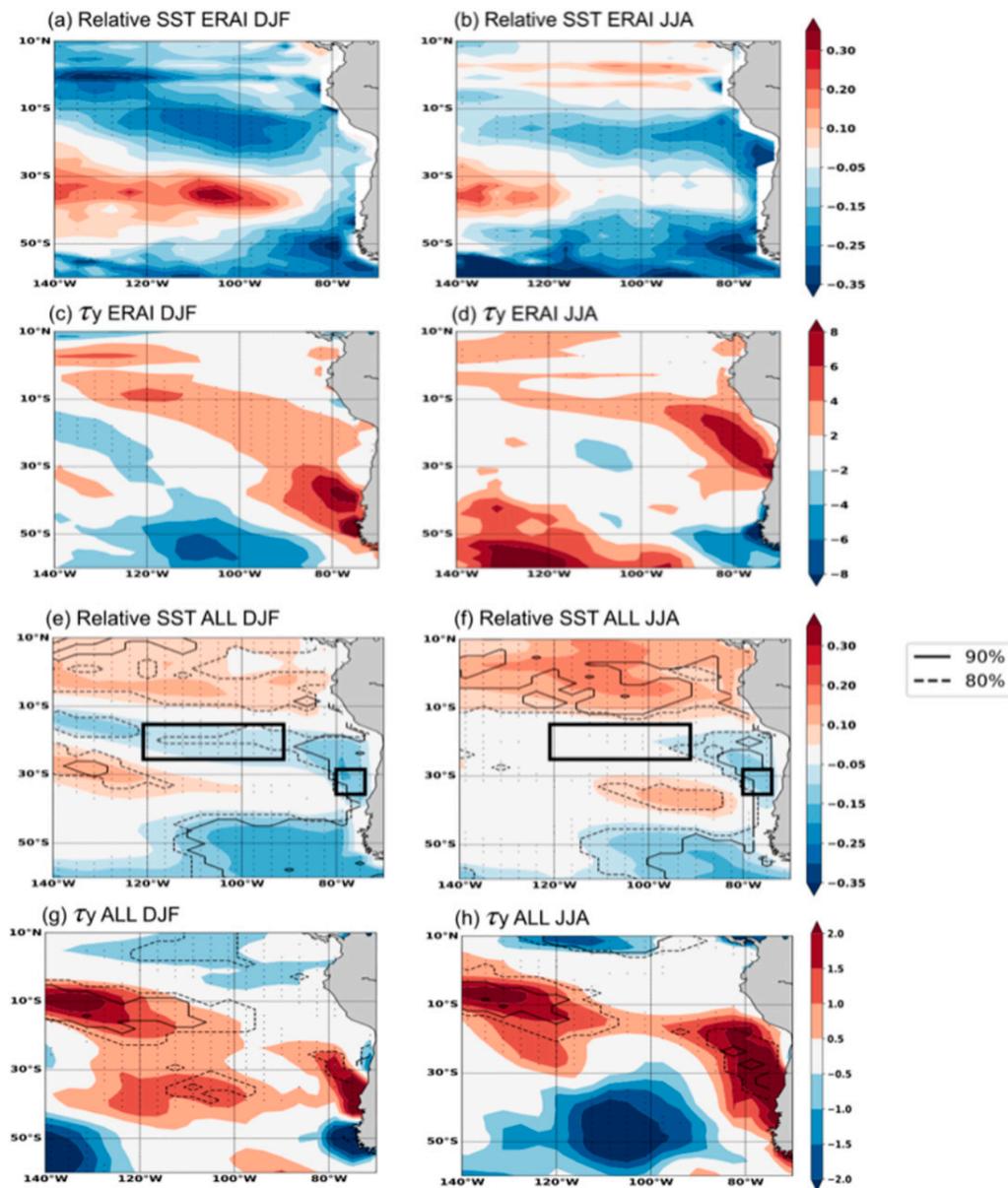


FIG. 5. ERA-Interim (a),(b) relative sea surface temperature (SST in °C per decade) and (c),(d) meridional surface wind stress (τ_y , in 10^{-1} Pa per decade) trends for the 1980–2014 period in DJF and JJA. (e)–(h) As in (a)–(d), but for the relative SST in ALL ensemble mean. Note that the meridional surface wind stress trends in ALL [(g),(h)] are four times weaker than ERA-I. Stippling denotes regions where the mean trends are significant at the 90% level. The black contour lines in (e)–(h) indicate where at least 80% (dashed lines) and 90% (solid lines) of the ALL member's trends have the same sign as the ensemble mean trend. The black boxes in (e),(f) correspond to the SEP and off-Chile regions used for the heat budget analyses in section 6.

regime, in agreement with other CMIP5 climate change projections (Ryckaczewski et al. 2015; Wang et al. 2015). Despite differences in amplitude and location that can be attributed to a combination of model biases and ensemble-averaging that damps the influence of internal variability in ALL, both observations and our model

results qualitatively agree on the southerly wind stress trend off Chile. Such trend may have enhanced the upwelling and drive the observed coastal and offshore cooling.

Figure 6 displays the time evolution of variables that characterize the variations of the upwelling off central

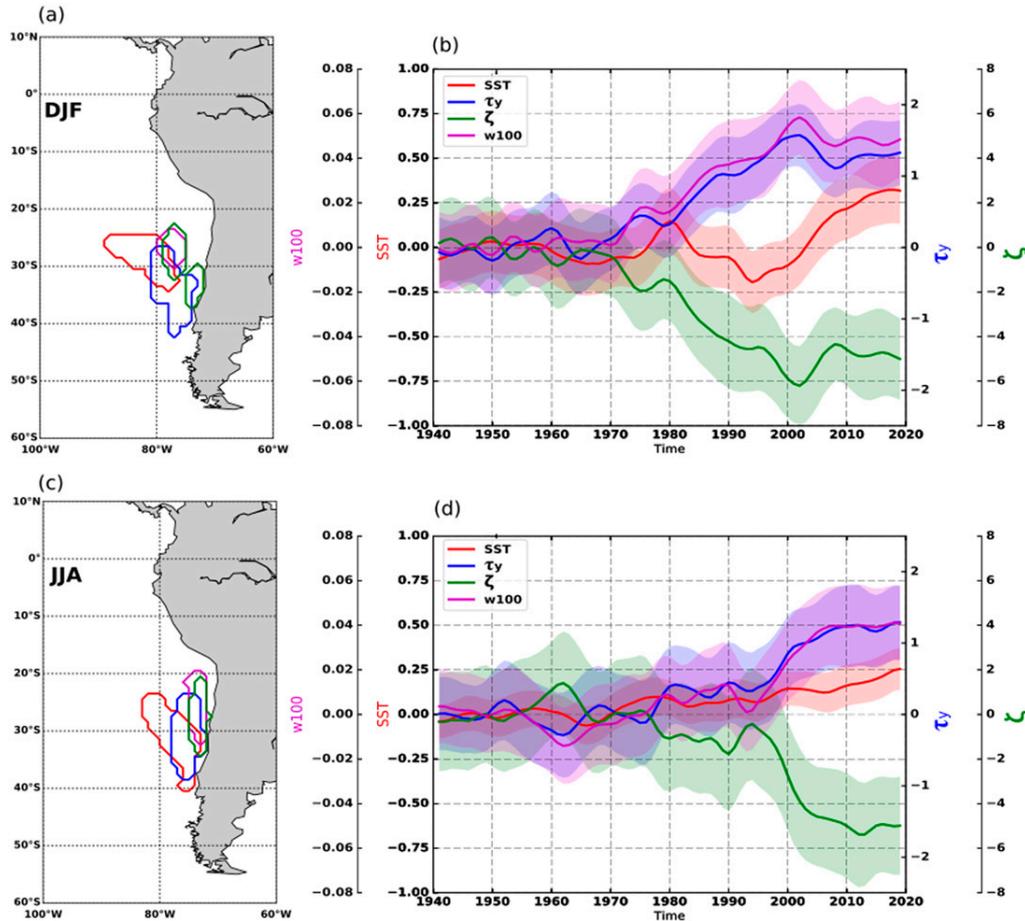


FIG. 6. Area averaged meridional surface wind stress (τ_y , blue, in 10^{-2} N m^{-2}), sea surface temperature (SST, red, in $^{\circ}\text{C}$), wind curl (ζ , green, in 10^{-9} N m^{-3}), and vertical velocity at 100-m depth in the ocean (W100, purple, in m day^{-1}) for ALL in (b) December–January–February (DJF) and (d) June–July–August (JJA). Thick solid lines indicate the ensemble means; shading shows the 95% confidence interval. The regions used for the area average are shown in (a),(c), colored, respectively, as the variables (see section 2c).

Chile. The dynamical variables have physically consistent variations during both seasons, with southerly (positive) wind stress, negative wind stress curl and positive ocean velocities all consistent with an enhanced wind-driven upwelling (Figs. 6b,d). In DJF, these variables start evolving around the mid-1970s to reach a new steady state with higher mean values around 2000 (Fig. 6b). While those dynamical parameters clearly point to an increased upwelling in response to the external forcing, the picture is less clear for SST that remains relatively steady in the 1980s but then warms from the mid-1990s onward. This may be due to the compensation between the cooling driven by the enhanced upwelling of cold subsurface water and warming due to positive radiative forcing. In austral winter (JJA), the upwelling enhances more clearly from the late 1990s onward, when SST has already started

rising due to the radiative forcing (Fig. 6d). It is also reasonable to hypothesize that cooling due to the increase of the upwelling is partly counterbalancing the radiatively forced SST increase during this season as well. We will investigate this hypothesis in more details in section 6.

b. Relative roles of forced versus internal variability on simulated regional trends

In this subsection, we characterize the relative contributions of externally forced and internal variability on decadal and multidecadal variations in the Pacific in our model. In ALL we rely on an empirical orthogonal function (EOF) analysis of SST, sea surface height (SSH), zonal (τ_x), and meridional (τ_y) wind stress in the Pacific sector between 60°N and 50°S (Fig. 7). Figure 7a displays the PC time series obtained from projecting each

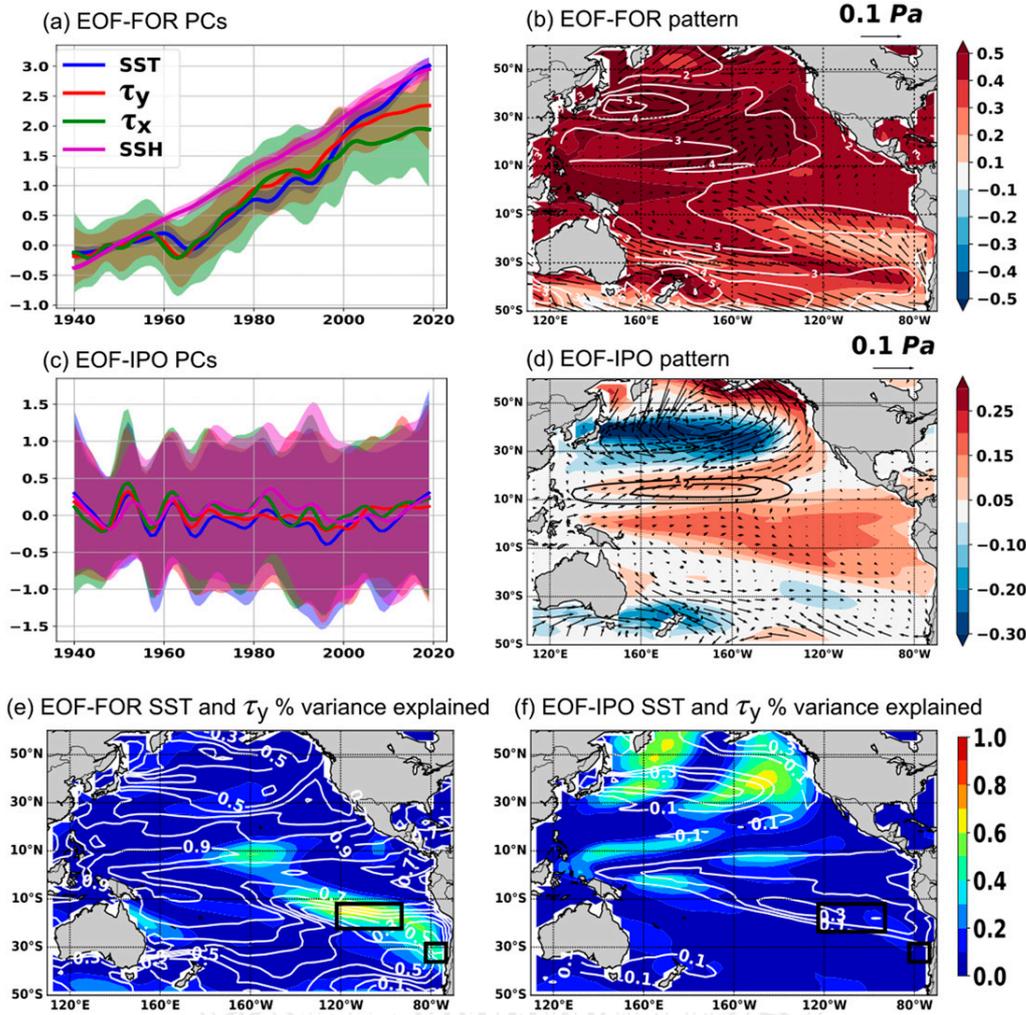


FIG. 7. EOFs and associated standardized PCs for ALL sea surface temperature (SST, in blue), meridional surface wind stress (τ_y , red), zonal surface wind stress (τ_x , in green), and sea surface height (SSH) derived from (a),(b) the leading ensemble mean EOF for each variable (EOF-FOR, capturing 80%–98% of the ensemble mean variations) and from (c),(d) the leading EOF of the residual (EOF-IPO) obtained after removal of the forced components. Shading in (a) and (c) shows the 90% confidence interval. (e) The regression pattern for the EOF-FOR and (f) the EOF-IPO PCs for each variable on the corresponding field anomalies of SST (color shading, in $^{\circ}\text{C}$), meridional surface wind stress (arrows, in 0.1 Pa), and SSH (in m; contours). The percentage of variance explained by the reconstructed τ_y (shading) and SST (white contours) field using the corresponding EOF-FOR and EOF-IPO PCs for the 1980–2014 period are shown in (e) and (f), respectively. The white and black contours in (b) and (d) for SSH, are in meters per standard deviation. The black boxes in (e),(f) correspond to the SEP and off-Chile regions used for the heat budget analyses in section 6.

ensemble member on the first EOF of ALL ensemble mean (Fig. 7b; hereafter EOF-FOR), that is, a robust estimate of the model externally forced response. EOF-FOR explains 98%, 78%, 88%, and 98% of the total variance of the ensemble-mean SST, τ_y , τ_x , and SSH, respectively (Fig. 7a; the ensemble mean and spread are displayed). We then isolated the leading mode of internal variability (EOF-IPO) by applying an EOF analysis of ALL individual members residuals, after removing

signals associated with EOF-FOR (Figs. 7c,d). The first PCs from EOF-IPO (Fig. 7c) explain 29% of the ensemble total variance for SST and τ_y , 26% for SSH, and 22% for τ_x , respectively. The related PCs exhibit a much larger spread than the first EOF-FOR PCs, as expected from internally generated climate variability. Their spatial pattern (Fig. 7d) strongly projects onto that of the IPO, the leading mode of observed internal variability at decadal time scale (Deser et al. 2010). This procedure is used to

separate the forced signal (EOF-FOR) from the leading mode of internal variability at decadal scale (EOF-IPO).

The forced surface wind stress, SST and SSH evolutions obtained from EOF-FOR (Figs. 7a,b) are consistent with Fig. 1d, and display enhanced trade winds over the central equatorial Pacific. Our model results contrast with the majority of CMIP3/5 scenario simulations that rather projects a weakened Walker circulation and an El Niño-like response pattern, with the central and eastern equatorial Pacific SST warming more than the western equatorial Pacific in response anthropogenic climate change (e.g., Collins et al. 2005; Held and Soden 2006; Kociuba and Power 2015). The SSH also rises more in the western Pacific in our model, in response to the strengthened equatorial trades.

In our region of focus, EOF-FOR reveals a weaker SST warming than in other regions and a wind strengthening in an oblique band from the Chilean coast to the central equatorial Pacific (Fig. 7b). In ALL, the forced component of the simulated variability accounts for more than 60% of the upwelling-favorable meridional wind stress (τ_y) increase (Fig. 7e) and about 20% of SST decadal variability locally off Chile (white contours in Fig. 7e). In contrast, internal variability related to IPO has no significant contribution to meridional wind stress and SST variability off Chile (Fig. 7f), but mostly contributes to these variables in the North Pacific. Farther offshore in the SEP box, the SST internal decadal variability is more important (Fig. 7f), the IPO explaining 10%–20% of SST variance while the forced component accounts for more than 40% (Fig. 7e). This analysis hence indicates that in our model IPO-related SST changes are weak relative to the forced changes in the SEP region, and very weak in the Chile upwelling region.

c. Relative roles of forced versus internal variability on observed regional trends

To evaluate the robustness of our model results and provide an estimate of the relative roles of external forcing and the IPO in observations, Figs. 8a–c display the regressed SST, surface winds, and sea level pressure (SLP) fields onto indices of the global warming signal (globally averaged SST) and IPO for the 1940–2014 period. The results reveal the imprint of the radiatively induced global warming except in North Pacific, southern Pacific, and off the Chilean coast where there is no warming signal. The global mean SST explains 1%, 32%, and 53% of the total variance for SST, meridional winds, and SLP, respectively, off Chile while the IPO explains 23% of SST variance and 1% of variance on average for both meridional winds and SLP variables (Figs. 8d–g). Farther offshore in the SEP region, the SST internal decadal variability related to IPO explains 37%,

30%, and 37% of SST, meridional winds and SLP variance, respectively, while the forced component accounts for 19%, 5%, and 17%. This analysis hence suggests that IPO-related upwelling-favorable meridional winds and related SLP changes in the Chile upwelling region are not significant, the coastal region being dominated by the forced changes while farther offshore in the SEP box, the SST internal decadal variability plays a more important role. These results are robust when the analysis is performed over a shorter, better-observed period (1960–2014; not shown). This analysis supports our conclusions based on our model large ensemble that the relative cooling off Chile and in the SEP might be tightly linked to the anthropogenic external forcing signal, which probably exerts a dominant role in reinforcing the coastal upwelling.

To further evaluate the role of the IPO phase on 1980–2014 regional SST trends in our model and observations, we show relative SST (RSST) trends in the SEP and off-Chile regions separately for ensemble members with a positive or negative IPO trend during the period (about half of the members for each, Fig. 9a). In observations, the negative IPO trend ($-0.29^\circ\text{C decade}^{-1}$; green vertical dashed line in Fig. 9a) is associated with a significant relative cooling trend both off Chile ($-0.17^\circ\text{C decade}^{-1}$) and in the SEP box ($-0.14^\circ\text{C decade}^{-1}$) regions (Figs. 9b,c; green vertical dashed line). In the model, however, the phase of the IPO does not have a clear fingerprint on local nor regional RSST trend. A significant RSST decrease off Chile ($-0.20^\circ\text{C decade}^{-1}$) and in SEP ($-0.12^\circ\text{C decade}^{-1}$) is simulated and its amplitude is consistent with observations, regardless of the IPO phase. This suggests again that the internally generated SST decadal variability is overwhelmed by the external forcing influence in ALL during the period 1980–2014. External forcing therefore likely influences the observed 1980–2014 trend.

Overall, this section demonstrates that the strengthening of the southerlies over recent decades is dominated by external forcing, both at the coast and offshore. The relative cooling trend is also dominated by external forcing in the Chile upwelling and SEP regions, with a larger contribution of internal variability (including the IPO) in the SEP. In the following section, we will relate this southerly wind anomaly to the Hadley cell extension and attribute this extension to specific external forcing.

5. Influence of the HC meridional extension and attribution to individual forcing

a. Connections between the HC and the PCUS

The strengthening of upwelling-favorable winds off Chile can be explained by large-scale atmospheric circulation changes. The climatological low-level atmospheric circulation in the PCUS region is associated with the

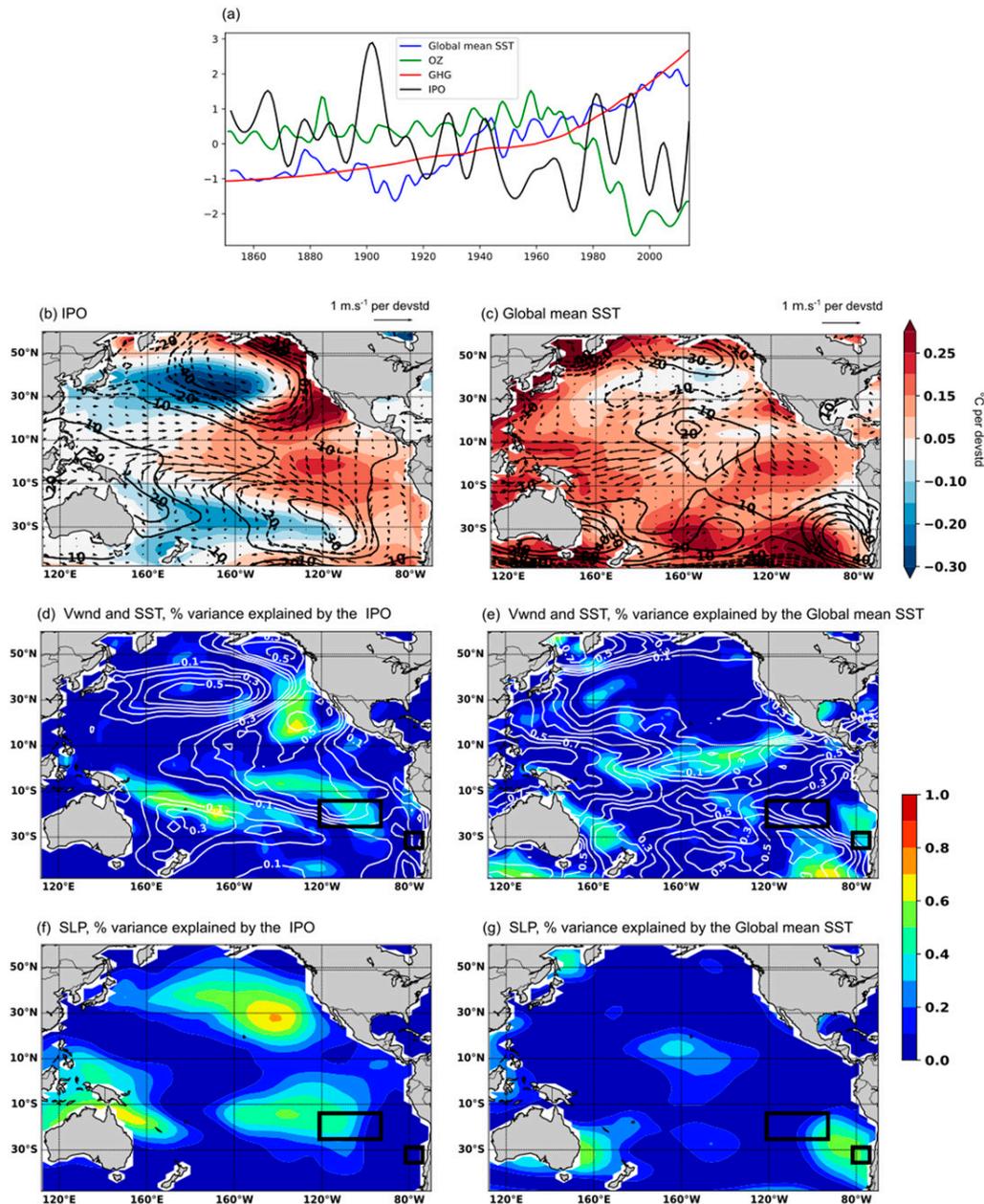


FIG. 8. (a) IPO (black curve; tripole index see section 2d), global warming signal (blue curve; globally averaged SST), atmospheric GHG (red curve; as provided by CMIP6), and stratospheric ozone (green curve; as provided by CMIP6) standardized indices for the 1940–2014 period. (b) The regression pattern on the IPO and (c) the global mean SST standardized indices on field anomalies of SST (ERSSTv4 dataset; color shading, in °C), 10m winds (from NOAA-20C reanalyses; vectors, in m s⁻¹), and sea level pressure (SLP, from NOAA-20C reanalyses; contours, in Pa). The percentage of variance explained for the meridional 10 m wind (color shading), SST (white contours), and SLP (color shading) field using the IPO and global mean SST indices are shown in (d),(f) and (e),(g), respectively. The black contours in (b) and (c) for SLP are in Pa per standard deviation. The black boxes in (d)–(g) locate the SEP and off-Chile regions.

eastern flank of the South Pacific anticyclone (SPA; Figs. 10a,b, black contours). We now evaluate the following hypotheses for explaining the development of externally forced southerly anomalies in the PCUS

region: either a southward expansion of the HC and southward shift of the SPA (e.g., Belmadani et al. 2014) or strengthening of the land-ocean sea level pressure contrast in the context of climate change (Bakun et al. 2010).

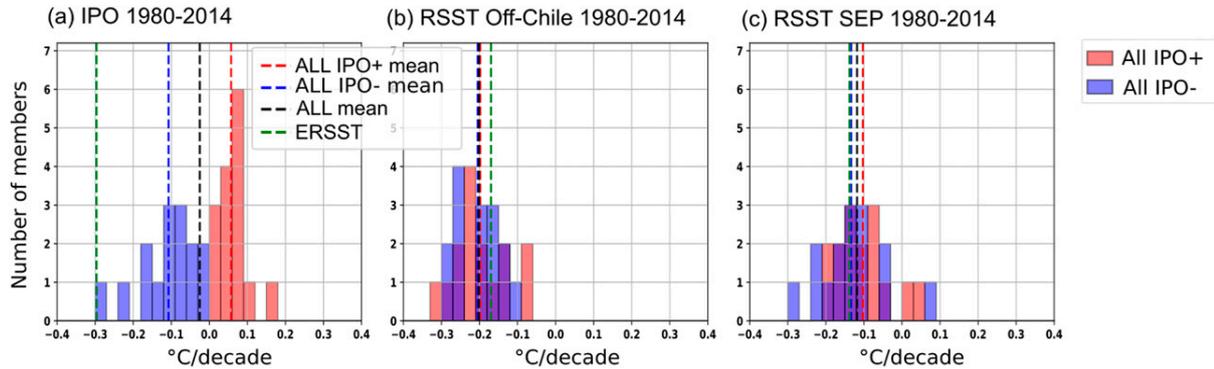


FIG. 9. Probability density function (pdf) for (a) the IPO index trends during 1980–2014 ($^{\circ}\text{C decade}^{-1}$). We computed the IPO index [relying on the tripole index (TPI) from Henley et al. (2015)], defined as the central equatorial Pacific SST ($10^{\circ}\text{S}–10^{\circ}\text{N}$, $170^{\circ}\text{E}–90^{\circ}\text{W}$) minus the average of northwest ($25^{\circ}–45^{\circ}\text{N}$, $140^{\circ}\text{E}–145^{\circ}\text{W}$) and southwest ($15^{\circ}–50^{\circ}\text{S}$, $150^{\circ}\text{E}–160^{\circ}\text{W}$) Pacific SST, smoothed using a 13-yr Chebyshev low-pass filter (see section 2d). We then use this index to build a pdf of 34-yr-long IPO index trends allowing to split between positive (red bars) and negative (blue bars) trends. (b) As in (a), but for the relative sea surface temperature (RSST) trends (in $^{\circ}\text{C decade}^{-1}$) off Chile. (c) As in (b), but in SEP. The off-Chile and SEP regions used for the SST area-averaged in (b) and (c) are shown as black rectangles in Figs. 5 and 7. The vertical dashed lines in (a)–(c) correspond to ALL ensemble mean (black) and the mean for positive (red bars) and negative (blue bars) IPO index trends. The green vertical dashed lines correspond to ERSSTv4 trends. Note we use relative SST anomalies, respectively, to the global mean SST for both ERSSTv4 and ALL (see text).

We isolate regions of homogeneous variability of meridional wind stress and vertical velocity at 100-m depth off Chile to build area-average low-pass-filtered time series for both variables (colored contours in Figs. 10a and 10b; see section 2d). Figures 10a and 10b display the regression of the SLP anomalies onto the low-pass-filtered meridional wind stress time series off Chile in ALL (Figs. 10a,b). Southerly wind anomalies off Chile since the 1970s (1990s) in DJF (JJA) are associated with positive SLP anomalies in the southern edge of the SPA, consistent with the HC poleward expansion in other CMIP5 models (e.g., Min and Son 2013; Choi et al. 2014; Nguyen et al. 2015). The ALL ensemble mean, however, does not reproduce a decline of SLP inland over South America, as would be expected from the Bakun et al. (2010) mechanism.

To further illustrate the relationship between the Southern Hemisphere HC edge (see section 3d) and the SPA, we define an area-averaged low-pass-filtered SLP index, using box average over the region outlined on Figs. 10a and 10b (yellow box). The high correlation (0.98 and 0.96 in DJF and JJA, respectively) between the latitudinal position of the southern edge of the HC (as defined in section 3d) and the SPA SLP further illustrates the strong link between the HC southward expansion and SPA strengthening (Fig. 10c). We also find strong relations between the SPA SLP and the low-pass-filtered time series of the velocity at 100-m depth (upwelling intensity) off Chile (correlation values of 0.97 and 0.98 in DJF and JJA, respectively, Fig. 10d).

Overall, analyses in this subsection illustrate a strong link between the simulated positions of the HC southern

boundary, the intensification of the SPA, the southerly winds and the strengthened coastal upwelling off the Chilean coast.

b. Attribution of HC variations to individual external forcing

Figure 11 displays the time series of the ensemble-mean latitude of the Southern Hemisphere HC poleward edge in ALL, GHG, and OZ during 1940–2020 and in twentieth-century NOAA–CIRES (Compo et al. 2011), NCEP (Kalnay et al. 1996), and ERA-I (Dee et al. 2011) reanalyses. The reanalyses and ALL ensemble mean all reveal (with the exception of NCEP in JJA) a significant poleward shift of the subsiding branch of the Hadley cell, from the 1970s onward with a stabilization during the last two decades of the simulated period. The shift is more pronounced in DJF ($\sim 0.8^{\circ}\text{S}$ in 1980–2014) than in JJA ($\sim 0.45^{\circ}\text{S}$). ALL results are within the range of reanalyses uncertainties, which indicate a 0.25° to 3° poleward expansion of the HC southern edge since the seventies (Figs. 11c,d; Seidel and Randel 2007; Staten et al. 2018; Grise et al. 2019).

Single-forcing experiments suggests that volcanic forcing does not contribute to the HC southward expansion (Figs. 11a,b). During austral winter (JJA), the ozone contribution is also weak, and the greenhouse gas forcing dominates the HC expansion, producing a linear trend in the HC poleward expansion, from 1970 to 2020 (Fig. 11b). GHG and stratospheric ozone depletion both act to expand the HC in austral summer, the impact of ozone depletion dominating before 1995 in agreement with previous studies (e.g., Min and Son 2013;

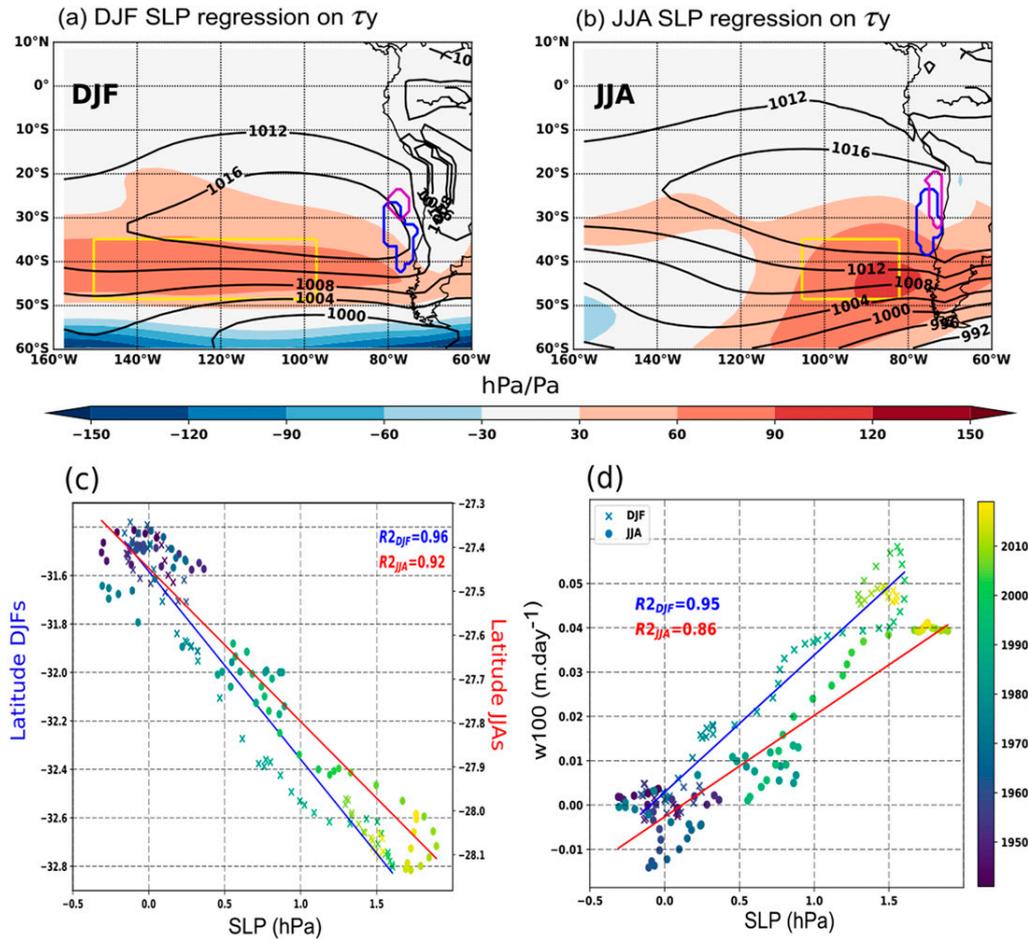


FIG. 10. Sea level pressure (SLP) 1941–61 climatology (black contours, every 4 hPa) and linear regression (shading, in hPa Pa^{-1}) of the SLP on the mean meridional surface wind stress off Chile in (a) December–January–February (DJF) and (b) June–July–August (JJA) for the 1980–2014 period using ALL Ensemble mean. (c) Scatterplot between the latitudinal positions of the Southern Hemisphere Hadley cell boundaries vs the area averaged South Pacific sea level pressure (SLP, in hPa) anomalies in DJF (DJFs, crosses) and JJA (JJAs, dots). (d) Scatterplot between vertical velocity at 100-m depth (w_{100} , in m day^{-1}) off the Chilean coast with the area averaged South Pacific SLP anomalies (in hPa) in DJF (crosses) and JJA (dots). The scatterplots are using values from ALL ensemble mean. The coefficient of determination (R^2) between each variable is given in (c),(d). The dots and crosses on (c),(d) are colored from the oldest (dark blue) to the most recent years (yellow) of the simulation for the whole 1940–2020 period. The blue and red lines show the regression lines in DJF (blue) and JJA (red). The yellow box in (a),(b) corresponds to the regions used for the SLP South Pacific anticyclone. The blue and purple contours in (a),(b) indicate the region used for the area average of meridional surface wind and w_{100} off Chile.

Waugh et al. 2015). After 1995, the ozone recovery (Szopa et al. 2013) yields a stabilization of the ozone contribution to the HC expansion, at a level similar to that of GHGs. Together, these two effects yield an austral summer HC meridional expansion from 1970 to 2000, followed by stabilization from 2000 onward.

c. Attribution of PCUS variations to various external forcing

Since all the variables that characterize the PCUS are highly correlated (see section 4a), Fig. 12 displays only

SST and w_{100} (hereafter upwelling) anomalies in single-forcing ensemble simulations. The (GHG + VOLC + OZ) mean SST and w_{100} time series fit well with the ALL curve, suggesting a linear additivity of each individual external forcings. Increased upwelling caused by the ozone depletion (Fig. 12a) explains most of the SST cooling visible in the early 1980s in DJF (Fig. 12c). The SST variations in OZ hence largely reflect wind-driven changes linked to stratospheric ozone changes. These effects of the ozone forcing start decreasing after 1995 in association with the ozone recovery. GHG contributes

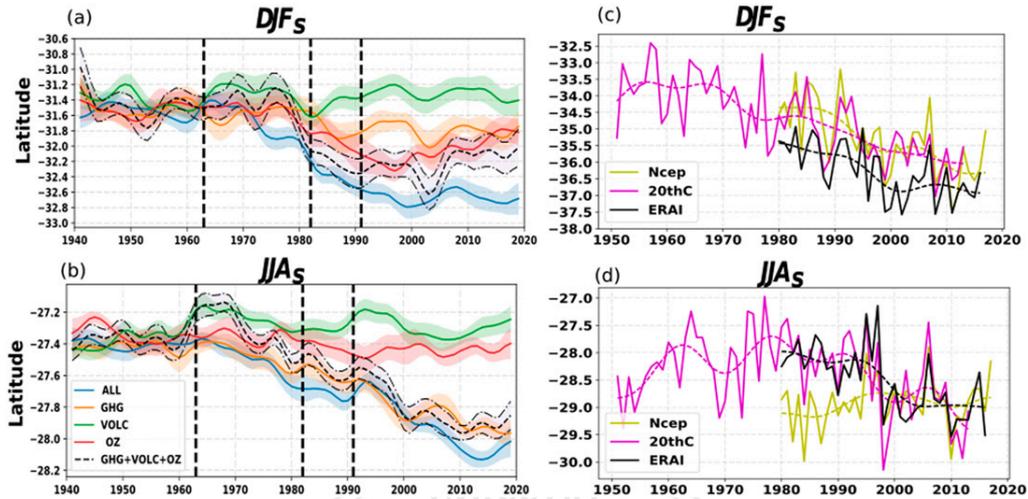


FIG. 11. Time series of the poleward edge of the Southern Hemisphere Hadley cell in (a) December–January–February (DJFs) and (b) June–July–August (JJAs) for ALL (blue), GHG (orange), VOLC (green), and OZ (red). Colored thick solid lines indicate the ensemble means, shading shows the 95% confidence interval. The sum between GHG, VOLC, and OZ is shown as a black dashed curve. The vertical dashed lines indicate the timing of Agung (March 1963), El Chichón (April 1982), and Pinatubo (June 1991) eruptions. (c),(d) As in (a),(b), but for NCEP (in yellow), NOAA-CIRES 20CR (in pink), and ERA-I (in black) reanalyses. Solid thick lines correspond to the annual time series and dashed lines after applying a 7-yr Hanning filter.

mainly to upwelling changes in JJA (Fig. 12a) through dynamically induced cooling but this is overwhelmed by radiatively induced warming (Fig. 12c). Hence the two competing effects between cooling through strengthened upwelling and warming through GHGs radiative effect sustain the relatively limited surface warming off Chile as compared to the surrounding regions (Fig. 1). Even if the volcanic forcing contributes through a small weakening after Agung (March 1963), El Chichón (April 1982), and Pinatubo (June, 1991), we clearly see a dynamical response with an upwelling intensification in response to the anthropogenic forcing during both seasons. GHG and OZ forcing within ALL combine to produce the upwelling intensification since the 1970s. This upwelling intensification first counteracts the effect of the GHGs radiative forcing, leading to a stable SST until ~2000. After the ozone recovery, the forcing from GHGs overwhelms the cooling effect of the enhanced upwelling, yielding a SST increase. Overall, this is consistent with the weaker SST increase off Chile than in surrounding regions, in response to radiative forcing. In the following section, we further investigate this surface warming using a simplified heat budget of the upper oceanic layer.

6. Why does the southeast Pacific warm less than the rest of Pacific Ocean?

We contrast the SST upper ocean budget of our two regions of interest, namely nearshore off Chile (32°–30°S,

80°–75°W) and farther offshore in the subtropical SEP (25°–17°S, 120°–90°W; Fig. 13a), with that of the entire Pacific (50°S–50°N, 120°E–75°W), in order to elucidate why these two regions warm at a reduced rate. The average Pacific warming starts 20 years earlier in the 1970s and is larger (+1.2°C in 2020) than that of the SEP (+0.6°C) and off-Chile (+0.2°C) regions (Fig. 13b). The contributions from the surface heat flux (labeled ATM; first term in the rhs of Eq. (1) in section 2d) and the oceanic processes (labeled OCE_{dyn}, second term in the rhs of Eq. (1), including lateral and vertical advection and mixing) are detailed in Fig. 13c.

At the scale of the Pacific Ocean (Fig. 13c, red curves), the GHG forcing induces a positive net air–sea flux (dominated by downward longwave radiation, not shown) that warms the ocean, as expected from the effect of greenhouse gases (e.g., Barnett et al. 2005). A cooling by oceanic processes largely compensates this warming. The climate change induced warming indeed occurs predominantly in the ocean surface layer (e.g., Barnett et al. 2005), and hence increases the temperature difference between the surface layer and the subsurface, and thus the cooling by turbulent fluxes at the bottom of the mixed layer. In the nearshore region, the warming is also due to atmospheric heat fluxes, but partly compensated by a larger cooling due to increased upwelling. This is supported by Fig. 13c (blue curves), which reveals a stronger cooling by oceanic processes in the off-Chile region (presumably largely because of

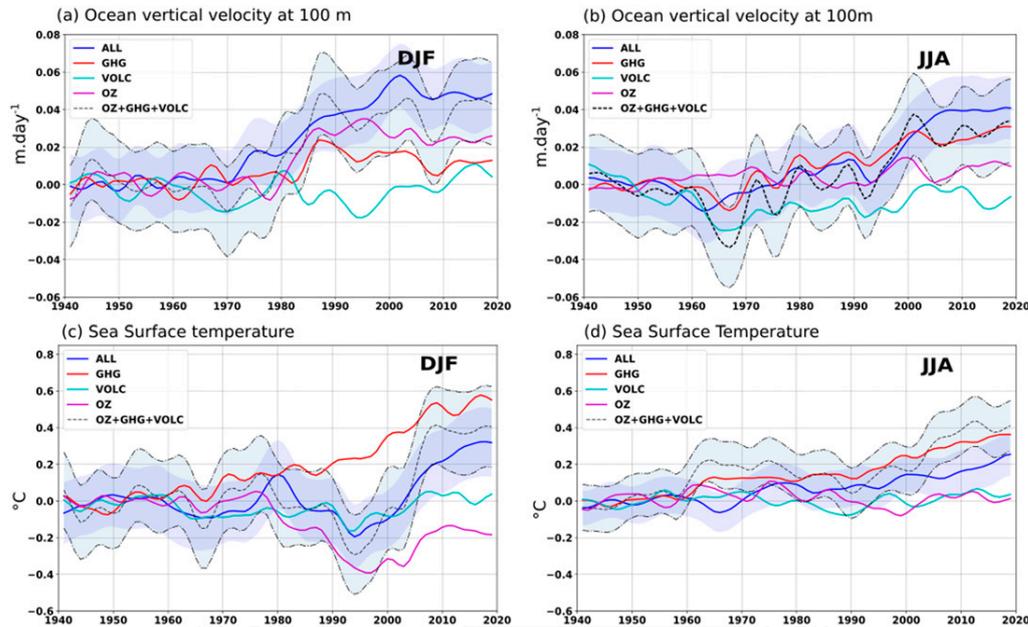


FIG. 12. (a),(b) Off-Chile vertical velocity at 100-m depth (in m day^{-1}) and (c),(d) sea surface temperature (in $^{\circ}\text{C}$) in December–January–February (DJF) and June–July–August (JJA) for ALL (blue), GHG (red), VOLC (green), and OZ (purple). The regions used for the area average off Chile are shown as purple and red contours in (a) and (c) of Fig. 6. The sum between GHG, VOLC, and OZ is shown as a black dashed curve and dashed-dotted lines. Colored thick solid lines indicate the ensemble means; shading shows the 95% confidence interval.

changes in upwelling) than over the entire Pacific (presumably largely because of changes in thermal stratification and vertical mixing of heat). Figure 13f displays upward velocity and cooling trends in the top 500 m near the coast off Chile that supports this view.

It is intriguing that in the offshore SEP region the warming rate is also less than for the Pacific average since it is not an upwelling region (Fig. 13b, green curve). There, the impact of atmospheric net air–sea heat fluxes trend [Q_{atm} in Eq. (2)] is almost nearly neutral (Fig. 13d, green dashed line) due to a balance between increased longwave radiations [Q_{lw} in Eq. (2)] contributing to warming ($6.28^{\circ}\text{C decade}^{-1}$ during 1980–2014) and the negative trend of net shortwave radiations at ocean surface [$Q_{\text{sw}} - Q_{\text{swh}}$ in Eq. (2); $-4.25^{\circ}\text{C decade}^{-1}$ during 1980–2014], assuming a fixed climatological mixed layer depth. Indeed, more low-level clouds are increasingly formed over the SEP region (not shown), due to relatively colder SST and boundary layer stratification under the synoptic HC descent over the SEP region (Bony et al. 2006), which results in a local planetary albedo increase. The strengthening of the surface wind anomalies also favors cooling through enhanced latent heat flux ($-4.03^{\circ}\text{C decade}^{-1}$ during 1980–2014), while sensible heat fluxes anomalies explain a small warming ($+0.63^{\circ}\text{C decade}^{-1}$ during 1980–2014). Near the coast,

however, low-level cloud cover decreases in response to a stronger southerly coastal jet in our simulations. This is consistent with the observed synoptic-scale covariability of surface winds and low cloud cover off Chile at 33°S (Garreaud and Muñoz 2005). Consequently, the downward shortwave flux is more intense near the coast than offshore, resulting in a larger Q_{atm} . Wind stress strengthening over SEP and coastal regions (Figs. 5g,h) also influences the mixed layer depth and hence could modulate the atmospheric heat fluxes impact on SST.

The respective roles of the mixed layer thickness (h) and net surface downward heat flux (Q) variations in ATM can be evaluated by fixing either h or Q to their climatological values (Fig. 13d). Whereas SST increases in the three considered regions, the mixed layer depth increases during 1960–2000 particularly in DJF at rates over 0.9 and $0.7 \text{ m decade}^{-1}$ in the SEP and nearshore regions, respectively, whereas no changes is evidenced on average over the Pacific basin (figures not shown). The annual-mean surface net heat flux in the SEP region being positive (4.80 W m^{-2} in ALL), a thickening of the mixed layer (and hence higher mixed layer heat capacity) results in a decrease of heating rate by surface heat fluxes ($-5.18^{\circ}\text{C decade}^{-1}$ during 1980–2014; green solid curve in Fig. 13c; green crossed line in Fig. 13d) (i.e., an anomalous cooling).

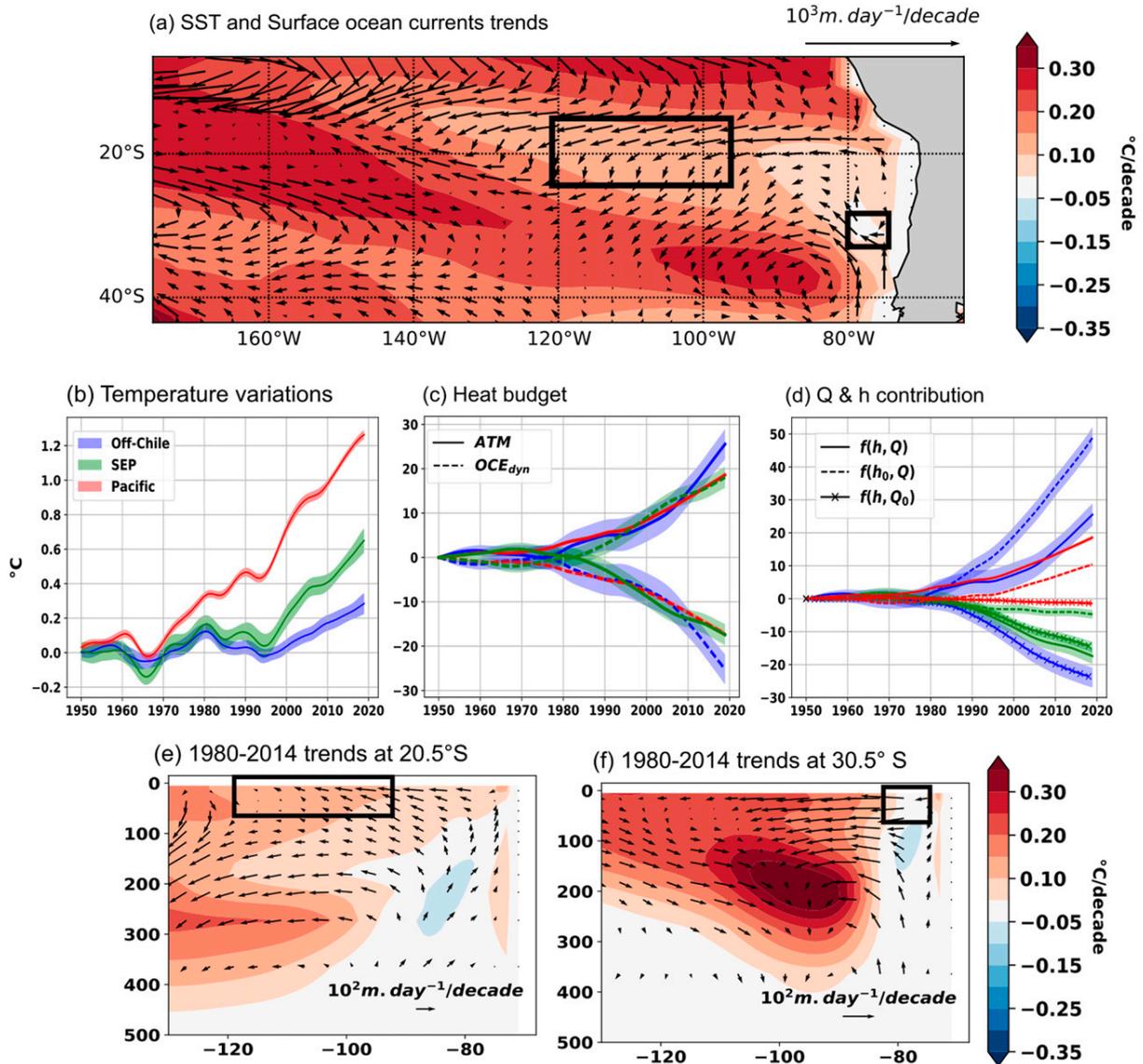


FIG. 13. ALL (a) ensemble mean trends of sea surface temperature (SST, shading, $^{\circ}\text{C} \text{ decade}^{-1}$) and surface current (arrows, $10^3 \text{ m} \text{ day}^{-1} \text{ decade}^{-1}$) over 1980–2014, (b) sea surface temperature anomalies, (c) integrated net mixed layer heat budget for the Pacific basin (red curves), southern Pacific (SEP, green curves), and off-Chile (blue curves) regions. The regions used for the SEP and off-Chile area averages are shown as black boxes in (a). In (c) the thick solid lines correspond to the influence of heat flux (ATM) and the dashed lines to the influence of oceanic processes (OCEdyn); see section 2e for details. (d) Influence of heat flux considering both heat flux (Q) and mixed layer (h) variations (thick solid lines), only Q (dashed lines), or h variations (crossed lines). The h_0 and Q_0 in the legend indicate that the climatological value is used for the calculation. Thick lines indicate the ensemble means while shading shows the 90% confidence interval. (e),(f) Vertical section of potential temperature (color shading, $^{\circ}\text{C} \text{ decade}^{-1}$) and velocity (arrows, $10^2 \text{ m} \text{ day}^{-1} \text{ decade}^{-1}$) trends over 1980–2014 period off Chile at 30.5°S and in SEP region at 29.5°S, respectively, for the ALL ensemble mean. A scaling factor of 1.4×10^4 off-Chile and 1.6×10^4 is applied on the vertical velocity.

In contrast, the ocean processes in the SEP contribute to an anomalous warming trend (green dashed line in Fig. 13c). This can neither be explained by changes in vertical velocities or mixing, because these changes (upwelling tendency and larger warming near the surface, Fig. 13e) would rather be conducive to

cooling. This suggests that changes in lateral advection drive the warming. The SEP reduced warming with respect to its surrounding (Fig. 13a) implies a warming trend due to the advection of anomalous SST by mean surface currents. The currents' changes also involve a southward component (due to Ekman transport,

Fig. 13a) that increases the southward advection of warmer water from farther north.

7. Discussion and conclusions

A SST cooling trend is observed off northern Chile since 1980 and off southern Peru since the mid-twentieth century. Bakun et al. (2010) argue that surface pressure decreases more on land than over the ocean in response to anthropogenic forcing, giving rise to upwelling-favorable winds in eastern boundary upwelling systems (EBUS), which would explain the PCUS cooling trend. The observed negative interdecadal Pacific oscillation (IPO) trend over the same period is, however, also associated with cooling in the southeastern Pacific and could contribute to the observed trend. In this study, we quantify the relative role of internal and external forcing in this trend, and the associated mechanisms. Our ensemble-mean historical simulation reproduces a weaker recent SST trend in the southeastern part of the basin than in the surroundings, with virtually no warming off Chile since the 1970s, and a southerly wind anomaly at the coast and offshore.

In line with other studies based on CMIP5 models (e.g., Wang et al. 2015; Rykaczewski et al. 2015), we find no clear low pressure trend over South America that could lead to a southerly wind and upwelling intensification. The externally forced component of the simulated variability accounts for more than 60% of upwelling-favorable meridional wind stress increase off Chile, while a negative IPO canonical internal variability pattern has no significant contribution to the meridional wind stress positive trend over the same period. In the real world the IPO has a different spatial structure, amplitude and time behavior to that in CMIP5-class models (Kravtsov 2017). Analyses of observed decadal trends in SST, wind stresses and SLP separated into the IPO and forced components are however consistent with our model results. Hence, regional and local SST cooling patterns in the SEP and off Chile might be tightly linked to the anthropogenic external forcing signal, which potentially exerts a dominant role in reinforcing the coastal upwelling. PCUS can also be modulated by coastally trapped waves remotely forced by equatorial Pacific zonal wind fluctuations (Montecinos et al. 2007). This process may also contribute to the coastal upwelling enhancement.

Previous modeling work show that south of 33°S, the poleward South Pacific anticyclone migration causes geostrophy to break down due to the Andes orographic barrier (Muñoz and Garreaud 2005; Belmadani et al. 2014), locally precluding the establishment of the geostrophic equilibrium in the alongshore direction. This

probably explains why the increase in modeled alongshore winds is rather associated with an intensification of the alongshore pressure gradient. We further link the development of these southerly wind anomalies to a poleward expansion of the Southern Hemisphere Hadley cell, and strengthened South Pacific anticyclone, from around 1980 until a new steady state is reached in austral summer 2000 and in austral winter 2010. Attribution analyses demonstrate that enhanced alongshore winds and upwelling near the Chilean coast are due to greenhouse gases and ozone anthropogenic forcings, as both contribute to shift poleward the Hadley cell southern boundary (e.g., Staten et al. 2018; Grise et al. 2019). The large tendency in the 1980s and the relatively stable steady state later are consistent with the reduced influence of stratospheric ozone after 1995. The upwelling intensification is the main driver of the negative relative SST (RSST) trend along the Chilean coast. Farther offshore, previous studies suggest that the main mechanism for the robust negative RSST trend in the SEP region is the wind–evaporation–SST (WES) feedback (e.g., Xie et al. 2010; Timmermann et al. 2010; Lu et al. 2012) (i.e., that stronger winds enhance evaporative cooling locally). Instead, we find a dominant role of the wind-forced mixed layer deepening, which leads to a decreased warming rate by surface air–sea fluxes.

A potential caveat of our study is the coarse spatial resolutions of both the atmospheric (1.87° in latitude and 3.75° longitude) and ocean (2° horizontal resolution of Chile) components of the IPSL-CM5A coupled model. Our coarse model resolution (similar to CMIP5 class models) does not resolve several small-scale coastal processes, such as those associated with mesoscale oceanic eddies that contribute to the offshore transport of the cold upwelled water (e.g., Colas et al. 2013), coastal capes (e.g., Renault et al. 2016), sea breeze (Franchito et al. 1998) and the intensified temperature gradient induced by the warming of the narrow plains between the coast and the high Andes off Peru and northern Chile (e.g., Garreaud and Falvey 2009). Our results are nevertheless consistent with those obtained by Belmadani et al. (2014) that used an atmospheric model at higher horizontal resolution (50 km). Relying on dynamical downscaling of global GCM climate scenarios, Belmadani et al. (2014) obtain a wind strengthening off central Chile around 30°–35°S and a weakening or no changes off Peru in agreement with regional atmospheric modeling (horizontal resolution ~45 km) works by Quijano-Vargas (2011) and Muñoz and Garreaud (2005) for central Peru and north-central Chile coastal regions, respectively. These studies did not however include coupling with the ocean. Future work with regional coupled models at higher spatial resolution will be needed to confirm the robustness of our results.

The model and observations are consistent in terms of relative SST (i.e., in displaying a cooling relative to the regional average in the SEP and PCUS). However, in terms of absolute SST, the observations indicate a slight cooling trend in the PCUS region, while the model indicates a slight warming trend. This is most likely associated with the too strong climate sensitivity of our climate model (Andrews et al. 2012; Forster et al. 2013), which results in an overestimated surface global warming that offsets the regional cooling trend. This is accentuated by the fact that aerosol forcing (which induces a global cooling trend) is omitted in our simulation. We however probably partly remove this bias by using relative SST (i.e., subtracting the overestimated global mean SST warming). Our approach relying on relative SST, observations, and large ensembles needed to explore the regional responses to internal variability and to forcing histories nevertheless allows shedding light on the intrinsic dynamics of the southeast Pacific in recent decades. Upcoming CMIP6 simulations would offer the possibility to evaluate those results in a multimodel context.

We did not discuss changes along the tropical Peruvian coast (4° – 15° S) in detail since this region does not display any clear relative SST cooling nor significant changes in wind intensity in the ALL ensemble mean. Previous studies however suggest that SST trends off northern Peru are dominated by more frequent central Pacific El Niños after 1980 (Dewitte et al. 2012). Note that the analysis of observed winds also suggests no clear trend in the northern part of the Peruvian upwelling (Belmadani et al. 2014) despite the negative IPO trend during 1980–2014 (Vuille et al. 2015). Other eastern boundary upwelling systems (EBUS) of the Southern Hemisphere such as the Benguela system off southern Africa (e.g., Lamont et al. 2018) and the one west of Australia (e.g., Rousseaux et al. 2012) do not appear to show any cooling (Fig. 1). Observed SST trends during 1980–2014 in the Atlantic Ocean mainly capture a large tropical Atlantic warming (Li et al. 2016) while Indian Ocean SST display a warming trend that may reflect the global warming signal (Dong and McPhaden 2016; Fig. 1). Dedicated studies are needed to attribute the ocean expression of the Southern Hemisphere HC expansion in other southern EBUSs, which largely depends on the local features such as bottom topography, mixed layer depth, net surface fluxes, strength of the stratification, and regional ocean currents.

To summarize, we find that GHG increase and ozone depletion over the last few decades project onto robust southeasterly wind stress and negative RSST trend patterns in the southeastern Pacific. Anthropogenic forcings dominate this negative RSST trend in our model, as

almost all the members display a weaker cooling in the southeastern Pacific than in surrounding regions (and the observed RSST trend lying within the simulated distribution). Our results suggest that GHG and ozone anthropogenic forcing are likely the key driving factors for the recent observed PCUS intensification, and that the recent negative IPO phase played little or no role in this regional trend. Considering the continuous increase in atmospheric GHG and the unexpected and persistent global emissions of stratospheric ozone-depleting CFC-11 since 2012 (Montzka et al. 2018), our results provide practical guidance for the future climate response in the PCUS.

Acknowledgments. This research was supported by the French National Research Agency under the program Facing Societal, Climate and Environmental Changes (MORDICUS Project, Grant ANR-13-SENV-0002). This work was granted access to the HPC resources of TGCC under the allocations 2016-017403 and A003017403 made by GENCI. This study also benefited from the IPSL mesocenter facility, which is supported by CNRS, UPMC, Labex L-IPSL (funded by the ANR Grant ANR-10-LABX-0018 and by the European FP7 IS-ENES2 Grant 312979). We thank the ECMWF for providing the ERA-Interim reanalysis. The OISST dataset was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado. Support for the Twentieth Century Reanalysis Project, version 2c, dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. NCEP reanalysis data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado (<https://www.esrl.noaa.gov/psd/>).

REFERENCES

- Adam, O., T. Schneider, and N. Harnik, 2014: Role of changes in mean temperatures versus temperature gradients in the recent widening of the Hadley circulation. *J. Climate*, **27**, 7450–7461, <https://doi.org/10.1175/JCLI-D-14-00140.1>.
- Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott, 2002: The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Climate*, **15**, 2205–2231, [https://doi.org/10.1175/1520-0442\(2002\)015<2205:TABTIO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2).
- Allen, R. J., and O. Ajoku, 2016: Future aerosol reductions and widening of the northern tropical belt. *J. Geophys. Res. Atmos.*, **121**, 6765–6786, <https://doi.org/10.1002/2016JD024803>.
- Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere–ocean climate models. *Geophys. Res. Lett.*, **39**, L09712, <https://doi.org/10.1029/2012GL051607>.
- Arakelian, A., and F. Codron, 2012: Southern Hemisphere jet variability in the IPSL GCM at varying resolutions. *J. Atmos. Sci.*, **69**, 3788–3799, <https://doi.org/10.1175/JAS-D-12-0119.1>.

- Aravena, G., B. Broitman, and N. C. Stenseth, 2014: Twelve years of change in coastal upwelling along the central-northern coast of Chile: Spatially heterogeneous responses to climatic variability. *PLoS ONE*, **9**, e90276, <https://doi.org/10.1371/journal.pone.0090276>.
- Arblaster, J. M., and G. A. Meehl, 2006: Contributions of external forcings to southern annular mode trends. *J. Climate*, **19**, 2896–2905, <https://doi.org/10.1175/JCLI3774.1>.
- Bakun, A., D. B. Field, A. Redondo-Rodríguez, and S. J. Weeks, 2010: Greenhouse gas, upwelling-favorable winds, and the future of coastal ocean upwelling ecosystems. *Global Change Biol.*, **16**, 1213–1228, <https://doi.org/10.1111/j.1365-2486.2009.02094.x>.
- Balmaseda, M. A., K. E. Trenberth, and E. Källén, 2013: Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.*, **40**, 1754–1759, <https://doi.org/10.1002/grl.50382>.
- Barnes, E. A., and L. Polvani, 2013: Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models. *J. Climate*, **26**, 7117–7135, <https://doi.org/10.1175/JCLI-D-12-00536.1>.
- Barnett, T. P., D. W. Pierce, K. M. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory, and W. M. Washington, 2005: Penetration of human-induced warming into the world's oceans. *Science*, **309**, 284–287, <https://doi.org/10.1126/science.1112418>.
- Bayr, T., and D. Dommenges, 2013: The tropospheric land–sea warming contrast as the driver of tropical sea level pressure changes. *J. Climate*, **26**, 1387–1402, <https://doi.org/10.1175/JCLI-D-11-00731.1>.
- Beaugrand, G., and Coauthors, 2019: Prediction of unprecedented biological shifts in the global ocean. *Nat. Climate Change*, **9**, 237–243, <https://doi.org/10.1038/s41558-019-0420-1>.
- Bellenger, H., É. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2014: ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dyn.*, **42**, 1999–2018, <https://doi.org/10.1007/s00382-013-1783-z>.
- Belmadani, A., V. Echevin, F. Codron, K. Takahashi, and C. Junquas, 2014: What dynamics drive future wind scenarios for coastal upwelling off Peru and Chile? *Climate Dyn.*, **43**, 1893–1914, <https://doi.org/10.1007/s00382-013-2015-2>.
- Bjostoch, A., C. W. Böning, F. U. Schwarzkopf, and J. Lutjeharms, 2009: Increase in Agulhas leakage due to poleward shift of Southern Hemisphere westerlies. *Nature*, **462**, 495–498, <https://doi.org/10.1038/nature08519>.
- Boé, J., A. Hall, F. Colas, J. C. McWilliams, X. Qu, J. Kurian, and S. B. Kapnick, 2011: What shapes mesoscale wind anomalies in coastal upwelling zones? *Climate Dyn.*, **36**, 2037–2049, <https://doi.org/10.1007/s00382-011-1058-5>.
- Boer, G., G. Flato, and D. Ramsden, 2000: A transient climate change simulation with greenhouse gas and aerosol forcing: Projected climate to the twenty-first century. *Climate Dyn.*, **16**, 427–450, <https://doi.org/10.1007/s003820050338>.
- Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate*, **19**, 3445–3482, <https://doi.org/10.1175/JCLI3819.1>.
- Bopp, L., and Coauthors, 2013: Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, **10**, 6225–6245, <https://doi.org/10.5194/bg-10-6225-2013>.
- Brient, F., and S. Bony, 2013: Interpretation of the positive low-cloud feedback predicted by a climate model under global warming. *Climate Dyn.*, **40**, 2415–2431, <https://doi.org/10.1007/s00382-011-1279-7>.
- Cai, W., 2006: Antarctic ozone depletion causes an intensification of the Southern Ocean super-gyre circulation. *Geophys. Res. Lett.*, **33**, L03712, <https://doi.org/10.1029/2005GL024911>.
- Capet, X. J., P. Marchesio, and J. C. McWilliams, 2004: Upwelling response to coastal wind profiles. *Geophys. Res. Lett.*, **31**, L13311, <https://doi.org/10.1029/2004GL020123>.
- Carton, J. A., and B. S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Wea. Rev.*, **136**, 2999–3017, <https://doi.org/10.1175/2007MWR1978.1>.
- Chavez, F. P., A. Bertrand, R. Guevara-Carrasco, P. Soler, and J. Csirke, 2008: The northern Humboldt Current system: Brief history, present status and a view towards the future. *Prog. Oceanogr.*, **79**, 95–105, <https://doi.org/10.1016/j.pocan.2008.10.012>.
- Chen, S., K. Wei, W. Chen, and L. Song, 2014: Regional changes in the annual mean Hadley circulation in recent decades. *J. Geophys. Res. Atmos.*, **119**, 7815–7832, <https://doi.org/10.1002/2014JD021540>.
- Choi, J., S. W. Son, J. Lu, and S. K. Min, 2014: Further observational evidence of Hadley cell widening in the Southern Hemisphere. *Geophys. Res. Lett.*, **41**, 2590–2597, <https://doi.org/10.1002/2014GL059426>.
- Choi, J.-W., I.-G. Kim, J.-Y. Kim, and C.-H. Park, 2016: The recent strengthening of walker circulation. *SOLA*, **12**, 96–99, <https://doi.org/10.2151/SOLA.2016-022>.
- Colas, F., X. Capet, J. C. McWilliams, and Z. Li, 2013: Mesoscale eddy buoyancy flux and eddy-induced circulation in Eastern Boundary Currents. *J. Phys. Oceanogr.*, **43**, 1073–1095, <https://doi.org/10.1175/JPO-D-11-0241.1>.
- Collins, M., and Coauthors, 2005: El Niño- or La Niña-like climate change? *Climate Dyn.*, **24**, 89–104, <https://doi.org/10.1007/s00382-004-0478-x>.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, <https://doi.org/10.1002/qj.776>.
- Cushing, D. H., 1971: Upwelling and the production of fish. *Adv. Mar. Biol.*, **9**, 255–334, [https://doi.org/10.1016/S0065-2881\(08\)60344-2](https://doi.org/10.1016/S0065-2881(08)60344-2).
- de Boissésou, E., M. A. Balmaseda, S. Abdalla, E. Källén, and P. E. M. Janssen, 2014: How robust is the recent strengthening of the tropical Pacific trade winds? *Geophys. Res. Lett.*, **41**, 4398–4405, <https://doi.org/10.1002/2014GL060257>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Deser, C., A. S. Phillips, and J. W. Hurrell, 2004: Pacific interdecadal climate variability: Linkages between the tropics and the North Pacific during boreal winter since 1900. *J. Climate*, **17**, 3109–3124, [https://doi.org/10.1175/1520-0442\(2004\)017<3109:PICVLB>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<3109:PICVLB>2.0.CO;2).
- , M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.*, **2**, 115–143, <https://doi.org/10.1146/annurev-marine-120408-151453>.
- Dewitte, B., and Coauthors, 2012: Change in El Niño flavours over 1958–2008: Implications for the long-term trend of the upwelling off Peru. *Deep-Sea Res. II*, **77–80**, 143–156, <https://doi.org/10.1016/j.dsr2.2012.04.011>.
- Dommenges, D., 2009: The ocean's role in continental climate variability and change. *J. Climate*, **22**, 4939–4952, <https://doi.org/10.1175/2009JCLI2778.1>.
- Dong, L., and M. J. McPhaden, 2016: Interhemispheric SST gradient trends in the Indian Ocean prior to and during the recent global warming hiatus. *J. Climate*, **29**, 9077–9095, <https://doi.org/10.1175/JCLI-D-16-0130.1>.

- Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 earth system model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, <https://doi.org/10.1007/s00382-012-1636-1>.
- England, M. H., and Coauthors, 2014: Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nat. Climate Change*, **4**, 222–227, <https://doi.org/10.1038/nclimate2106>.
- Falvey, M., and R. D. Garreaud, 2009: Regional cooling in a warming world: Recent temperature trends in the southeast Pacific and along the west coast of subtropical South America (1979–2006). *J. Geophys. Res.*, **114**, D04102, <https://doi.org/10.1029/2008JD010519>.
- Fichefet, T., and M. Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res.*, **102**, 12 609–12 646, <https://doi.org/10.1029/97JC00480>.
- Fleming, L. E., and K. J. Anchukaitis, 2016: North Pacific decadal variability in the CMIP5 last millennium simulations. *Climate Dyn.*, **47**, 3783–3801, <https://doi.org/10.1007/s00382-016-3041-7>.
- Folland, C., D. Parker, A. Colman, and R. Washington, 1999: Large scale modes of ocean surface temperature since the late nineteenth century. *Beyond El Niño: Decadal and Interdecadal Climate Variability*, A. Navarra, Ed., Springer Science & Business Media, 73–102.
- Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res. Atmos.*, **118**, 1139–1150, <https://doi.org/10.1002/JGRD.50174>.
- Foster, G., and S. Rahmstorf, 2011: Global temperature evolution 1979–2010. *Environ. Res. Lett.*, **6**, 044022, <https://doi.org/10.1088/1748-9326/6/4/044022>.
- Franchito, S. H., V. B. Rao, J. L. Stech, and J. A. Lorenzetti, 1998: The effect of coastal upwelling on the sea-breeze circulation at Cabo Frio, Brazil: A numerical experiment. *Ann. Geophys.*, **16**, 866–871, <https://doi.org/10.1007/S00585-998-0866-3>.
- Fréon, P., M. Barange, and J. Aristegui, 2009: Eastern boundary upwelling ecosystems: Integrative and comparative approaches. *Prog. Oceanogr.*, **83**, 1–14, <https://doi.org/10.1016/j.pocean.2009.08.001>.
- Garreaud, R., and R. Muñoz, 2005: The low-level jet off the west coast of South America: Structure and variability. *Mon. Wea. Rev.*, **133**, 2246–2261, <https://doi.org/10.1175/MWR2972.1>.
- , and M. Falvey, 2009: The coastal winds off western subtropical South America in future climate scenarios. *Int. J. Climatol.*, **29**, 543–554, <https://doi.org/10.1002/joc.1716>.
- Gastineau, G., H. Le Treut, and L. Li, 2008: Hadley circulation changes under global warming conditions indicated by coupled climate models. *Tellus*, **60A**, 863–884, <https://doi.org/10.1111/j.1600-0870.2008.00344.x>.
- , A. Friedman, M. Khodri, and J. Vialard, 2018: Global ocean heat content redistribution during the 1998–2012 Interdecadal Pacific Oscillation negative phase. *Climate Dyn.*, **53**, 1187–1208, <https://doi.org/10.1007/S00382-018-4387-9%20>.
- Gillett, N. P., and D. W. Thompson, 2003: Simulation of recent Southern Hemisphere climate change. *Science*, **302**, 273–275, <https://doi.org/10.1126/science.1087440>.
- Grise, K. M., and Coauthors, 2019: Recent tropical expansion: Natural variability or forced response? *J. Climate*, **32**, 1551–1571, <https://doi.org/10.1175/JCLI-D-18-0444.1>.
- Gutiérrez, D., and Coauthors, 2011: Coastal cooling and increased productivity in the main upwelling zone off Peru since the mid-twentieth century. *Geophys. Res. Lett.*, **38**, L07603, <https://doi.org/10.1029/2010GL046324>.
- Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, <https://doi.org/10.1175/JCLI3990.1>.
- Henley, B. J., J. Gergis, D. J. Karoly, S. Power, J. Kennedy, and C. K. Folland, 2015: A tripole index for the Interdecadal Pacific Oscillation. *Climate Dyn.*, **45**, 3077–3090, <https://doi.org/10.1007/s00382-015-2525-1>.
- Hourdin, F., and Coauthors, 2013: LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dyn.*, **40**, 2193–2222, <https://doi.org/10.1007/s00382-012-1343-y>.
- Hu, Y., C. Zhou, and J. Liu, 2011: Observational evidence for poleward expansion of the Hadley circulation. *Adv. Atmos. Sci.*, **28**, 33–44, <https://doi.org/10.1007/s00376-010-0032-1>.
- , L. Tao, and J. Liu, 2013: Poleward expansion of the Hadley circulation in CMIP5 simulations. *Adv. Atmos. Sci.*, **30**, 790–795, <https://doi.org/10.1007/s00376-012-2187-4>.
- Huang, B., and Coauthors, 2015: Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4). Part I: Upgrades and intercomparisons. *J. Climate*, **28**, 911–930, <https://doi.org/10.1175/JCLI-D-14-00006.1>.
- Ibarra, A. A., C. Reid, and A. Thorpe, 2000: The political economy of marine fisheries development in Peru, Chile and Mexico. *J. Latin Amer. Stud.*, **32**, 503–527, <https://doi.org/10.1017/S0022216X00005824>.
- Johanson, C. M., and Q. Fu, 2009: Hadley cell widening: Model simulations versus observations. *J. Climate*, **22**, 2713–2725, <https://doi.org/10.1175/2008JCLI2620.1>.
- Joshi, M. M., J. M. Gregory, M. J. Webb, D. M. H. Sexton, and T. C. Johns, 2008: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Climate Dyn.*, **30**, 455–465, <https://doi.org/10.1007/s00382-007-0306-1>.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYP>2.0.CO;2).
- Kämpf, J., and P. Chapman, Eds., 2016: The Peruvian-Chilean coastal upwelling system. *Upwelling Systems of the World*, Springer, 161–194.
- Kim, Y. H., S. K. Min, S. W. Son, and J. Choi, 2017: Attribution of the local Hadley cell widening in the Southern Hemisphere. *Geophys. Res. Lett.*, **44**, 1015–1024, <https://doi.org/10.1002/2016GL072353>.
- Knutson, T., J. P. Kossin, C. Mears, J. Perlwitz, and M. F. Wehner, 2017: Detection and attribution of climate change. *Climate Science Special Report: Fourth National Climate Assessment*, D. J. Wuebbles et al., Eds., Vol. I, U.S. Global Change Research Program, 114–132.
- Kociuba, G., and S. B. Power, 2015: Inability of CMIP5 models to simulate recent strengthening of the walker circulation: Implications for projections. *J. Climate*, **28**, 20–35, <https://doi.org/10.1175/JCLI-D-13-00752.1>.
- Kohonen, T., 1998: The self-organizing map. *Neurocomputing*, **21**, 1–6, [https://doi.org/10.1016/S0925-2312\(98\)00030-7](https://doi.org/10.1016/S0925-2312(98)00030-7).
- Kohyama, T., and D. L. Hartmann, 2017: Nonlinear ENSO Warming Suppression (NEWS). *J. Climate*, **30**, 4227–4251, <https://doi.org/10.1175/JCLI-D-16-0541.1>.
- Kravtsov, S., 2017: Pronounced differences between observed and CMIP5-simulated multidecadal climate variability in the twentieth century. *Geophys. Res. Lett.*, **44**, 5749–5757, <https://doi.org/10.1002/2017GL074016>.

- Krinner, G., and Coauthors, 2005: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochem. Cycles*, **19**, GB1015, <https://doi.org/10.1029/2003GB002199>.
- Lamont, T., M. García-Reyes, S. J. Bograd, C. D. van der Lingen, and W. J. Sydeman, 2018: Upwelling indices for comparative ecosystem studies: Variability in the Benguela upwelling system. *J. Mar. Syst.*, **188**, 3–16, <https://doi.org/10.1016/j.jmarsys.2017.05.007>.
- L'Heureux, M., S. Lee, and B. Lyon, 2013: Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nat. Climate Change*, **3**, 571–576, <https://doi.org/10.1038/nclimate1840>.
- Li, G., and S.-P. Xie, 2014: Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems. *J. Climate*, **27**, 1765–1780, <https://doi.org/10.1175/JCLI-D-13-00337.1>.
- Li, X., S. P. Xie, S. T. Gille, and C. Yoo, 2016: Atlantic-induced pan-tropical climate change over the past three decades. *Nat. Climate Change*, **6**, 275–279, <https://doi.org/10.1038/nclimate2840>.
- Lu, J., and B. Zhao, 2012: The role of oceanic feedback in the climate response to doubling CO₂. *J. Climate*, **25**, 7544–7563, <https://doi.org/10.1175/JCLI-D-11-00712.1>.
- , G. A. Vecchi, and T. Reichler, 2007: Expansion of the Hadley cell under global warming. *Geophys. Res. Lett.*, **34**, L06805, <https://doi.org/10.1029/2006GL028443>.
- Luo, J. J., W. Sasaki, and Y. Masumoto, 2012: Indian Ocean warming modulates Pacific climate change. *Proc. Natl. Acad. Sci. USA*, **109**, 18 701–18 706, <https://doi.org/10.1073/pnas.1210239109>.
- Ma, S., and T. Zhou, 2016: Robust strengthening and westward shift of the tropical Pacific Walker circulation during 1979–2012: A comparison of 7 sets of reanalysis data and 26 CMIP5 models. *J. Climate*, **29**, 3097–3118, <https://doi.org/10.1175/JCLI-D-15-0398.1>.
- Madec, G., 2008: NEMO Ocean engine. Institut Pierre-Simon Laplace (IPSL), Note du Pole de Modélisation 27, 300 pp.
- McLandress, C., T. G. Shepherd, J. F. Scinocca, D. A. Plummer, M. Sigmond, A. I. Jonsson, and M. C. Reader, 2011: Separating the dynamical effects of climate change and ozone depletion. Part II: Southern Hemisphere troposphere. *J. Climate*, **24**, 1850–1868, <https://doi.org/10.1175/2010JCLI3958.1>.
- McPhaden, M. J., S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an integrating concept in earth science. *Science*, **314**, 1740–1745, <https://doi.org/10.1126/science.1132588>.
- Min, S. K., and S. W. Son, 2013: Multimodel attribution of the Southern Hemisphere Hadley cell widening: Major role of ozone depletion. *J. Geophys. Res. Atmos.*, **118**, 3007–3015, <https://doi.org/10.1002/JGRD.50232>.
- Minobe, S., 1997: A 50–70 year climatic oscillation over the North Pacific and North America. *Geophys. Res. Lett.*, **24**, 683–686, <https://doi.org/10.1029/97GL00504>.
- , 1999: Resonance in bidecadal and pentadecadal climate oscillations over the North Pacific: Role in climatic regime shifts. *Geophys. Res. Lett.*, **26**, 855–858, <https://doi.org/10.1029/1999GL001119>.
- Mitas, C. M., and A. Clement, 2005: Has the Hadley cell been strengthening in recent decades? *Geophys. Res. Lett.*, **32**, L03809, <https://doi.org/10.1029/2004GL021765>.
- Montecinos, A., S. Purca, and O. Pizarro, 2003: Interannual-to-interdecadal sea surface temperature variability along the western coast of South America. *Geophys. Res. Lett.*, **30**, 1570, <https://doi.org/10.1029/2003GL017345>.
- , O. Leth, and O. Pizarro, 2007: Wind-driven interdecadal variability in the eastern tropical South Pacific. *J. Geophys. Res.*, **112**, C04019, <https://doi.org/10.1029/2006JC003571>.
- Montzka, S. A., and Coauthors, 2018: An unexpected and persistent increase in global emissions of ozone-depleting CFC-11. *Nature*, **557**, 413–417, <https://doi.org/10.1038/s41586-018-0106-2>.
- Muñoz, R. C., and R. Garreaud, 2005: Dynamics of the low-level jet off the west coast of subtropical South America. *Mon. Wea. Rev.*, **133**, 3661–3677, <https://doi.org/10.1175/MWR3074.1>.
- Newman, M., and Coauthors, 2016: The Pacific decadal oscillation, revisited. *J. Climate*, **29**, 4399–4427, <https://doi.org/10.1175/JCLI-D-15-0508.1>.
- Nguyen, H., A. Evans, C. Lucas I. Smith, and B. Timbal, 2013: The Hadley circulation in reanalyses: Climatology, variability, and change. *J. Climate*, **26**, 3357–3376, <https://doi.org/10.1175/JCLI-D-12-00224.1>.
- , C. Lucas, A. Evans, B. Timbal, and L. Hanson, 2015: Expansion of the Southern Hemisphere Hadley cell in response to greenhouse gas forcing. *J. Climate*, **28**, 8067–8077, <https://doi.org/10.1175/JCLI-D-15-0139.1>.
- Nidheesh, A., M. Lengaigne, J. Vialard, T. Izumo, A. Unnikrishnan, and C. Cassou, 2017: Influence of ENSO on the Pacific decadal oscillation in CMIP models. *Climate Dyn.*, **49**, 3309–3326, <https://doi.org/10.1007/s00382-016-3514-8>.
- Oyarzún, D., and C. M. Brierley, 2019: The future of coastal upwelling in the Humboldt current from model projections. *Climate Dyn.*, **52**, 599–615, <https://doi.org/10.1007/s00382-018-4158-7>.
- Pérez, F. F., and Coauthors, 2010: Plankton response to weakening of the Iberian coastal upwelling. *Global Change Biol.*, **16**, 1258–1267, <https://doi.org/10.1111/j.1365-2486.2009.02125.x>.
- Plesca, E., V. Grützun, and S. A. Buehler, 2018: How robust is the weakening of the Pacific Walker circulation in CMIP5 idealized transient climate simulations? *J. Climate*, **31**, 81–97, <https://doi.org/10.1175/JCLI-D-17-0151.1>.
- Power, S., T. Casey, C. Folland, A. Colman, and V. Mehta, 1999: Interdecadal modulation of the impact of ENSO on Australia. *Climate Dyn.*, **15**, 319–324, <https://doi.org/10.1007/s003820050284>.
- Quan, X.-W., M. P. Hoerling, J. Perlwitz, H. F. Diaz, and T. Xu, 2014: How fast are the tropics expanding? *J. Climate*, **27**, 1999–2013, <https://doi.org/10.1175/JCLI-D-13-00287.1>.
- Quijano-Vargas, J. J., 2011: Simulación de la dinámica del viento superficial sobre la costa de Ica utilizando el modelo numérico de la atmósfera de mesoescala MM5. M. S. thesis in Engineering in Fluid Mechanics, Universidad Nacional Mayor de San Marcos, 172 pp.
- Renault, L., C. Deutsch, J. C. McWilliams, H. Frenzel, J. H. Liang, and F. Colas, 2016: Partial decoupling of primary productivity from upwelling in the California Current system. *Nat. Geosci.*, **9**, 505–508, <https://doi.org/10.1038/ngeo2722>.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, [https://doi.org/10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2).
- Riahi, K., and Coauthors, 2011: RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, **109**, 33–57, <https://doi.org/10.1007/s10584-011-0149-y>.
- Roemmich, D., J. Gilson, R. Davis, P. Sutton, S. Wijffels, and S. Riser, 2007: Decadal spinup of the South Pacific subtropical gyre. *J. Phys. Oceanogr.*, **37**, 162–173, <https://doi.org/10.1175/JPO3004.1>.
- Rousseaux, C. S. G., R. J. Lowe, M. Feng, A. M. Waite, and P. A. Thompson, 2012: The role of the Leeuwin Current and mixed

- layer depth on the autumn phytoplankton bloom off Ningaloo Reef, Western Australia. *Cont. Shelf Res.*, **32**, 22–35, <https://doi.org/10.1016/j.csr.2011.10.010>.
- Ryckaczewski, R. R., J. P. Dunne, W. J. Sydeman, M. García-Reyes, B. A. Black, and S. J. Bograd, 2015: Poleward displacement of coastal upwelling-favorable winds in the ocean's eastern boundary currents through the 21st century. *Geophys. Res. Lett.*, **42**, 6424–6431, <https://doi.org/10.1002/2015GL064694>.
- Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack, 1993: Stratospheric aerosol optical depths, 1850–1990. *J. Geophys. Res.*, **98**, 22 987–22 994, <https://doi.org/10.1029/93JD02553>.
- Seidel, D. J., and W. J. Randel, 2007: Recent widening of the tropical belt: Evidence from tropopause observations. *J. Geophys. Res.*, **112**, D20113, <https://doi.org/10.1029/2007JD008861>.
- Sen Gupta, A., and M. H. England, 2006: Coupled ocean–atmosphere–ice response to variations in the southern annular mode. *J. Climate*, **19**, 4457–4486, <https://doi.org/10.1175/JCLI3843.1>.
- Small, R. J., E. Curchitser, K. Hedstrom, B. Kauffman, and W. G. Large, 2015: The Benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model. *J. Climate*, **28**, 9409–9432, <https://doi.org/10.1175/JCLI-D-15-0192.1>.
- Solomon, A., and M. Newman, 2012: Reconciling disparate twentieth century Indo-Pacific Ocean temperature trends in the instrumental record. *Nat. Climate Change*, **2**, 691–699, <https://doi.org/10.1038/nclimate1591>.
- Son, S. W., and Coauthors, 2010: Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment. *J. Geophys. Res.*, **115**, D00M07, <https://doi.org/10.1029/2010JD014271>.
- Staten, P. W., J. J. Rutz, T. Reichler, and J. Lu, 2012: Breaking down the tropospheric circulation response by forcing. *Climate Dyn.*, **39**, 2361–2375, <https://doi.org/10.1007/s00382-011-1267-y>.
- , J. Lu, K. M. Grise, S. M. Davis, and T. Birner, 2018: Re-examining tropical expansion. *Nat. Climate Change*, **8**, 768–775, <https://doi.org/10.1038/s41558-018-0246-2>.
- Stouffer, R., S. Manabe, and K. Bryan, 1989: Interhemispheric asymmetry in climate response to a gradual increase of atmospheric CO₂. *Nature*, **342**, 660–662, <https://doi.org/10.1038/342660a0>.
- Strub, P. T., J. Mesías, V. Montecino, J. Ruttlant, and S. Salinas, 1998: Coastal ocean circulation off western South America Coastal Segment (6, E). *The Global Coastal Ocean: Regional Studies and Syntheses*, Robinson, R. Allan, and K. H. Brink, Eds., *The Sea—Ideas and Observations on Progress in the Study of the Seas*, Vol. 11, John Wiley and Sons, 273–314.
- Sverdrup, H. U., 1938: On the process of upwelling. *J. Mar. Res.*, **1**, 155–164, <https://doi.org/10.1357/002224038806440638>.
- Szopa, S., and Coauthors, 2013: Aerosol and ozone changes as forcing for climate evolution between 1850 and 2100. *Climate Dyn.*, **40**, 2223–2250, <https://doi.org/10.1007/s00382-012-1408-y>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Thompson, D. W., and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate change. *Science*, **296**, 895–899, <https://doi.org/10.1126/science.1069270>.
- Timmermann, A., S. McGregor, and F. Jin, 2010: Wind effects on past and future regional sea level trends in the southern Indo-Pacific. *J. Climate*, **23**, 4429–4437, <https://doi.org/10.1175/2010JCLI3519.1>.
- Tokinaga, H., S.-P. Xie, A. Timmermann, S. McGregor, T. Ogata, H. Kubota, and Y. M. Okumura, 2011: Regional patterns of tropical Indo-Pacific climate change: Evidence of the Walker circulation weakening. *J. Climate*, **25**, 1689–1710, <https://doi.org/10.1175/JCLI-D-11-00263.1>.
- Vargas, G., S. Pantoja, J. A. Ruttlant, C. B. Lange, and L. Ortlieb, 2007: Enhancement of coastal upwelling and interdecadal ENSO-like variability in the Peru–Chile current since late 19th century. *Geophys. Res. Lett.*, **34**, L13607, <https://doi.org/10.1029/2006GL028812>.
- Vecchi, G. A., and B. J. Soden, 2007: Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature*, **450**, 1066–1070, <https://doi.org/10.1038/nature06423>.
- , A. Clement, and B. J. Soden, 2008: Examining the tropical Pacific's response to global warming. *Eos, Trans. Amer. Geophys. Union*, **89**, 81–83, <https://doi.org/10.1029/2008EO090002>.
- Vigneau, E., and E. M. Qannari, 2003: Clustering of variables around latent components. *Commun. Stat. Simul. Comput.*, **32**, 1131–1150, <https://doi.org/10.1081/SAC-120023882>.
- Vuille, M., E. Franquist, R. Garreaud, W. Lavado, and C. Bolívar Cáceres, 2015: Impact of the global warming hiatus on Andean temperature. *J. Geophys. Res. Atmos.*, **120**, 3745–3757, <https://doi.org/10.1002/2015JD023126>.
- Wang, D., T. C. Gouhier, B. A. Menge, and A. R. Ganguly, 2015: Intensification and spatial homogenization of coastal upwelling under climate change. *Nature*, **518**, 390–394, <https://doi.org/10.1038/nature14235>.
- Waugh, D. W., C. I. Garfinkel, and L. M. Polvani, 2015: Drivers of the recent tropical expansion in the Southern Hemisphere: Changing SSTs or ozone depletion? *J. Climate*, **28**, 6581–6586, <https://doi.org/10.1175/JCLI-D-15-0138.1>.
- Xie, S.-P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global warming pattern formation: Sea surface temperature and rainfall. *J. Climate*, **23**, 966–986, <https://doi.org/10.1175/2009JCLI3329.1>.
- , and G. Philander, 1994: A coupled ocean–atmosphere model of relevance to the ITCZ in the eastern Pacific. *Tellus*, **46A**, 340–350, <https://doi.org/10.3402/tellusa.v46i4.15484>.

3.5 Expressions océaniques de l'expansion de la cellule de Hadley

Les résultats discutés dans les sections précédentes posent également la question de la réponse océanique à l'expansion des cellules Hadley et au mouvement vers les pôles des vents d'ouest à travers l'Hémisphère Sud. En particulier, nous n'avons pas discuté en détail des changements le long de la côte tropicale péruvienne (4°S-15°S), car cette région ne présente pas de refroidissement relatif clair des SSTs ni de changement significatif de l'intensité du vent (Dewitte et al., 2012; Espinoza-Morriberón et al., 2017). Il est à noter que l'analyse des vents observés ne suggère pas non plus de tendance claire dans la partie nord de l'upwelling péruvien (Belmadani et al., 2014) malgré la tendance négative de l'IPO entre 1980-2014 (Vuille et al., 2015). Cela est également vrai pour d'autres systèmes d'upwelling de bord Est (en anglais, Eastern Boundary Upwelling Systems, EBUS) de l'Hémisphère Sud, tels que le système de Benguela au large de l'Afrique australe (par exemple Lamont et al., 2018) et celui à l'ouest de l'Australie (par exemple Rousseaux et al., 2012), qui ne semblent montrer de tendance au refroidissement. Les tendances observées des SSTs au cours de la période 1980-2014 dans l'océan Atlantique, capturent principalement le signal à grande échelle de la phase de l'AMV (Deser et al., 2010) tandis que les SSTs de l'océan Indien indiquent une tendance au réchauffement qui peut refléter le signal du réchauffement climatique (Dong and McPhaden, 2016).

La figure 3.9 fournit une tentative d'estimer dans différents jeux observations de la période 1980-2014, comment l'expansion vers les pôles de la cellule Hadley peut influencer les d'autres EBUS de l'Hémisphère Sud. Les figures 3.9b-c indiquent un déplacement vers les régions polaires des centres de haute pression et des gyres océaniques subtropicaux du Pacifique Sud, de l'Atlantique Sud et de l'océan Indien. Ces résultats sont cohérents avec des études précédentes qui ont signalé le déplacement vers les pôles des courants de bord Ouest (Yang et al., 2016; Seager and Simpson, 2016) et des changements de la circulation océanique dans l'Hémisphère Sud en réponse à la tendance positive du Mode Annulaire Sud (en anglais Southern Annular Mode, SAM ; Sen Gupta and England, 2006; Cai, 2006; Roemmich et al., 2007). Etant donné que l'expansion de la cellule de Hadley et la tendance positive du SAM sont corrélées (Nguyen et al., 2013) et toutes deux induites par l'appauvrissement de la couche d'ozone antarctique (par exemple, Thompson and Solomon, 2002; Gillett and Thompson, 2003; Arblaster and Meehl, 2006; Son et al., 2010), il n'est pas surprenant qu'ils projettent de façon similaire un déplacement vers les pôles et un spin-up des gyres océaniques subtropicaux.

Malgré certaines différences entre les jeux de données, des caractéristiques cohérentes émergent aux moyennes latitudes, avec un renforcement significatif du stress du vent méridien au sud de 30°S, sur les bordures polaires des EBUS de tous les bassins océaniques (Fig. 3.9c-d). Sur les bordures équatoriales des EBUS, les vents sont en revanche affaiblis au nord de 30°S. Ceci pourrait expliquer pourquoi aucun vent favorable aux upwellings n'est observé au nord du Pérou (4°S-15°S) et au nord de Benguela (15°S – 29°S) malgré les tendances de l'IPO et AMV pendant cette période.

Plus au sud, au niveau de l'upwelling du sud Benguela (29–34,3° S; 12–18,5° E) et

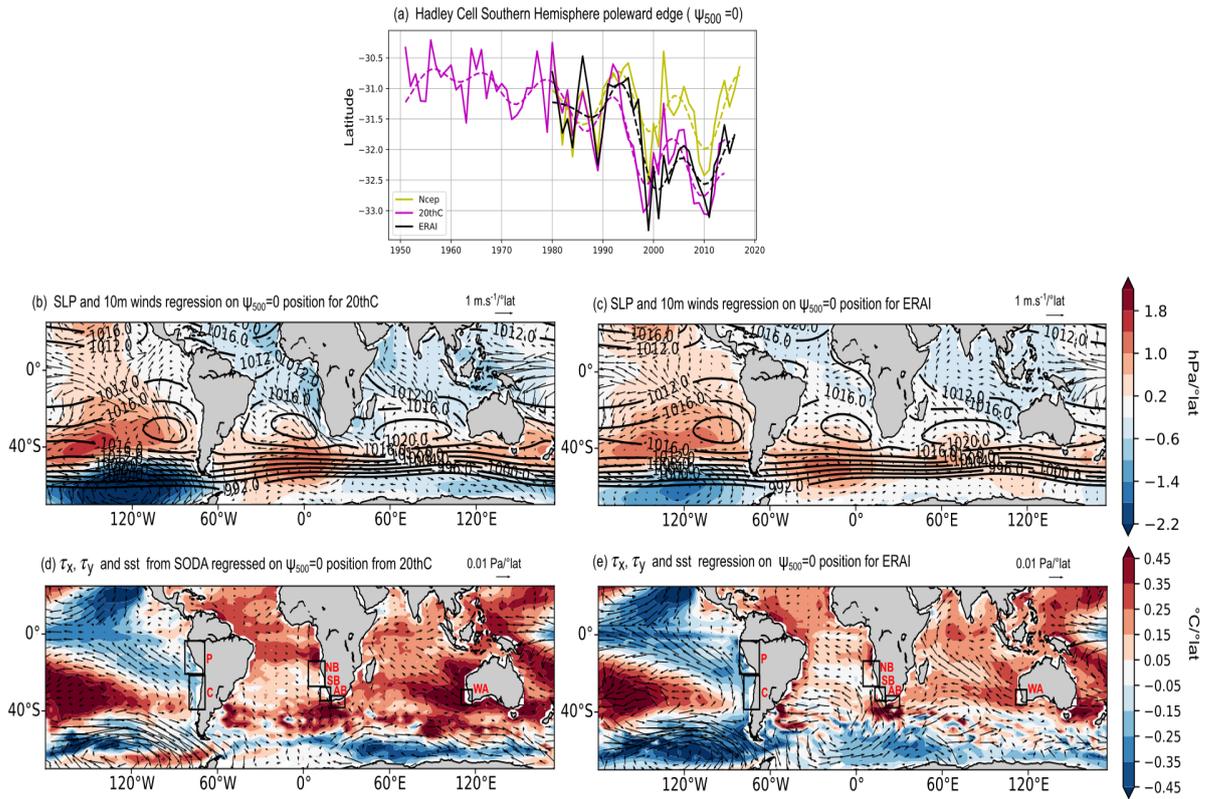


Figure 3.9: (a) Séries annuelles de la position de la limite sud de la cellule de Hadley dans les réanalyses NCEP (en jaune), NOAA-CIRES 20CR (en violet) et ERA-Interim (ERAI; en noir) pour la période 1950-2014. (b) Régression linéaire de la pression au niveau de la mer (SLP, en couleur; en hPa/degrés de latitude) et des vents à 10 mètres (vecteurs, en m.s^{-1} /degrés de latitude) sur la position de la limite sud de la cellule de Hadley de NOAA-CIRES 20CR. La climatologie moyenne de la SLP pour la période 1980-2014 est indiquée par les contours noirs (tous les 4 hPa). (c) comme (b) mais pour ERAI. (d) Régression linéaire de la SST (en couleur; en $^{\circ}\text{C}$ /degrés de latitude) et du stress du vent à la surface (vecteurs; en 10^{-2} Pa/degrés de latitude) de la réanalyse océanique SODA3.3.1 avec la position de la limite sud de la cellule de Hadley de NOAA-CIRES 20CR. (e) Comme (d) mais pour ERAI. Les régressions des panneaux (b-e) sont réalisées sur la période 1980-2014. Les boîtes noires sur les panneaux (d-e) localisent les upwelling de bord Est de l'Hémisphère Sud, **P** pour Pérou [4°S – 25°S]; **C** pour Chili [25°S – 40°S], **NB** pour Northern Benguella [15 – 29°S ; 8 – 16.8°E]; **SB** pour Southern Benguella [15°S – 29°S]; **AB** pour Agulhas Benguella [33.3 – 37.5°S ; 18.5 – 27°E] et le long des côtes Australiennes **WA** pour West Australia [30°S – 35°S].

l'Agulhas Benguella ($33,3$ – $37,5^{\circ}\text{S}$; $18,5$ – 27°E), un renforcement du stress du vent méridien est observé le long du plateau continental de l'Afrique du Sud (Fig. 3.9c-d). Le faible refroidissement relatif voir l'augmentation des SSTs au large de l'Afrique australe pourrait donc être dominé par le renforcement du courant des Auguilles (en anglais, Agulhas leakage) en provenance de l'océan Indien chaud, en réponse aux déplacements latitudinaux des vents d'ouest au cours de la même période (Biaستoch et al., 2009). Le réchauffement au large de la région à l'Ouest de l'Australie pourrait également être influencé par des courants océaniques régionaux tels que l'« Indian throughflow » (Dong and McPhaden, 2016) ou les courants est-australiens (Yang et al., 2016), qui se sont tous deux renforcés au cours des dernières décennies. Des études dédiées sont toutefois nécessaires pour confirmer ces hypothèses et attribuer l'expression océanique de l'expansion de la Cellule

de Hadley dans d'autres EBUS de l'Hémisphère Sud et qui dépendra de la topographie régionale, de la profondeur de la couche de mélange, des flux nets de surface, de la stratification et des courants océaniques régionaux.

Partie III

Reconstruction de la variabilité climatique des 2000 dernières années

Chapitre 4

Modélisation et méthodes d'assimilation de données paléoclimatiques

Sommaire

Notations	89
4.1 Préambule	90
4.2 PAGES2k Consortium	90
4.3 L'assimilation de données en contexte paléoclimatique	91
4.3.1 Approches et méthodes d'assimilations de données	91
4.3.2 Approches séquentielles et variationnelles	93
4.3.3 La règle de Bayes	94
4.3.4 Filtrage Bayésien	94
4.4 Les filtres particuliers	95
4.4.1 Le filtre particuliers séquentiel pondéré	95
4.4.2 Le filtre particuliers avec ré-échantillonnage	97
4.4.3 Méthode de ré-échantillonnage	98
4.4.4 Taille effective de l'échantillon N_{eff}	99
4.4.5 En pratique	100
4.5 Le LIM	100
4.5.1 Présentation	100
4.5.2 Définition du modèle LIM	101
4.5.3 Le LIM couplé	102
4.5.4 Le LIM en tant que modèle d'intégration	103
4.6 Indices de performances	104

Notations

\mathcal{R}	système dynamique réel
χ	système dynamique réel
\mathbf{x}	vecteur d'état d'un sous-système dynamique de \mathcal{R} de dimension \mathcal{H}
\mathbf{y}	vecteur d'observations
\mathcal{H}	dimension du vecteur d'état \mathbf{x}
\mathbf{x}_k	vecteur d'état \mathbf{x} au pas de temps k
\mathcal{M}	modèle numérique reproduisant la dynamique des \mathbf{x}_k
p^i	particule i
N	nombre de particules
\mathbf{x}_k^i	vecteur d'état de la particule i à l'instant k
w_k^i	poids associé à la particules i à l'instant k
$\delta_x(\cdot)$	mesure de Dirac centré au point x
\mathbf{y}_k	observations disponibles au pas de temps k
$\mathbf{y}_{1:k}$	suite des observations jusqu'au pas de temps k
$\beta(\cdot)$	composante linéaire
$\eta(\cdot)$	composante non linéaire
$f(\cdot)$	forçage externe
\mathbf{B}	matrice de covariance
ε	bruit Gaussien
\mathbf{Q}	matrice de covariance du bruit
\mathbf{G}	propagateur linéaire
$\sigma(t)$	bruit Gaussien généré à l'instant t
$\mathbf{C}(\tau)$	matrice de covariance au lag τ
q_{ik}	i^{eme} coefficient du k^{eme} vecteur propre de la matrice \mathbf{Q}
n_k	k^{eme} valeur propre de la matrice \mathbf{Q}
N_{eff}	taille effective de l'échantillon

4.1 Préambule

Les observations instrumentales permettent de documenter la variabilité climatique sur une période relativement courte et soumise à l'effet des forçages anthropiques. Comme nous l'avons vu dans le chapitre précédent, contraindre le réalisme des modèles, démêler le rôle relatif des forçages externes sur les modes décennaux et les tendances récentes à l'échelle régionale demeure encore aujourd'hui un défi. Il est toutefois possible d'augmenter de façon significative l'échantillonnage et la compréhension physique de la variabilité naturelle à l'aide de reconstructions paléoclimatiques dans les modèles.

Une des approches utilisées consiste à combiner de façon optimale via des méthodes d'assimilation de données, les connaissances actuelles du climat et sa variabilité avec les informations climatiques contenues dans les archives naturelles, tout en tenant compte de leurs incertitudes (Van Leeuwen, 2009; Hakim et al., 2013). Bien que couramment utilisées sur la période instrumentale (Talagrand, 1997), les progrès de nos connaissances dans le domaine des sciences du climat, des méthodes statistiques ainsi que l'augmentation de la puissance de calcul rendent aujourd'hui envisageable le développement de ce type d'approches en paléoclimats.

Les adaptations nécessaires pour l'assimilation de proxys, dont la précision est généralement limitée (Jones and Mann, 2004) et la répartition sur le globe inégale, progressent rapidement (Widmann et al., 2010; Hakim et al., 2013). Elles ont déjà permis la reconstruction d'un certains nombres d'indices climatiques, tels que des moyennes globales ou hémisphériques, (Moberg et al., 2005; Mann et al., 2009; Goosse et al., 2010; Annan and Hargreaves, 2013; PAGES2k Consortium, 2013) voire même plus récemment des reconstructions spatiales (Hakim et al., 2016; Perkins and Hakim, 2017). Ces développements récents souffrent néanmoins d'un certain nombre de limitations qu'il convient d'évaluer.

L'objectif de ce chapitre est d'introduire les outils et principes théoriques qui seront utilisés dans le chapitre 5 suivant, qui présente et teste une nouvelle méthode d'assimilation de données paléoclimatiques dans le modèle de l'IPSL. Le présent chapitre est organisé de la façon suivante: nous présentons dans un premier temps la base de données Pages2k de proxys disponibles pour l'assimilation (section 4.2). Nous dressons ensuite un bref état de l'art des principales méthodes utilisées actuellement dans le cadre de reconstructions paléoclimatiques et introduisons le concept des méthodes d'assimilation de données (section 4.3). Nous détaillons en particulier les algorithmes des méthodes de filtres particulières (section 4.4). Nous introduisons ensuite les modèles stochastiques linéaires (LIM pour « Linear Inverse Model ») (section 4.5) et pour finir (section 4.6) nous présentons les critères de performances utilisés dans le chapitre suivant.

4.2 PAGES2k Consortium

Les proxys documentant les 2000 dernières années et compilés par le Consortium PAGES2k (avec PAGES pour Past Global Change ; PAGES2k Consortium, 2013; PAGES2k Consortium, 2019) constituent des preuves de la gamme très étendue de la variabilité naturelle du

climat et largement sous observée par la fenêtre instrumentale des 100 dernières années. Il est par exemple admis, malgré quelques incertitudes sur l'amplitude et la période exacte, que le climat de l'Hémisphère Nord a connu des conditions relativement douces durant l'« optimum médiéval » entre 950 et 1250, suivi d'une période plus froide le « petit âge de glace » entre 1450 et 1850 (Mann et al., 2009; Gebbie and Huybers, 2019). Le réseau PAGES2k constitué de 9 groupes de travail distincts a pour objectif de permettre la création et la distribution d'une base de données publique de proxys utiles pour l'étude de la variabilité des températures de surface, précipitation et autres données environnementales au cours des 2000 dernières années. La version 2 regroupe un ensemble de 692 enregistrements provenant de 648 emplacements dans 49 pays différents. Les données proviennent de 11 types d'archives distincts (arbres, glace, sédiments, coraux, etc.) avec une majorité issue de cernes d'arbre (415 sur les 692 enregistrements). Avec une moyenne de 760 ans, les données peuvent couvrir des fenêtres temporelles allant de quelques décennies à 2000 ans. Leur résolution temporelle varie également selon les différents types d'archive avec une majorité d'archives possédant une résolution temporelle de 1 an (Fig. 4.1). Le contrôle qualité effectué pour la sélection des données est décrit dans (PAGES2k Consortium, 2017). En particulier un des critères primordial repose sur une bonne retranscription (selon le type de variable considérée) de la variabilité climatique observée sur la période instrumentale (par exemple Morice et al., 2012).

4.3 L'assimilation de données en contexte paléoclimatique

4.3.1 Approches et méthodes d'assimilations de données

Deux types d'approche peuvent être utilisés en contexte paléoclimatique: les approches dites «offline» et «online» (Matsikaris et al., 2015). Tandis que les approches offline utilisent un ensemble de simulations déjà réalisées avant l'analyse pour reconstruire à posteriori le champ souhaité (Goosse et al., 2006b; Hakim et al., 2016), les approches online orientent quant à elles la trajectoire de la ou des simulations durant l'analyse en fonction des observations disponibles (Fig. 4.2; Goosse et al., 2010; Liu et al., 2017)

Dans les approches online, les simulations sont régulièrement ramenées vers la cible, c'est à dire les observations disponibles. Elles sont ensuite au pas de temps suivant propagées à partir de ces états corrigés (Fig. 4.2). En théorie les approches online sont préférables pour permettre une continuité et une cohérence physique le long de l'assimilation. En pratique le choix entre ces deux approches dépendra du système étudié (Matsikaris et al., 2015). Les approches online sont conseillées dans le cas de systèmes prédictibles, alors que les approches offline, plus simple à réaliser, sont utiles dans le cas de système peu ou pas prédictible.

Parmi les nombreuses méthodes testées (Fierro et al., 1997; Mann et al., 1999; Storch et al., 2000; Barkmeijer et al., 2003; Smerdon et al., 2010; Annan and Hargreaves, 2013; Guillot et al., 2015), les plus populaires sont les méthodes de filtres de Kalman d'ensemble et filtres particuliers. Testées aussi bien en version online (Liu et al., 2017) qu'en version

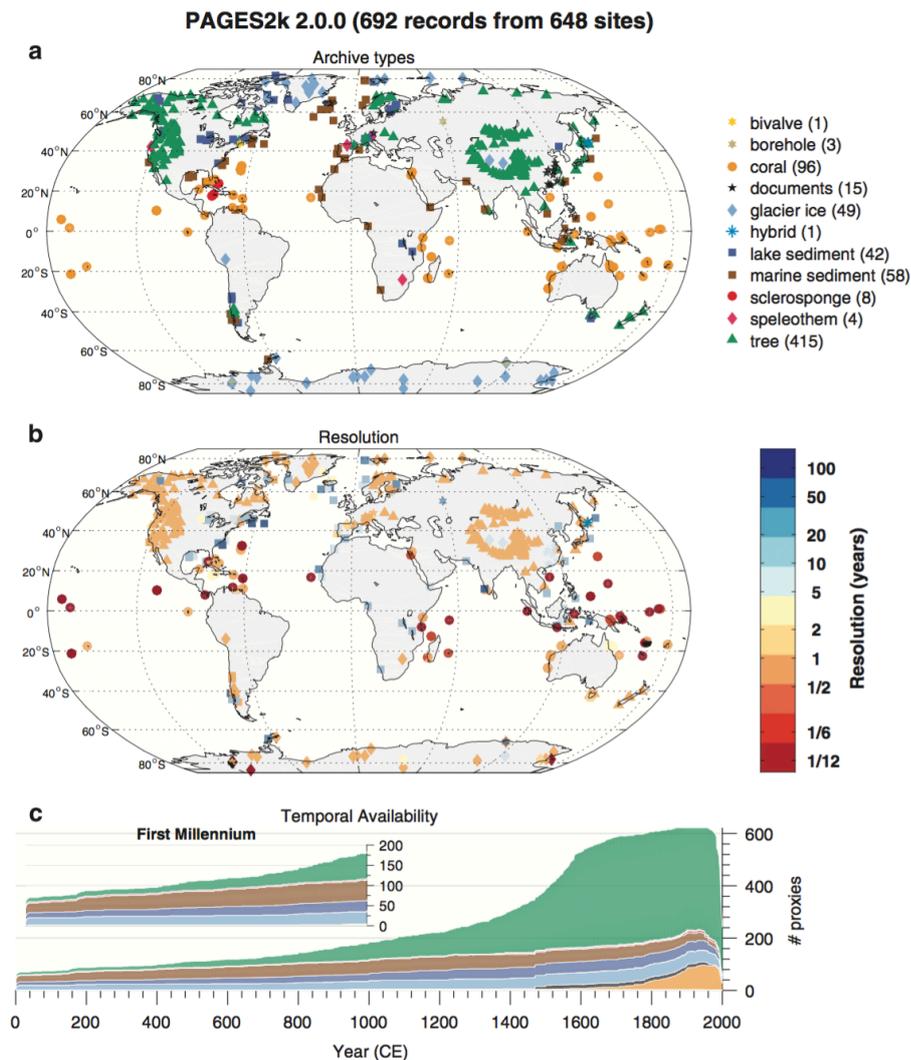


Figure 4.1: Disponibilité spatio-temporelle des données PAGES2K. (a) Localisation géographique des différentes archives naturelles, (b) résolution temporelle des données PAGES2k définie comme la médiane de l'espace entre les observations consécutives. (c) Nombre de données cumulées disponibles en fonction du temps avec les couleurs correspondant au type d'archive de (a). (source: PAGES2k Consortium, 2017)

offline (Crespin et al., 2009; Hakim et al., 2016), elles utilisent dans tous les cas un ensemble de simulations (ou particules) pour estimer de manière empirique l'état du système réel. Celles-ci sont toutefois régulièrement confrontées au même problème de dégénérescence associé à la résolution d'un problème de grande dimension avec un nombre de particules restreint. La solution théorique étant d'augmenter le nombre de particules utilisées de façon significative, des solutions pratiques ont dû être développées en raison du coût des modèles climatiques utilisés.

Deux solutions possibles, pouvant être appliquées simultanément ou non, sont généralement utilisées pour les filtres de Kalman d'ensemble (ou ENKF pour en anglais, Ensemble Kalman Filter): la localisation et l'inflation. Tandis la localisation compense le nombre restreint de particules en limitant spatialement le rayon d'influence d'une observation (Houtekamer and Mitchell, 2001; Hamill et al., 2001), l'inflation compense artificiellement

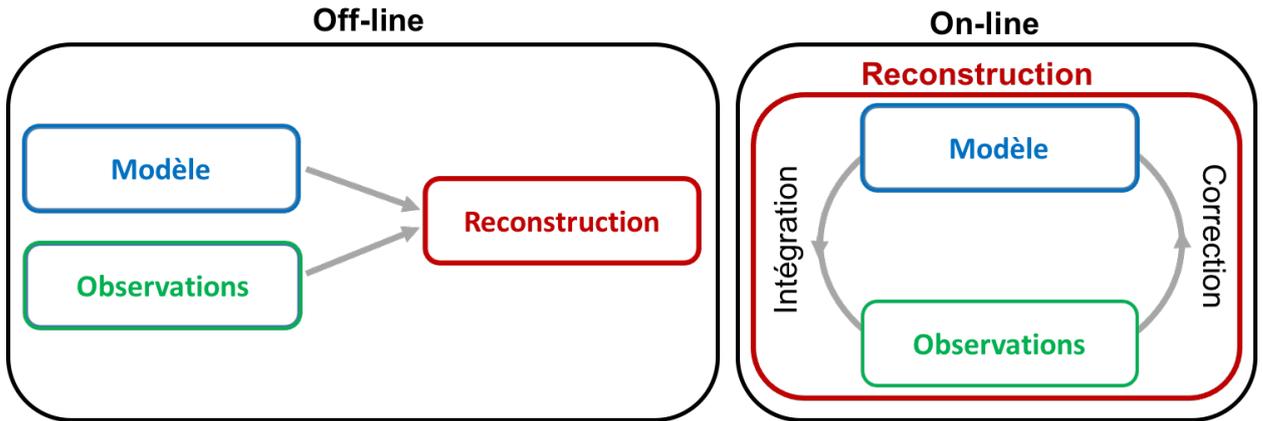


Figure 4.2: Illustration des approches offlines et online. Dans les approches offlines les simulations du modèle sont réalisées indépendamment des observations et la reconstruction est réalisée à posteriori. Dans les approches online le modèle est guidé par les observations au fur et à mesure de la reconstruction.

la sous-estimation de la variance de l'ensemble durant la phase d'analyse (Anderson and Anderson, 1999). Les différences et les avantages des approches basées sur un filtre de Kalman et filtre particulaires sont discutés par la suite. Nous avons dans le cadre de cette thèse fait le choix d'utiliser des méthodes de filtres particulaires, pour lesquelles d'autres solutions ont été proposées (Van Leeuwen, 2009).

Pour introduire les notions théoriques qui seront nécessaires pour la suite, nous noterons \mathcal{R} le système dynamique représentant la réalité et décrit par le vecteur d'état χ tel que pour tout t : $\chi(t+1) = \mathcal{R}(\chi)$. Les vecteurs d'état décrivant une/ou plusieurs variables d'état de \mathcal{R} seront notés $(\mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{x}(t_1) = \mathbf{x}_1, \dots, \mathbf{x}(t_k) = \mathbf{x}_k)$ et représenteront l'évolution temporelle de la ou des variables d'état choisies (comme par exemple les températures de surface). \mathcal{M} est le modèle numérique tentant de reproduire la dynamique des vecteurs d'état \mathbf{x}_k . Ces vecteurs d'états \mathbf{x}_k seront décrits par un ensemble d'observations \mathbf{y}_k . L'objectif des méthodes d'assimilation de données présentées par la suite sera alors d'estimer au mieux \mathbf{x}_k , à partir de la combinaison des connaissances issues des observations $p_y(\mathbf{y}_k/\mathbf{x})$ avec celles fournies par le modèle \mathcal{M} .

4.3.2 Approches séquentielles et variationnelles

D'un point de vue algorithmique, il existe également deux formes d'assimilation de données : l'assimilation séquentielle et l'assimilation variationnelle (Talagrand, 1997; Widmann et al., 2010):

- L'assimilation séquentielle consiste à intégrer le modèle dans le temps et à corriger l'état prédit par celui-ci à chaque fois qu'une observation est disponible. Les corrections sont pour cela réalisées au fur et à mesure de l'assimilation, les observations étant considérées individuellement à chaque pas de temps.
- L'assimilation variationnelle détermine la solution optimale en fonction de toutes les observations disponibles à la fois (présent, passé et même futur si disponible). Le

principe est de minimiser la fonction de coût définissant la distance entre l'état réel du système (les observations) et les solutions possibles (prédites par le modèle).

Nous travaillerons par la suite avec des algorithmes de type séquentiels et nous noterons $\mathbf{y}_{1:k} = (\mathbf{y}_1, \dots, \mathbf{y}_k)$ la suite des observations assimilées au fur et à mesure jusqu'au pas de temps k .

4.3.3 La règle de Bayes

Si l'on note $p_x(\mathbf{x})$ la probabilité à priori du vecteur d'état \mathbf{x} de dimension \mathcal{H} , alors le but des méthodes d'assimilations probabilistes est de fournir la meilleure estimation de la fonction de densité conditionnelle $p_x(\mathbf{x}/\mathbf{y})$ de \mathbf{x} à partir des connaissances fournies par les observations \mathbf{y} sur l'état réel du système. Celle-ci peut, selon la règle de Bayes, être calculée de la façon suivante :

$$p_x(\mathbf{x}/\mathbf{y}) = \frac{p_y(\mathbf{y}/\mathbf{x}) * p_x(\mathbf{x})}{p_y(\mathbf{y})} \quad (4.1)$$

En pratique, lors de l'assimilation à un instant k d'un ensemble d'observations notées \mathbf{y}_k , le dénominateur $p_y(\mathbf{y} = \mathbf{y}_k)$ est une constante exprimée de la façon suivante :

$$p_y(\mathbf{y} = \mathbf{y}_k) = \int_{\mathbb{R}^H} p_y(\mathbf{y} = \mathbf{y}_k/\mathbf{x}) * p_x(\mathbf{x}) d\mathbf{x} \quad (4.2)$$

Mais comme la connaissance de cette constante n'est pas nécessaire et utile pour connaître l'allure de la fonction de densité $p_x(\mathbf{x}/\mathbf{y} = \mathbf{y}_k)$, il est courant de ne considérer par la suite que la relation de proportionnalité suivante :

$$p_x(\mathbf{x}/\mathbf{y} = \mathbf{y}_k) \propto p_y(\mathbf{y} = \mathbf{y}_k/\mathbf{x}) * p_x(\mathbf{x}) \quad (4.3)$$

avec $p_y(\mathbf{y} = \mathbf{y}_k/\mathbf{x})$ la fonction de vraisemblance de l'observation \mathbf{y}_k supposée connue. Celle-ci est en pratique souvent considérée comme Gaussienne mais étant donnée que l'expression de cette vraisemblance dépend en réalité des jeux d'observations utilisés, nous ne donnerons que dans le chapitre suivant l'expression de la fonction de vraisemblance utilisée lors de nos expériences. La relation de proportionnalité précédente est primordiale, celle-ci étant à la base de la majorité des méthodes d'assimilation de données dont notamment les filtres bayésiens.

4.3.4 Filtrage Bayésien

Le filtrage bayésien (ou l'estimation récursive bayésienne) permet d'estimer de façon récursive la densité de probabilité $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$. Cette approche séquentielle est à l'origine des méthodes de filtre de Kalman et filtres particulaires. Dans le cas particulier des systèmes linéaires où le bruit est supposé Gaussien, le filtre de Kalman est la solution explicite du filtre bayésien. Dans les cas plus généraux, de systèmes non-linéaires avec des bruits non nécessairement gaussiens, des méthodes de filtres particulaires sont utilisées. Le filtre bayésien se décompose en 2 étapes (prédiction et analyse) et permet de calculer par

réurrence la probabilité conditionnelle $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k+1})$ à partir de $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$.

L'initialisation du filtrage bayésien est réalisée en tirant l'ensemble des particules à l'instant $k = 0$ selon la loi de distribution $p_x(\mathbf{x}_0)$.

Dès lors, si l'on suppose $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$ connu :

- L'étape de prévision relit la densité de probabilité $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k})$ aux densités de probabilité $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$ et $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$ par la formule de Chapman-Kolmogorov :

$$p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k}) = \int_{\mathbb{R}^H} p_x(\mathbf{x}_{k+1}/\mathbf{x}_k) * p_x(\mathbf{x}_k/\mathbf{y}_{1:k}) d\mathbf{x}_k \quad (4.4)$$

où $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$ est connu par hypothèse et $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$ est estimé par l'utilisation du modèle numérique \mathcal{M} .

- L'étape d'analyse relit la densité de probabilité $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k+1})$ aux densités $p_y(\mathbf{y}_{k+1}/\mathbf{x}_{k+1})$ et $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k})$ par la relation de proportionnalité (eq. 4.3):

$$p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k+1}) \propto p_y(\mathbf{y}_{k+1}/\mathbf{x}_{k+1}) * p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k}) \quad (4.5)$$

où $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k})$ est estimé lors de l'étape de prévision et où $p_y(\mathbf{y}_{k+1}/\mathbf{x}_{k+1})$ est la fonction de vraisemblance de l'observation \mathbf{y}_{k+1} .

4.4 Les filtres particulières

Les filtres à particules sont fondés sur des méthodes de Monte Carlo pour approximer à tout instant la loi de densité conditionnelle $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$ par un ensemble de N particules p^0, p^1, \dots, p^N . Ces particules sont respectivement décrites à tout instant k par les vecteurs d'états $\mathbf{x}_k^0, \mathbf{x}_k^1, \dots, \mathbf{x}_k^N$ et pondérés par les poids $w_k^0, w_k^1, \dots, w_k^N$:

$$p_x(\mathbf{x}_k/\mathbf{y}_{1:k}) = \sum_{i=1}^N w_k^i * \delta_{\mathbf{x}_k}(\mathbf{x}_k^i) \quad \text{avec} \quad \sum_{i=1}^N w_k^i = 1 \quad (4.6)$$

où $\delta_{\mathbf{x}_k}(\cdot)$ représente la mesure de Dirac centrée au point \mathbf{x}_k .

4.4.1 Le filtre particulières séquentiel pondéré

L'algorithme d'échantillonnage séquentiel pondéré (noté SIS pour en anglais Sequential Importance Sampling) consiste à mettre à jour les w_k^i en fonction des observations introduites au fur et à mesure de l'assimilation. Les nouveaux poids w_{k+1}^i sont estimés à partir des poids w_k^i au pas de temps précédents et de l'adéquation de chaque particule avec les nouvelles observations introduites.

Lors de l'étape de prévision, la loi de densité $p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k})$ est estimée par intégration du modèle numérique selon la loi de densité $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$ de l'ensemble des particules décrivant la distribution $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$.

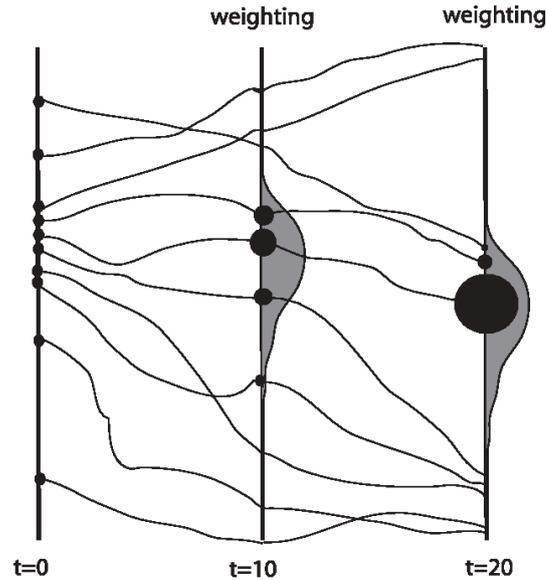


Figure 4.3: Représentation du filtre particulaire SIS pour 2 itérations $t = 10$ et $t = 20$. Les courbes en trait plein représentent la trajectoire des particules dans l'espace des états, le diamètre des cercles noirs représente le poids des particules, et les distributions (en gris) représentent les fonctions de vraisemblance des observations (source : Van Leeuwen, 2009).

$$p_x(\mathbf{x}_{k+1}/\mathbf{y}_{1:k}) = \sum_{i=1}^N w_k^i * \delta_{x_{k+1}}(\mathbf{x}_{k+1}^i) \quad \text{avec} \quad \sum_{i=1}^N w_k^i = 1 \quad (4.7)$$

Lors de l'étape d'analyse, les nouveaux poids w_{k+1}^i sont calculés à l'aide de la relation de proportionnalité (eq. 4.3) :

$$w_{k+1}^i \propto w_k^i * p_y(\mathbf{y}_{k+1}/\mathbf{x}_{k+1}) \quad \text{avec} \quad \sum_{i=1}^N w_k^i = 1 \quad \text{et} \quad \sum_{i=1}^N w_{k+1}^i = 1 \quad (4.8)$$

Algorithme SIS:

Initialisation:

- Générer N particules selon la distribution $p_x(\mathbf{x}_0)$
- $\forall i, 0 < i < N : w_0^i = \frac{1}{N}$

Pour chaque \mathbf{y}_k , avec $k > 0$:

Pour chaque particule i , avec $0 < i < N$:

- Intégrer la particule p^i tel que : $\mathbf{x}_k^i = \mathcal{M}(\mathbf{x}_{k-1}^i)$
- Calculer $w_k^i = w_{k-1}^i * p_y(\mathbf{y}_k/\mathbf{x}_k^i)$

Fin Pour

Fin Pour

Comme dans l'algorithme SIS seuls les poids w_k^i sont mis à jour lors de chaque itération, aucune correction/modification n'est réalisée sur l'ensemble des particules. L'aspect «

offline » de cette approche permet donc l'utilisation d'un ensemble de membres déjà réalisés. Mais les filtres particuliers SIS possèdent en revanche l'inconvénient de dégénérer rapidement au bout de quelques itérations. Illustré à $t = 20$ dans la Figure 4.3, la majorité des particules étant loin des observations, la dégénérescence du filtre particuliers se traduit au bout de quelques itérations par une accumulation de la distribution des poids sur une unique particule (Goosse et al., 2006b; Goosse et al., 2006a). Ce problème peut être en théorie résolu par l'utilisation d'un très grand nombre de particules. En pratique le nombre de particules nécessaires augmente de façon exponentielle avec le nombre d'itérations et la dimension du problème (Van Leeuwen, 2009). Les deux options envisageables sont alors soit de diminuer drastiquement la dimension du problème et/ou d'optimiser l'utilisation d'un nombre restreint de particules.

4.4.2 Le filtre particuliers avec ré-échantillonnage

Parmi les nombreuses variantes de filtres particuliers proposées pour pallier le problème de dégénérescence du SIS (Doucet et al., 2000; Van Leeuwen, 2009), nous avons décidé d'utiliser le filtre particuliers avec ré-échantillonnage (noté SIR pour en anglais Sequential Importance Resampling) qui a déjà été appliqué en contexte paléoclimatique avec le modèle LOVECLIM (Dubinkina et al., 2011; Goosse et al., 2012; Mairesse et al., 2013; Mathiot et al., 2013). Le principe de ce filtre est d'éviter l'effondrement des poids observés dans le filtre SIS (Fig. 4.3) en introduisant une étape de ré-échantillonnage lors de l'étape d'analyse. Cette nouvelle étape, illustrée à $t = 10$ dans la Figure 4.4, permet d'obtenir un nouvel ensemble de particules de poids égaux en dupliquant les particules de poids forts et abandonnant les particules de poids faibles. Cet aspect online du filtre particuliers SIR permet de conserver l'ensemble des particules proches des observations. Afin de rester représentatif de la distribution estimée, l'ensemble des N nouvelles particules est généré suivant la distribution décrite par les w_k^i avant ré-échantillonnage puis fixé tous égaux à $\frac{1}{N}$ une fois le ré-échantillonnage réalisé. L'introduction de cet aspect online dans le filtre particuliers SIR permet de réduire en partie l'aspect dégénérative des filtres particuliers et évite l'intégration inutile de particules loin des observations.

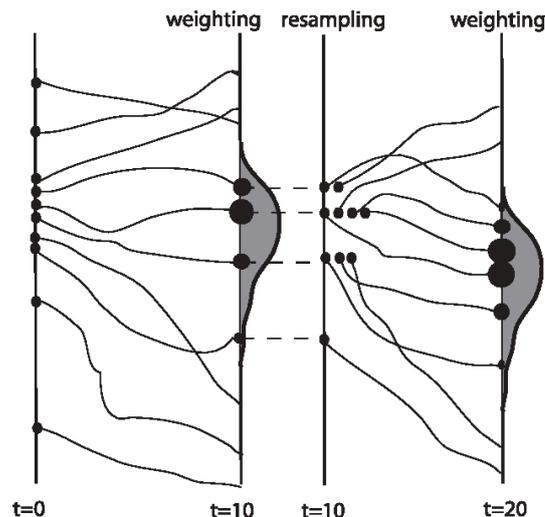


Figure 4.4: Représentation du filtre particulaire SIR pour une itération à $t = 10$. Les courbes en trait plein représentent la trajectoire des particules dans l'espace des états, le diamètre des cercles noirs représente le poids des particules, et les distributions (en gris) représentent les fonctions de vraisemblance des observations. L'étape de ré-échantillonnage du SIR est illustrée à $t = 10$ où les particules de poids fort sont dupliquées et les particules de poids faible sont abandonnées (source : Van Leeuwen, 2009)

Algorithme SIR

Initialisation:

- Générer N particules selon la distribution $p_x(\mathbf{x}_0)$
- $\forall i, 0 < i < N : w_0^i = \frac{1}{N}$

Pour chaque \mathbf{y}_k , avec $k > 0$:

Pour chaque particule i , avec $0 < i < N$:

- Intégrer la particule p^i tel que : $\mathbf{x}_k^i = \mathcal{M}(\mathbf{x}_{k-1}^i)$
- Calculer $w_k^i = p_y(\mathbf{y}_k / \mathbf{x}_k^i)$

Fin Pour

- Normaliser les w_k^i , tel que :

$$\forall i, 0 < i < N : w_k^i = \frac{w_k^i}{\sum_{k=1}^N w_k^i}$$

- Ré-échantillonner selon la distribution des w_k^i puis fixer :

$$\forall i, 0 < i < N : w_k^i = \frac{1}{N}$$

Fin Pour

4.4.3 Méthode de ré-échantillonnage

Plusieurs méthodes de ré-échantillonnage peuvent être utilisées par le SIR (Van Leeuwen, 2009). La méthode la plus courante (Gordon et al., 1993) consiste à tirer aléatoirement les N nouvelles particules à chaque pas de temps k selon la densité de probabilité décrite par les w_k^i (eq. 4.8). Bien qu'en théorie cette méthode soit optimale, en pratique le nombre de copies d'une particule p^i peut fortement varier d'une réalisation à une autre. Ce « bruit d'échantillonnage » peut induire une sous-représentation des particules de poids forts et une surreprésentation des particules de poids faibles. D'autres méthodes, notamment

la méthode de ré-échantillonnage résiduel proposé par [Liu and Chen, 1998](#) permet de réduire l'effet de ce bruit d'échantillonnage. Cette méthode pouvant être considérée comme "élitiste" a l'avantage d'assurer une bonne duplication des particules de poids forts en décomposant l'étape de ré-échantillonnage à l'instant k en deux sous étapes:

- La première appelée étape de duplication/échantillonnage minimal consiste à dupliquer chaque particule p^i un nombre de fois d^i égal à la partie entière de $N * w_k^i$.
- Et la seconde appelée étape d'échantillonnage résiduel consiste à tirer les $N - \sum_{i=1}^N d^i$ particules restantes selon la nouvelle distribution f^i où f^i est égale à la partie décimale de $N * w_k^i$.

Considérons le cas où l'on a un nombre de $N=5$ particules (p^1, p^2, p^3, p^4, p^5) de poids respectifs à l'instant k ($w_k^1 = 0.11, w_k^2 = 0.42, w_k^3 = 0.08, w_k^4 = 0.30, w_k^5 = 0.09$) :

On multiplie les w_k^i par N : ($N * w_k^1 = 0.55, N * w_k^2 = 2.1, N * w_k^3 = 0.4, N * w_k^4 = 1.5, N * w_k^5 = 0.45$) et on prend la partie entière ($d^1 = 0, d^2 = 2, d^3 = 0, d^4 = 1, d^5 = 0$). Cette première étape assure la duplication d'au moins 2 fois la particule p^2 et 1 fois la particule p^4 . Dans la seconde étape les $N - 3 = 2$ particules restantes sont alors tirées aléatoirement selon la nouvelle densité de probabilité suivante ($f^1 = 0.55, f^2 = 0.1, f^3 = 0.4, f^4 = 0.5, f^5 = 0.45$), les particules p^2 et p^4 pouvant être de nouveaux dupliquées.

4.4.4 Taille effective de l'échantillon N_{eff}

Bien que la redistribution régulière des poids entre les particules dans le cadre d'un SIR réduise le risque de dégénérescence au fils des itérations, il reste toujours possible selon la dimension du problème, qu'entre deux itérations l'ensemble des poids s'accumule sur un nombre très restreint de particules. Pour éviter la duplication massive lors de l'étape de ré-échantillonnage de quelques particules suite à cette dégénérescence, [Kong et al., 1994](#) propose de contrôler la dégénérescence des filtres particulières à l'aide de l'évolution de la taille effective de l'échantillon N_{eff} . Cet indice relié à la variance des poids peut être estimé facilement durant l'étape d'analyse par :

$$N_{eff} \approx \frac{1}{\sum_{i=1}^N (w_k^i)^2} \quad (4.9)$$

Le coefficient N_{eff} tendra vers 1 lorsque la variance des poids normalisés w_k^i est élevée, et tendra vers N lorsqu'au contraire la variance des w_k^i est faible. Les deux cas extrêmes étant :

- $N_{eff} = N$ si tous les w_k^i sont égaux à $\frac{1}{N}$
- $N_{eff} = 1$ si une seule particule à un poids non nul égale à 1

A l'aide d'un seuil fixé N_T , cet indice permet de détecter la dégénérescence du filtre. Si N_{eff} est inférieur à ce seuil, des méthodes du type « try again » peuvent être utilisées pour rééquilibrer le poids des particules [Spiller et al., 2008](#); [Van Leeuwen, 2009](#).

4.4.5 En pratique

Comme nous l'avons vu dans les sections précédentes pour les filtres bayésiens et en particulier les méthodes de filtres particulaires, les modèles numériques sont utilisés, lors de l'étape de prédiction, pour l'intégration entre les pas de temps k et $k + 1$ de l'ensemble des particules selon la loi de densité $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$. L'estimation empirique de la distribution $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$ nécessitant un nombre de particules suffisant, [Liu et al., 2017](#) ont montré, que dans le cadre de reconstructions spatiales online, les résultats des SIR s'avèrent moins bons que ceux obtenus par des filtres de Kalman d'ensemble ajusté (une variante du filtre de Kalman d'ensemble). Les auteurs ayant conclu que dans le cadre d'un nombre de particules limitées (96 pour leur expérience), les méthodes du type ENKF et l'utilisation de la localisation donnent de meilleurs résultats que les filtres particulaires malgré leur supériorité théorique. Face à ce constat et pour permettre l'intégration d'un très grand nombre de particules, nous présentons dans la section suivante un modèle simplifié et très peu coûteux, le LIM (pour en anglais Linear Inverse Model) comme modèle de substitution aux modèles climatiques généralement utilisés. L'utilisation de ce modèle dans un filtre particulaires SIR permettra par la suite de générer à faible coût un grand nombre de particules.

4.5 Le LIM

4.5.1 Présentation

L'utilisation de plus en plus répandue de modèles numériques, l'augmentation du nombre d'observations et l'augmentation de la puissance de calcul informatique disponible, ont permis d'énormes progrès dans la compréhension et la modélisation de notre système climatique. On peut distinguer 2 catégories de modèles, les modèles climatiques comme les GCMs qui se basent sur les lois fondamentales de la physique (voir Chapitre 2 section 2.5), et les modèles statistiques qui cherchent à déterminer la dynamique d'un système à partir de l'observation de certaines variables d'état. Le LIM, qui utilise une approche statistique, a montré son utilité aussi bien dans la prédiction et l'étude des variables océaniques ([Penland and Sardeshmukh, 1995](#); [Penland and Matrosova, 1998](#); [Johnson et al., 2000](#)) que dans celles de variables atmosphériques ([Winkler et al., 2001](#)). Le LIM s'est également avéré utile dans l'étude des interactions air-mer sur la variabilité tropicale ([Newman et al., 2009](#)) et dans la séparation de la variabilité forcée et interne ([Frankignoul et al., 2017](#)). [Newman, 2013](#) a montré que dans le cadre des prédictions décennales des températures de surface, les prédictions obtenues avec un LIM aux horizons 2-5 ans et 6-9 ans sont aussi bonnes (voir meilleurs dans certains cas) que celles réalisées durant

l'exercice de prédiction décennale CMIP5 (Taylor et al., 2012).

En se basant sur de nombreux articles et en particulier les articles de (Penland, 1989; Battisti and Hirst, 1989; Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995; Penland, 1996; Winkler et al., 2001), les sections suivantes introduisent et résument les hypothèses et concepts théoriques qui justifient l'utilisation du LIM comme modèle numérique \mathcal{M} pour représenter la dynamiques de \mathbf{x} .

4.5.2 Définition du modèle LIM

Tout vecteur d'anomalies $\chi(t)$ décrivant l'état d'un système dynamique \mathcal{R} à un instant t peut s'écrire sous la forme suivante :

$$\frac{d\chi(t)}{dt} = \beta(\chi(t)) + \eta(\chi(t)) + f(t) \quad (4.10)$$

où $\beta(\chi(t))$ constitue la partie linéaire de \mathcal{R} , $\eta(\chi(t))$ la composante non-linéaire de \mathcal{R} et $f(t)$ les forçages externes. Pour un vecteur d'état $\mathbf{x}(t)$ de dimension \mathcal{H} décrivant par exemple la dynamique des températures de surface et correspondant à un sous-système dynamique de $\chi(t)$, le LIM est construit sur l'hypothèse que ce sous-système dynamique peut être représenté par l'addition d'une partie linéaire prédictible et d'une partie non-linéaire représentée par des fluctuations (bruit Gaussien) corrélées spatialement mais non temporellement. En approximant l'évolution du vecteur d'état \mathbf{x} par un processus de Markov, le LIM extrait l'opérateur linéaire \mathbf{B} du système dynamique décrit par \mathbf{x} .

$$\frac{d\mathbf{x}}{dt} = \mathbf{B}\mathbf{x} + \varepsilon \quad (4.11)$$

et approxime l'ensemble des composantes non-linéaires et des forçages par un bruit Gaussien ε de matrice de covariance \mathbf{Q} . L'intégration de cette équation pour un décalage temporel τ donne (Penland, 1989) :

$$\mathbf{x}(t + \tau) = \mathbf{G}(\tau) * \mathbf{x}(t) + \sigma(t + \tau) \quad (4.12)$$

où la matrice de Green $\mathbf{G}(\tau)$ vaut :

$$\mathbf{G}(\tau) = \exp(\mathbf{B}\tau) \quad (4.13)$$

et où l'incertitude prédite $\sigma(t + \tau)$ est de la forme:

$$\sigma(t + \tau) = \mathbf{G}(\tau) \int_t^{t+\tau} \mathbf{G}(t')^{-1} \varepsilon(t') dt' \quad (4.14)$$

La matrice \mathbf{B} peut alors être calculée pour n'importe quel $\tau = \tau_0$ de la façon suivante:

$$\mathbf{B} = \tau_0^{-1} \ln[\mathbf{G}(\tau_0)] = \tau_0^{-1} \ln\left[\frac{\mathbf{C}(\tau_0)}{\mathbf{C}(0)}\right] \quad (4.15)$$

avec $\mathbf{C}(\tau_0)$ et $\mathbf{C}(0)$ les matrices de covariances au lag τ_0 et 0 :

$$\mathbf{C}(\tau_0) = \langle \mathbf{x}(t + \tau_0), \mathbf{x}(t)^T \rangle \quad (4.16)$$

$$\mathbf{C}(0) = \langle \mathbf{x}(t), \mathbf{x}(t)^T \rangle \quad (4.17)$$

la densité de probabilité de transition d'un processus de Markov devant satisfaire l'équation de Fokker-Planck. En supposant la stationnarité de \mathbf{x} , et en résolvant (eq. 4.11) comme indiqué dans (Penland and Matrosova, 1994), on obtient la relation de Fluctuation-Dissipation suivante :

$$\mathbf{Q} + \mathbf{B}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{B}^T = 0 \quad (4.18)$$

qui nous permet de calculer la matrice de covariance $\mathbf{Q} = \langle \boldsymbol{\varepsilon}^T \cdot \boldsymbol{\varepsilon} \rangle$.

Une fois les matrices \mathbf{Q} et \mathbf{B} déterminées pour un τ_0 fixé, celles-ci sont utilisées pour intégrer eq. 4.11. L'intégration est réalisée selon un pas de temps Δt , par une procédure en 2 étapes à l'aide d'un schéma de Runge-Kutta d'ordre 2 (Kloeden and Platen, 1992; Penland and Matrosova, 1994):

$$\begin{cases} z_i(t + \Delta t) &= z_i(t) + \sum_{j=1}^H b_{ij} * z_j(t) * \Delta t + \sum_{k=1}^H q_{ik} \sqrt{n_k * \Delta t * r_k} \\ x_i(t + \frac{\Delta t}{2}) &= \frac{1}{2}[z_i(t) + z_i(t + \Delta t)] \end{cases} \quad (4.19)$$

où les q_k et n_k sont les k^{eme} vecteurs propres et valeurs propres de la matrice \mathbf{Q} , les b_{ij} les coefficients de la matrice \mathbf{B} , et \mathbf{r} un vecteur de dimension H contenant des valeurs tirées aléatoirement selon une distribution gaussienne de variance 1. Les conditions initiales sont par la suite prises en compte en fixant $\mathbf{z}(0) = \mathbf{x}(0)$

4.5.3 Le LIM couplé

Supposons $\mathbf{x}_1(t)$ et $\mathbf{x}_2(t)$ deux vecteurs d'anomalies qui représentent l'état de deux variables d'état à un instant t . Si le vecteur $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$ peut-être approximé par un AR1 multivarié alors le LIM de l'équation 4.11 peut s'écrire :

$$\frac{d}{dt} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{21} \\ \mathbf{B}_{12} & \mathbf{B}_{22} \end{pmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix} \quad (4.20)$$

avec $\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{21} \\ \mathbf{B}_{12} & \mathbf{B}_{22} \end{pmatrix}$ représentant la partie déterministe du système et $\begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{bmatrix}$ représentant la partie stochastique. Les sous matrices \mathbf{B}_{12} et \mathbf{B}_{21} représentent respectivement l'influence de \mathbf{x}_1 sur \mathbf{x}_2 et \mathbf{x}_2 sur \mathbf{x}_1 . Les matrices de covariances \mathbf{Q}_1 et \mathbf{Q}_2 respectives des bruits gaussien $\boldsymbol{\varepsilon}_1$ et $\boldsymbol{\varepsilon}_2$ se calculent par :

$$\mathbf{Q}_1 + \underbrace{[\mathbf{B}_{11}\mathbf{C}_{11}(0) + \mathbf{C}_{11}(0)\mathbf{B}_{11}^T]}_{x_1} + \underbrace{[\mathbf{B}_{12}\mathbf{C}_{21}(0) + \mathbf{C}_{12}(0)\mathbf{B}_{21}^T]}_{\text{influence de } x_2} = 0 \quad (4.21)$$

$$\mathbf{Q}_2 + \underbrace{[\mathbf{B}_{22}\mathbf{C}_{22}(0) + \mathbf{C}_{22}(0)\mathbf{B}_{22}^T]}_{x_2} + \underbrace{[\mathbf{B}_{21}\mathbf{C}_{12}(0) + \mathbf{C}_{21}(0)\mathbf{B}_{12}^T]}_{\text{influence de } x_1} = 0 \quad (4.22)$$

Ce type de modélisation très utile permet d'étudier par l'analyse des sous matrices \mathbf{B}_{ij} la dynamique inter variables (Newman et al., 2011). Dans l'exemple de l'eq 4.20 d'un LIM à 2 variables, les matrices \mathbf{B}_{11} et \mathbf{B}_{22} représentent la dynamique non couplée de chaque variable, c'est à dire après avoir supprimé l'influence de la seconde variable. Le modèle peut être intégré en fixant certains coefficients à zéros pour retirer par exemple l'influence de la variable 2 sur 1 ($\mathbf{B}_{12} = 0$), de la variable 1 sur 2 ($\mathbf{B}_{21} = 0$), ou bien encore en gardant uniquement le couplage entre variables et supprimant toute dynamique intrinsèque ($\mathbf{B}_{11} = \mathbf{B}_{22} = 0$).

Bien que très utile pour étudier la dynamique d'un système, l'introduction de variables supplémentaires dans un LIM à une variable n'améliore pas nécessairement la représentation de celle-ci par le LIM (Johnson et al., 2000; Xue et al., 2000). Néanmoins (Newman et al., 2011) ont montré, en accord avec des études antérieures (Jin, 1997a), qu'une meilleure connaissance de l'état de l'océan profond et en particulier de la profondeur de la thermocline, améliorerait (principalement à long terme) les prédictions et la modélisation de la SST tropicale par le LIM.

4.5.4 Le LIM en tant que modèle d'intégration

Les résultats obtenus avec des LIM sont très convaincants et souvent comparables à ceux obtenus avec des modèles climatiques (Newman, 2013). De plus, le LIM a montré son utilité lors d'études antérieures dans de nombreux cas d'applications. Il a notamment été utilisé dans le cadre de prédiction de champs de température de surface (Newman, 2007; Alexander et al., 2008; Newman, 2013) ou bien de champs océaniques (Tziperman et al., 2008; Hawkins and Sutton, 2009). De plus, son application dans le cadre de reconstructions paléoclimatiques par (Perkins and Hakim, 2017) a montré que des approches du type ENKF intégrant des prédictions faites par un LIM déterministe permettait d'améliorer les résultats obtenus lors de l'assimilation.

Liu et al., 2017 conclut, comme de nombreuses autres études, que dans le cadre de reconstructions paléoclimatiques, la principale limite des SIR réside dans le nombre restreint de particules utilisées. Dans la mesure où un modèle peu coûteux est capable de reproduire de façon convaincante la loi de densité observée $p_x(\mathbf{x}_{k+1}/\mathbf{x}_k)$, il est alors possible d'obtenir, avec un très grand nombre de particules, une estimation empirique plus fine de la loi de densité $p_x(\mathbf{x}_k/\mathbf{y}_{1:k})$. Nous testons dans le chapitre 5 un SIR utilisant un LIM comme auxiliaire aux modèles d'intégration généralement utilisés, dans notre cas un OAGCM. Les tests sont réalisés en « modèle parfait », pour la reconstruction spatiale des températures de surface simulées par le modèle IPSL-CM5A-LR.

4.6 Indices de performances

Nous présentons dans cette section les indices et tests de performance utilisés par la suite pour évaluer les reconstructions obtenues. Nous avons choisi 4 indices quantitatifs ayant des propriétés complémentaires et le diagramme de rang qui permet d'évaluer de manière visuelle les prévisions réalisées. Soit deux séries temporelles \mathbf{v} et \mathbf{r} de dimension N , avec \mathbf{v} les valeurs réelles de moyenne \bar{v} et de déviation standard σ_v et \mathbf{r} les valeurs reconstruites de moyenne \bar{r} et de déviation standard σ_r .

- Les deux premiers indices couramment utilisés sont la racine carrée de l'erreur quadratique moyenne $RMSE$ et la corrélation $Corr$:

$$RMSE(\mathbf{r}, \mathbf{v}) = \sqrt{\sum_{t=1}^N \left[\frac{1}{N} (r_t - v_t)^2 \right]} \quad (4.23)$$

$$Corr(\mathbf{r}, \mathbf{v}) = \frac{1}{N} \sum_{t=1}^N \frac{(r_t - \bar{r})(v_t - \bar{v})}{\sigma_r \sigma_v} \quad (4.24)$$

- Le coefficient d'efficacité (CE) de (Nash and Sutcliffe, 1970) qui en plus d'évaluer la covariabilité en terme de phase comme les corrélations, considère également l'amplitude et les biais de la série reconstruite :

$$CE(\mathbf{r}, \mathbf{v}) = 1 - \frac{\sum_{t=1}^N (r_t - v_t)^2}{\sum_{t=1}^N (v_t - \bar{v})^2} \quad (4.25)$$

avec des valeurs comprises entre 1 et moins l'infini. Plus les valeurs du CE sont élevées, meilleure est la reconstruction, une valeur de $CE = 1$ indique une reconstruction parfaite. Des valeurs négatives n'impliquent en revanche pas nécessairement de mauvais résultats (Cook et al., 1999; Hakim et al., 2016), c'est pourquoi il reste généralement intéressant de mettre en perspective les valeurs obtenues avec des valeurs de références.

Les indices précédents n'étant pas utilisables pour évaluer la distribution décrite par un ensemble de séries temporelles, nous utilisons l'indice continu de probabilité (ou $CRPS$ pour en anglais Continuous rank probability score) tel qu'il est défini dans (Tipton et al., 2016; Perkins and Hakim, 2017), comme critère de score pour évaluer la distribution d'un ensemble $r_t^1, r_t^2, \dots, r_t^K$ de K séries temporelles reconstruites :

- $CRPS$ (Gneiting and Raftery, 2007) :

$$CRPS = \sum_{t=1}^N \left(\frac{1}{K} \sum_i [r_t^i - v_t] - \frac{1}{2K^2} \sum_i \sum_j [r_t^i - r_t^j] \right) \quad (4.26)$$

Tandis qu'un $CRPS$ de 0 reflète un ensemble parfait de membre tous égaux à la valeur réelle (pas d'enveloppe), plus l'erreur moyenne ou la déviation standard de la distribution augmente, plus la valeur du $CRPS$ augmente.

- Le Diagramme de rang

Le diagramme de rang aussi appelé diagramme de Talagrand (Anderson, 1996; Hamill and Colucci, 1997; Talagrand, 1999) permet de valider de façon visuelle, la distribution des prédictions d'un modèle. Le principe consiste à visualiser et comparer l'histogramme des valeurs observées selon la distribution des valeurs prédites. L'ensemble des valeurs prédites par le modèle est pour cela divisée en M quantiles notés q_0, q_1, \dots, q_M qui seront utilisés comme distribution de référence. La probabilité d'appartenance pour une prédiction au quantile $\llbracket q_j, q_{j+1} \rrbracket$ étant par définition de $\frac{1}{M}$, on vérifie si cette probabilité d'appartenance est similaire pour l'ensemble des observations. Dans le cas de distributions prédites et observées similaires, la fréquence d'apparition des observations dans un intervalle $\llbracket q_j, q_{j+1} \rrbracket$ sera d'environ $\frac{1}{M}$.

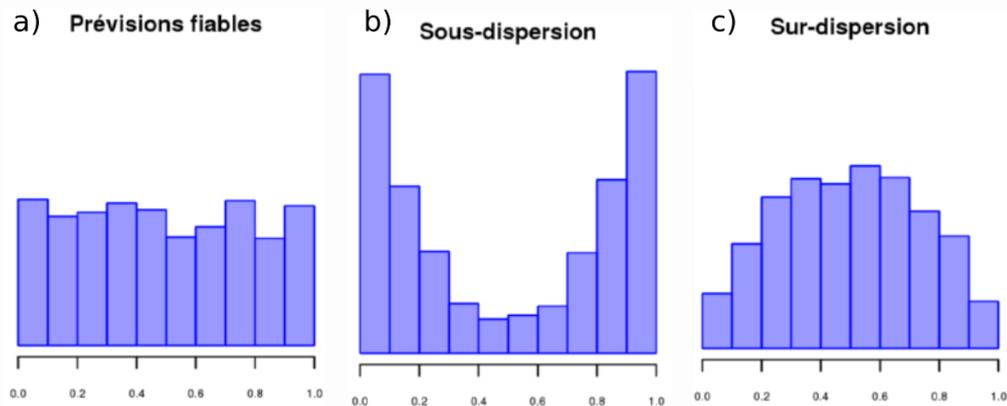


Figure 4.5: Représentation de la forme du diagramme de rang dans le cas (a) de prévisions fiables, (b) de prévisions sous estimant la variance et (c) de prévisions surestimant la variance. Avec en ordonnée la fréquence d'apparition des observations $\frac{o_j}{\sum_{j=0}^M o_j}$ en fonction des intervalles réguliers $\llbracket i_j, i_{j+1} \rrbracket$.

La construction d'un diagramme de rang peut être réalisée de la façon suivante :

- Diviser l'intervalle $\llbracket 0, 1 \rrbracket$ en M intervalles réguliers notés i_0, i_1, \dots, i_M .
- Diviser l'ensemble des prévisions du modèle en M quantiles notés q_0, q_1, \dots, q_M .
- Compter le nombre d'observations o_j comprises dans chaque intervalle $\llbracket q_j, q_{j+1} \rrbracket$, et associer la fréquence d'apparition $\frac{o_j}{\sum_{j=0}^M o_j}$ à l'intervalle $\llbracket i_j, i_{j+1} \rrbracket$.
- Tracer sous forme de diagramme en bâton, la fréquence d'apparition des observations $\frac{o_j}{\sum_{j=0}^M o_j}$ en fonction des effectifs des intervalles $\llbracket i_j, i_{j+1} \rrbracket$ associés.

Une fois tracée, la forme du diagramme révèle les différences entre les distributions prédites et observées. Un diagramme plat (Fig. 4.5a) signifie que la distribution des prévisions est fiable, la probabilité qu'une observation tombe dans un intervalle plutôt qu'un autre est équivalente à $\frac{1}{M}$. Un diagramme en forme de «U» (Fig 4.5b)

traduit une sous-estimation de la variance par le modèle car la distribution prédite est trop peu dispersée et les classes centrales contiennent moins d'observations qu'elles devraient. Si au contraire le diagramme est en forme de cloche (Fig 4.5c), cela traduit une surestimation de la variance par le modèle avec une distribution prédite trop dispersée et des classes périphériques qui contiennent moins d'observations qu'elles devraient. Un diagramme asymétrique traduit, quant à lui, la présence de biais entre les deux distributions.

Chapitre 5

Reconstruction de la variabilité des températures de surface en contexte paléoclimatique avec le SIR-LIM

Sommaire

Notations	109
5.1 Préambule	110
5.2 Protocole expérimental	110
5.3 Le SIR-LIM	111
5.4 Expériences et validations	114
5.5 Calcul des EOFs	116
5.6 Construction et Validation du LIM	119
5.7 Le LIM en tant qu'émulateur des modèles climatiques	123
5.8 Reconstructions spatiales de la variabilité des températures de surface	129
5.9 Reconstructions des principaux modes de variabilité pour le modèle IPSL-CM5A-LR	133
5.10 Sensibilité aux paramètres d'assimilation	138
5.10.1 Nombre de particules et nombre d'EOFs	138
5.10.2 Estimation de la matrice d'erreur C	140

Notations

\mathcal{R}	système dynamique réel
χ	vecteur d'état décrivant \mathcal{R}
\mathbf{x}	vecteur d'état d'un sous-système dynamique de \mathcal{R} de dimension \mathcal{H}
\mathbf{y}	vecteur d'observations
\mathcal{H}	dimension du vecteur d'état \mathbf{x}
\mathbf{x}_k	vecteur d'état \mathbf{x} au pas de temps k
\mathcal{M}	modèle numérique reproduisant la dynamique des \mathbf{x}_k
p^i	particule i
N	nombre de particules
\mathbf{x}_k^i	vecteur d'état de la particule i à l'instant k
w_k^i	poids associé à la particules i à l'instant k
\mathbf{y}_k	observations disponibles au pas de temps k
$\mathbf{y}_{1:k}$	suite des observations jusqu'au pas de temps k
\mathbf{B}	matrice de covariance
ε	bruit Gaussien
\mathbf{Q}	matrice de covariance du bruit
\mathbf{G}	propagateur linéaire
$\sigma(t)$	bruit Gaussien généré à l'instant t
$\mathbf{C}(\tau)$	matrice de covariance au lag τ
N_{eff}	taille effective de l'échantillon

5.1 Préambule

Dans ce chapitre, nous présentons et testons une méthode d'assimilation de donnée de filtre particulières avec ré-échantillonnage (noté SIR-LIM où SIR pour en anglais Sequential Importance Resampling) utilisant un LIM (pour pour en anglais Linear Inverse Model) comme modèle d'intégration. L'évaluation de ce type de méthodes est faite en utilisant comme jeux de données réelles, les variables simulées par le modèle IPSL-CM5A-LR. Ces expériences de validation en « modèle parfait » très courantes (Smerdon, 2012; Werner et al., 2013; Gómez-Navarro et al., 2017) permettent de pallier aux manques d'observations disponibles en transposant le problème réel dans le monde des modèles climatiques où toutes les variables sont connues partout et à tout instant. Afin dans un premier temps de s'affranchir de l'influence des forçages externes, la simulation de contrôle du modèle IPSL-CM5A-LR (c'est-à-dire avec forçage constant) a été prise comme représentative de la réalité. Notre objectif est de reconstruire la variabilité inter annuelle des températures de surface de cette simulation appelée par la suite « cible » à partir d'un nombre restreint de pseudo-proxys simulés. Longue de 2700 ans avec des sorties mensuelles, cette simulation a été divisé en deux parties, la première partie servant de base d'apprentissage pour construire notre LIM et la seconde partie de base de validation pour tester le SIR-LIM. Plusieurs sous-échantillonnage spatiaux pouvant être testés pour la localisation des pseudo-proxys (Dubinkina and Goosse, 2013), 2 expériences notées « Dense » et « Pages2k » utilisant 2 distributions spatiales différentes, ont été testé pour la localisation de nos pseudo-proxys. Dans les sections suivantes, nous présentons le protocole expérimental utilisé (section 5.2). Nous décrivons ensuite la construction du LIM et celle du SIR-LIM (section 5.3) et les expériences réalisées (section 5.4). Nous validons la bonne reproduction de la dynamique du système réel par le LIM (section 5.6 et 5.7) et nous évaluons la reconstruction obtenue avec le SIR-LIM (section 5.8). Pour finir nous testons la sensibilité du SIR-LIM aux paramètres utilisés (section 5.10.1), avant de discuter l'estimation de l'erreur de modélisation (section 5.10.2).

5.2 Protocole expérimental

Pour réaliser nos expériences nous disposons d'une simulation de contrôle du modèle IPSL-CM5A-LR de 2700 ans (cf. Chapitre 2 section 2.5). Une moyenne glissante de 12 mois a été appliquée aux sorties mensuelles de cette simulation pour s'affranchir de la saisonnalité. Seules les 1700 dernières années de cette simulation sont exploitées dans ce chapitre pour s'affranchir des problèmes de dérive du modèle suite au spin-up de la simulation. Sur les 1700 ans, les 1000 premières années constituent l'ensemble d'apprentissage (APP) et les 700 années restantes l'ensemble de validation (VAL). L'ensemble APP est utilisé pour construire notre LIM et l'ensemble VAL est réservé aux expériences de validation du LIM et du SIR-LIM. Nous avons choisi dans un premier temps de travailler sur la reconstruction de la température de surface, variable la plus communément déduite de l'analyse des proxys. Les cibles à reconstruire dans les sections suivantes sont donc prises

sur VAL et les valeurs des pseudo-proxys sont simulées par l'ajout d'un bruit gaussien de déviation standard 0.5 aux températures de surface afin de prendre compte l'incertitude présente dans les proxys réels. A la différence d'expériences antérieures (Dubinkina et al., 2011; Goosse et al., 2012) aucun filtre spatial n'a été utilisé. La grille utilisée est celle du LMDZ5A (la composante atmosphérique de l'IPSL-CM5A-LR) de résolution 1.87° pour les latitudes et 3.75° pour les longitudes. Les climatologies et la calibration du LIM ont été faites sur APP, tandis que les figures et tests présentés ci-après sont, sauf indication contraire dans la légende ou le texte, réalisés sur VAL.

5.3 Le SIR-LIM

La méthode testée SIR-LIM est un filtre particulaire SIR utilisant le LIM comme modèle d'intégration. Pour reprendre les notations du Chapitre 4, la simulation de contrôle du modèle IPSL-CM5A-LR correspond dans notre cas à la réalité \mathcal{R} et le LIM au modèle d'intégration \mathcal{M} servant à approcher la dynamique des températures de surface. Le LIM étant un modèle stochastique linéaire, celui-ci est généralement construit à partir des principaux modes de variabilité représentés dans les premières EOFs (Penland, 1989). Les EOFs utilisées pour notre LIM sont présentées dans la section 5.5. Elles sont calculées sur les températures de surface (filtrées avec une moyenne glissante 12 mois) de APP après avoir centré-réduit et pondéré chaque point de grille de façon à ce que chaque unité de surface soit équitablement considérée. Ces EOFs seront utilisées pour projeter nos particules dans l'espace des EOFs, et permettront de reconstruire ensuite des champs de températures de surface à partir de PCs prédites par le LIM et/ou reconstruites par le SIR-LIM (Fig. 5.1). En d'autres termes, toutes les PCs manipulées par la suite correspondent à la projection des températures de surface dans l'espace des EOFs décrits dans la section 5.5.

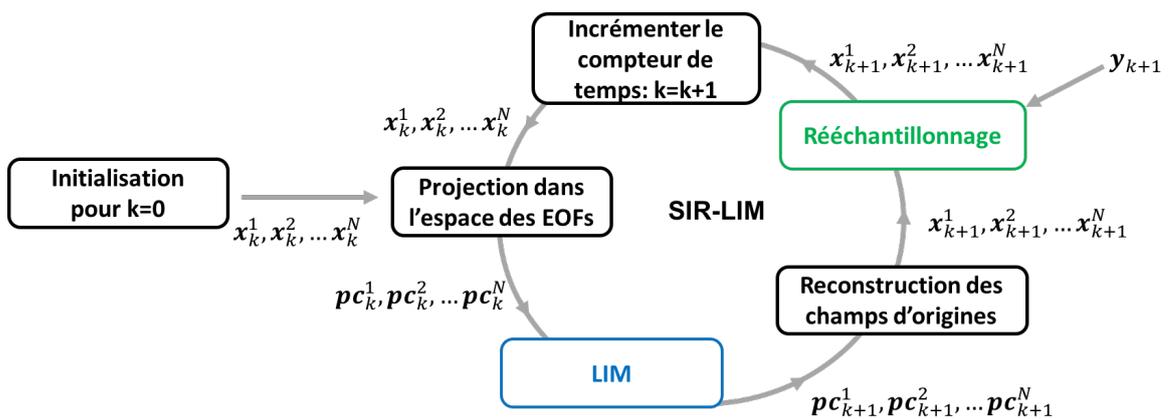


Figure 5.1: Représentation conceptuelle du SIR-LIM pour N particules et du passage des particules de l'espace des états réels noté x_k^i à celui des EOFs noté pc_k^i . Le vecteur des observations disponibles au pas de temps k est noté y_k .

Les champs d'origines, c'est à dire les températures de surface dans notre cas, seront notées x tandis que leurs projections dans l'espace des EOFs seront notés pc . Les vecteurs

d'observations (ou pseudo-proxys) seront notés \mathbf{y}_k à l'instant k . Les vecteurs \mathbf{x}_k^{cible} et \mathbf{pc}_k^{cible} représentent les vecteurs d'états du système réel à l'instant k dans respectivement l'espace d'origine et l'espace des EOFs. Les N particules du SIR-LIM seront notées p^0, p^1, \dots, p^N , leur vecteur d'états à l'instant k seront notés $\mathbf{x}_k^0, \mathbf{x}_k^1, \dots, \mathbf{x}_k^N$ dans l'espace d'origine et $\mathbf{pc}_k^0, \mathbf{pc}_k^1, \dots, \mathbf{pc}_k^N$ dans l'espace des EOFs. Alors que l'étape de ré-échantillonnage du filtre particulaire SIR-LIM est réalisée dans l'espace d'origine, la construction du LIM et son intégration sont réalisées dans l'espace des EOFs (Fig. 5.1). Le SIR-LIM étant, comme tous les filtres particulaire, basé sur des approximations du type Monte Carlo, nous obtenons en sortie de l'algorithme un ensemble de N particules décrivant de manière empirique l'état du système à approximer. A chaque pas de temps la moyenne de ces particules sera considérée comme notre reconstruction tandis que les incertitudes liées à cette reconstruction seront estimées par l'enveloppe décrite par l'ensemble des particules.

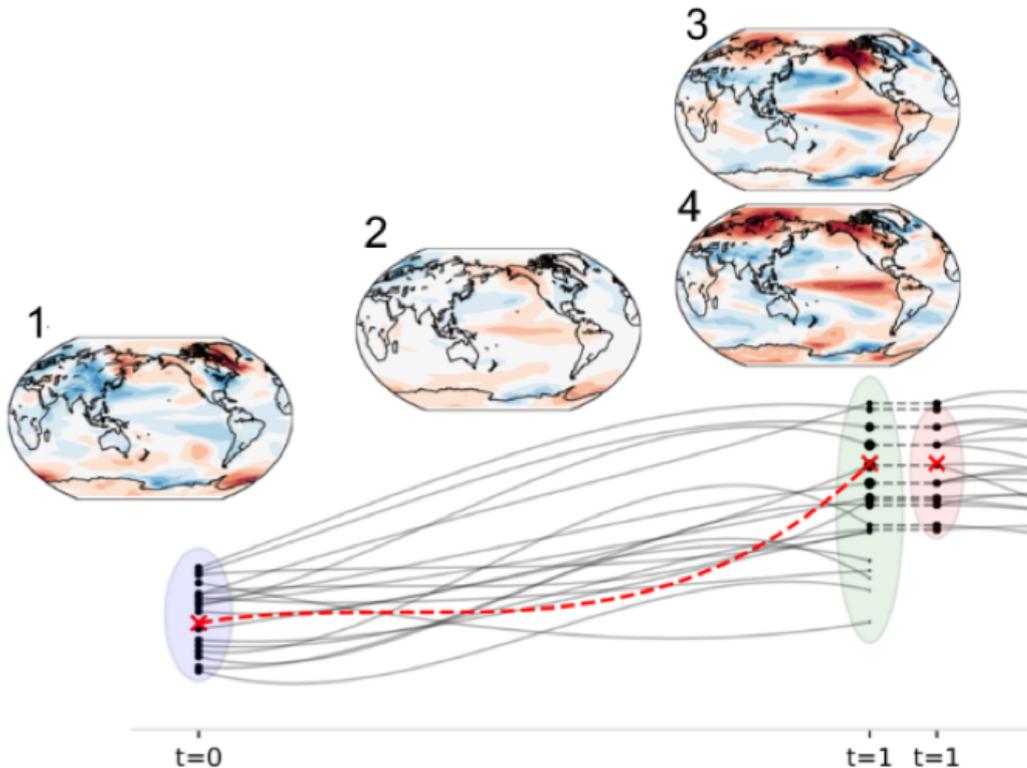


Figure 5.2: Représentation schématique du SIR-LIM avec en bleu l'état initial, en vert l'ensemble des particules après intégration avec le LIM et en rouge après ré-échantillonnage. Les croix rouges correspondent à la position des pseudo-proxys. Les cartes d'anomalies correspondent à la moyenne d'ensemble des particules (1) à l'initialisation, (2) après intégration avec le LIM, (3) après ré-échantillonnage et (4) la cible.

Si l'on reprend l'équation 4.11 avec les notations introduites précédemment, le LIM s'écrit donc :

$$\frac{d\mathbf{pc}}{dt} = \mathbf{B} * \mathbf{pc} + \varepsilon \quad (5.1)$$

Avec \mathbf{pc} les P premières PCs des températures de surface, \mathbf{B} l'opérateur linéaire du

système dynamique et ε le bruit Gaussien de matrice de covariance \mathbf{Q} . Une fois les matrices \mathbf{B} et \mathbf{Q} estimées sur APP comme décrits dans la section 4.5, le modèle d'intégration (eq. 5.1) sera utilisé dans le SIR-LIM pour intégrer l'ensemble des particules dans le temps. Lors de chaque étape de ré-échantillonnage les poids des particules seront estimés grâce à la relation d'équivalence 4.3. La fonction de vraisemblance utilisée dans (eq. 4.3) pour estimer l'adéquation de chaque particule avec les observations, sera pour la suite considérée gaussienne de la forme :

$$p_y(\mathbf{y}_k / \mathbf{x}_k^i, \mathbf{y}_{0:k-1}) = A * \exp\left(\frac{-1}{2} [\mathbf{y}_k - T(\mathbf{x}_k^i)]^T \mathbf{C}^{-1} [\mathbf{y}_k - T(\mathbf{x}_k^i)]\right) \quad (5.2)$$

où \mathbf{y}_k est le vecteur d'observations contenant l'ensemble des pseudo-proxys à l'instant k , A est une constante de normalisation, $T(\cdot)$ est un opérateur de transfert de l'espace des états vers les observations qui consiste dans notre cas à appliquer un simple masque pour récupérer les valeurs du modèle correspondant à la localisation des observations \mathbf{y}_k et \mathbf{C} la matrice de covariance d'erreur des observations.

La matrice de covariance d'erreur \mathbf{C} inclut à la fois l'erreur instrumentale \mathbf{C}_{instru} et l'erreur de représentativité \mathbf{C}_r . Nous l'avons estimé comme dans Dubinkina et al., 2011 par la somme de deux erreurs :

$$\mathbf{C} = \mathbf{C}_{instru} + \mathbf{C}_r \quad (5.3)$$

où l'erreur instrumentale \mathbf{C}_{instru} est prise égale à la déviation standard du bruit blanc Gaussien ajouté lors de la simulation des pseudo-proxys (c'est-à-dire 0.5) et où l'erreur de représentativité \mathbf{C}_r prend en compte les différences de résolution et d'échelles spatiales entre le modèle utilisé et les observations (Dubinkina et al., 2011). Une partie de la variabilité réelle des températures de surface n'étant pas considérée lors de l'étape « Reconstruction des champs d'origines » de la figure 5.1, l'erreur de représentativité permet dans notre cas de prendre en compte l'approximation faite par l'utilisation d'un nombre restreint d'EOFs pour décrire l'évolution des champs de températures de surface réels. Il en effet évident que l'utilisation d'un sous-ensemble d'EOFs tendra à omettre une partie de la variabilité réelle, et qu'il est donc nécessaire de prendre en compte cette erreur dite de représentativité lors de l'estimation de la vraisemblance $p_y(\mathbf{y}_k / \mathbf{x}_k^i)$ entre les pseudo-proxys de températures réels et les reconstructions de températures de surface associées à un sous ensemble d'EOFs. Cette erreur de représentativité a pour sa part été considérée telle que suggéré par Sansó et al., 2008 et Dubinkina et al., 2011 de la forme :

$$\mathbf{C}_r = \sigma \mathbf{C}_{mod} \quad (5.4)$$

où le terme \mathbf{C}_{mod} est estimé sur APP comme la variabilité des températures de surface, et où l'estimation du terme σ n'est quant à elle pas triviale (Ueno et al., 2010). Nous l'avons pris égale à 1.5 dans un premier temps sachant qu'une série d'expériences est effectuée par la suite (section 5.10.2) pour évaluer l'influence de ce coefficient dans la qualité de

reconstruction.

Un pas de temps de ré-échantillonnage de 1 an a été appliqué pour le filtre à particule SIR-LIM et un pas d'intégration similaire à (Penland and Matrosova, 1994) de 6 heures ($\Delta t = 1/120$ mois) a été utilisé pour le LIM. Comme le LIM intègre par construction une composante aléatoire, nous n'avons pas comme c'est généralement le cas pour des modèles déterministes (Dubinkina et al., 2011; Liu et al., 2017) introduit de perturbations à la suite de l'étape de ré-échantillonnage pour dissocier les particules identiques. La figure 5.2 illustre la première itération (de $k=0$ à $k=1$) de l'algorithme du SIR-LIM et le rôle du LIM dans l'intégration de l'ensemble des particules. Soit un ensemble de K vecteurs d'observations $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k$ et N particules p^1, p^2, \dots, p^N , l'algorithme du SIR-LIM illustré dans la figure 5.1 se résume de la façon suivante:

Algorithme SIR-LIM

Initialisation:

- Calculer les P premières EOFs sur APP.
- Initialiser les états initiaux $\mathbf{x}_0^1, \mathbf{x}_0^2, \dots, \mathbf{x}_0^N$ des particules p^1, p^2, \dots, p^N en prenant aléatoirement N états sur les 1000 ans de APP.
- Projeter $\mathbf{x}_0^1, \mathbf{x}_0^2, \dots, \mathbf{x}_0^N$ dans l'espace décrit par les P premières EOFs pour obtenir les positions des particules p^1, p^2, \dots, p^N dans l'espace des EOFs à $k = 0$.

Pour chaque y_k , avec $k > 0$:

Pour chaque particule i , avec $1 < i < N$:

- Intégrer la particule p^i avec le LIM pour obtenir \mathbf{pc}_k^i à partir de \mathbf{pc}_{k-1}^i
- Reconstruire le champ initial \mathbf{x}_k^i à partir de \mathbf{pc}_k^i .
- Calculer les poids associés:

$$w_k^i = A * \exp\left(\frac{-1}{2} [\mathbf{y}_k - T(\mathbf{x}_k^i)]^T \mathbf{C}^{-1} [\mathbf{y}_k - T(\mathbf{x}_k^i)]\right)$$

Fin Pour

- Normaliser les w_k^i , tel que :

$$\forall i, 0 < i < N : w_k^i = \frac{w_k^i}{\sum_{k=1}^N w_k^i}$$

- Ré-échantillonner les particules selon la distribution des w_k^i (Liu and Chen, 1998).
- Mettre à jour les poids :

$$\forall i, 0 < i < N : w_k^i = \frac{1}{N}$$

Fin Pour

5.4 Expériences et validations

Nous introduisons dans la section suivante l'ensemble des expériences réalisées pour évaluer l'efficacité du LIM et du SIR-LIM. Comme expliqué précédemment la ou les cibles à reconstruire sont toujours présent sur VAL et le Tableau 5.1 récapitule l'ensemble de ces

expériences en indiquant les sections où celles-ci sont réalisées.

Persistence:

Pour valider l'utilisation du LIM comme modèle dynamique, nous comparons dans la section 5.6 les prédictions du LIM aux horizons 12 et 18 mois avec une approche très simple notée Persistence. Cette approche consiste à tout simplement prédire pour le futur la même valeur que le présent, ceci donnant pour n'importe quel horizon h et à n'importe quel pas de temps t la prédiction suivante $x_{t+h} = x_t$. La comparaison des prédictions du LIM avec la Persistence permet d'évaluer les capacités de prédiction du LIM par rapport à une méthode de prédiction très basique.

Dense et Pages2k:

Comme la qualité de reconstruction dépend de la distribution spatiale des proxys disponibles (Evans et al., 1998; Comboul et al., 2015), nous avons réalisé 2 expériences avec différents sous-échantillonnages spatiaux (Fig. 5.3). Pour la première expérience (notée Pages2k), la localisation des pseudo-proxys a été prise similaire à la localisation réelle des proxys de températures disponibles à l'an 1000 dans la base PAGES2K-v2 (PAGES2k Consortium). Cette expérience est représentative du problème réel avec un sous-échantillonnage des océans par rapport aux continents et un sous-échantillonnage de l'Hémisphère Sud par rapport à l'Hémisphère Nord (avec environ 77% des proxys dans l'Hémisphère Nord). Pour la seconde expérience (notée Dense), un échantillonnage régulier de 167 proxys a été considéré sur l'ensemble du globe (Fig. 5.3). Pour les 2 expériences les pseudo-proxys ont été simulés en ajoutant un bruit gaussien de déviation standard 0.5 aux températures de surface de VAL.

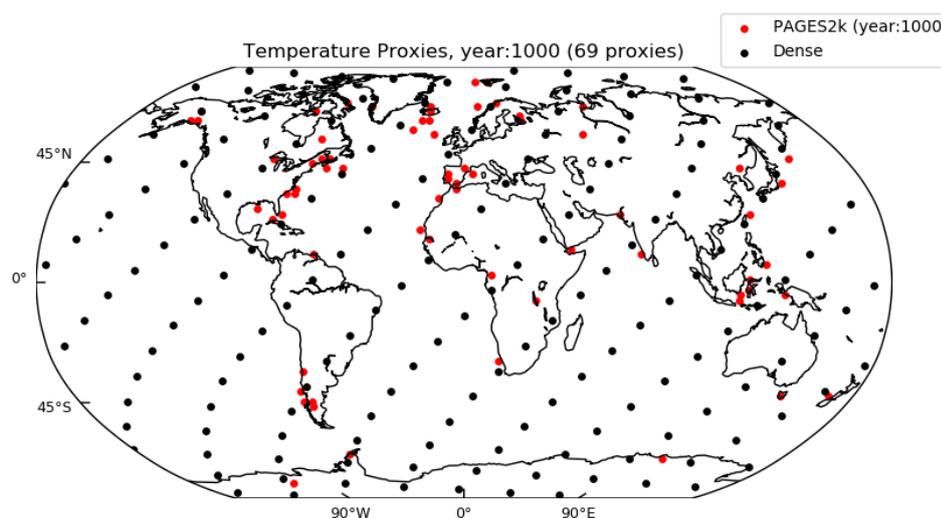


Figure 5.3: Localisation des 69 pseudo-proxys de températures utilisés pour l'expérience Pages2K en rouge et des 167 pseudo-proxys de températures utilisés pour l'expérience Dense en noir.

Contribution du LIM dans le SIR-LIM:

Afin d'évaluer la contribution du LIM dans les reconstructions obtenues, il semble intéressant de comparer les résultats obtenus par le SIR-LIM avec celle d'un SIR utilisant un modèle d'intégration basique. Il est en effet évident qu'indépendamment du modèle d'intégration utilisé, l'étape de ré-échantillonnage du SIR-LIM explique à elle seule une partie des scores obtenus lors de reconstructions. Ce score dit de « ré-échantillonnage » est dû à la sélection des particules proches des observations et dépend de nombreux paramètres comme la fréquence de ré-échantillonnage, la dimension du problème, le nombre de particules utilisé ou encore la fonction de vraisemblance utilisée. Une méthode simple pour estimer ce score de « ré-échantillonnage » du SIR-LIM est de remplacer le modèle d'intégration, c'est-à-dire le LIM, par un bruit blanc Gaussien. Ce filtre particuliers noté SIR-0 aura donc pour modèle d'intégration $\mathbf{p}\mathbf{c} = \varepsilon$ avec ε un bruit blanc Gaussien de même déviation standard que les PCs de APP. L'auto-covariance étant nulle dans le cas d'un bruit blanc, les scores obtenus par le SIR-0 seront attribués à l'étape de ré-échantillonnage indépendamment de tout aspect physique/dynamique lié au modèle d'intégration. La contribution du LIM comme modèle d'intégration dans un filtre particuliers SIR sera par la suite évalué comme la différence entre le SIR-LIM et le SIR-0.

noDA comme référence:

Dans la section 5.10.1, nous évaluons la sensibilité du SIR-LIM à certains paramètres. Les valeurs de RMSE et de CE et leur ordre de grandeur n'étant pas toujours très parlant, il semble intéressant d'introduire une valeur de référence afin de juger de la relative sensibilité de la méthode aux paramètres utilisés. Nous introduisons pour cela une expérience appelée noDA, qui représente le cas sans assimilation de données. Le score de noDA sera calculé en prenant comme reconstruction un bout de simulation pris arbitrairement sur APP. Cette simulation de même taille que la cible représente le cas d'une reconstruction où aucune correction n'a été apportée à la simulation du modèle. Dans le cas de notre expérience avec forçages constants où la cible et la période prise arbitrairement sur APP sont indépendantes, la RMSE de noDA est en théorie égale à la racine carrée de la somme des deux variances. Ces deux scores seront utilisés comme référence pour calculer une RMSE dite normalisée et un écart de CE noté ΔCE définie respectivement comme la RMSE divisée par la RMSE de noDA, et le CE moins le CE de noDA. Les scores de noDA utilisés dans les figures 5.23 et 5.24 correspondent à la RMSE et au CE obtenus entre la cible à reconstruire d'une période de 100 ans prise sur VAL et une période de 100 ans prise arbitrairement sur APP.

5.5 Calcul des EOFs

Utilisé dans le cadre d'application géophysique, le LIM est généralement construit à partir des principaux modes de variabilité retenus dans les premières EOFs (Penland, 1989).

Table 5.1: Liste des ensembles d'expériences réalisées. La colonne Sections indique dans quelles sections de ce chapitre les expériences sont utilisées.

Nom de la méthode ou de l'expérience	Principe et Objectif	Sections
Persistence	Evaluer la qualité de prédiction du LIM par rapport à une méthode de prédiction basique prédisant pour le futur la même valeur que le présent.	5.6
Pages2k	Reconstruction effectuée en prenant un sous-échantillonnage spatial de pseudo-proxys similaire aux proxys de températures disponibles à l'an 1000 dans la base de donnée Pages2k (Fig. 5.3).	5.8; 5.10
Dense	Même chose que Pages2k mais avec un échantillonnage spatial de pseudo-proxys dense et régulier (Fig. 5.3)	5.8; 5.10
SIR-LIM	Filtre particuliers SIR utilisant le LIM comme modèle d'intégration (Fig. 5.1).	5.8; 5.10
SIR-0	Filtre particuliers SIR similaire au SIR-LIM mais utilisant un bruit gaussien plutôt qu'un LIM comme modèle d'intégration.	5.8
noDA	Cas d'une reconstruction sans assimilation de donnée utilisée comme score de référence.	5.10.1

Les EOFs sont calculées à partir de critères purement statistiques et sont incapables de détecter la chronologie temporelle de phénomènes géophysiques. Souvent normées par construction, les EOFs représentent des patterns stationnaires inadaptés pour décrire à travers le temps l'évolution et la propagation spatiale de certains phénomènes physiques. Les PCs associées donnent le signe et l'amplitude de la variabilité décrite au cours du temps par chacune d'elles. Les patterns physiques réels pouvant être différents de ceux décrits par les EOFs, toute interprétation physique de celles-ci se doit donc d'être prudente. La base décrite par les 30 premières EOFs expliquent 66.2% de la variance des anomalies de températures de surface (Fig. 5.4a).

La première EOF et sa PC associée très corrélées à l'indice Niño3.4 (0.88) et au TPI non filtré (0.92), explique à elles seules 18,8% de la variance des températures de surface globale après avoir retiré (en filtrant) le cycle saisonnier. Les secondes et troisièmes EOFs représentent quant à elles respectivement 5.1% et 4.4% de la variance totale. La seconde EOF semble avoir capturé le principal mode de variabilité des températures de surface de l'Hémisphère Sud avec une empreinte spatiale pouvant correspondre à l'expression du SAM (pour en anglais Southern Annular Mode) sur les températures de surface (voir Fig. 12 de [Sen Gupta and England, 2006](#)). La troisième EOF donne quant à elle une description complémentaire de la variabilité tropicale dans l'Océan Pacifique décrite par la première EOF mais avec des températures de signe opposé entre l'océan Pacifique et le reste des tropiques. La corrélation lagée entre les PCs de la première et troisième EOF est de 0.46

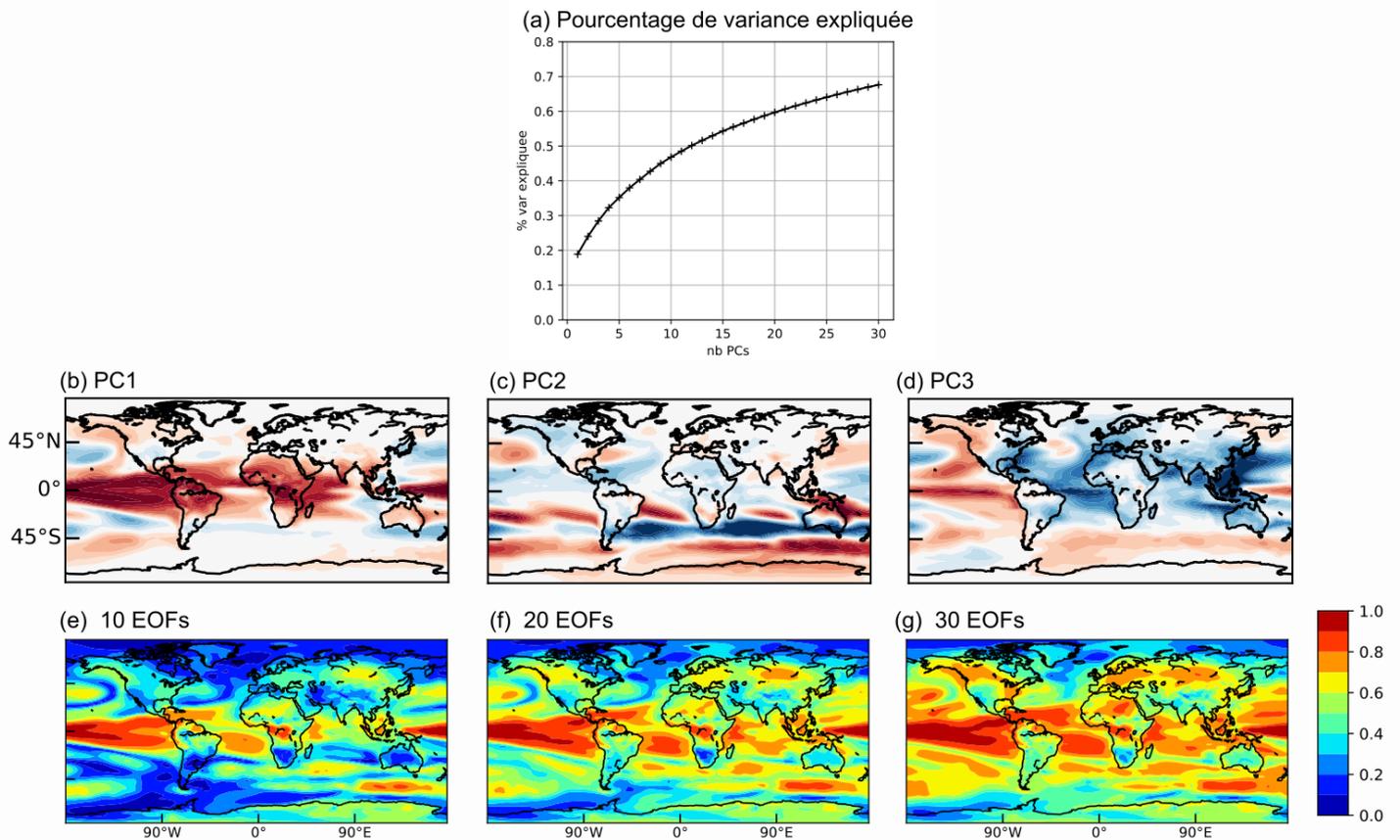


Figure 5.4: (a) Pourcentage cumulé de variance expliquée en fonction du nombre d'EOFs pour APP, (b-d) 3 premières EOFs et (e-g) Pourcentage cumulé de variance expliquée sur APP par les (e) 10, (f) 20 et (g) 30 premières EOFs.

quand la PC1 précède de 12 mois la PC3 et de 0.38 quand la PC3 précède de 12 mois la PC1. Ce résultat révèle la présence d'une dynamique potentiellement prédictible dans les tropiques entre l'EOF1 et l'EOF3. Alors qu'une bonne partie de la variabilité tropicale (+ de 50%) est capturée par les 10 premières EOFs (Fig. 5.4c et Tableau 5.2), ce n'est qu'à partir d'une vingtaine d'EOFs qu'un bon pourcentage (+ de 40%) de la variance extratropicale est également représentée. Nous noterons qu'en revanche le pourcentage de variance expliquée aux hautes latitudes de l'Hémisphère Nord (au-dessus de 45°N) reste faible même pour 30 EOFs.

Table 5.2: Pourcentage moyen de variance expliquée par hémisphère en fonction du nombre d'EOFs.

	10 EOFs	20 EOFs	30 EOFs
NH (au-dessus de 25°N)	28%	42%	50%
Tropique (entre 25°S-25°N)	55%	64%	70%
SH (en dessous de 25°S)	33%	46%	54%

5.6 Construction et Validation du LIM

Comme nous l'avons vu dans le chapitre 4, la modélisation par le LIM d'un système dynamique suppose que celui-ci est décomposable en une partie linéaire prédictible et une partie non-linéaire temporellement pas ou peu prédictible. Parmi les tests proposés par (Penland and Sardeshmukh, 1995) pour valider l'utilisation du LIM dans la modélisation d'un système, on trouve:

- la matrice de covariance \mathbf{Q} est définie positive (t1)
- la matrice \mathbf{B} est indépendante du τ_0 utilisé pour la calculer dans eq. 4.15 (t2)
- les prédictions faites par le LIM sont bonnes (t3)
- la croissance de l'erreur de prédiction du LIM est tel que prédite par Penland, 1989 (t4)
- le système est Gaussien (t5)

Nous vérifions ci-dessous l'ensemble de ces points pour le LIM construit à partir des $P = 30$ premières PCs et utilisé dans la section 5.7, 5.8 et 5.9.

(t1) Sous condition que la modélisation de \mathbf{pc} par (eq. 5.1) soit correcte, alors la matrice \mathbf{Q} représente la matrice de covariance du bruit (Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995). Les vecteurs propres de la matrice \mathbf{Q} généralement appelés « EOFs du bruit » traduisent la prise en compte lors de l'application du bruit gaussien ε par le LIM de corrélations spatiales (mais non temporelles) (Penland and Sardeshmukh, 1995). La matrice \mathbf{Q} comme toute matrice de covariance doit être définie positive et ses valeurs propres représenter le pourcentage de variance expliquée par chaque « EOFs du bruit » (Penland and Matrosova, 1994). La présence de valeurs propres négatives pouvant s'expliquer soit par l'utilisation d'une période d'apprentissage trop courte ou soit par la présence de processus non linéaire, le modèle reste toutefois utilisable si l'ordre de grandeur de ces valeurs propres négatives reste faible (Penland and Matrosova, 1994; Winkler et al., 2001). Dans le cadre de notre étude, des valeurs propres négatives n'apparaissant que pour des LIMs construits avec plus de 36 PCs ($P > 36$), la matrice \mathbf{Q} est donc bien définie positive pour tous les LIM des sections suivantes.

(t2) Le test d'indépendance de \mathbf{B} au τ_0 utilisé pour l'estimer est généralement appelé « τ test ». Si l'équation (eq. 5.1) est bien représentative de la dynamique du système étudié, alors en théorie l'estimation de la matrice \mathbf{B} est indépendante du τ_0 utilisé pour l'estimer dans (eq. 4.15). Bien qu'un échec à ce test d'indépendance n'implique pas la non-linéarité du système (Penland and Sardeshmukh, 1995), il constitue un argument très solide dans la validation du LIM en tant que modèle. Comme il s'avère qu'en pratique des bornes supérieures et inférieures existent dans le choix du τ_0 notamment à cause du problème de Nyquist (DelSole, 2000; Winkler et al., 2001), cette vérification de l'indépendance de \mathbf{B} au τ_0 utilisé est problématique en pratique. Une des astuces proposées par (Winkler et al., 2001) pour simplifier la validation du « τ test » est de valider celui-ci non pas par le calcul de \mathbf{B} pour différent τ_0 mais plutôt par la vérification de la bonne estimation de la matrice

de Green $\mathbf{G}(\tau)$ de (eq. 4.13) pour différents τ . Avec \mathbf{B} estimé pour un $\tau_0 = 11$ mois, si le système réel peut être correctement approché par un LIM la matrice de Green $\mathbf{G}(\tau)$ réelle devrait être bien estimée pour tout lag τ par $\exp(\mathbf{B}\tau)$. Dans la figure 5.5 similaire à la figure 6 de Winkler et al., 2001, la comparaison de la matrice de Green $\mathbf{G}(\tau)$ réelle avec celle reproduite par le LIM pour différent lag τ montre la bonne reproduction par le LIM de la matrice de Green $\mathbf{G}(\tau)$ réelle aux différents lags τ . Ceci valide d'une part que le système est principalement linéaire et d'autre part que le τ_0 utilisé pour le calcul de la matrice \mathbf{B} est adapté.

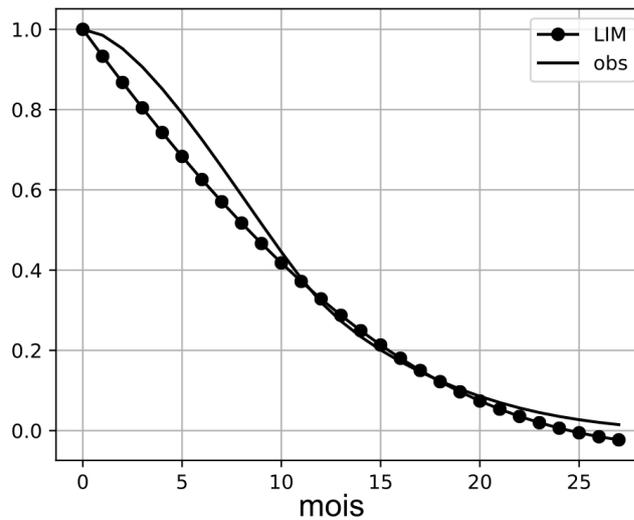


Figure 5.5: Validation du « τ test » avec la trace de la matrice de Green $\mathbf{G}(\tau)$ prédite par le LIM divisée par la trace de $\mathbf{G}(0)$ (courbes à points) et la trace de la matrice de Green $\mathbf{G}(\tau)$ observée normalisée par la trace de $\mathbf{G}(0)$ observée (courbe sans points) selon le nombre de mois (en abscisse).

(t3) Afin de vérifier que le LIM reproduit bien la dynamique du système souhaité, il est important d'évaluer la qualité des prédictions faites par celui-ci. En tant que modèle stochastique décrit par eq. 5.1, la meilleure prédiction (dans le sens la plus probable) de $\mathbf{pc}(t + \tau)$ à partir de $\mathbf{pc}(t)$ pour tout lag τ est:

$$\mathbf{pc}(t + \tau) = \mathbf{G}(\tau)\mathbf{pc}(t) \quad (5.5)$$

Les capacités de prédictions du LIM sont évaluées sur les 700 ans de température de surface de VAL aux horizons 12 et 18 mois relativement à la Persistance (Fig. 5.6, 5.7, 5.8). Le LIM surpasse clairement la Persistance sur l'ensemble du globe avec respectivement aux horizons 12 et 18 mois +0.11 et + 0.15 de corrélations globales et +0.35 et +0.51 de CE globale par rapport à la Persistance. Les scores obtenus par la persistance devenant très faibles au-delà de 18 mois, la différence de niveau avec le LIM est d'autant plus marquée. Les meilleurs scores de prédictions du LIM se situent dans les tropiques et le Pacific Nord avec des corrélations moyennes de 0.26, 0.31 et 0.44 à l'horizon 12 mois et 0.19, 0.22 et 0.31 à l'horizon 18 mois pour respectivement l'Hémisphère Nord, l'Hémisphère Sud et les tropiques. Les différences notables obtenues dans la localisation des régions les mieux

prédites par notre LIM et celles trouvées par [Newman, 2013](#) s'expliquent certainement d'une part, par la présence de forçages contribuant de façon significative à la construction de leur LIM (voir Figure 4 de [Newman, 2013](#)) et d'autre part, par les différences entre la dynamique observée sur la période historique dans un contexte de changement climatique et celle reproduite par le modèle de l'IPSL-CM5A-LR en contexte préindustrielle.

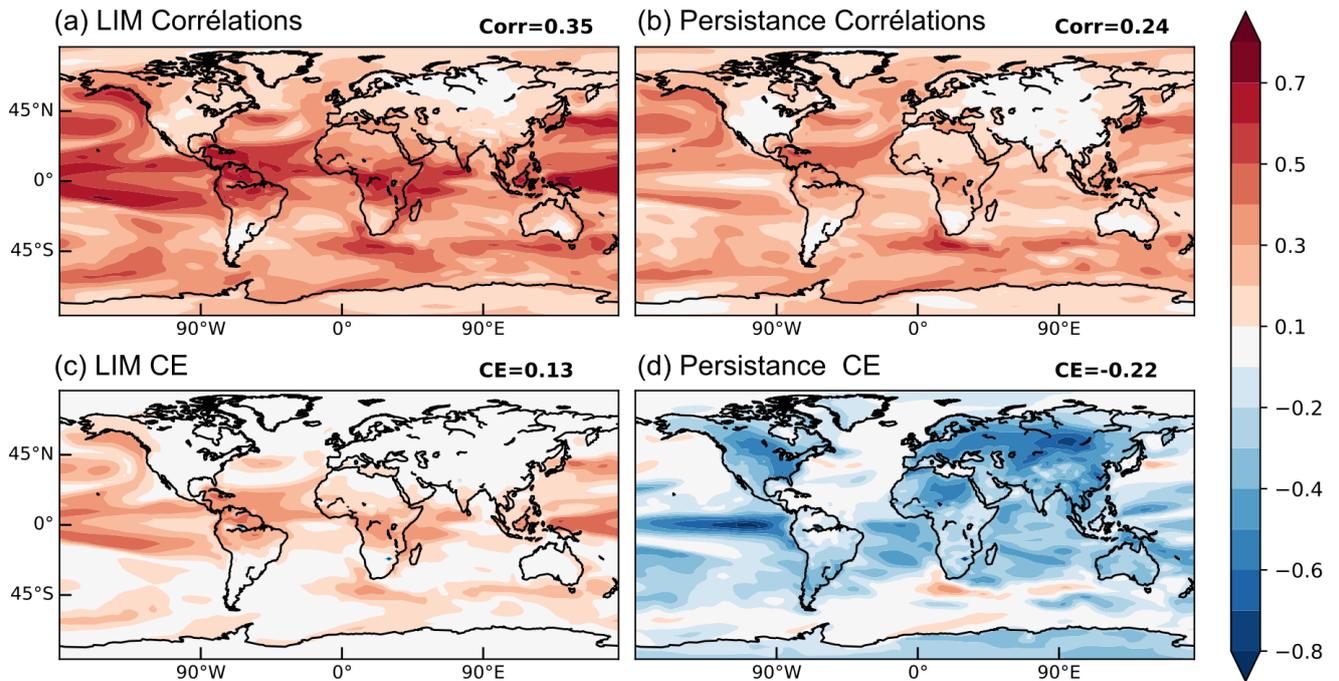


Figure 5.6: (a-b) Corrélations et (c-d) coefficients CE entre les températures de surface réelles de VAL et les températures de surface prédites par (a,c) le LIM et (b,d) la Persistence à l'horizon 12 mois. Les valeurs moyennes de corrélations et CE globales sont indiquées en haut à droite de chaque panneau.

(t4) [Penland, 1989](#) ayant montré que dans le cas de système purement linéaire, la matrice de covariance de l'erreur de prédiction σ croit telle que :

$$\langle \sigma(\tau), \sigma(\tau)^T \rangle = \mathbf{C}(0) - \mathbf{G}(\tau)\mathbf{C}(0)\mathbf{G}(\tau)^T \quad (5.6)$$

le test (t4) constitue une façon supplémentaire de valider l'utilisation d'un modèle linéaire en comparant la croissance de l'erreur de prédiction réelle du LIM avec la croissance théorique prédite par (eq. 5.6). Cette croissance théorique étant inférieure à la croissance réelle (Fig. 5.9), la différence entre les deux courbes reflète la présence de composantes non linéaires. Néanmoins le fait que la courbe du LIM soit à la fois inférieure à celles de la Persistence, de la climatologie et de même allure que la courbe théorique, confirme la dynamique principalement linéaire du système décrits par les 30 premières PCs. L'allure de la courbe serait pour sa part très différente (voir Figure B1 de [Penland and Sardeshmukh, 1995](#)) dans le cas de système non linéaire.

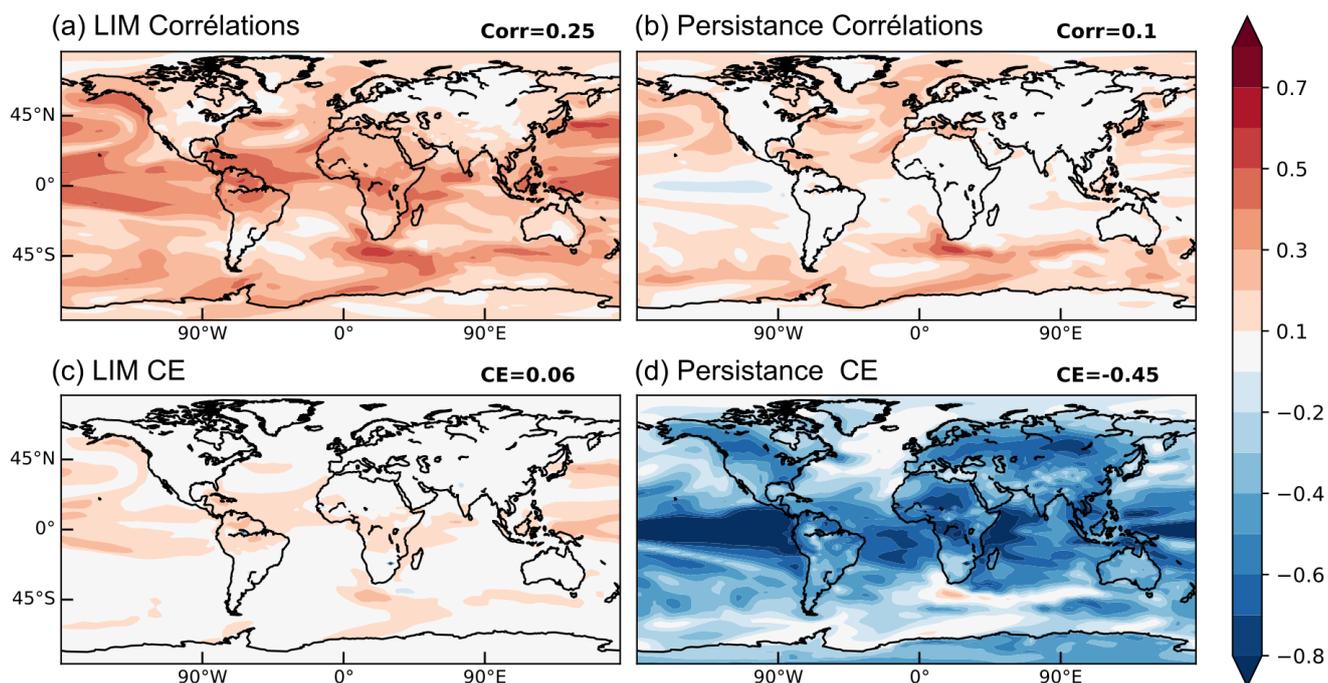


Figure 5.7: Comme la Figure 5.6 mais à l'horizon 18 mois.

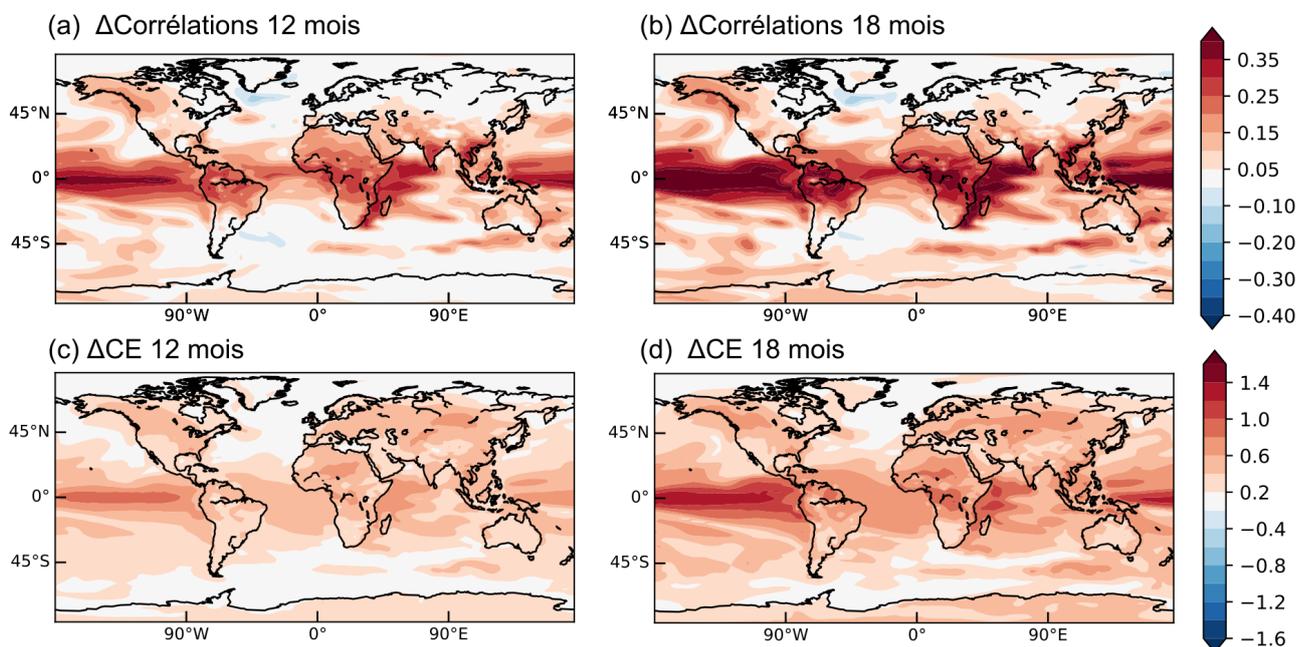


Figure 5.8: Différence de (a-b) Corrélations et (c-d) d'indice CE entre le LIM et la Persistence aux horizons (a,c) 12 mois et (b,d) 18 mois.

(t5) Le LIM ne pouvant reproduire que la dynamique de système Gaussien (Penland and Sardeshmukh, 1995; Martinez-Villalobos et al., 2018), nous comparons dans la figure 5.10, les distributions des 9 premières PCs réelles avec celles simulées par le LIM. Nous avons pour cela intégré le LIM constitué des 30 premières PCs sur une période de 10

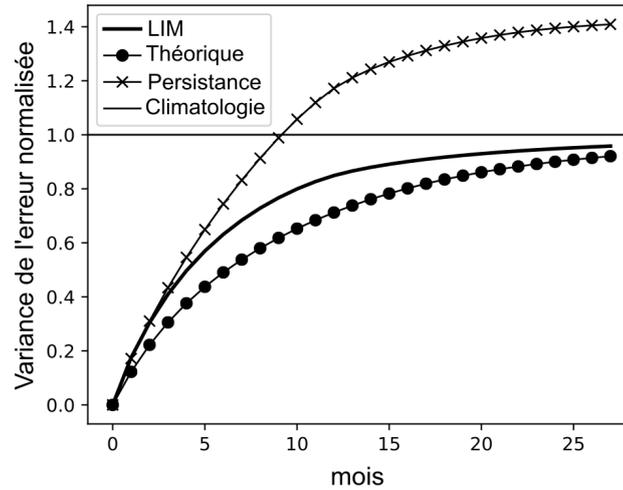


Figure 5.9: Croissance de la variance de l'erreur normalisée en fonction du nombre de mois (abscisse) pour la croissance théorique (courbe avec des points), les prédictions du LIM (courbe épaisse), la Persistance (courbe avec des croix) et la climatologie (ligne horizontale).

000 ans et comparé les distributions obtenues pour les 9 premières PCs avec celles des 9 premières PCs de APP (d'une période de 1 000 ans) et de VAL (d'une période de 700 ans). Dans l'ensemble le LIM reproduit de façon convaincante les statistiques observées avec des distributions similaires à celles de APP et VAL pour les 6 premières PCs et des biais visibles à partir des PCs 7, 8 et 9.

En plus de valider les tests (t1-5) proposés par [Penland and Sardeshmukh, 1995](#) nous évaluons également les distributions des incréments du LIM entre k et $k+h$ pour différents horizons h . Nous comparons pour cela dans la figure 5.11 à l'aide de diagrammes de rangs, les distributions $\mathbf{p}_x(\mathbf{pc}_{k+h} - \mathbf{pc}_k / \mathbf{pc}_k)$ prédites par le LIM aux horizons $h=4$, $h=8$ et $h=12$ mois avec les distributions réelles pour les 3 premières PCs (Fig. 5.11). Nous avons pour cela réalisé la construction de 100 diagrammes de rangs pour chaque PC aux différents horizons (Fig. 5.11). Comme on peut le voir les distributions $\mathbf{p}_x(\mathbf{pc}_{k+h} - \mathbf{pc}_k / \mathbf{pc}_k)$ semblent bien reproduites pour les 3 premières PCs à l'horizon $h=12$ mois. Mais la forme en « U » des diagrammes de rangs des PC2 et PC3 à l'horizon 4 mois montrent que le LIM sous-estime la variabilité de ces 2 PCs à court terme en sous-estimant les distributions $\mathbf{p}_x(\mathbf{pc}_{k+h} - \mathbf{pc}_k / \mathbf{pc}_k)$.

5.7 Le LIM en tant qu'émulateur des modèles climatiques

Ayant discuté dans la section précédente de la bonne reproduction par le LIM du système décrit par les 30 premières PCs, nous souhaitons maintenant examiner les capacités du LIM à reproduire certaines caractéristiques du climat. Nous examinons pour cela les prédictions du LIM pour les indices des principaux modes de variabilité climatiques dont l'ENSO, l'IPO via l'indice TPI et l'AMV. Dans la représentation de ces indices par le LIM nous distinguerons d'une part l'erreur dite de « réduction » ou « représentativité

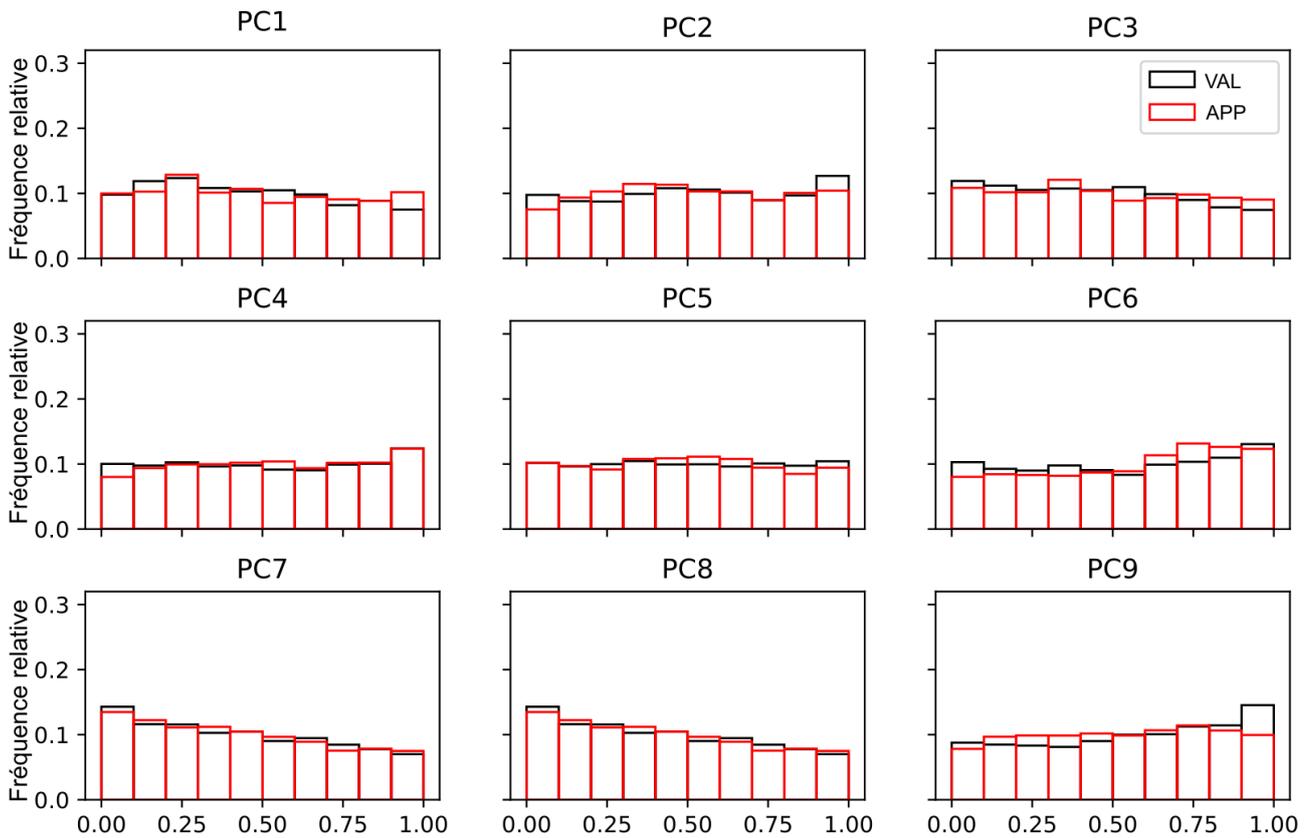


Figure 5.10: Diagramme de rang des 9 premières PCs entre les distributions prédites par un LIM de 10 000 ans et les distributions des PCs de APP (en rouge) et de VAL (en noir).

» liée à l'utilisation d'un sous ensemble de 30 EOFs pour représenter la variabilité des températures de surface et de l'erreur dite de « modélisation » liée à la représentation par le LIM du système des 30 premières PCs par l'éq 5.1. Nous évaluons dans un premier temps la qualité des prédictions pour les différents indices climatiques faites par la composante linéaire-déterministe du LIM, dans un second temps nous illustrons l'intérêt de la composante stochastique non-linéaire du LIM pour estimer l'incertitude lié à ces prédictions et pour finir nous discutons des limites du LIM pour reproduire certaines caractéristiques du climat.

Les prédictions du LIM sont évaluées dans la figure 5.12 jusqu'à l'horizons 3 ans et comparées à d'une part la persistance des températures de surface réelles (noté Persistance dans la Figure 5.12) et d'autre part à la persistance des températures de surface reconstruites à partir des 30 premières PCs (noté EOFs Persistance dans la Figure 5.12). Comme on peut le voir les prédictions associées à la Persistance et à EOFs Persistance sont similaires et se superposent pour le TPI et l'ENSO alors qu'elles sont différentes pour l'AMV. Les prédictions du LIM surpassent clairement les deux types de persistances pour les indices du TPI et de l'ENSO mais pas pour l'AMV. On peut en effet observer que pour l'AMV les prédictions du LIM surpassent celles de EOFs Persistance mais sous-performe la Persistance aux horizons inférieurs à 29 mois (Fig. 5.12e). En d'autre terme la modélisation

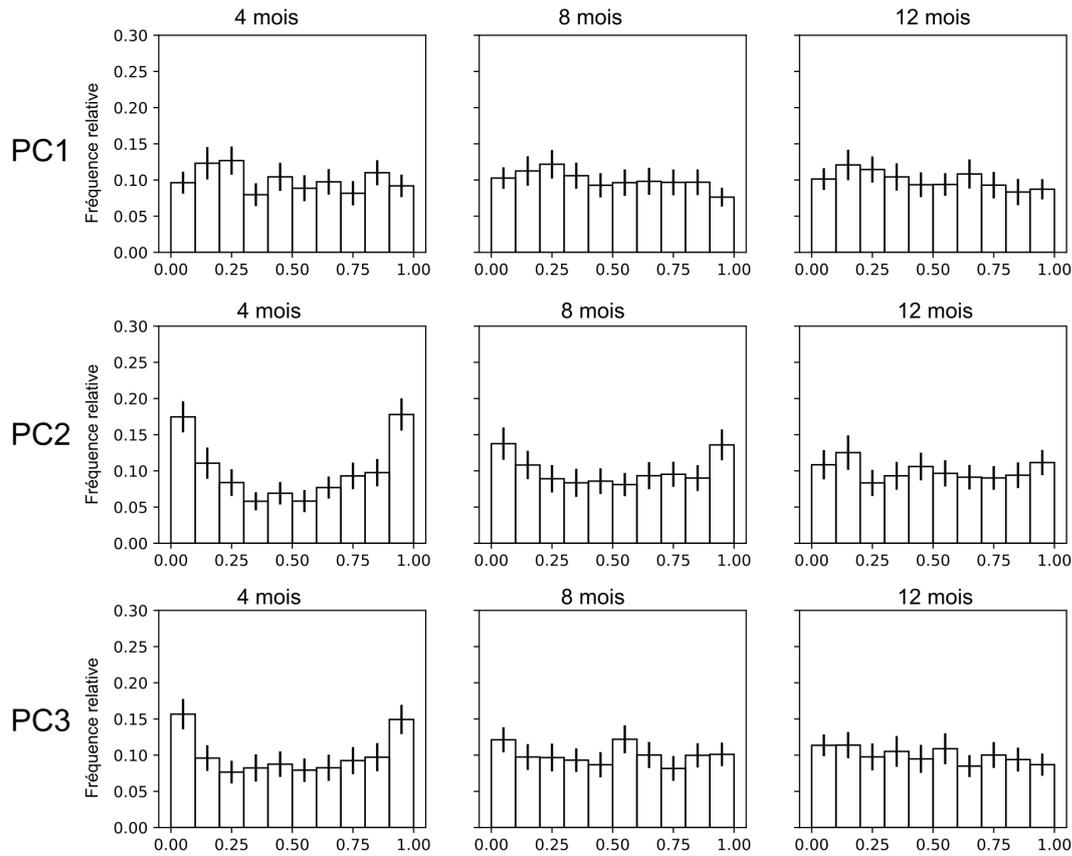


Figure 5.11: Diagrammes de rang moyens entre les distributions $p_x(\mathbf{pc}_{k+h} - \mathbf{pc}_k / \mathbf{pc}_k)$ réelles et prédites par le LIM aux horizons $h=4$ mois (colonne de gauche), $h=8$ mois (colonne du milieu) et $h=12$ mois (colonne de droite). Pour chaque PC le diagramme de rang moyen est estimé aux différents horizons comme la moyenne de 100 diagrammes de rang et la barre d'erreur représente 2 fois la déviation standard à travers l'ensemble de ces réalisations. Chacun des 100 diagrammes de rang aux différents horizons a été estimé pour 1000 prédictions prises aléatoirement sur VAL.

par le LIM du système des 30 première PCs sous forme eq 5.1 n'est pas à l'origine de cette sous performance de l'AMV, celle-ci ayant plutôt pour cause les différences marquées entre l'indice d'AMV réel et l'indice d'AMV des températures de surface reconstruites à partir des 30 premières PCs (AMV-30EOFs). Une simple correction dans un LIM dit « corrigé » dans la figure 5.12 permet de prendre en compte ce biais en ajoutant aux prédictions du LIM le biais initial présent entre les températures de surface réelles et reconstruites. La prise en compte de ce biais de représentativité permet aux prédictions du LIM « corrigé » de surpasser la Persistance pour l'indice d'AMV (Fig. 5.12c). Ces différences marquées entre l'indice d'AMV réel et l'indice AMV-30EOFs sont dues au faible pourcentage de variance expliquée au-dessus de 45° nord dans l'Atlantique Nord (Fig. 5.4). Avec des corrélations de 0.61 et 0.47 entre la variabilité tropicale et respectivement l'indice d'AMV-30EOFs et l'indice d'AMV réel, l'indice AMV-30EOFs représente d'avantage la partie tropicale de la variabilité de l'Atlantique Nord. En d'autre terme nous retiendrons que le LIM surpasse de façon convaincante les prédictions obtenues par la persistance associée à nos 30 premières PCs pour les indices du TPI, ENSO et de

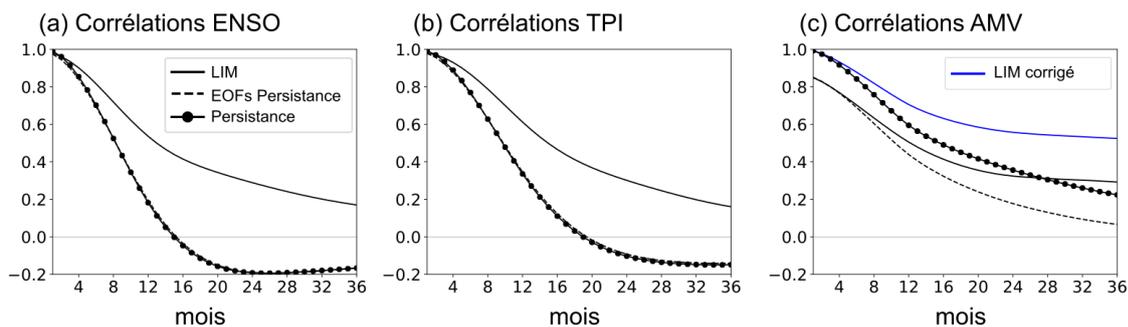


Figure 5.12: Reproduction par le LIM de la partie prédictible des indices (a) d'ENSO, (b) du TPI et (c) de l'AMV. Corrélations obtenues pour différents horizons sur les 1000 ans de VAL pour les prédictions du LIM (courbe noire), la persistance des températures de surface réelles (courbe noire avec des points) et la persistance des températures de surface reconstruites à partir des 30 premiers PCs (courbe pointillée). Pour (c) le LIM corrigé (courbe en bleu) représente le score obtenu par les prédictions du LIM aux différents horizons après ajout aux prédictions du LIM du biais initial entre les températures de surface reconstruites à partir des 30 premières PCs et les températures de surface réelles.

l'AMV mais que notre sous ensemble d'EOFs ne donne qu'une description partielle de la variabilité des températures de surface dans l'Atlantique Nord. Ne faisant pas partie du cadre de notre étude, nous noterons toutefois que dans le cas d'expériences spécifiques, c'est-à-dire avec un sous-ensemble d'EOFs dédié au bassin Atlantique, le LIM a montré aussi bien à l'échelle saisonnière que décennale des capacités de prédictions comparable à celle des GCMs dans l'Atlantique Nord (Hawkins and Sutton, 2009; Zanna, 2012; Huddart et al., 2017).

Le LIM étant composé d'une partie prédictible supposée linéaire et d'une composante non linéaire modélisable à l'aide d'un bruit gaussien de covariance \mathbf{Q} , nous illustrons dans les figures 5.13, 5.14 et 5.15 la combinaison de ces deux composantes pour les différents indices discutés précédemment (TPI, ENSO, AMV). Un ensemble de 100 particules a pour cela été intégré sur une période de 2 ans pour différents états initiaux de manière à illustrer le comportement du LIM dans différentes situations. Cette figure donne un aperçu du comportement du LIM pour différentes situations et permet d'illustrer l'intérêt de la composante stochastique du LIM pour propager dans le temps un ensemble de particules de manière à décrire l'ensemble des états futurs probables à partir de la connaissance du système à $t=0$. Comme ceci a déjà été observé dans la figure 5.12 et discuté précédemment, on note des différences marquées entre l'indice d'AMV réel et l'indice d'AMV-30EOFs (Fig. 5.15). La composante stochastique du LIM donne une bonne estimation de l'incertitude liée aux prédictions du TPI (Fig. 5.13) et de l'ENSO (Fig. 5.14) par le LIM. L'AMV-30EOFs étant principalement associé à la variabilité tropicale dans l'Atlantique, la bonne reproduction de l'AMV par notre sous ensemble d'EOFs est pour sa part plus discutable (Fig. 5.15).

N'ayant jusqu'ici considéré les performances du LIM qu'en moyenne temporelle, nous souhaitons à présent distinguer les performances obtenues par le LIM selon les circonstances et l'état du système à la date initiale. Diverses études ont montré que la prévisibilité du climat varie dans le temps en fonction de son état de départ (Brune et al.,

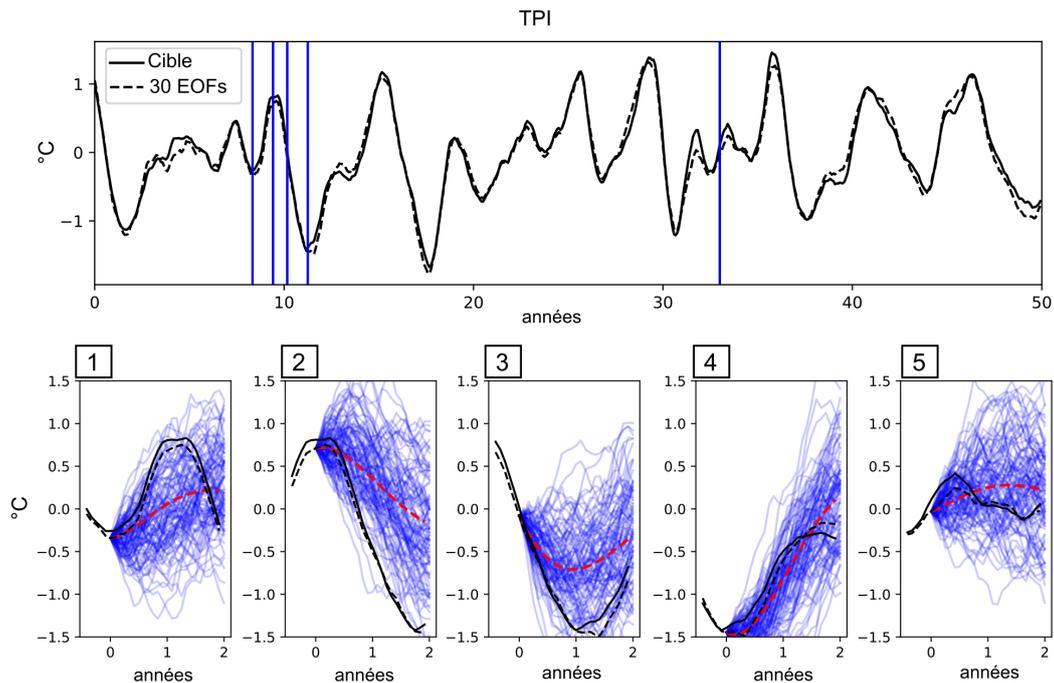


Figure 5.13: Panneau du haut : Indice du TPI pour les 50 premières années de VAL calculé avec les températures de surface réelles en trait pleins et avec les températures de surface reconstruites avec les 30 premières PCs en pointillés noirs. Les traits verticaux indiquent les dates de départs des 5 panneaux du bas dans lesquelles la courbe en pointillés rouges représente la composante déterministe linéaire du LIM et les trajectoires en bleu montrent l'ensemble des 100 particules décrivant la composante stochastique non-linéaire du LIM.

2018) et que la crédibilité d'une prédiction dépend de l'état du climat à la date initiale (Planton et al., 2018; Borchert et al., 2019). Il a été notamment montré que le contenu en chaleur dans l'océan Pacifique tropical Ouest donnait de meilleure prédiction pour les événements du type El Niño que les événements La Niña (Planton et al., 2018) et que le niveau de confiance associée aux prédictions de l'AMV était meilleur lorsque des anomalies de transport de chaleur dans l'océan Atlantique Nord subpolaire étaient observées à la date initiale (Borchert et al., 2019). Nous évaluons dans la figure 5.16 l'évolution des erreurs de prédictions des indices d'ENSO, du TPI et de l'AMV par le LIM en fonction de leur phase négative, neutre ou positive à la date de départ. Ne pouvant produire, dans le cas d'un LIM, que des prédictions et distributions symétriques (Martinez-Villalobos et al., 2018), la figure 5.16 permet en plus d'évaluer les performances du LIM pour les 3 indices, d'évaluer la présence plus ou moins importante d'une dynamique non-linéaire de ces indices aux différents horizons en fonction de leur valeur initiale.

Nous savons par exemple de nos jours que l'ENSO possède des caractéristiques asymétriques, avec des événements El Niño pouvant atteindre des amplitudes plus importantes que celles d'événement La Niña mais que ces derniers tendent à être plus persistants dans le temps (Hoerling et al., 1997; Okumura and Deser, 2010). Les transitions d'une phase à l'autre sont également asymétriques, leur déclenchement et succession impliquant des dynamiques différentes (Larkin and Harrison, 2002; Planton et al., 2018).

L'erreur de prédiction du LIM étant relativement plus faible à court terme lors de La

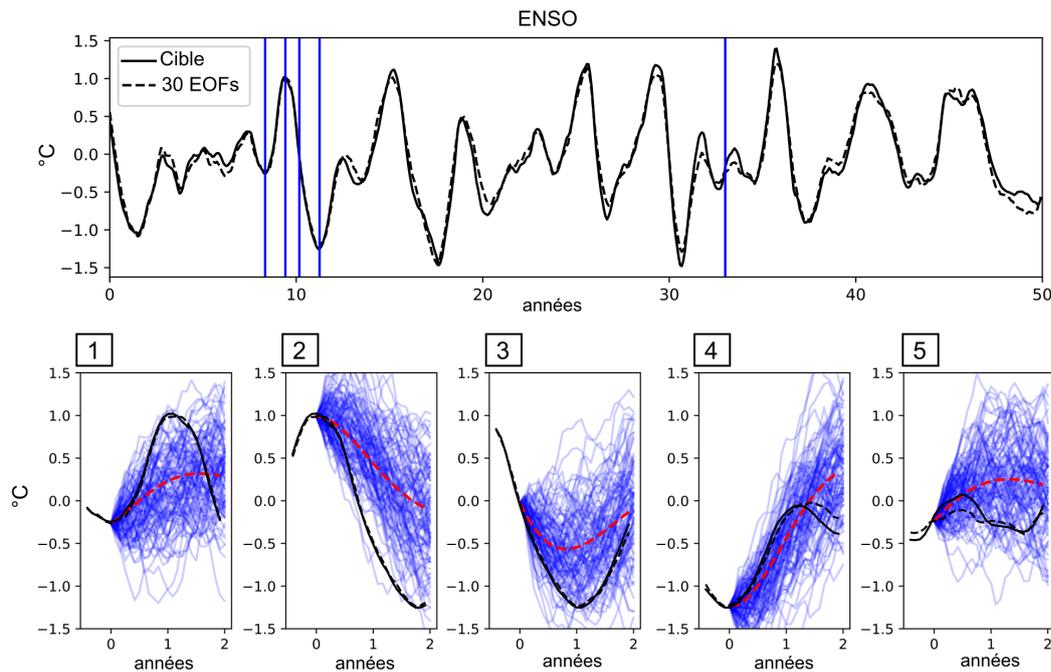


Figure 5.14: Panneau du haut : Indice d'ENSO pour les 50 premières années de VAL calculé avec les températures de surface réelles en trait pleins et avec les températures de surface reconstruites avec les 30 premières PCs en pointillés noirs. Les traits verticaux indiquent les dates de départs des 5 panneaux du bas dans lesquelles la courbe en pointillés rouges représente la composante déterministe linéaire du LIM et les trajectoires en bleu montrent l'ensemble des 100 particules décrivant la composante stochastique non-linéaire du LIM.

Niña et à long terme lors d'événement El Niño (Fig. 5.16), la composante linéaire associée à la dynamique de l'ENSO (du modèle de l'IPSL) est donc relativement plus marquée à court terme dans le cadre d'événement La Niña et plus prononcée à long terme dans le cadre d'événement El Niño. Ces résultats suggèrent que les événements La Niña moins sujets au forçage stochastique des vents, d'une durée plus longue et d'une amplitude moins variable (Menkes et al., 2014; Puy et al., 2017) donne à court terme une dynamique plus régulière/linéaire que les événements El Niño. En revanche à long terme les événements El Niño étant plus souvent suivis par des La Niña que les événements La Niña par des El Niño, les prédictions à long terme lors d'événements El Niño tendent à être meilleures que celles réalisées lors d'événement La Niña (Larkin and Harrison, 2002). Bien que moins prononcé, un raisonnement, similaire, semble pouvoir s'appliquer à l'AMV avec en particulier des prédictions à court et moyen terme qui semblent meilleures dans la cadre de phases neutres de l'AMV que de phases non-neutres (positives ou négatives). Sortant du cadre de notre étude, nous soulignerons que potentiellement chacun de ces modes de variabilités peut influencer la dynamique des autres bassins océaniques et que la qualité des prédictions obtenue pour chaque bassin dépend certainement de l'état des autres bassins à la date initiale (Zanchettin et al., 2016; Levine et al., 2017).

Malgré les hypothèses fortes du LIM lors de la représentation du système dynamique des 30 premières PCs des températures de surface comme la somme d'un terme prédictible

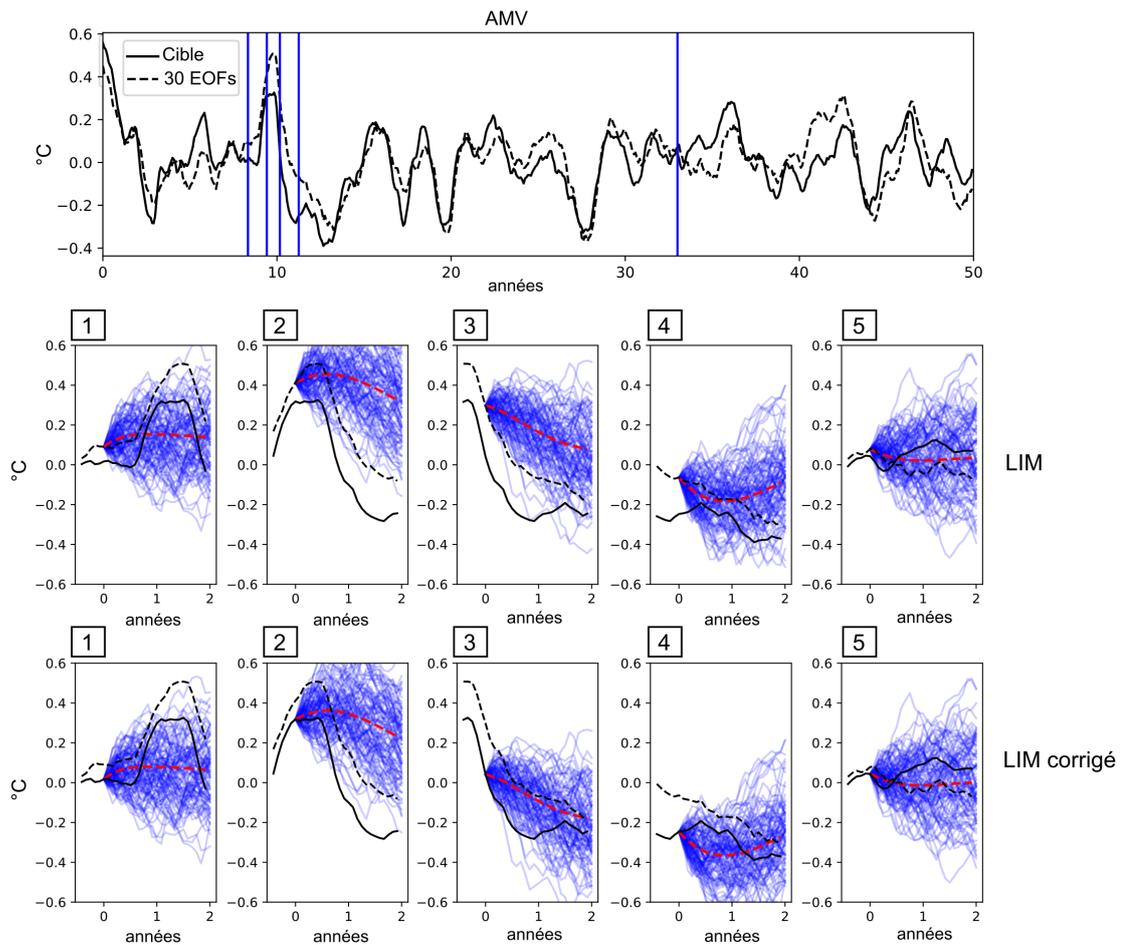


Figure 5.15: Indice de l'AMV pour les 50 premières années de VAL calculé avec les températures de surface réelles en trait pleins et avec les températures de surface reconstruites avec les 30 premières PCs en pointillés noirs. Les traits verticaux indiquent les dates de départs des 10 panneaux numérotés de 1 à 5 du bas dans lesquelles la courbe en pointillés rouges représente la composante déterministe linéaire du LIM et les trajectoires en bleu montrent l'ensemble des 100 particules décrivant la composante stochastique non-linéaire du LIM (ligne du milieu) et du LIM corrigé (ligne du bas)

linéaire et d'un terme stochastique non-linéaire modélisable à l'aide d'un bruit Gaussien ϵ , le LIM constitue au vu des résultats précédents un bon candidat pour représenter la dynamique des températures de surface à la fois en terme de prévisions qu'en terme d'estimation de l'incertitude et sera utilisé dans les sections suivantes comme un émulateur des modèles climatiques permettant l'intégration d'un grand nombre de particules.

5.8 Reconstructions spatiales de la variabilité des températures de surface

Après avoir validé l'utilisation du LIM comme modèle dynamique du système décrit par les 30 premières PCs des températures de surface, nous souhaitons maintenant évaluer la qualité de reconstructions des températures de surface obtenues avec le SIR-LIM. Les

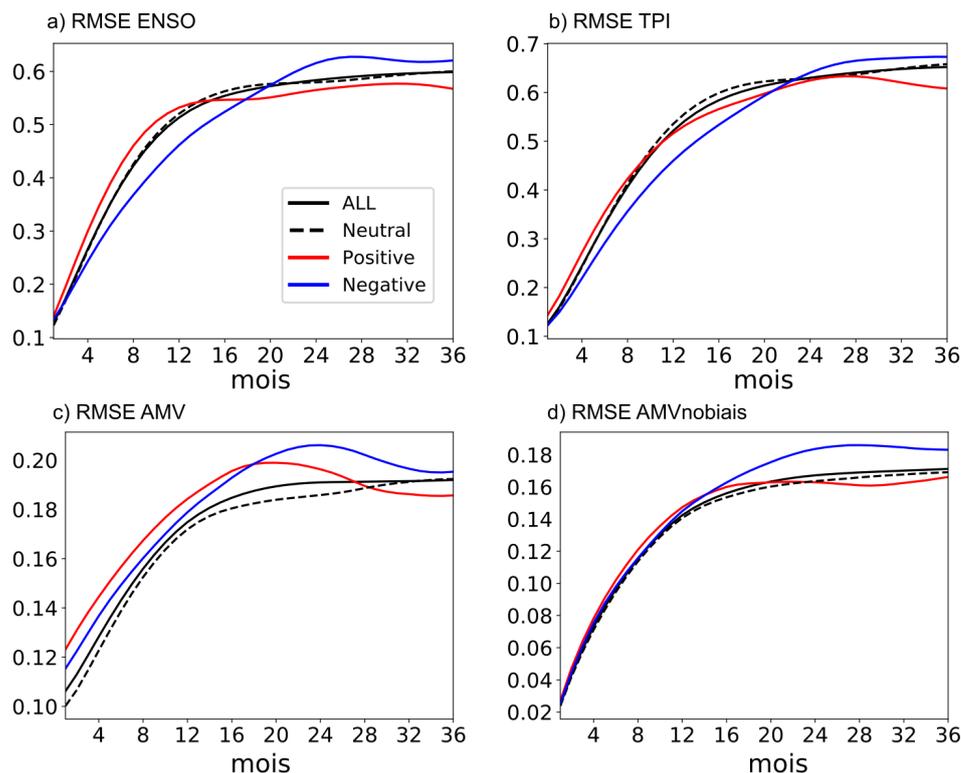


Figure 5.16: RMSE des prédictions du LIM pour l'ENSO (en haut à gauche), le TPI (en haut à droite), l'AMV (en bas à gauche) et les prédictions du LIM corrigé pour l'AMV (en bas à droite). En abscisse l'horizon en mois. La courbe noire représente la RMSE générale des prédictions, la courbe en rouge représente la RMSE des prédictions lorsque l'indice dépasse 1 déviation standard à la date initiale, la courbe en bleu lorsque l'indice est inférieur à -1 déviation standard à la date initiale et la courbe en pointillés lorsque l'indice est compris entre -1 et 1 déviation standard à la date initiale.

expériences de cette section ont été réalisées pour une période de 150 ans prise arbitrairement sur VAL. Le LIM utilisé correspond à celui construit à partir des 30 premières EOFs des températures de surface et testé dans la section 5.6 et 5.7. Le SIR-LIM et SIR-0 utilisent un pas de ré-échantillonnage de 1 ans et un ensemble de 900 particules. Nous analysons ci-dessous les reconstructions du SIR-LIM et SIR-0 obtenues pour les 3 premières PCs ainsi que les températures de surface reconstruites à partir des 30 premières EOFs dans les expériences Dense et Pages2k.

Contribution du LIM dans le SIR-LIM:

Les scores obtenus par le SIR-LIM et SIR-0 dans les expériences Dense et Pages2K pour les 3 premières PCs sont indiquées dans le Tableau 5.3. Comme on peut le voir les scores de corrélations du SIR-0 étant non nul, le score dit de « ré-échantillonnage » contribue de façon non négligeable dans les scores du SIR-LIM. Les différences de scores entre les deux SIR prouvent néanmoins de façon convaincante l'utilité du LIM dans l'intégrations des particules. Aucune composante dynamique n'intervenant dans le modèle d'intégration du SIR-0, les différences de scores entre les PCs du SIR-0 reflètent l'importance relative accordée à chaque PCs lors de l'étape de ré-échantillonnage et de sélection des particules

(Tableau 5.3).

Table 5.3: Corrélations et CE de la reconstruction des PC1 à PC3 obtenus avec le SIR-LIM et le SIR-0 sur une période de 150 ans prise sur VAL pour les expériences Pages2k et Dense.

	Pages2k				Dense			
	SIR-LIM		SIR-0		SIR-LIM		SIR-0	
	Corr	CE	Corr	CE	Corr	CE	Corr	CE
PC1	0.87	0.70	0.15	0.01	0.92	0.84	0.12	-0.11
PC2	0.48	0.23	0.09	-0.06	0.72	0.51	0.09	-0.29
PC3	0.62	0.35	0.08	-0.05	0.72	0.47	0.07	-0.22

Impact du sous-échantillonnage spatial sur les reconstructions:

Bien que visible dans les 3 PCs, les différences de scores obtenus entre les expériences Pages2k et Dense varient d'une PC à l'autre (Tableau 5.3). La 1^{ère} PC qui décrit principalement la variabilité tropicale (Fig. 5.4) semble la mieux reconstruite pour les deux expériences. La PC2 représentative de la variabilité dans l'Hémisphère Sud semble quant à elle la plus impactée des 3 premières PCs par la différence de sous-échantillonnage spatial entre Pages2k et Dense (Tableau 5.3). Ces écarts reflètent le manque de proxys réels disponibles pour décrire la variabilité dans l'Hémisphère Sud et en particulier la variabilité décrite par la PC2 (Fig. 5.17).

La taille effective de l'échantillon N_{eff} introduite (eq. 4.9) dans le Chapitre 4 permet de suivre la dégénérescence du filtre particulaire dans le temps. Comme on peut le voir sur la figure 5.17, le ré-échantillonnage régulier des particules permet d'éviter la dégénérescence du filtre SIR-LIM dans le temps. Toutefois le nombre de particules nécessaire augmentant de façon exponentielle avec la dimension du problème (Snyder et al., 2008; Van Leeuwen, 2009), celui-ci augmente également lorsque l'incertitude sur le système observé diminue (Dubinkina et al., 2011). Ce risque de dégénérescence plus élevé est visible pour Dense avec des valeurs de N_{eff} en moyenne 4.6 fois plus petites que dans Pages2k (Fig. 5.17d). Cependant lorsque la description plus fine du système observé n'engendre pas la dégénérescence du filtre, celle-ci permet d'obtenir comme dans l'expérience Dense de meilleures reconstructions (Tableau 5.3, Fig. 5.17). Les incertitudes liées à ces reconstructions semblent correctement représentées par l'ensemble des particules avec la cible comprises dans l'enveloppe des particules 98.8%, 99.7% et 98.9% du temps pour respectivement les PC1, PC2 et PC3 de Pages2k et 97.2%, 99.3% et 94.1% du temps pour respectivement les PC1, PC2 et PC3 de Dense (Fig. 5.17).

La qualité de reconstruction des différentes PCs se retrouvent dans la reconstruction spatiale des températures de surface (Fig. 5.6), avec au première ordre une reconstruction meilleure dans les Tropiques et le pacifique Nord que le reste du globe et des reconstructions moins bonnes dans l'Hémisphère Sud pour Pages2k. Des scores tout à fait

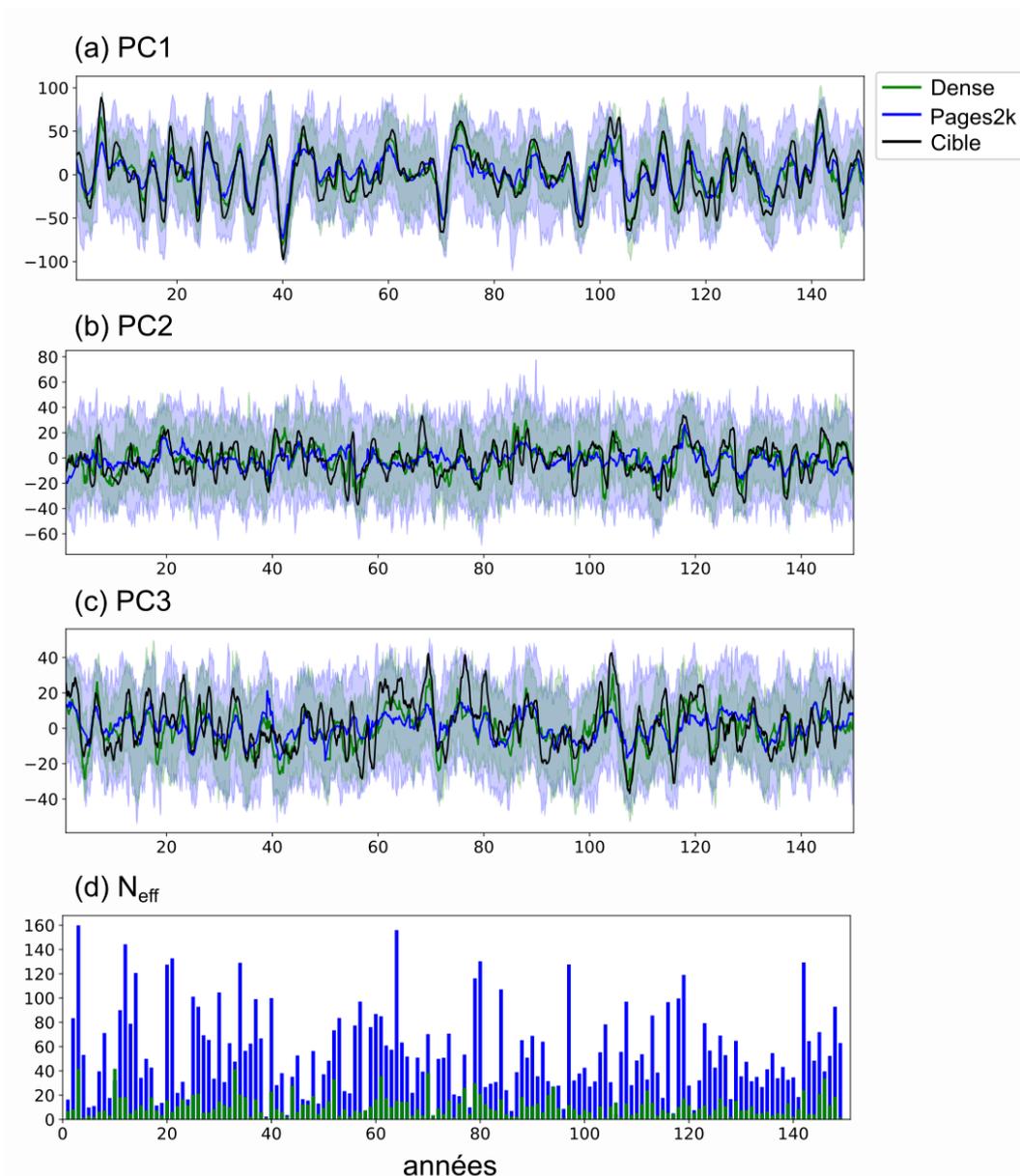


Figure 5.17: (a-c) Reconstruction par le SIR-LIM des 3 premières PCs des températures de surface pour les expériences Pages2k et Dense sur une période de 150 ans prise dans VAL. Les traits pleins noirs indiquent la cible à reconstruire. Les traits pleins bleus et verts indiquent les reconstructions obtenues par la moyenne de l'ensemble des particules et les enveloppes bleues et vertes décrivent la dispersion totale de l'ensemble des particules. (d) Evolution temporelle de la taille effective de l'échantillon pour le SIR-LIM estimée par eq 4.9 pour l'expérience Pages2k (en bleu) et Dense (en vert).

satisfaisants sont visibles dans les régions à la frontière des océans Atlantique et Arctique, et plus généralement au-dessus de 50°N , malgré le faible pourcentage de variance expliquée par les 30 premières EOFs (Fig. 5.4).

Malgré des reconstructions globalement meilleures du SIR-LIM pour Dense que Pages2k, on remarque pour certaines régions des reconstructions meilleures pour Pages2k (Fig. 5.8). Ces régions correspondent aux zones avec une densité de pseudo-proxys plus

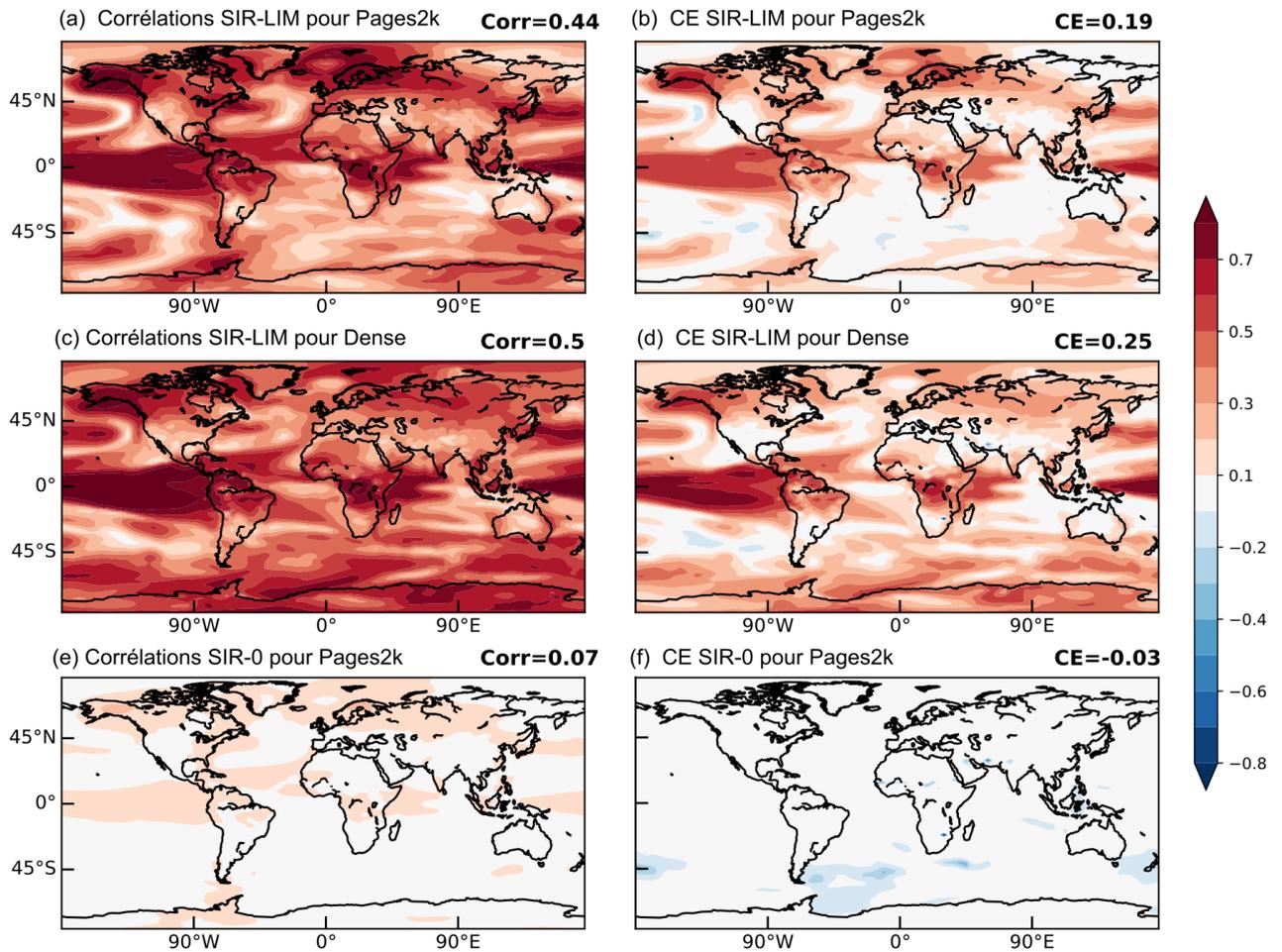


Figure 5.18: (a,c,e) Corrélations et (b,d,f) CE pour (a,b) le SIR-LIM dans l'expérience Pages2k, (c,d) le SIR-LIM dans l'expérience Dense et (e,f) le SIR-0 dans l'expérience Pages2k. Les scores sont estimés sur les reconstructions des températures de surface d'une période de 150 ans prise sur VAL. Les moyennes globales de corrélations et CE sont indiquées en haut à droite de chaque panneau.

importante dans Pages2k que Dense (Fig. 5.3), comme au sud du Chili et de l'Argentine, au sud de l'Alaska, la région de l'Atlantique nord-ouest avec un grand nombre de proxys le long des côtes est d'Amérique du Nord, vers l'Islande à la frontière de l'océan Atlantique et de l'océan Arctique, ou encore la région du nord de l'Espagne jusqu'à la côte ouest d'Afrique du Nord.

5.9 Reconstructions des principaux modes de variabilité pour le modèle IPSL-CM5A-LR

Ayant effectué dans la section 5.8 une première analyse spatiale des reconstructions obtenues par le SIR-LIM nous nous intéressons dans cette section à la reconstruction spécifique de l'ENSO, de l'AMV et de la variabilité à l'échelle du bassin Pacifique à l'aide du TPI. La figure 5.20 montre les reconstructions de ces indices avec le SIR-LIM pour les expériences Dense et Pages2k, l'ensemble des particules donne une estimation de

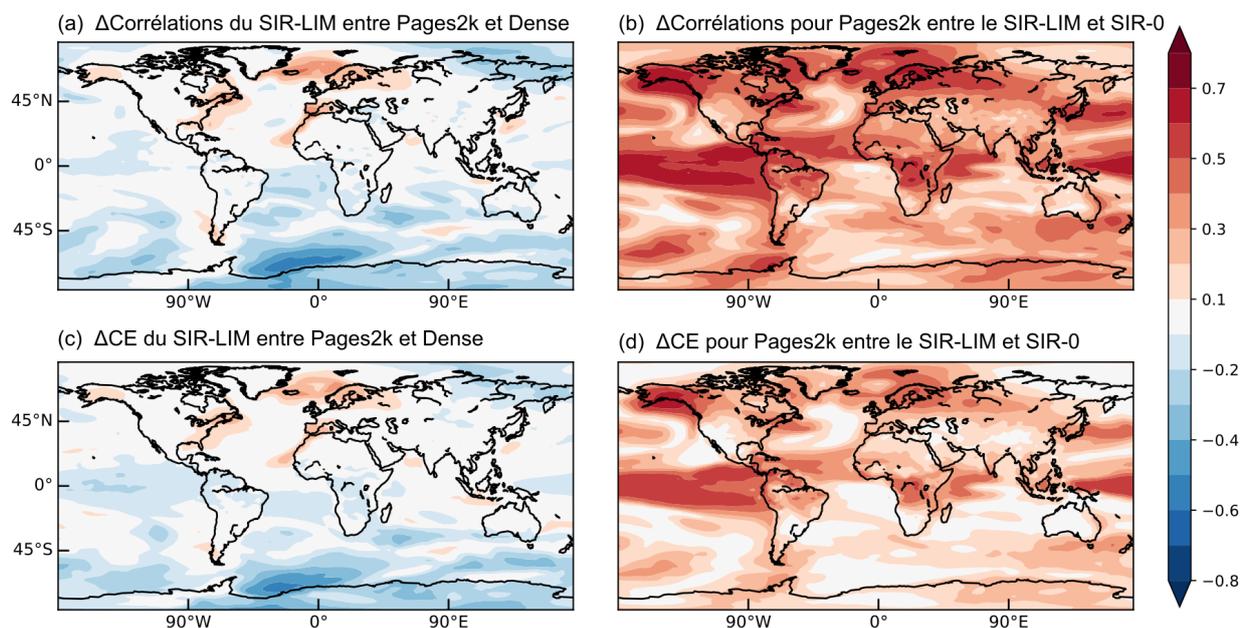


Figure 5.19: Différences de (a-b) corrélations et (c-d) CE (a,c) entre les reconstructions du SIR-LIM pour l'expérience Pages2k et l'expérience Dense et (b,d) entre les reconstructions obtenus par le SIR-LIM et le SIR-0 pour l'expérience Pages2k.

l'incertitude associée à ces reconstructions avec un intervalle de confiance de 99%, 98.4% et 97.3% (98%, 97.3% et 93.8%) dans l'expérience Pages2k (Dense) pour respectivement les indices d'ENSO, du TPI et de l'AMV.

Comme pour les prédictions du LIM dans la section 5.7, en plus de comparer les indices reconstruits aux indices réels, nous comparons également ceux-ci aux indices reconstruits à partir des 30 premières PCs (Fig. 5.21). Le SIR-LIM tend à sous-estimer, comme c'est souvent le cas avec ce type de reconstructions (Liu et al., 2017), la variabilité réelle avec un pourcentage de variance expliquée dans la figure 5.21 de 30% pour l'AMV ($R^2=0.3$), de 70% pour le TPI ($R^2=0.7$) et de 63% pour l'ENSO ($R^2=0.63$). La reconstruction de l'AMV associée à l'ensemble des 30 premières PCs explique 74% de variance de la variabilité réelle tandis que la reconstruction des indices d'ENSO et du TPI des 30 premières PCs explique jusqu'à 97% de variance. Ce pourcentage plus faible pour l'AMV semble néanmoins ne pas expliquer totalement les différences de scores significatives obtenues pour la reconstruction de l'AMV avec les autres modes de variabilités.

La figure 5.22 montre que ces difficultés seraient en réalité surtout liées d'une part à une empreinte de l'AMV sur les températures de surface moins marquée que le TPI et l'ENSO et d'autre part à une distribution spatiale des proxys peu favorable à sa reconstruction. Avec uniquement 1 et 5 pseudo-proxys corrélés à plus de 0.5 dans respectivement l'expérience Pages2k et Dense pour l'AMV en comparaison de l'ENSO et du TPI qui compte tous les deux 10 proxys corrélés à plus de 0.5 dans Pages2k et respectivement 19 et 28 dans Dense, le manque de pseudo-proxys significativement corrélés (c'est à dire >0.5) avec l'indice d'AMV explique les différences de scores obtenues avec les autres indices (Fig. 5.22a,b).

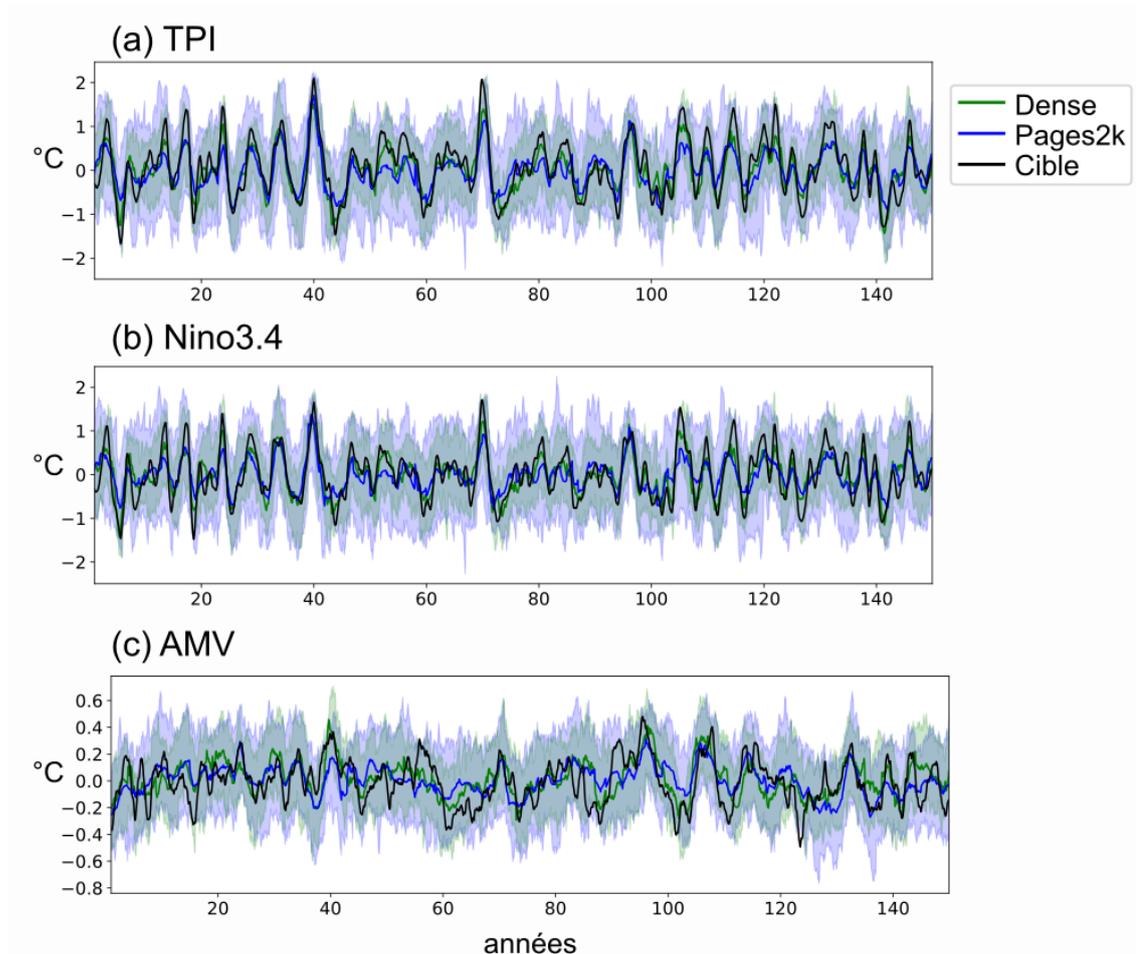


Figure 5.20: Reconstruction par le SIR-LIM des indices (a) Niño3.4, (b) TPI non filtré et (c) AMV pour les expériences Pages2k (en bleu) et Dense (en vert). Les traits pleins noirs indiquent la cible à reconstruire. Les traits pleins bleus et verts indiquent les reconstructions obtenues par la moyenne de l'ensemble des particules et les enveloppes bleues et vertes décrivent la dispersion totale de l'ensemble des particules.

Chaque modèle climatique reproduisant plus ou moins fidèlement l'empreinte spatiale observée des différents modes de variabilité sur les températures de surface, il est important de noter que tout comme la localisation des proxys joue un rôle essentiel sur la qualité des reconstructions envisageables (Evans et al., 1998; Comboul et al., 2015), la bonne reproduction des différents modes de variabilité par le modèle climatique utilisé est essentielle pour déterminer le type de variabilité pouvant être associé aux différents proxys selon leur localisation. En particulier dans le cadre d'expérience en modèle parfait la bonne reproduction ou non par le modèle IPSL-CM5A-LR de l'amplitude et de l'empreinte spatiale associée aux différents modes de variabilité joue un rôle déterminant sur l'estimation de la qualité des reconstructions en réalité envisageables. Ayant déjà discuté dans le chapitre 2 et 3 de certains des biais associés à l'IPSL-CM5A-LR, nous estimons ci-dessous leur impact sur la qualité des reconstructions obtenues et de l'importance plus ou moins vraisemblable accordée aux différents pseudo-proxys lors de la reconstruction selon leur localisation géographique.

Le modèle de l'IPSL-CM5A-LR fait par exemple partie des nombreux modèles simulant

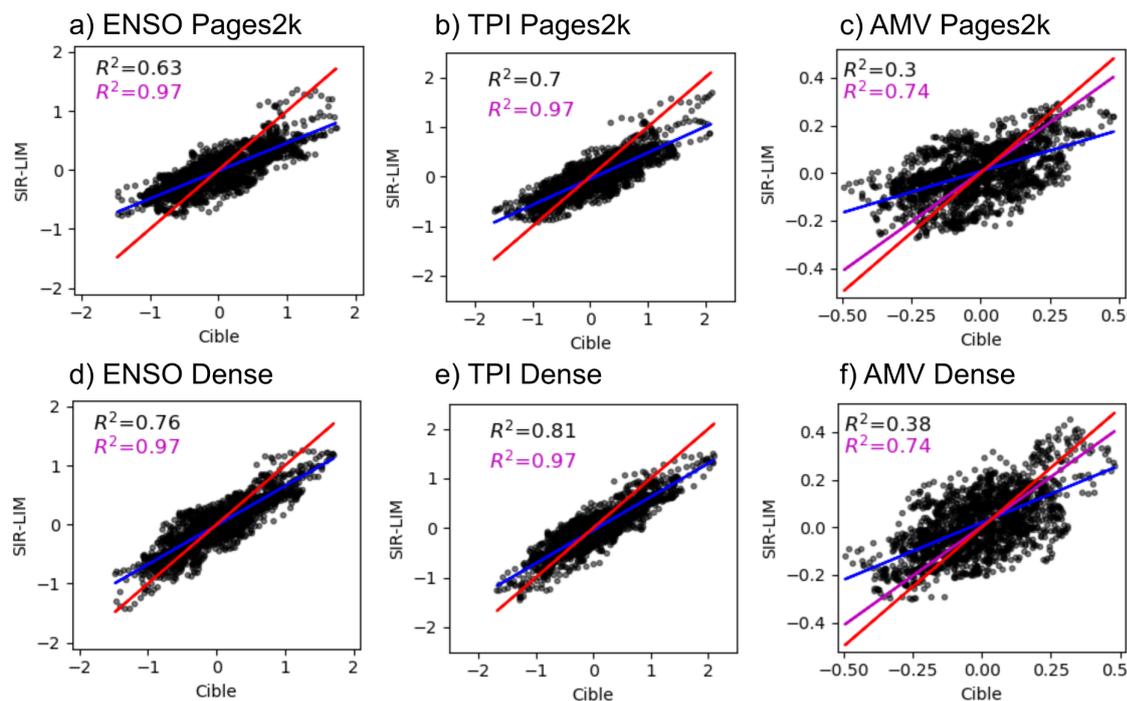


Figure 5.21: (a-b) Diagramme de dispersion des indices reconstruits sur une période de 150 ans de VAL avec le SIR-LIM en ordonnée en fonction des indices réels en abscisse dans les expériences (a-c) Pages2k et (d-f) Dense pour (a,d) l'ENSO, (b,e) le TPI et (c,f) l'AMV. Le coefficient de détermination noté entre les indices reconstruits et réels sont indiqués dans chaque panneau en noir, tandis que le coefficient de détermination entre l'indice réel et l'indice reconstruit à partir des 30 premières PCs sont indiqués en violets. Les droites rouges indiquent le cas d'une reconstruction parfaite, les droites bleues indiquent les régressions entre les reconstructions du SIR-LIM et les indices réels et pour (c,f) les droites violettes indiquent les régressions entre les indices reconstruits à partir des 30 premières PCs et les indices réels.

des anomalies de SST tropicales s'étendant trop loin vers l'ouest lors d'événement El Niño (Li et al., 2016). Ce biais tendrait dans nos expériences à mal estimer la contribution des proxys localisés sur le continent maritime dans la reconstruction de l'ENSO et du TPI. En revanche les téléconnexions atmosphériques observées de l'ENSO dans les régions du Canada, au nord-ouest des Etats-Unis, au sud des Etats-Unis, en Afrique du Sud semblent bien reproduites par l'IPSL-CM5A-LR suggérant une bonne utilisation des pseudo-proxys disponibles dans ces régions. Le TPI décrit pour sa part la variabilité à l'échelle du bassin Pacifique avec des corrélations plus marquées dans les extratropiques que l'ENSO (Fig. 5.22). Sa variabilité décennale connue sous le nom d'IPO tend à être sous-estimée (amplitude et persistance) dans l'IPSL-CM5A-LR par rapport aux observations (comme dans tous les modèles de classe CMIP, cf 3). Les modèles tendent également à simuler une variabilité extratropicale plus faible dans l'Hémisphère Sud que dans l'Hémisphère Nord. Ce biais tendrait à d'une part sous-estimer l'utilité des proxys disponibles le long des côtes du Pérou-Chili pour reconstruire le TPI, et d'autre part à sous-estimer la qualité des reconstructions en réalité possible dans l'Hémisphère Sud notamment dans le Pacifique.

Chaque modèle reproduisant des empreintes spatiales très distincts pour l'AMV (Zhang and Wang, 2013; Ba et al., 2014) on retrouve néanmoins dans l'IPSL-CM5A-LR

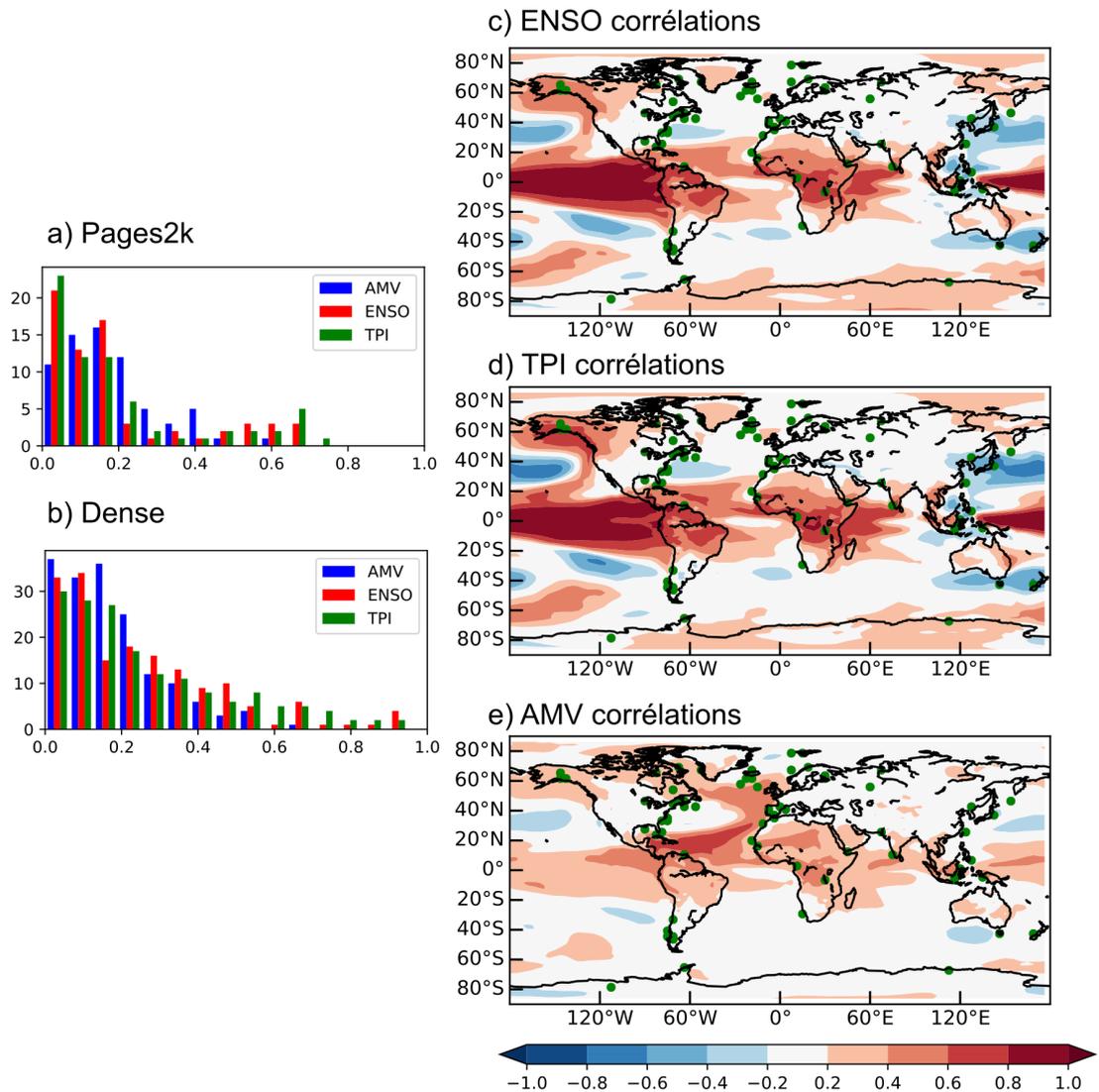


Figure 5.22: (a-b) Histogramme avec en ordonnée le nombre de pseudo-proxys selon leurs corrélations en valeur absolues pour (a) Pages2k et (b) Dense. (c-e) Corrélations des indices (c) d'ENSO, (d) TPI et (e) d'AMV sur les températures de surface globale et où les points verts indiquent la localisation des pseudo-proxys utilisés dans l'expérience Pages2k.

comme dans les observations et de nombreux modèles, une AMV caractérisée par une «comma-shape» (Fig. 5.22) avec une empreinte spatiale plus prononcée de l'AMV dans l'Atlantique Nord tropical, aux environs de 20°N, et une extension le long du côté Est du bassin Atlantique jusqu'à 60°N. Malgré une covariance plus importante aux alentours de 60°N (Fig. A.2), l'AMV montre en terme de corrélations des maximums aux alentours de 20°N (Fig. 5.22). L'impact de l'AMV étant en terme de corrélations moins marquée à l'échelle interannuelle sur les températures de surface que celui du TPI ou de l'ENSO, ceci ne facilite pas sa reconstruction. De plus l'absence de corrélations notables (> 0.2) entre l'indice d'AMV et les températures de surface le long des côtes Atlantiques des Etats-Unis ne permet pas de mettre à contribution les nombreux proxys présents dans ces régions (Fig. 5.22). Les proxys les plus corrélés avec l'AMV dans l'expérience Pages2k se situent dans le cadre de notre expérience plutôt vers l'Islande, le long des côtes Est

d’Afrique du Nord voir en Afrique centrale. De manière plus globale malgré les biais du modèle IPSL-CM5A-LR qui tend à ne pas reproduire certaines relations inter-bassins observées comme le lien IPO-AMV (Marini and Frankignoul, 2014; Swingedouw et al., 2017), à sous-estimer la variabilité associée à l’AMV et à sous-estimer son impact sur les températures de surface notamment sur le continent Nord-Américain (Ba et al., 2014; Ruprich-Robert et al., 2018), il est très probable que des reconstructions plus fines de l’AMV soient en réalité envisageables avec des modèles représentant de manière plus fidèle la variabilité associée à l’AMV et reproduisant de manière plus convaincante les liens inter-bassin observés.

5.10 Sensibilité aux paramètres d’assimilation

Dans les deux sections précédentes, nous avons évalué les performances du SIR-LIM pour reconstruire les températures de surface à partir des 30 premières PCs et d’un ensemble de 900 particules. Nous discutons dans la section suivante de la sensibilité du SIR-LIM à ces paramètres et de leur influence sur la qualité des reconstructions obtenues. Les tests de cette section sont effectués sur une période de 100 ans prise sur VAL. Le calcul du score noDA décrit dans la section 5.4 représente le cas sans assimilation de données. Pour les figures 5.23 et 5.24, les RMSE dites normalisées sont les RMSE divisées par la RMSE de noDA, et de même les ΔCE sont la différence entre les CE et le CE de noDA. Les scores de noDA étant indiqués en haut à droite de chaque panneau, il est possible de retrouver les RMSE d’origines (non normalisées) en multipliant par les RMSE de noDA, et d’obtenir les CE d’origines en ajoutant les CE de noDA.

5.10.1 Nombre de particules et nombre d’EOFs

A l’échelle globale:

L’analyse des scores obtenus dans la figure 5.23 pour la RMSE et la figure A.3 pour les corrélations et CE, indique une cohérence entre eux et amène aux conclusions suivantes: pour un nombre restreint de particules il est nécessaire de réduire de façon drastique la dimension du problème (les meilleurs scores étant obtenus avec peu d’EOFs). La convergence des scores avec l’augmentation du nombre de particules pour les différentes courbes montre la meilleure reconstruction possible pour un nombre d’EOFs fixé. Les courbes pour 20, 30 et 36 EOFs semblent converger vers les mêmes scores que ce soit pour $\sigma = 1.5$ (Fig. 5.23), $\sigma = 1$ ou $\sigma = 2$ (Fig. A.4). Cette convergence commune pour des nombres d’EOFs différents révèle la présence d’un nombre limite d’EOFs. Ce nombre d’EOFs dit « utile » sera défini comme le nombre minimal d’EOFs à partir duquel pour un nombre de particules suffisamment grand l’introduction d’EOFs supplémentaires n’apporte pas d’amélioration à la reconstruction. Dans le cadre de l’expérience Pages2k, les améliorations devenant globalement moins visibles au-delà de 20 PCs et 700 particules, les paramètres utilisés pour la reconstruction de la section 5.8 et 5.9 sont donc parmi les meilleurs possibles.

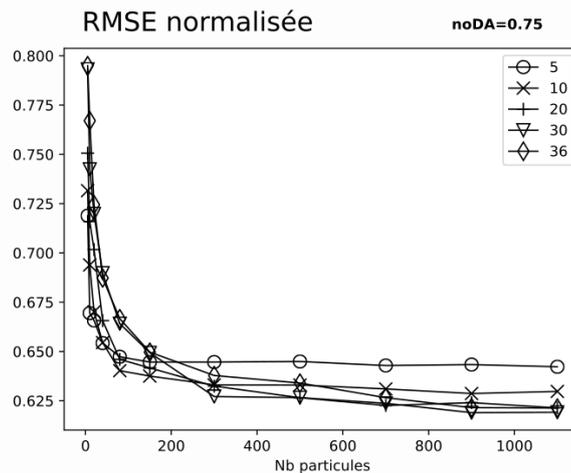


Figure 5.23: Expériences de sensibilité du SIR-LIM au nombre d'EOFs (de 5 à 36 EOFs) et nombre de particules utilisées pour $\sigma=1.5$. Les expériences sont réalisées avec le sous-échantillonnage spatial Pages2k pour des reconstructions d'une période de 100 ans prise sur VAL. Les scores correspondent aux moyennes globales de RMSE divisées par la RMSE de noDA. La RMSE de noDA est indiqué en haut à droite et permet de retrouver les valeurs de RMSE d'origines en multipliant par la RMSE de noDA.

Par hémisphères:

L'analyse des scores moyens de RMSE par hémisphères révèle des différences notables entre les expériences Dense et Pages2k selon l'hémisphère (Fig. 5.24). Comme on peut le voir de part sa variance plus importante, la RMSE noDA de l'Hémisphère Nord (constitué à environ 61% d'océans) est plus élevée par rapport à celle de l'Hémisphère Sud (constitué à environ 81% d'océans). Lorsque l'on regarde l'influence du nombre d'EOFs sur la qualité des reconstructions dans les différents hémisphères (Fig. 5.24), on observe que le nombre d'EOFs « utiles » varie selon le sous-échantillonnage spatial utilisé d'un hémisphère à un autre. Il est en effet logique que plus le nombre de proxys disponibles dans une région est grand plus la description de sa variabilité est fine et que par conséquent son nombre d'EOFs « utiles » puisse être potentiellement élevé. Ceci est par exemple le cas dans l'Hémisphère Sud où le nombre d'EOFs « utiles » passe de 20 pour Dense à 5 pour Pages2k. Dans d'autres cas comme par exemple dans les tropiques, la majorité de la variabilité étant expliquée par les toutes premières EOFs (Tableau 5.2 et Fig. 5.4), l'ajout d'EOFs n'apporte pas d'amélioration notable dans les 2 expériences.

Cette différence marquée entre le nombre d'EOFs utiles dans Pages2k et Dense de l'Hémisphère Sud est révélatrice des difficultés rencontrées pour reconstruire la variabilité naturelle de l'Hémisphère Sud à partir d'un nombre de proxys si restreint. Seule la variabilité décrite par les 5 premières EOFs étant dans notre cas utile pour la reconstruction de la variabilité des températures de surface dans l'Hémisphère Sud, ces résultats bien que réalisés avec une simulation ne prenant pas en compte le rôle de la variabilité forcée, poussent à réfléchir quant aux bénéfices apportés par l'utilisation potentielle de modèles

complexes et coûteux pour la reconstruction de la variabilité des températures de surface dans l’Hémisphère Sud. En d’autres termes, un filtre particulaire SIR réalisé avec un modèle simplifié capable de reproduire la dynamique des 5 premières EOFs des températures de surface globale donnerait dans le cadre de nos expériences d’aussi bons résultats dans la reconstruction des températures de surface de l’Hémisphère Sud (en dessous de 25°S) qu’un filtre particulaire SIR réalisé avec le modèle IPSL-CM5A-LR comme modèle d’intégration.

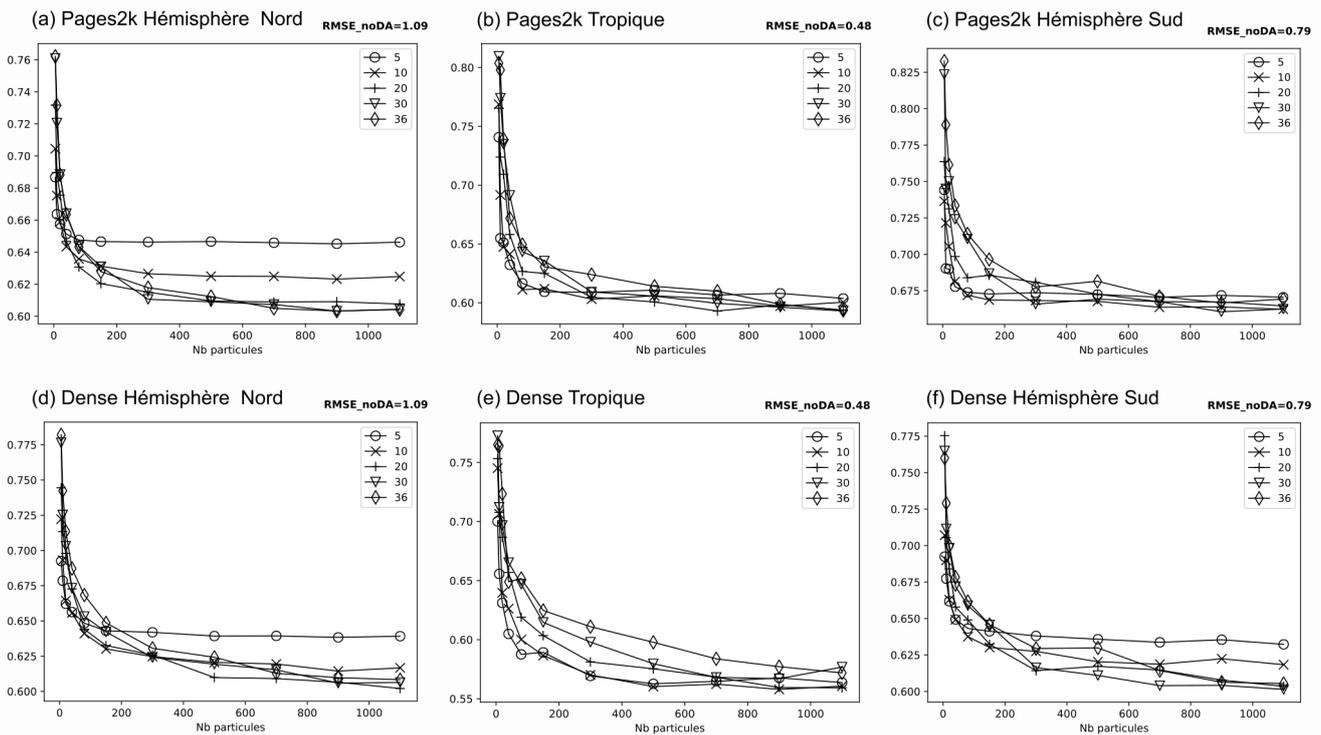


Figure 5.24: Même chose que (Fig 5.23) pour (a-c) Pages2k et (d-f) Dense mais moyennée sur (a,d) l’Hémisphère Nord, (b,e) les tropiques et (c,f) l’Hémisphère Sud. La RMSE noDA utilisée pour la normalisation est indiquée en haut à droite de chaque panneau.

5.10.2 Estimation de la matrice d’erreur C

Comme nous avons pu le voir précédemment dans les figures (Fig. 5.17) et (Fig. 5.20), les méthodes d’ensemble fournissent naturellement une quantification de l’incertitude des reconstructions via l’ensemble des membres/particules. Or le coefficient σ utilisé dans l’estimation de la matrice C (eq. 5.3, 5.4) influence la distribution de cet ensemble en intervenant sur l’écart type de la fonction de vraisemblance des observations (eq. 5.2). Tandis qu’un σ trop faible donnerait un ensemble de poids w_i homogène et un ré-échantillonnage des particules peu sélectif, un σ trop élevé donnerait pour sa part des poids w_i hétérogène et un ré-échantillonnage des particules trop sélectif.

Bien que les scores obtenus dans les Figure 5.23 et A.4 pour différents σ ne semblent pas varier, celui-ci influence la distribution d’écrite par l’ensemble des particules et donc

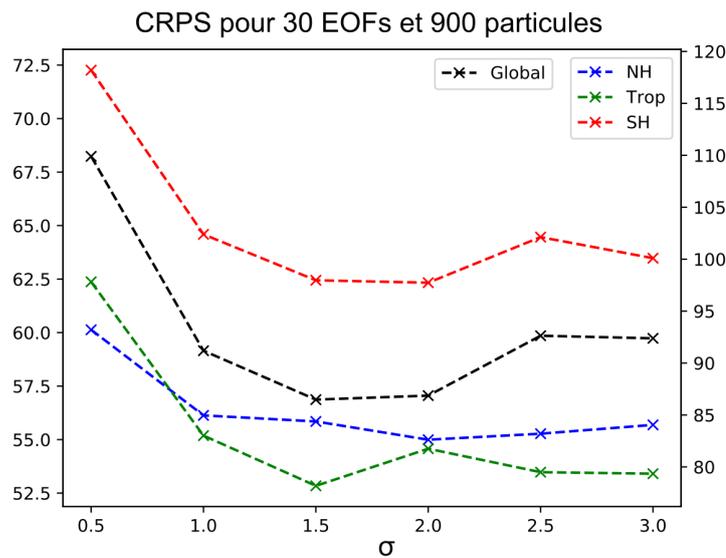


Figure 5.25: Sensibilité au σ utilisé pour l'expérience Pages2k sur une période de 100 ans prise sur VAL et estimée par le CRPS pour les températures moyennes globales (en noir), températures de l'Hémisphère Nord (en bleu), et de l'Hémisphère Sud (en rouge) et des tropiques (en vert). L'axe des ordonnées de gauche correspond au CRPS des températures moyennes globales et celui de droite aux autres courbes.

l'estimation des incertitudes associées à ces reconstructions. La bonne quantification de ces incertitudes faisant partie intégrante de l'objectif de ce type de méthode, nous estimons dans la figure 5.25 la qualité de la distribution des particules en fonction du σ utilisé. Comme on peut voir un σ de 0.5 donne les moins bons résultats, le sigma optimal étant dans notre cas $\sigma = 1.5$ pour la reconstruction des températures moyennes globales et tropicales et $\sigma = 2$ pour l'Hémisphère Sud et Nord. De ce fait la valeur $\sigma = 1.5$ utilisée dans les sections précédentes semble adaptée. S'agissant d'un domaine d'étude à part entière (Ueno et al., 2010), nous nous contenterons ici de dire que le fait que le σ optimal ne soit pas le même selon les hémisphères laisse envisager que la modélisation de la matrice d'erreur de représentativité C_r telle que définie par eq 5.4 n'est pas optimale et qu'il est possible que l'utilisation d'un σ variable sur le globe améliore les résultats.

Chapitre 6

Conclusions et Perspectives

Sommaire

6.1 Synthèse des résultats et Discussion	143
6.2 Perspectives	147
6.2.1 Améliorations du SIR-LIM	147
6.2.2 Reconstructions paléoclimatiques avec des GCMs	149

6.1 Synthèse des résultats et Discussion

Dans le contexte actuel de changement climatique, pour pouvoir assurer le suivi de l'évolution de notre climat en transition, et plus généralement notre environnement, et pouvoir interpréter les changements observés et en attribuer la ou les causes probables, il est tout d'abord nécessaire de pouvoir s'appuyer sur des réseaux de données d'observations et des observatoires pérennes et pluridisciplinaires. La combinaison d'observations instrumentales aux variables climatiques déduits d'archives naturelles carbonatés résolvant l'échelle annuelle à pluri-annuelle, telles que les coraux en domaine marin ou les spéléothèmes en domaine continental, permet d'étendre la fenêtre d'observation au delà de la période instrumentale et une meilleure caractérisation de la variabilité naturelle. La phase de compréhension du fonctionnement du système climatique et environnemental Terre peut passer, d'une part par l'utilisation d'outils statistiques appliqués à l'analyse de ces données d'observations (instrumentales et proxy), d'autre part par des approches diagnostiques mettant en jeu certains concepts ou théories, et plus largement par des outils de modélisation représentant la complexité des processus et mécanismes physiques en jeu dans le système Terre, tout en appréhendant les différentes échelles spatiales et temporelles. Ces modèles de climat sont très élaborés, cependant ils comportent encore de forts biais et incertitudes, ce qui nécessite des travaux de validation à l'aide d'observations adaptées. Ceci est indispensable pour évaluer ensuite le niveau de confiance et d'incertitude des projections climatiques du futur fournies par ces modèles.

Dans le chapitre 3, j'ai essayé de développer ce types d'approches afin d'aborder la question suivante : Est-il possible de distinguer à l'échelle régionale l'action des forçages naturels de celle des forçages anthropiques, liées par exemple à l'évolution des concentrations atmosphériques des gaz à effet de serre ? Il est possible de répondre en partie à

cette question en utilisant des observations directes du climat récent (depuis 1850 environ). L'utilisation de modèles climatiques permet aussi d'apporter des éléments de réponse. Aujourd'hui, grâce à l'augmentation récente des capacités de calcul numérique il est possible d'effectuer des grands ensembles de simulations avec des modèles globaux sur plusieurs centaines d'années. Ces modèles peuvent être validés notamment en confrontant les résultats des simulations climatiques avec des observations ou des reconstructions. Ils fournissent notamment des estimations de la réponse du système climatique à l'évolution des différents forçages connus, et permettent éventuellement d'identifier la signature spatiale à l'échelle régionale de la réponse climatique à certains de ces forçages.

L'expansion de la cellule de Hadley (HC) observée et simulée par le modèle IPSL-CM5A-LR et les modèles CMIP5 en réponse au forçage anthropique constitue par exemple un résultat robuste tandis que de nombreuses incertitudes persistent sur les tendances observées de la circulation de Walker dans le Pacifique équatorial. Une émergence aussi claire de l'empreinte anthropique sur la tendance de la HC dans l'Hémisphère Sud pourrait être due à l'amplitude plus faible de la variabilité interne par rapport à l'influence combinée des GES et de l'appauvrissement de la couche d'ozone. A l'échelle régionale celle-ci se traduit par un renforcement des anticyclones subtropicaux de l'Hémisphère Sud dont l'expression océanique sur les tendances des SSTs observées peuvent néanmoins être masquée selon le contexte régional par l'influence de la variabilité climatique interne.

D'après nos résultats et dans le cadre d'expériences spécifiques avec le modèle IPSL-CM5A-LR, les tendances observées depuis 1980 le long des côtes péruviennes chiliennes, émergent du bruit de la variabilité naturelle en réponse à l'expansion des cellules de Hadley, et se traduisent par une intensification de l'anticyclone subtropical dans le Pacifique Sud et des alizés qui s'étendent depuis la côte Sud Américaine dans le Pacifique sud vers le Pacifique équatorial. Cette expansion des alizés favorise ainsi l'approfondissement de la couche de mélange océanique au large dans le Pacifique Sud Est, une remontée de la thermocline et la résurgence d'eau froide près des côtes Chiliennes. L'analyse du bilan de chaleur de la couche de mélange océanique simulée à l'échelle de tout le Pacifique, dans le Pacifique Sud Est et proche de la côte, montre que le pompage d'Ekman s'oppose au signal régional du réchauffement climatique et l'annule près de la côte. En revanche, plus au large, loin de la zone d'upwelling, le refroidissement relatif régional dans le Pacifique Sud Est résulte d'un approfondissement de la couche mélange par le renforcement des alizés en réponse au forçage anthropiques. Ces résultats indiquent le rôle dominant des forçages externes dans les tendances au refroidissement observées dans le sud-est du Pacifique quelque soit la phase de l'IPO.

A travers cette étude nous avons pu réaliser les limites associées d'une part à l'utilisation d'observations instrumentales sur une période relativement courte et soumise à l'effet des forçages anthropiques et d'autre part aux nombreux biais non négligeables présents dans les modèles climatiques. Si des efforts existent afin de traiter les biais persistant des modèles (comme celui de la double ITCZ ou de la « cold tong »), il est aujourd'hui possible d'augmenter la fenêtre d'échantillonnage de la variabilité naturelle de façon significative. Les nombreuses synthèses récentes de proxys comme celles issues de PAGES2K (PAGES2k

(Consortium, 2017) révèle la gamme très étendue de la variabilité naturelle et démontre qu'elle est très largement sous observée sur la période instrumentale (Fig. 6.1). Développer ce type de reconstructions à l'échelle régionale permettrait de mettre en perspective les changements récents et de comprendre les réponses dynamiques associées par exemple aux conditions relativement douces qu'a connu le climat de l'Hémisphère Nord durant « optimum médiéval » entre 950 et 1250, et de la période plus froide du « petit âge de glace » entre 1450 et 1850 (Mann et al., 2009).

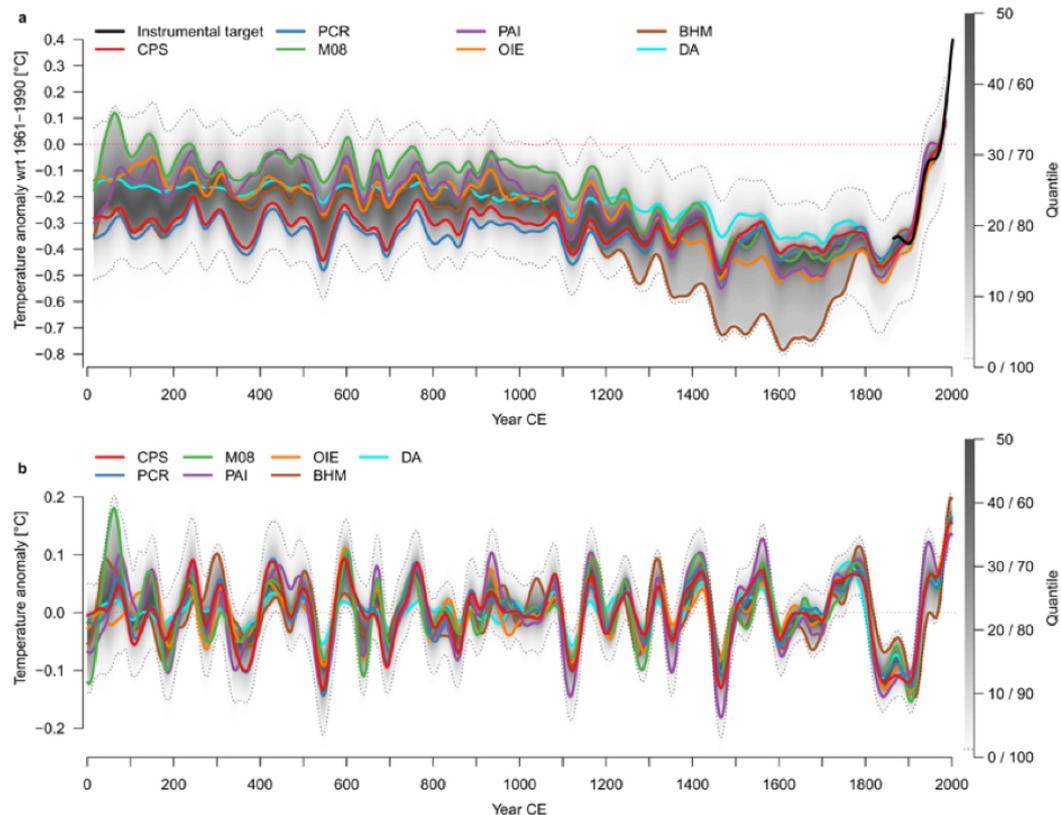


Figure 6.1: Reconstruction des températures de surface globales par les méthodes CPS (pour en anglais Composite Plus Scaling), PCR (pour en anglais Principal Component Regression), OIE (pour en anglais Optimal Information Extraction), BHM (pour en anglais Bayesian Hierarchical Model) ainsi que M08 (pour Regularized errors in variables), PAI (pour Pairwise comparison) et DA (pour Offline data assimilation). Les courbes en couleurs représentent la reconstruction médiane obtenue par chacune des méthodes après application (a) d'un filtre passe bas 30 ans et (b) d'un filtre passe bande d'une fenêtre de 30-200 ans. La période instrumentale est indiquée dans (a) par la courbe noire. Le dégradé de gris quantifie les quantiles des reconstructions obtenues à travers les différentes méthodes avec les 2.5^{ème} et 97.5^{ème} centiles indiqués en traits noirs pointillés (Source : PAGES2k Consortium, 2019).

Dans cet objectif nous avons mis au point et testé dans les chapitre 4 et 5 une méthode d'assimilation de données « online » appelée SIR-LIM (avec SIR pour Sequential Importance Resampling et LIM pour Linear Inverse Model) pour la reconstruction de la variabilité naturelle interannuelle à décennale des températures de surface en contexte paléoclimatique. Le SIR-LIM est un filtre particuliers avec ré-échantillonnage SIR utilisant un LIM comme modèle d'intégration. Cette approche est testée en contexte « modèle parfait » à l'aide d'une simulation de contrôle du modèle IPSL-CM5A-LR considérée

comme représentative de la réalité. Cette étude a été motivée d'une part par les études antérieures prouvant la capacité du LIM à reproduire la dynamique de certaines variables atmosphériques et océaniques (Penland and Sardeshmukh, 1995; Penland and Matrosova, 1998; Johnson et al., 2000; Winkler et al., 2001) et d'autre part par les résultats prometteurs obtenus par l'utilisation des filtres particulières dans le cadre de reconstructions paléoclimatiques (Dubinkina et al., 2011).

Afin de pallier aux contraintes liées à l'utilisation de modèles climatiques coûteux (Dubinkina et al., 2011; Liu et al., 2017), nous testons l'utilisation du LIM comme modèle de substitution dans un filtre particulières SIR noté SIR-LIM. Une première expérience notée Pages2k simule une répartition des pseudo-proxys similaire à la localisation des proxys de température à l'an 1000 dans la base de données PAGES2k, et une seconde expérience Dense utilise quant à elle une répartition de pseudo-proxys régulière sur l'ensemble du globe. Les résultats obtenus lors de la validation du LIM en tant que modèle de substitution (section 5.6 et 5.7) montrent que celui-ci reproduit de façon vraisemblable la dynamique des principaux modes de variabilité des températures de surface et permet une propagation cohérente des particules dans le temps. On a en effet pu observer que les prédictions du LIM surpassent la persistance aux horizons 12 et 18 mois (Fig. 5.8) et que la composante stochastique du LIM permet une bonne représentation de l'incertitude associée à ces prédictions (Fig. 5.14, 5.13, 5.15). L'utilisation du LIM comme modèle d'intégration dans le SIR-LIM a permis l'intégration d'un grand ensemble de 900 particules décrivant la dynamique des 30 premières EOFs des températures de surface.

En comparaison de certaines expériences réalisées par Liu et al., 2017 avec un GCM couplé et un ensemble de 96 particules, le SIR-LIM montre des résultats très prometteurs. Les reconstructions obtenues par le SIR-LIM étant meilleures pour Dense que Pages2k, on observe pour Pages2k une incertitude plus importante et des reconstructions globalement moins bonnes que pour Dense avec en particulier dans l'Hémisphère Sud +0.25 de corrélation pour la moyenne des températures dans l'Hémisphère Sud de Dense par rapport à Pages2k (Fig. 5.18, Tableau A.5). L'étude de sensibilité au nombre d'EOFs et au nombre de particules utilisées dans le SIR-LIM (Fig. 5.23) a montré un comportement tout à fait attendu, avec de meilleures reconstructions pour les expériences de faible dimension (avec peu d'EOFs) dans le cas d'un nombre réduit de particules et une dégénérescence du SIR-LIM de plus en plus prononcée pour les problèmes de grandes dimensions (avec beaucoup d'EOFs) et un nombre réduit de particules. L'estimation d'un nombre d'EOFs dit « utile » a permis une bonne estimation de la qualité des reconstructions potentiellement possibles par hémisphère en fonction du sous-échantillonnage spatial des proxys (Fig. 5.24). Nous retiendrons que pour le sous-échantillonnage spatial de pseudo-proxys réalistes (Pages2k) et les incertitudes associées à ces pseudo-proxys, seule la variabilité décrite par les 20 premières EOFs contribue à la reconstruction de la variabilité des températures de surface à l'échelle globale et que cette valeur descend à cause du faible nombre de proxys disponibles, à 5 EOFs dans l'Hémisphère Sud (Fig. 5.24). Nous avons également pu constater qu'à partir d'un certain nombre de particules (environ 700 pour Pages2k) la qualité des reconstructions tend à converger vers une reconstruction dite optimale (Fig.

5.23). Ces conclusions n'étant évidemment vraies que dans le cadre de nos expériences avec forçage constant, nous soulignerons l'importance du rôle de la variabilité forcée dans le cadre de reconstructions réelles. Or comme la prise en compte de cette composante forcée ne peut à ce jour se faire qu'à l'aide de modèles climatiques capables de reproduire la réponse dynamique associée aux variations du forçage externe au cours du dernier millénaire, il est évident que les réponses dynamiques de ces modèles aux forçages externes se doit d'être prises en compte et intégrées lors de reconstructions paléoclimatiques.

Nous soulignerons d'autre part que chaque modèle reproduisant plus ou moins bien certains éléments de la dynamique du climat réel, l'ensemble de ces résultats peut varier en fonction du modèle utilisé comme référence (section 5.9). Il est en effet important de rappeler que bien que très utile pour évaluer la robustesse des méthodes d'assimilation de données ces approches en modèle parfait supposent une bonne représentation de la dynamique du climat par le modèle supposé comme représentatif de la réalité. Et de ce fait cette hypothèse très discutée, notamment sur la bonne reproduction par les modèles climatiques du lien tropique-extratropique dans le Pacifique Sud (Henley et al., 2017 et Chapitre 3 section 3.2), peut potentiellement induire une surestimation ou sous-estimation des reconstructions possibles notamment de la variabilité extratropicale dans l'Hémisphère Sud et du rôle des proxys tropicaux dans ces reconstructions.

6.2 Perspectives

Les résultats obtenus dans le cadre de ces expériences en modèle parfait offrent deux axes de progressions futures avec d'une part des améliorations possibles du SIR-LIM en tant que méthode de reconstructions paléoclimatiques peu coûteux et d'autre part le perfectionnement des méthodes de filtre particulières avec l'utilisation de GCMs de nouvelles générations.

6.2.1 Améliorations du SIR-LIM

Intégration du rôle des forçages:

Le principal frein à l'utilisation du SIR-LIM dans le cadre de reconstructions paléoclimatiques réelles, est l'utilisation des matrices \mathbf{B} et \mathbf{Q} constantes et l'hypothèse de stationnarité faite lors de la calibration du LIM (Penland, 1996; Newman et al., 2011). Bien que dans le cadre de forçages constants représentatifs de l'époque préindustrielle la dynamique du système étudié puisse être considérée comme stationnaire, cette hypothèse n'est pas vérifiée dans la réalité. Il est en effet aujourd'hui évident que les forçages externes en plus d'avoir un impact sur l'état moyen du climat modifient également sa dynamique (Power et al., 2013; Khodri et al., 2017; cf. Chapitre 3). Il est donc peu probable voir certain qu'un LIM et des EOFs construits sur une période donnée dans un contexte climatique particulier ne produisent pas dans un contexte différent la bonne dynamique. Bien qu'il soit possible d'envisager l'ajout d'un terme de forçage dans le LIM (Sardeshmukh and Sura, 2009), la meilleure approche consisterait à construire différents ensembles d'EOFs et

LIMs selon les différents forçages et ou contextes climatiques. En calibrant par exemple le LIM par période ou contexte climatique sur un ensemble de simulations offline soumises aux forçages du dernier millénaire, les matrices \mathbf{B} et \mathbf{Q} ne seraient alors plus constantes et seraient modifiées à chaque pas de temps afin de prendre en compte l'influence du forçage sur la dynamique du système. Ce type d'approches déjà réalisé dans le cadre de variations saisonnières des matrices \mathbf{B} et \mathbf{Q} (Johnson et al., 2000; Newman et al., 2011) pourrait être reproduit dans le cadre de reconstructions paléoclimatiques. Néanmoins le nombre de simulations nécessaires pour avoir une bonne estimation des matrices \mathbf{B} et \mathbf{Q} aux différents pas de temps reste toutefois à déterminer, celui-ci pouvant devenir rapidement important en fonction du nombre d'EOFs considérées.

Intégration des dynamiques non-linéaires:

Malgré le succès du LIM pour reproduire la dynamique de nombreux champs, celui-ci est en tant que modèle stochastique additif construit par l'addition d'un terme linéaire représentant la composante prédictible du système et d'un bruit blanc Gaussien représentant sa composante non linéaire. Les filtres particuliers pouvant être utilisés dans le cadre de systèmes non linéaires pour des distributions de densité non nécessairement Gaussiennes, le CAM-LIM proposé par (Martinez-Villalobos et al., 2018) constitue une solution intéressante pour étendre et/ou améliorer les capacités du LIM dans le cas de variables géophysiques possédant des distributions non Gaussiennes et/ou asymétrique. Peu d'études ayant cependant testées les résultats obtenus avec le CAM-LIM, son utilité dans le cadre de reconstructions paléoclimatiques reste à prouver.

Reconstruction de plusieurs variables:

Bien que la majorité des LIMs soient généralement construits à partir d'une unique variable (Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995; Penland, 1996; Penland and Matrosova, 1998; Johnson et al., 2000; Alexander et al., 2008), il est possible de construire des LIM dit couplé à plusieurs variables (Newman et al., 2009; Winkler et al., 2001). En plus de permettre la reconstruction d'autres variables par le SIR-LIM, l'introduction de variables supplémentaires pourrait potentiellement améliorer les prédictions du LIM. Il serait par exemple intéressant de considérer les couplages à l'origine des principaux modes de variabilité des températures de surface tropicale en mettant à profit dans le SIR-LIM la théorie de recharge-décharge de (Jin, 1997a; Jin, 1997b) qui suggère que le contenu en chaleur de l'océan évoluera de façon différente selon la SST tout comme la SST évoluera de façon différente selon le contenu en chaleur de l'océan à la date de départ, le rôle de la salinité dans le déclenchement d'événement El Niño avec un LIM couplé salinité-SST (Vialard and Delecluse, 1998), ou bien encore le couplage océan-atmosphère primordial dans les tropiques et pour les zones d'upwelling avec un LIM couplé vents-SST (Ekman, 1905; Bjerknes, 1969; Newman et al., 2009).

LIM Multi modèles:

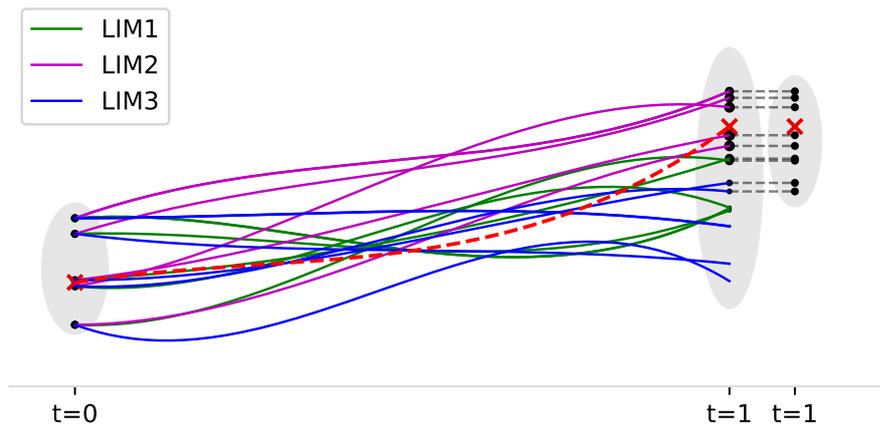


Figure 6.2: Illustration d'un filtre particulaire SIR pour 3 LIM calibrés sur 3 modèles différents. Les croix rouges correspondent à la position des proxys, les courbes bleues, violettes et rouges indiquent la trajectoire du même ensemble de N particules pour les 3 LIM différents. Chaque particule étant intégrée avec les 3 LIM un ensemble de $3 \times N$ particules est générée à $t=1$. L'étape de ré-échantillonnage illustrée à $t=1$ permet de revenir à un ensemble de N particules en dupliquant et/ou conservant les particules de poids forts et abandonnant les particules de poids faibles.

Le LIM étant un modèle statistique celui-ci nécessite une période de calibration (notée précédemment APP) pour déterminer les matrices \mathbf{B} et \mathbf{Q} . Cette période de calibration pouvant aussi bien se faire sur une base de modèles climatiques que d'observations, il serait judicieux de tester les différentes possibilités. Chaque modèle climatique étant plus ou moins bon dans la reproduction de la dynamique de certaines composantes du climat, un SIR-LIM constitué de plusieurs LIM permettrait de combiner la connaissance issue des différents modèles. Il est par exemple possible de construire un LIM basé sur un ensemble de modèle (comme la base de modèles CMIP5) et d'utiliser le LIM comme un outil pour harmoniser et extraire la dynamique commune de l'ensemble de ces modèles qui possèdent chacun leur propre résolution spatiale, temporelle et dynamique. Une autre possibilité serait de construire différents LIM à partir de différents modèles comme illustré dans la figure 6.2. Une reconstruction basée sur un ensemble de 3 LIM calibrés sur 3 modèles couplés différents serait par exemple possible. L'intégration de l'ensemble des N particules par chacun des LIM au pas de temps k donnerait un nouvel ensemble de $3 \times N$ particules au pas de temps $k+1$. Ces $3 \times N$ particules pouvant être ré-échantillonnées en un nouvel ensemble de N particules, la sélection des particules issues des LIMs les plus représentatives de la dynamique du système réel se fera donc naturellement en fonction des observations disponibles. Des reconstructions pouvant être testées à différentes échelles de temps et spatiales, l'utilisation du LIM comme émulateur des modèles dynamiques offre de nombreuses perspectives dans le cadre de reconstructions paléoclimatiques.

6.2.2 Reconstructions paléoclimatiques avec des GCMs

Bien que l'utilisation de GCMs dans des filtres particulaire SIR reste de nos jours compliquée, ces modèles constituent actuellement les meilleurs outils à disposition pour

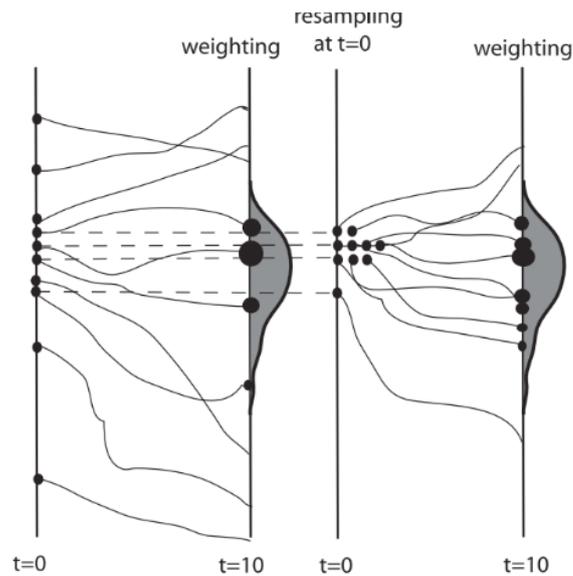


Figure 6.3: Représentation du filtre particulaire auxiliaire pour une itération. Les traits représentent la trajectoire des particules dans l'espace des états, la largeur des cercles noirs le poids des particules et les distributions en gris la fonction de vraisemblance des observations. Dans un premier temps un modèle peu coûteux est utilisé pour prédire à partir de $t=0$ la position des particules à $t=10$. Les particules du GCM sont alors ré-échantillonnées de manière à favoriser les particules les plus susceptibles de correspondre aux observations à $t=10$. Le modèle coûteux est alors utilisé pour intégrer cette nouvelle ensemble de particules (source: [Van Leeuwen, 2009](#)).

reproduire et comprendre les processus dynamiques associés aux variations climatiques des 2000 dernières années (incluant la période historique). La prise en compte dans ces GCMs de l'ensemble des processus physiques globaux et régionaux constitue l'unique façon de comprendre le rôle des forçages externes et des modes décennaux sur la variabilité du climat passé. La poursuite de ce travail consiste donc dans l'immédiat à mettre en place des méthodes de reconstructions « online », capable d'utiliser la physique de ces modèles. La mise en place en online de filtres particulaires de plus en plus sophistiqués ([Van Leeuwen et al., 2019](#)) et/ou filtre d'ensemble de Kalman sur de longues périodes est de nos jours possible grâce à l'augmentation des capacités de calculs et le développement de modèle massivement parallélisé, comme par exemple la version CM6 du modèle de l'IPSL en cours de finalisation permettant de simuler jusqu'à cent ans par jour.

L'association d'un émulateur tel que le LIM en tant qu'auxiliaire à un GCM dans un filtre particulaires comme introduit par ([Pitt and Shephard, 1999](#)) peut permettre de réduire de façon judicieuse le nombre de particules nécessaires et donc améliorer les résultats à ce jour obtenus avec des GCM dans le cadre de filtres particulaires ([Liu et al., 2017](#)). Le principe du filtre particulaires auxiliaire consiste à pré sélectionner les simulations du GCM à propager, en prenant en compte les observations au pas de temps $k+1$ lors du ré-échantillonnage des particules au temps k par une première intégration des particules de l'auxiliaire peu coûteux afin d'estimer la probabilité qu'une particule corresponde aux observations au pas de temps $k+1$ (Fig. 6.3). Cette première étape d'intégration permet d'éviter l'intégration coûteuse de particules peu susceptibles de correspondre aux observations par la suite. Le LIM pouvant être mis à contribution

dans ce type d'approche, il permettrait de réduire le nombre de particules généralement nécessaires pour les filtres particulaires du type SIR. L'efficacité de ce filtre particulaires reste néanmoins dépendante de la dynamique du système et peut selon le problème n'aboutir qu'à une réduction peu ou pas significative du nombre de particules nécessaires. Parmi les autres variantes de filtres particulaires potentiellement capables de réduire le nombre de particules nécessaires, nous citerons également les filtres particulaires à densité proposée (ou en anglais proposal density particle filter) qui semble ces derniers temps connaître un certain succès (Van Leeuwen et al., 2019; Pinheiro et al., 2019).

Bibliography

- Adler, Robert F, George J Huffman, Alfred Chang, Ralph Ferraro, Ping-Ping Xie, John Janowiak, Bruno Rudolf, Udo Schneider, Scott Curtis, David Bolvin, et al. (2003). "The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present)". In: *Journal of hydrometeorology* 4.6, pp. 1147–1167.
- Alexander, Michael A, Ludmila Matrosova, Cécile Penland, James D Scott, and Ping Chang (2008). "Forecasting Pacific SSTs: Linear inverse model predictions of the PDO". In: *Journal of Climate* 21.2, pp. 385–402.
- Alexander, Michael A, K Halimeda Kilbourne, and Janet A Nye (2014). "Climate variability during warm and cold phases of the Atlantic Multidecadal Oscillation (AMO) 1871–2008". In: *Journal of Marine Systems* 133, pp. 14–26.
- Anderson, Jeffrey L (1996). "A method for producing and evaluating probabilistic forecasts from ensemble model integrations". In: *Journal of climate* 9.7, pp. 1518–1530.
- Anderson, Jeffrey L and Stephen L Anderson (1999). "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts". In: *Monthly Weather Review* 127.12, pp. 2741–2758.
- Andreoli, Rita V and Mary T Kayano (2005). "ENSO-related rainfall anomalies in South America and associated circulation features during warm and cold Pacific decadal oscillation regimes". In: *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25.15, pp. 2017–2030.
- Andrews, Timothy, Jonathan M Gregory, Mark J Webb, and Karl E Taylor (2012). "Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models". In: *Geophysical Research Letters* 39.9.
- Annan, JD and Julia Catherine Hargreaves (2013). "A new global reconstruction of temperature changes at the Last Glacial Maximum". In: *Climate of the Past* 9.1, pp. 367–376.
- Aravena, Guillermo, Bernardo Broitman, and Nils Christian Stenseth (2014). "Twelve years of change in coastal upwelling along the central-northern coast of Chile: spatially heterogeneous responses to climatic variability". In: *PLoS One* 9.2, e90276.
- Arblaster, JMGA, G Meehl, and A Moore (2002). "Interdecadal modulation of Australian rainfall". In: *Climate Dynamics* 18.6, pp. 519–531.
- Arblaster, Julie M and Gerald A Meehl (2006). "Contributions of external forcings to southern annular mode trends". In: *Journal of climate* 19.12, pp. 2896–2905.
- Aumont, Oliver and Laurent Bopp (2006). "Globalizing results from ocean in situ iron fertilization studies". In: *Global Biogeochemical Cycles* 20.2.

- Ba, Jin, Noel S Keenlyside, Mojib Latif, Wonsun Park, Hui Ding, Katja Lohmann, Juliette Mignot, Matthew Menary, Odd Helge Otterå, Bert Wouters, et al. (2014). "A multi-model comparison of Atlantic multidecadal variability". In: *Climate dynamics* 43.9-10, pp. 2333–2348.
- Balmaseda, Magdalena Alonso, Kristian Mogensen, and Anthony T Weaver (2013). "Evaluation of the ECMWF ocean reanalysis system ORAS4". In: *Quarterly Journal of the Royal Meteorological Society* 139.674, pp. 1132–1161.
- Barkmeijer, J, T Iversen, and TN Palmer (2003). "Forcing singular vectors and other sensitive model structures". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 129.592, pp. 2401–2423.
- Battisti, David S and Anthony C Hirst (1989). "Interannual variability in a tropical atmosphere–ocean model: Influence of the basic state, ocean geometry and nonlinearity". In: *Journal of the atmospheric sciences* 46.12, pp. 1687–1712.
- Beer, Jürg, W Mende, and R1 Stellmacher (2000). "The role of the sun in climate forcing". In: *Quaternary Science Reviews* 19.1-5, pp. 403–415.
- Bellenger, Hugo, Éric Guilyardi, Julie Leloup, Matthieu Lengaigne, and Jérôme Vialard (2014). "ENSO representation in climate models: From CMIP3 to CMIP5". In: *Climate Dynamics* 42.7-8, pp. 1999–2018.
- Belmadani, Ali, Vincent Echevin, Francis Codron, Ken Takahashi, and Clémentine Junquas (2014). "What dynamics drive future wind scenarios for coastal upwelling off Peru and Chile?" In: *Climate dynamics* 43.7-8, pp. 1893–1914.
- Bengtsson, Lennart and Jagadish Shukla (1988). "Integration of space and in situ observations to study global climate change". In: *Bulletin of the American Meteorological Society* 69.10, pp. 1130–1143.
- Bhend, Jonas, Jörg Franke, Doris Folini, Martin Wild, and Stefan Brönnimann (2012). "An ensemble-based approach to climate reconstructions". In: *Climate of the Past* 8.3, pp. 963–976.
- Biastoch, Arne, Claus W Böning, Franziska U Schwarzkopf, and JRE Lutjeharms (2009). "Increase in Agulhas leakage due to poleward shift of Southern Hemisphere westerlies". In: *Nature* 462.7272, p. 495.
- Bindoff, Nathaniel L, Jürgen Willebrand, Vincenzo Artale, Anny Cazenave, Jonathan M Gregory, Sergey Gulev, Kimio Hanawa, Corinne Le Quere, Sydney Levitus, Yukihiro Nojiri, et al. (2007). "Observations: oceanic climate change and sea level". In:
- Bindoff, Nathaniel L, Peter A Stott, Krishna Mirle AchutaRao, Myles R Allen, Nathan Gillett, David Gutzler, Kabumbwe Hansingo, G Hegerl, Yongyun Hu, Suman Jain, et al. (2013). "Detection and attribution of climate change: from global to regional". In:
- Bjerknes, Jakob (1969). "Atmospheric teleconnections from the equatorial Pacific". In: *Monthly weather review* 97.3, pp. 163–172.
- Boisséson, Eric de, Magdalena Alonso Balmaseda, and Michael Mayer (2018). "Ocean heat content variability in an ensemble of twentieth century ocean reanalyses". In: *Climate Dynamics* 50.9-10, pp. 3783–3798.

- Booth, Ben BB, Nick J Dunstone, Paul R Halloran, Timothy Andrews, and Nicolas Bellouin (2012). "Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability". In: *Nature* 484.7393, pp. 228–232.
- Borchert, Leonard F, André Düsterhus, Sebastian Brune, Wolfgang A Müller, and Johanna Baehr (2019). "Forecast-Oriented Assessment of Decadal Hindcast Skill for North Atlantic SST". In: *Geophysical Research Letters* 46.20, pp. 11444–11454.
- Boucher, O, D Randall, P Artaxo, C Bretherton, G Feingold, P Forster, VM Kerminen, Y Kondo, H Liao, U Lohmann, et al. (2013). "Clouds and aerosols. Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change". In: K., Tignor, M., Allen, SK, Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, PM, *Cambridge University Press, Cambridge, UK and New York, NY, USA, <https://www.ipcc.ch/report/ar5/>* (last access: 5 October 2017).
- Brient, Florent and Sandrine Bony (2013). "Interpretation of the positive low-cloud feedback predicted by a climate model under global warming". In: *Climate Dynamics* 40.9-10, pp. 2415–2431.
- Brune, Sebastian, André Düsterhus, Holger Pohlmann, Wolfgang A Müller, and Johanna Baehr (2018). "Time dependency of the prediction skill for the North Atlantic subpolar gyre in initialized decadal hindcasts". In: *Climate Dynamics* 51.5-6, pp. 1947–1970.
- Buckley, Brendan M, Kevin J Anchukaitis, Daniel Penny, Roland Fletcher, Edward R Cook, Masaki Sano, Aroonrut Wichienkeo, Ton That Minh, Truong Mai Hong, et al. (2010). "Climate as a contributing factor in the demise of Angkor, Cambodia". In: *Proceedings of the National Academy of Sciences* 107.15, pp. 6748–6752.
- Büntgen, Ulf, Willy Tegel, Kurt Nicolussi, Michael McCormick, David Frank, Valerie Trouet, Jed O Kaplan, Franz Herzig, Karl-Uwe Heussner, Heinz Wanner, et al. (2011). "2500 years of European climate variability and human susceptibility". In: *Science* 331.6017, pp. 578–582.
- Cai, Wenju (2006). "Antarctic ozone depletion causes an intensification of the Southern Ocean super-gyre circulation". In: *Geophysical Research Letters* 33.3.
- Cai, Wenju, Agus Santoso, Guojian Wang, Sang-Wook Yeh, Soon-Il An, Kim M Cobb, Mat Collins, Eric Guilyardi, Fei-Fei Jin, Jong-Seong Kug, et al. (2015). "ENSO and greenhouse warming". In: *Nature Climate Change* 5.9, pp. 849–859.
- Carton, James A and Benjamin S Giese (2008). "A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA)". In: *Monthly Weather Review* 136.8, pp. 2999–3017.
- Cassou, Christophe, Yochanan Kushnir, Ed Hawkins, Anna Pirani, Fred Kucharski, In-Sik Kang, and Nico Caltabiano (2018). "Decadal climate variability and predictability: challenges and opportunities". In: *Bulletin of the American Meteorological Society* 99.3, pp. 479–490.
- Chavez, Francisco P, Arnaud Bertrand, Renato Guevara Carrasco, Pierre Soler, and Jorge Csirke (2008). "The northern Humboldt Current System: Brief history, present status and a view towards the future". In: *Progress in Oceanography* 79.special issue 2-4, pp. 95–105.

- Cherchi, Annalisa, Simona Masina, and Antonio Navarra (2008). "Impact of extreme CO₂ levels on tropical climate: a CGCM study". In: *Climate dynamics* 31.7-8, pp. 743–758.
- Christensen, Jens Hesselbjerg, Krishna Kumar Kanikicharla, Edvin Aldrian, Soon Il An, Iracema Fonseca Albuquerque Cavalcanti, Manuel De Castro, Wenjie Dong, Prashant Goswami, Alex Hall, Joseph Katongo Kanyanga, et al. (2013). "Climate phenomena and their relevance for future regional climate change". In: *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, pp. 1217–1308.
- Chung, Eui-Seok, Axel Timmermann, Brian J Soden, Kyung-Ja Ha, Lei Shi, and Viju O John (2019). "Reconciling opposing Walker circulation trends in observations and model projections". In: *Nature Climate Change* 9.5, pp. 405–412.
- Church, John A and Neil J White (2011). "Sea-level rise from the late 19th to the early 21st century". In: *Surveys in geophysics* 32.4-5, pp. 585–602.
- Clement, Amy, Katinka Bellomo, Lisa N Murphy, Mark A Cane, Thorsten Mauritsen, Gaby Rädel, and Bjorn Stevens (2015). "The Atlantic Multidecadal Oscillation without a role for ocean circulation". In: *Science* 350.6258, pp. 320–324.
- Comboul, Maud, Julien Emile-Geay, Gregory J Hakim, and Michael N Evans (2015). "Paleoclimate sampling as a sensor placement problem". In: *Journal of Climate* 28.19, pp. 7717–7740.
- Compo, Gilbert P, Jeffrey S Whitaker, Prashant D Sardeshmukh, Nobuki Matsui, Robert J Allan, Xungang Yin, Byron E Gleason, Russell S Vose, Glenn Rutledge, Pierre Bessemoulin, et al. (2011). "The twentieth century reanalysis project". In: *Quarterly Journal of the Royal Meteorological Society* 137.654, pp. 1–28.
- Cook, Edward R, David M Meko, David W Stahle, and Malcolm K Cleaveland (1999). "Drought reconstructions for the continental United States". In: *Journal of Climate* 12.4, pp. 1145–1162.
- Crespin, Elisabeth, Hugues Goosse, Thierry Fichefet, and Michael E Mann (2009). "The 15th century Arctic warming in coupled model simulations with data assimilation." In: *Climate of the Past* 5.3.
- Crutzen, Paul J (2016). "Geology of mankind". In: *Paul J. Crutzen: A Pioneer on Atmospheric Chemistry and Climate Change in the Anthropocene*. Springer, pp. 211–215.
- Cushing, DH (1971). "Upwelling and the production of fish". In: *Advances in marine biology*. Vol. 9. Elsevier, pp. 255–334.
- Dai, Aiguo (2006). "Precipitation characteristics in eighteen coupled climate models". In: *Journal of climate* 19.18, pp. 4605–4630.
- Dai, Aiguo (2013). "Increasing drought under global warming in observations and models". In: *Nature climate change* 3.1, pp. 52–58.
- Dai, Aiguo, John C Fyfe, Shang-Ping Xie, and Xingang Dai (2015). "Decadal modulation of global surface temperature by internal climate variability". In: *Nature Climate Change* 5.6, p. 555.
- Dee, Dick P, SM Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, d P Bauer, et al. (2011). "The ERA-Interim reanalysis:

- Configuration and performance of the data assimilation system". In: *Quarterly Journal of the royal meteorological society* 137.656, pp. 553–597.
- DelSole, Timothy (2000). "A fundamental limitation of Markov models". In: *Journal of the atmospheric sciences* 57.13, pp. 2158–2168.
- DelSole, Timothy, Michael K Tippett, and Jagadish Shukla (2011). "A significant component of unforced multidecadal variability in the recent acceleration of global warming". In: *Journal of Climate* 24.3, pp. 909–926.
- Delworth, Thomas L and Richard J Greatbatch (2000). "Multidecadal thermohaline circulation variability driven by atmospheric surface flux forcing". In: *Journal of Climate* 13.9, pp. 1481–1495.
- Delworth, Thomas L, Rong Zhang, and Michael E Mann (2007). "Decadal to centennial variability of the Atlantic from observations and models". In: *Geophysical Monograph-American Geophysical Union* 173, p. 131.
- DeMenocal, Peter B (2001). "Cultural responses to climate change during the late Holocene". In: *Science* 292.5517, pp. 667–673.
- Deser, Clara, Michael A Alexander, Shang-Ping Xie, and Adam S Phillips (2010). "Sea surface temperature variability: Patterns and mechanisms". In: *Annual review of marine science* 2, pp. 115–143.
- Dewitte, Boris, J Vazquez-Cuervo, Katerina Goubanova, Sérena Illig, Ken Takahashi, Gildas Cambon, Sara Purca, D Correa, Dimitri Gutiérrez, Abdelfettah Sifeddine, et al. (2012). "Change in El Nino flavours over 1958–2008: Implications for the long-term trend of the upwelling off Peru". In: *Deep Sea Research Part II: Topical Studies in Oceanography* 77, pp. 143–156.
- Dijkstra, Henk A and J David Neelin (1995). "Ocean-atmosphere interaction and the tropical climatology. Part II: Why the Pacific cold tongue is in the east". In: *Journal of climate* 8.5, pp. 1343–1359.
- Ding, Ruiqiang, Jianping Li, and Yu-heng Tseng (2015). "The impact of South Pacific extratropical forcing on ENSO and comparisons with the North Pacific". In: *Climate Dynamics* 44.7-8, pp. 2017–2034.
- Doney, Scott C, William M Balch, Victoria J Fabry, and Richard A Feely (2009). "Ocean acidification: a critical emerging problem for the ocean sciences". In: *Oceanography* 22.4, pp. 16–25.
- Dong, Bo and Aiguo Dai (2015). "The influence of the interdecadal Pacific oscillation on temperature and precipitation over the globe". In: *Climate dynamics* 45.9-10, pp. 2667–2681.
- Dong, Lu and Michael J McPhaden (2016). "Interhemispheric SST gradient trends in the Indian Ocean prior to and during the recent global warming hiatus". In: *Journal of Climate* 29.24, pp. 9077–9095.
- Doucet, Arnaud, Simon Godsill, and Christophe Andrieu (2000). "On sequential Monte Carlo sampling methods for Bayesian filtering". In: *Statistics and computing* 10.3, pp. 197–208.

- Dubinkina, Svetlana and Hugues Goosse (2013). "An assessment of particle filtering methods and nudging for climate state reconstructions". In: *Climate of the Past* 9.3, p. 1141.
- Dubinkina, Svetlana, Hugues Goosse, Yoann Sallaz-Damaz, Elisabeth Crespin, and Michel Crucifix (2011). "Testing a particle filter to reconstruct climate changes over the past centuries". In: *International Journal of Bifurcation and Chaos* 21.12, pp. 3611–3618.
- Dufresne, J-L, M-A Foujols, Sébastien Denvil, Arnaud Caubel, Olivier Marti, Olivier Aumont, Yves Balkanski, Slimane Bekki, Hugo Bellenger, Rachid Benshila, et al. (2013). "Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5". In: *Climate Dynamics* 40.9-10, pp. 2123–2165.
- Dufresne, Jean-Louis and Sandrine Bony (2008). "An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models". In: *Journal of Climate* 21.19, pp. 5135–5144.
- Edwards, Paul N (2011). "History of climate modeling". In: *Wiley Interdisciplinary Reviews: Climate Change* 2.1, pp. 128–139.
- Ekman, Vagn Walfrid (1905). "On the influence of the earth's rotation on ocean-currents." In:
- Enfield, David B, Alberto M Mestas-Nuñez, and Paul J Trimble (2001). "The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US". In: *Geophysical Research Letters* 28.10, pp. 2077–2080.
- England, Matthew H, Shayne McGregor, Paul Spence, Gerald A Meehl, Axel Timmermann, Wenju Cai, Alex Sen Gupta, Michael J McPhaden, Ariaan Purich, and Agus Santoso (2014). "Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus". In: *Nature climate change* 4.3, pp. 222–227.
- Espinoza-Morriberón, D, Vincent Echevin, François Colas, J Tam, J Ledesma, L Vásquez, and M Graco (2017). "Impacts of E I N ño events on the P eruvian upwelling system productivity". In: *Journal of Geophysical Research: Oceans* 122.7, pp. 5423–5444.
- Evan, Amato T, Daniel J Vimont, Andrew K Heidinger, James P Kossin, and Ralf Ben-nartz (2009). "The role of aerosols in the evolution of tropical North Atlantic Ocean temperature anomalies". In: *Science* 324.5928, pp. 778–781.
- Evans, Michael N, Alexey Kaplan, and Mark A Cane (1998). "Optimal sites for coral-based reconstruction of global sea surface temperature". In: *Paleoceanography* 13.5, pp. 502–516.
- Falvey, Mark and René D Garreaud (2009). "Regional cooling in a warming world: Recent temperature trends in the southeast Pacific and along the west coast of subtropical South America (1979–2006)". In: *Journal of Geophysical Research: Atmospheres* 114.D4.
- Fang, Miao and Xin Li (2016). "Paleoclimate data assimilation: Its motivation, progress and prospects". In: *Science China Earth Sciences* 59.9, pp. 1817–1826.
- Farneti, Riccardo, Franco Molteni, and Fred Kucharski (2014). "Pacific interdecadal variability driven by tropical–extratropical interactions". In: *Climate dynamics* 42.11-12, pp. 3337–3355.

- Fichefet, Thierry and MA Morales Maqueda (1997). "Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics". In: *Journal of Geophysical Research: Oceans* 102.C6, pp. 12609–12646.
- Fichefet, Thierry and MA Morales Maqueda (1999). "Modelling the influence of snow accumulation and snow-ice formation on the seasonal cycle of the Antarctic sea-ice cover". In: *Climate Dynamics* 15.4, pp. 251–268.
- Fierro, Ricardo D, Gene H Golub, Per Christian Hansen, and Dianne P O'Leary (1997). "Regularization by truncated total least squares". In: *SIAM Journal on Scientific Computing* 18.4, pp. 1223–1241.
- Figuroa, Dante and Carlos Moffat (2000). "On the influence of topography in the induction of coastal upwelling along the Chilean coast". In: *Geophysical Research Letters* 27.23, pp. 3905–3908.
- Fleming, Laura E and Kevin J Anchukaitis (2016). "North Pacific decadal variability in the CMIP5 last millennium simulations". In: *Climate Dynamics* 47.12, pp. 3783–3801.
- Folland, CK and DE Parker (1995). "Correction of instrumental biases in historical sea surface temperature data". In: *Quarterly Journal of the Royal Meteorological Society* 121.522, pp. 319–367.
- Folland, CK, JA Renwick, MJ Salinger, and AB Mullan (2002). "Relative influences of the interdecadal Pacific oscillation and ENSO on the South Pacific convergence zone". In: *Geophysical Research Letters* 29.13, pp. 21–1.
- Forster, P, V Ramaswamy, P Artaxo, T Berntsen, R Betts, DW Fahey, et al. (2007a). *Chapter 2: Changes in atmospheric constituents and in radiative forcing.-contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, 2007.*
- Forster, Piers, Venkatachalam Ramaswamy, Paulo Artaxo, Terje Berntsen, Richard Betts, David W Fahey, James Haywood, Judith Lean, David C Lowe, Gunnar Myhre, et al. (2007b). "Changes in atmospheric constituents and in radiative forcing. Chapter 2". In: *Climate Change 2007. The Physical Science Basis.*
- Forster, Piers M, Timothy Andrews, Peter Good, Jonathan M Gregory, Lawrence S Jackson, and Mark Zelinka (2013). "Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models". In: *Journal of Geophysical Research: Atmospheres* 118.3, pp. 1139–1150.
- Frankignoul, Claude and Klaus Hasselmann (1977). "Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability". In: *Tellus* 29.4, pp. 289–305.
- Frankignoul, Claude, Guillaume Gastineau, and Young-Oh Kwon (2017). "Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic multidecadal oscillation and the Pacific decadal oscillation". In: *Journal of Climate* 30.24, pp. 9871–9895.
- Fréon, Pierre, Manuel Barange, and Javier Aristegui (2009). "Eastern boundary upwelling ecosystems: integrative and comparative approaches preface". In: *Progress in Oceanography* 83, Jan–14.

- Ganachaud, Alexandre and Carl Wunsch (2003). "Large-scale ocean heat and freshwater transports during the World Ocean Circulation Experiment". In: *Journal of Climate* 16.4, pp. 696–705.
- Garreaud, Rene D and Mark Falvey (2009). "The coastal winds off western subtropical South America in future climate scenarios". In: *International Journal of Climatology: A Journal of the Royal Meteorological Society* 29.4, pp. 543–554.
- Gastineau, Guillaume and Claude Frankignoul (2012). "Cold-season atmospheric response to the natural variability of the Atlantic meridional overturning circulation". In: *Climate dynamics* 39.1-2, pp. 37–57.
- Gastineau, Guillaume, Andrew R Friedman, Myriam Khodri, and Jérôme Vialard (2019). "Global ocean heat content redistribution during the 1998–2012 Interdecadal Pacific Oscillation negative phase". In: *Climate dynamics* 53.1-2, pp. 1187–1208.
- Gebbie, G and P Huybers (2019). "The little ice age and 20th-century deep Pacific cooling". In: *science* 363.6422, pp. 70–74.
- Gill, Adrian E (1980). "Some simple solutions for heat-induced tropical circulation". In: *Quarterly Journal of the Royal Meteorological Society* 106.449, pp. 447–462.
- Gillett, Nathan P and David WJ Thompson (2003). "Simulation of recent Southern Hemisphere climate change". In: *Science* 302.5643, pp. 273–275.
- Gleissberg, W (1939). "A long-periodic fluctuation of the sun-spot numbers". In: *The Observatory* 62, pp. 158–159.
- Gneiting, Tilmann and Adrian E Raftery (2007). "Strictly proper scoring rules, prediction, and estimation". In: *Journal of the American statistical Association* 102.477, pp. 359–378.
- Gómez-Navarro, Juan José, Eduardo Zorita, Christoph Raible, and Raphael Neukom (2017). "Pseudo-proxy tests of the analogue method to reconstruct spatially resolved global temperature during the Common Era". In: *Climate of the Past* 13.6, pp. 629–648.
- Gong, Zhanqiu, Cheng Sun, Jianping Li, Juan Feng, Fei Xie, Ruiqiang Ding, Yun Yang, and Jiaqing Xue (2019). "An inter-basin teleconnection from the North Atlantic to the subarctic North Pacific at multidecadal time scales". In: *Climate Dynamics*, pp. 1–16.
- Goosse, Hugues, Olivier Arzel, Jürg Luterbacher, ME Mann, H Renssen, Nadja Riedwyl, A Timmermann, Elena Xoplaki, and Heinz Wanner (2006a). "The origin of the European Medieval Warm Period". In:
- Goosse, Hugues, Hans Renssen, Axel Timmermann, Raymond S Bradley, and Michael E Mann (2006b). "Using paleoclimate proxy-data to select optimal realisations in an ensemble of simulations of the climate of the past millennium". In: *Climate Dynamics* 27.2-3, pp. 165–184.
- Goosse, Hugues, Elisabeth Crespin, Anne de Montety, ME Mann, H Renssen, and A Timmermann (2010). "Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation". In: *Journal of Geophysical Research: Atmospheres* 115.D9.
- Goosse, Hugues, Elisabeth Crespin, Svetlana Dubinkina, Marie-France Loutre, Michael E Mann, Hans Renssen, Yoann Sallaz-Damaz, and Drew Shindell (2012). "The role of

- forcing and internal dynamics in explaining the “Medieval Climate Anomaly””. In: *Climate dynamics* 39.12, pp. 2847–2866.
- Gordon, Neil J, David J Salmond, and Adrian FM Smith (1993). “Novel approach to nonlinear/non-Gaussian Bayesian state estimation”. In: *IEEE proceedings F (radar and signal processing)*. Vol. 140. 2. IET, pp. 107–113.
- Graham, Nicholas E and Warren B White (1988). “The El Niño cycle: A natural oscillator of the Pacific ocean—atmosphere system”. In: *Science* 240.4857, pp. 1293–1302.
- Grise, Kevin M, Sean M Davis, Isla R Simpson, Darryn W Waugh, Qiang Fu, Robert J Allen, Karen H Rosenlof, Caroline C Ummerhofer, Kristopher B Karnauskas, Amanda C Maycock, et al. (2019). “Recent tropical expansion: natural variability or forced response?” In: *Journal of Climate* 32.5, pp. 1551–1571.
- Guillot, Dominique, Bala Rajaratnam, Julien Emile-Geay, et al. (2015). “Statistical paleoclimate reconstructions via Markov random fields”. In: *The Annals of Applied Statistics* 9.1, pp. 324–352.
- Gutiérrez, Dimitri, Ioanna Bouloubassi, Abdelfettah Sifeddine, Sara Purca, Katerina Goubanova, Michelle Graco, David Field, Laurence Méjanelle, Federico Velazco, Anne Lorre, et al. (2011). “Coastal cooling and increased productivity in the main upwelling zone off Peru since the mid-twentieth century”. In: *Geophysical Research Letters* 38.7.
- Haigh, Joanna D (2007). “The Sun and the Earth’s climate”. In: *Living reviews in solar physics* 4.1, p. 2.
- Hakim, Gregory, James Annan, Stefan Brönnimann, Michel Crucifix, Tamsin Edwards, Hugues Goosse, André Paul, Gerhard van der Schrier, and Martin Widmann (2013). “Overview of data assimilation methods”. In: *PAGES news* 21.2.
- Hakim, Gregory J, Julien Emile-Geay, Eric J Steig, David Noone, David M Anderson, Robert Tardif, Nathan Steiger, and Walter A Perkins (2016). “The last millennium climate reanalysis project: Framework and first results”. In: *Journal of Geophysical Research: Atmospheres* 121.12, pp. 6745–6764.
- Hale, George E, Ferdinand Ellerman, Seth Barnes Nicholson, and Alfred Harrison Joy (1919). “The magnetic polarity of sun-spots”. In: *The Astrophysical Journal* 49, p. 153.
- Hamill, Thomas M and Stephen J Colucci (1997). “Verification of Eta-RSM short-range ensemble forecasts”. In: *Monthly Weather Review* 125.6, pp. 1312–1327.
- Hamill, Thomas M, Jeffrey S Whitaker, and Chris Snyder (2001). “Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter”. In: *Monthly Weather Review* 129.11, pp. 2776–2790.
- Hasselmann, Klaus (1997). “Multi-pattern fingerprint method for detection and attribution of climate change”. In: *Climate dynamics* 13.9, pp. 601–611.
- Hathaway, David H (2015). “The solar cycle”. In: *Living Reviews in Solar Physics* 12.1, p. 4.
- Haug, Gerald H, Detlef Günther, Larry C Peterson, Daniel M Sigman, Konrad A Hughen, and Beat Aeschlimann (2003). “Climate and the collapse of Maya civilization”. In: *Science* 299.5613, pp. 1731–1735.

- Hawkins, Ed and Rowan Sutton (2009). "Decadal predictability of the Atlantic Ocean in a coupled GCM: Forecast skill and optimal perturbations using linear inverse modeling". In: *Journal of Climate* 22.14, pp. 3960–3978.
- Haywood, James and Olivier Boucher (2000). "Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review". In: *Reviews of geophysics* 38.4, pp. 513–543.
- Hegerl, Gabriele and Francis Zwiers (2011). "Use of models in detection and attribution of climate change". In: *Wiley interdisciplinary reviews: climate change* 2.4, pp. 570–591.
- Held, Isaac M and Brian J Soden (2006). "Robust responses of the hydrological cycle to global warming". In: *Journal of climate* 19.21, pp. 5686–5699.
- Henley, Benjamin J, Joelle Gergis, David J Karoly, Scott Power, John Kennedy, and Chris K Folland (2015). "A tripole index for the interdecadal Pacific oscillation". In: *Climate Dynamics* 45.11-12, pp. 3077–3090.
- Henley, Benjamin J, Gerald Meehl, Scott B Power, Chris K Folland, Andrew D King, Jaclyn N Brown, David J Karoly, Francois Delage, Ailie JE Gallant, Mandy Freund, et al. (2017). "Spatial and temporal agreement in climate model simulations of the Interdecadal Pacific Oscillation". In: *Environmental Research Letters* 12.4, p. 044011.
- Hoerling, Martin P, Arun Kumar, and Min Zhong (1997). "El Niño, La Niña, and the nonlinearity of their teleconnections". In: *Journal of Climate* 10.8, pp. 1769–1786.
- Holton, James R, Peter H Haynes, Michael E McIntyre, Anne R Douglass, Richard B Rood, and Leonhard Pfister (1995). "Stratosphere-troposphere exchange". In: *Reviews of geophysics* 33.4, pp. 403–439.
- Hourdin, Frédéric, Ionela Musat, Sandrine Bony, Pascale Braconnot, Francis Codron, Jean-Louis Dufresne, Laurent Fairhead, Marie-Angèle Filiberti, Pierre Friedlingstein, Jean-Yves Grandpeix, et al. (2006). "The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection". In: *Climate Dynamics* 27.7-8, pp. 787–813.
- Hourdin, Frédéric, Marie-Alice Foujols, Francis Codron, Virginie Guemas, Jean-Louis Dufresne, Sandrine Bony, Sébastien Denvil, Lionel Guez, François Lott, Josefine Ghattas, et al. (2013a). "Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model". In: *Climate Dynamics* 40.9-10, pp. 2167–2192.
- Hourdin, Frédéric, Jean-Yves Grandpeix, Catherine Rio, Sandrine Bony, Arnaud Jam, Frédérique Cheruy, Nicolas Rochetin, Laurent Fairhead, Abderrahmane Idelkadi, Ionela Musat, et al. (2013b). "LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection". In: *Climate Dynamics* 40.9-10, pp. 2193–2222.
- Houtekamer, Peter L and Herschel L Mitchell (2001). "A sequential ensemble Kalman filter for atmospheric data assimilation". In: *Monthly Weather Review* 129.1, pp. 123–137.

- Huang, Boyin, Viva F Banzon, Eric Freeman, Jay Lawrimore, Wei Liu, Thomas C Peterson, Thomas M Smith, Peter W Thorne, Scott D Woodruff, and Huai-Min Zhang (2015). "Extended reconstructed sea surface temperature version 4 (ERSST. v4). Part I: Upgrades and intercomparisons". In: *Journal of climate* 28.3, pp. 911–930.
- Huddart, Benjamin, Aneesh Subramanian, Laure Zanna, and Tim Palmer (2017). "Seasonal and decadal forecasts of Atlantic Sea surface temperatures using a linear inverse model". In: *Climate Dynamics* 49.5-6, pp. 1833–1845.
- Hughes, MALCOLM K, J Guiot, and CM Ammann (2010). "An emerging paradigm: Process-based climate reconstructions". In: *Inside PAGES* 6.
- Huybers, Peter and William Curry (2006). "Links between annual, Milankovitch and continuum temperature variability". In: *Nature* 441.7091, pp. 329–332.
- IPCC (2018). "Impacts of 1.5 C global warming on natural and human systems". In: *Global warming of 1.5 C.: An IPCC Special Report*. IPCC Secretariat, pp. 175–311.
- Jeandel, Catherine and Rémy Mosseri (2011). "Le climat à découvert". In: *CNRS editions*.
- Jebri, Beyrem, Myriam Khodri, Vincent Echevin, Guillaume Gastineau, Sylvie Thiria, Jérôme Vialard, and Nicolas Lebas (2020). "Contributions of internal variability and external forcing to the recent trends in the Southeastern Pacific and Peru-Chile upwelling system". In: *Journal of Climate*.
- Jin, Fei-Fei (1997a). "An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model". In: *Journal of the atmospheric sciences* 54.7, pp. 811–829.
- Jin, Fei-Fei (1997b). "An equatorial ocean recharge paradigm for ENSO. Part II: A stripped-down coupled model". In: *Journal of the Atmospheric Sciences* 54.7, pp. 830–847.
- Jin, Fei-Fei, Seon Tae Kim, and Luis Bejarano (2006). "A coupled-stability index for ENSO". In: *Geophysical Research Letters* 33.23.
- Johnson, Scot D, David S Battisti, and ES Sarachik (2000). "Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies". In: *Journal of climate* 13.1, pp. 3–17.
- Jones, Philip D and Michael E Mann (2004). "Climate over past millennia". In: *Reviews of Geophysics* 42.2.
- Joos, Fortunat and Renato Spahni (2008). "Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years". In: *Proceedings of the National Academy of Sciences* 105.5, pp. 1425–1430.
- Kalnay, Eugenia, Masao Kanamitsu, Robert Kistler, William Collins, Dennis Deaven, Lev Gandin, Mark Iredell, Suranjana Saha, Glenn White, John Woollen, et al. (1996). "The NCEP/NCAR 40-year reanalysis project". In: *Bulletin of the American meteorological Society* 77.3, pp. 437–472.
- Kämpf, Jochen and Piers Chapman (2016). "The Peruvian-Chilean Coastal Upwelling System". In: *Upwelling Systems of the World*. Springer, pp. 161–201.
- Kanzow, Torsten, SA Cunningham, William E Johns, J JM Hirschi, J Marotzke, MO Baringer, CS Meinen, MP Chidichimo, C Atkinson, LM Beal, et al. (2010). "Seasonal variability of the Atlantic meridional overturning circulation at 26.5 N". In: *Journal of Climate* 23.21, pp. 5678–5698.

- Kaufman, Darrell S, David P Schneider, Nicholas P McKay, Caspar M Ammann, Raymond S Bradley, Keith R Briffa, Gifford H Miller, Bette L Otto-Bliesner, Jonathan T Overpeck, Bo M Vinther, et al. (2009). "Recent warming reverses long-term Arctic cooling". In: *Science* 325.5945, pp. 1236–1239.
- Keeling, Ralph F and Charles D Keeling (2017). "Atmospheric Monthly In Situ CO2 Data-Mauna Loa Observatory, Hawaii". In: *Scripps CO2 program data. UC San Diego Library Digital Collections*.
- Kent, Elizabeth C, Scott D Woodruff, and David I Berry (2007). "Metadata from WMO publication no. 47 and an assessment of voluntary observing ship observation heights in ICOADS". In: *Journal of Atmospheric and Oceanic Technology* 24.2, pp. 214–234.
- Khodri, Myriam, Takeshi Izumo, Jérôme Vialard, Serge Janicot, Christophe Cassou, Matthieu Lengaigne, Juliette Mignot, Guillaume Gastineau, Eric Guilyardi, Nicolas Lebas, et al. (2017). "Tropical explosive volcanic eruptions can trigger El Niño by cooling tropical Africa". In: *Nature communications* 8.1, pp. 1–13.
- Kim, Hye-Mi, Peter J Webster, and Judith A Curry (2012). "Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts". In: *Geophysical Research Letters* 39.10.
- Kirtman, Ben, Scott B Power, Akintayo John Adedoyin, George J Boer, Roxana Bojariu, Ines Camilloni, Francisco Doblas-Reyes, Arlene M Fiore, Masahide Kimoto, Gerald Meehl, et al. (2013). "Near-term climate change: projections and predictability". In: Kistler, Robert, Eugenia Kalnay, William Collins, Suranjana Saha, Glenn White, John Woollen, Muthuvel Chelliah, Wesley Ebisuzaki, Masao Kanamitsu, Vernon Kousky, et al. (2001). "The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation". In: *Bulletin of the American Meteorological society* 82.2, pp. 247–268.
- Kloeden, Peter E and Eckhard Platen (1992). "Higher-order implicit strong numerical schemes for stochastic differential equations". In: *Journal of statistical physics* 66.1-2, pp. 283–314.
- Knight, Jeff R, Robert J Allan, Chris K Folland, Michael Vellinga, and Michael E Mann (2005). "A signature of persistent natural thermohaline circulation cycles in observed climate". In: *Geophysical Research Letters* 32.20.
- Kociuba, Greg and Scott B Power (2015). "Inability of CMIP5 models to simulate recent strengthening of the Walker circulation: Implications for projections". In: *Journal of Climate* 28.1, pp. 20–35.
- Kong, Augustine, Jun S Liu, and Wing Hung Wong (1994). "Sequential imputations and Bayesian missing data problems". In: *Journal of the American statistical association* 89.425, pp. 278–288.
- Krinner, Gerhard, Nicolas Viovy, Nathalie de Noblet-Ducoudré, Jérôme Ogée, Jan Polcher, Pierre Friedlingstein, Philippe Ciais, Stephen Sitch, and I Colin Prentice (2005). "A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system". In: *Global Biogeochemical Cycles* 19.1.
- Kushnir, Yochanan, Adam A Scaife, Raymond Arritt, Gianpaolo Balsamo, George Boer, Francisco Doblas-Reyes, Ed Hawkins, Masahide Kimoto, Rupa Kumar Kolli, Arun

- Kumar, et al. (2019). "Towards operational predictions of the near-term climate". In: *Nature Climate Change* 9.2, pp. 94–101.
- Lamont, T, M García-Reyes, SJ Bograd, CD van der Lingen, and WJ Sydeman (2018). "Upwelling indices for comparative ecosystem studies: Variability in the Benguela Upwelling System". In: *Journal of Marine Systems* 188, pp. 3–16.
- Larkin, Narasimhan K and DE Harrison (2002). "ENSO warm (El Niño) and cold (La Niña) event life cycles: Ocean surface anomaly patterns, their symmetries, asymmetries, and implications". In: *Journal of climate* 15.10, pp. 1118–1140.
- Lean, Judith (1991). "Variations in the Sun's radiative output". In: *Reviews of Geophysics* 29.4, pp. 505–535.
- Levine, Aaron FZ, Michael J McPhaden, and Dargan MW Frierson (2017). "The impact of the AMO on multidecadal ENSO variability". In: *Geophysical Research Letters* 44.8, pp. 3877–3886.
- Levine, Aaron FZ, Dargan MW Frierson, and Michael J McPhaden (2018). "AMO forcing of multidecadal Pacific ITCZ variability". In: *Journal of Climate* 31.14, pp. 5749–5764.
- Levitus, Sydney, John I Antonov, Tim P Boyer, Olga K Baranova, Hernan Eduardo Garcia, Ricardo Alejandro Locarnini, Alexey V Mishonov, JR Reagan, Dan Seidov, Evgeney S Yarosh, et al. (2012). "World ocean heat content and thermocline sea level change (0–2000 m), 1955–2010". In: *Geophysical Research Letters* 39.10.
- Li, Gen and Shang-Ping Xie (2014). "Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems". In: *Journal of Climate* 27.4, pp. 1765–1780.
- Li, Gen, Shang-Ping Xie, Yan Du, and Yiyong Luo (2016). "Effects of excessive equatorial cold tongue bias on the projections of tropical Pacific climate change. Part I: The warming pattern in CMIP5 multi-model ensemble". In: *Climate dynamics* 47.12, pp. 3817–3831.
- Li, Yang, Quanliang Chen, Xiaoran Liu, Jianping Li, Nan Xing, Fei Xie, Juan Feng, Xin Zhou, Hongke Cai, and Zhenglin Wang (2019). "Long-term trend of the tropical Pacific trade winds under global warming and its causes". In: *Journal of Geophysical Research: Oceans* 124.4, pp. 2626–2640.
- Liguori, Giovanni and Emanuele Di Lorenzo (2018). "Meridional modes and increasing Pacific decadal variability under anthropogenic forcing". In: *Geophysical Research Letters* 45.2, pp. 983–991.
- Liu, Huaran, Zhengyu Liu, and Feiyu Lu (2017). "A Systematic Comparison of Particle Filter and EnKF in Assimilating Time-Averaged Observations". In: *Journal of Geophysical Research: Atmospheres* 122.24, pp. 13–155.
- Liu, Jun S and Rong Chen (1998). "Sequential Monte Carlo methods for dynamic systems". In: *Journal of the American statistical association* 93.443, pp. 1032–1044.
- Liu, Yonggang, X San Liang, and Robert H Weisberg (2007). "Rectification of the bias in the wavelet power spectrum". In: *Journal of Atmospheric and Oceanic Technology* 24.12, pp. 2093–2102.

- Liu, Zhengyu and Mike Alexander (2007). "Atmospheric bridge, oceanic tunnel, and global climatic teleconnections". In: *Reviews of Geophysics* 45.2.
- Lockwood, Mike, Richard Giles Harrison, T Woollings, and Sami K Solanki (2010). "Are cold winters in Europe associated with low solar activity?" In: *Environmental Research Letters* 5.2, p. 024001.
- Luo, Jing-Jia, Wataru Sasaki, and Yukio Masumoto (2012). "Indian Ocean warming modulates Pacific climate change". In: *Proceedings of the National Academy of Sciences* 109.46, pp. 18701–18706.
- Luo, Jing-Jia, Gang Wang, and Dietmar Dommenges (2018). "May common model biases reduce CMIP5's ability to simulate the recent Pacific La Niña-like cooling?" In: *Climate dynamics* 50.3-4, pp. 1335–1351.
- Lyu, Kewei, Xuebin Zhang, John A Church, and Jianyu Hu (2016). "Evaluation of the interdecadal variability of sea surface temperature and sea level in the Pacific in CMIP3 and CMIP5 models". In: *International Journal of Climatology* 36.11, pp. 3723–3740.
- L'Heureux, Michelle L, Sukyoung Lee, and Bradfield Lyon (2013). "Recent multidecadal strengthening of the Walker circulation across the tropical Pacific". In: *Nature Climate Change* 3.6, pp. 571–576.
- MacDonald, Glen M and Roslyn A Case (2005). "Variations in the Pacific Decadal Oscillation over the past millennium". In: *Geophysical Research Letters* 32.8.
- Macfarling Meure, Cecelia, David Etheridge, Cathy Trudinger, P Steele, R Langenfelds, T Van Ommen, A Smith, and J Elkins (2006). "Law Dome CO₂, CH₄ and N₂O ice core records extended to 2000 years BP". In: *Geophysical Research Letters* 33.14.
- Madec, G (2008). "NEMO ocean engine: Notes du Pole de Modélisation 27". In: *Paris: Institut Pierre-Simon Laplace (IPSL)*.
- Maher, Nicola, Alexander Sen Gupta, and Matthew H England (2014). "Drivers of decadal hiatus periods in the 20th and 21st centuries". In: *Geophysical Research Letters* 41.16, pp. 5978–5986.
- Mairesse, Aurélien, Hugues Goosse, Pierre Mathiot, Heinz Wanner, and Svetlana Dubinkina (2013). "Investigating the consistency between proxy-based reconstructions and climate models using data assimilation: a mid-Holocene case study". In: *Climate of the Past* 9.6, pp. 2741–2757.
- Mann, Michael E, Raymond S Bradley, and Malcolm K Hughes (1999). "Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations". In: *Geophysical research letters* 26.6, pp. 759–762.
- Mann, Michael E, Zhihua Zhang, Scott Rutherford, Raymond S Bradley, Malcolm K Hughes, Drew Shindell, Caspar Ammann, Greg Faluvegi, and Fenbiao Ni (2009). "Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly". In: *Science* 326.5957, pp. 1256–1260.
- Mantua, Nathan J and Steven R Hare (2002). "The Pacific decadal oscillation". In: *Journal of oceanography* 58.1, pp. 35–44.
- Marini, Camille and Claude Frankignoul (2014). "An attempt to deconstruct the Atlantic multidecadal oscillation". In: *Climate dynamics* 43.3-4, pp. 607–625.

- Martinez-Villalobos, Cristian, Daniel J Vimont, Cécile Penland, Matthew Newman, and J David Neelin (2018). "Calculating state-dependent noise in a linear inverse model framework". In: *Journal of the Atmospheric Sciences* 75.2, pp. 479–496.
- Mathiot, Pierre, Hugues Goosse, Xavier Crosta, Barbara Stenni, Martina Braida, Hans Renssen, Cédric J Van Meerbeeck, Valérie Masson-Delmotte, Aurélien Mairesse, and Svetlana Dubinkina (2013). "Using data assimilation to investigate the causes of Southern Hemisphere high latitude cooling from 10 to 8 ka BP". In:
- Matsikaris, Anastasios, Martin Widmann, and Johann H Jungclaus (2015). "On-line and off-line data assimilation in palaeoclimatology: a case study". In: *Climate of the Past* 11, pp. 81–93.
- McCabe, Gregory J, Michael A Palecki, and Julio L Betancourt (2004). "Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States". In: *Proceedings of the National Academy of Sciences* 101.12, pp. 4136–4141.
- McGregor, Shayne and Axel Timmermann (2011). "The effect of explosive tropical volcanism on ENSO". In: *Journal of climate* 24.8, pp. 2178–2191.
- McGregor, Shayne, Axel Timmermann, Malte F Stuecker, Matthew H England, Mark Merrifield, Fei-Fei Jin, and Yoshimitsu Chikamoto (2014). "Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming". In: *Nature Climate Change* 4.10, pp. 888–892.
- McKittrick, Ross (2010). "A critical review of global surface temperature data products". In: *Available at SSRN 1653928*.
- McMichael, Anthony J (2000). "The urban environment and health in a world of increasing globalization: issues for developing countries". In: *Bulletin of the World Health Organization* 78, pp. 1117–1126.
- Medhaug, Iselin and Tore Furevik (2011). "North Atlantic 20th century multidecadal variability in coupled climate models: Sea surface temperature and ocean overturning circulation". In:
- Meehl, Gerald A and Aixue Hu (2006). "Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multidecadal Pacific sea surface temperature anomalies". In: *Journal of Climate* 19.9, pp. 1605–1623.
- Meehl, Gerald A, Julie M Arblaster, Grant Branstator, and Harry Van Loon (2008). "A coupled air–sea response mechanism to solar forcing in the Pacific region". In: *Journal of Climate* 21.12, pp. 2883–2897.
- Meehl, Gerald A, Julie M Arblaster, Katja Matthes, Fabrizio Sassi, and Harry van Loon (2009a). "Amplifying the Pacific climate system response to a small 11-year solar cycle forcing". In: *Science* 325.5944, pp. 1114–1118.
- Meehl, Gerald A, Lisa Goddard, James Murphy, Ronald J Stouffer, George Boer, Gokhan Danabasoglu, Keith Dixon, Marco A Giorgetta, Arthur M Greene, ED Hawkins, et al. (2009b). "Decadal prediction: can it be skillful?" In: *Bulletin of the American Meteorological Society* 90.10, pp. 1467–1486.
- Menkes, Christophe E, Matthieu Lengaigne, Jérôme Vialard, Martin Puy, Patrick Marchesiello, Sophie Cravatte, and Gildas Cambon (2014). "About the role of Westerly Wind

- Events in the possible development of an El Niño in 2014". In: *Geophysical research letters* 41.18, pp. 6476–6483.
- Misios, Stergios, Daniel M Mitchell, Lesley J Gray, K Tourpali, Katja Matthes, L Hood, Hauke Schmidt, G Chiodo, R Thiéblemont, E Rozanov, et al. (2016). "Solar signals in CMIP-5 simulations: effects of atmosphere–ocean coupling". In: *Quarterly Journal of the Royal Meteorological Society* 142.695, pp. 928–941.
- Moberg, Anders, Dmitry M Sonechkin, Karin Holmgren, Nina M Datsenko, and Wibjörn Karlén (2005). "Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data". In: *Nature* 433.7026, pp. 613–617.
- Mohino, Elsa, Serge Janicot, and Juergen Bader (2011). "Sahel rainfall and decadal to multi-decadal sea surface temperature variability". In: *Climate dynamics* 37.3-4, pp. 419–440.
- Morice, Colin P, John J Kennedy, Nick A Rayner, and Phil D Jones (2012). "Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set". In: *Journal of Geophysical Research: Atmospheres* 117.D8.
- Muller, Richard A, Judith Curry, Donald Groom, Robert Jacobsen, Saul Perlmutter, Robert Rohde, Arthur Rosenfeld, Charlotte Wickham, and Jonathan Wurtele (2013). "Decadal variations in the global atmospheric land temperatures". In: *Journal of Geophysical Research: Atmospheres* 118.11, pp. 5280–5286.
- Muñoz, Ricardo C and RenéD Garreaud (2005). "Dynamics of the low-level jet off the west coast of subtropical South America". In: *Monthly weather review* 133.12, pp. 3661–3677.
- Myhre, Gunnar, Drew Shindell, François-Marie Bréon, William Collins, Jan Fuglestedt, Jianping Huang, Dorothy Koch, Jean-François Lamarque, David Lee, Blanca Mendoza, et al. (2013). "Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change". In: K., Tignor, M., Allen, SK, Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, PM, Cambridge University Press Cambridge, United Kingdom and New York, NY, USA.
- Nash, J Eamonn and Jonh V Sutcliffe (1970). "River flow forecasting through conceptual models part I—A discussion of principles". In: *Journal of hydrology* 10.3, pp. 282–290.
- Navarra, Antonio and Valeria Simoncini (2010). *A guide to empirical orthogonal functions for climate data analysis*. Springer Science & Business Media.
- Newman, Matthew (2007). "Interannual to decadal predictability of tropical and North Pacific sea surface temperatures". In: *Journal of climate* 20.11, pp. 2333–2356.
- Newman, Matthew (2013). "An empirical benchmark for decadal forecasts of global surface temperature anomalies". In: *Journal of Climate* 26.14, pp. 5260–5269.
- Newman, Matthew, Gilbert P Compo, and Michael A Alexander (2003). "ENSO-forced variability of the Pacific decadal oscillation". In: *Journal of Climate* 16.23, pp. 3853–3857.
- Newman, Matthew, Prashant D Sardeshmukh, and Cécile Penland (2009). "How important is air–sea coupling in ENSO and MJO evolution?" In: *Journal of Climate* 22.11, pp. 2958–2977.

- Newman, Matthew, Michael A Alexander, and James D Scott (2011). "An empirical model of tropical ocean dynamics". In: *Climate dynamics* 37.9-10, p. 1823.
- Newman, Matthew, Michael A Alexander, Toby R Ault, Kim M Cobb, Clara Deser, Emanuele Di Lorenzo, Nathan J Mantua, Arthur J Miller, Shoshiro Minobe, Hisashi Nakamura, et al. (2016). "The Pacific decadal oscillation, revisited". In: *Journal of Climate* 29.12, pp. 4399–4427.
- Nguyen, H, A Evans, C Lucas, I Smith, and B Timbal (2013). "The Hadley circulation in reanalyses: Climatology, variability, and change". In: *Journal of Climate* 26.10, pp. 3357–3376.
- Nidheesh, AG, Matthieu Lengaigne, Jérôme Vialard, Takeshi Izumo, AS Unnikrishnan, and Christophe Cassou (2017). "Influence of ENSO on the Pacific decadal oscillation in CMIP models". In: *Climate Dynamics* 49.9-10, pp. 3309–3326.
- Okumura, Yuko M (2013). "Origins of tropical Pacific decadal variability: Role of stochastic atmospheric forcing from the South Pacific". In: *Journal of climate* 26.24, pp. 9791–9796.
- Okumura, Yuko M and Clara Deser (2010). "Asymmetry in the duration of El Niño and La Niña". In: *Journal of Climate* 23.21, pp. 5826–5843.
- Oldenborgh, Geert Jan van, Francisco J Doblas-Reyes, Bert Wouters, and Wilco Hazeleger (2012). "Decadal prediction skill in a multi-model ensemble". In: *Climate dynamics* 38.7-8, pp. 1263–1280.
- O'Neill, Brian C, Claudia Tebaldi, Detlef P Van Vuuren, Veronika Eyring, Pierre Friedlingstein, George Hurtt, Reto Knutti, Elmar Kriegler, Jean-Francois Lamarque, Jason Lowe, et al. (2016). "The scenario model intercomparison project (ScenarioMIP) for CMIP6". In:
- Otterå, Odd Helge, Mats Bentsen, Helge Drange, and Lingling Suo (2010). "External forcing as a metronome for Atlantic multidecadal variability". In: *Nature Geoscience* 3.10, pp. 688–694.
- Oudar, Thomas, Paul J Kushner, John C Fyfe, and Michael Sigmond (2018). "No Impact of Anthropogenic Aerosols on Early 21st Century Global Temperature Trends in a Large Initial-Condition Ensemble". In: *Geophysical Research Letters* 45.17, pp. 9245–9252.
- PAGES2k Consortium (2013). "Continental-scale temperature variability during the past two millennia". In: *Nature geoscience* 6.5, pp. 339–346.
- PAGES2k Consortium (2017). "A global multiproxy database for temperature reconstructions of the Common Era". In: *Scientific data* 4, p. 170088.
- PAGES2k Consortium (2019). "Consistent multi-decadal variability in global temperature reconstructions and simulations over the common era". In: *Nature geoscience* 12.8, p. 643.
- Pendergrass, Angeline G, Reto Knutti, Flavio Lehner, Clara Deser, and Benjamin M Sanderson (2017). "Precipitation variability increases in a warmer climate". In: *Scientific reports* 7.1, pp. 1–9.
- Penland, Cecile (1989). "Random forcing and forecasting using principal oscillation pattern analysis". In: *Monthly Weather Review* 117.10, pp. 2165–2185.

- Penland, Cécile (1996). "A stochastic model of IndoPacific sea surface temperature anomalies". In: *Physica D: Nonlinear Phenomena* 98.2-4, pp. 534–558.
- Penland, Cécile and Ludmila Matrosova (1994). "A balance condition for stochastic numerical models with application to the El Niño-Southern Oscillation". In: *Journal of climate* 7.9, pp. 1352–1372.
- Penland, Cécile and Ludmila Matrosova (1998). "Prediction of tropical Atlantic sea surface temperatures using linear inverse modeling". In: *Journal of Climate* 11.3, pp. 483–496.
- Penland, Cécile and Prashant D Sardeshmukh (1995). "The optimal growth of tropical sea surface temperature anomalies". In: *Journal of climate* 8.8, pp. 1999–2024.
- Penner, Joyce E, MO Andreae, H Annegarn, L Barrie, J Feichter, D Hegg, A Jayaraman, R Leitch, D Murphy, J Nganga, et al. (2001). "Aerosols, their direct and indirect effects". In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, pp. 289–348.
- Perkins, Walter A and Gregory J Hakim (2017). "Reconstructing paleoclimate fields using online data assimilation with a linear inverse model". In: *Climate of the Past* 13.5, p. 421.
- Philander, S George (1989). "El Niño, La Niña, and the southern oscillation". In: *International geophysics series* 46, pp. X–289.
- Picaut, Joël, François Masia, and Y Du Penhoat (1997). "An advective-reflective conceptual model for the oscillatory nature of the ENSO". In: *Science* 277.5326, pp. 663–666.
- Pielke Sr, Roger A, Gregg Marland, Richard A Betts, Thomas N Chase, Joseph L Eastman, John O Niles, Dev Dutta S Niyogi, and Steven W Running (2002). "The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases". In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 360.1797, pp. 1705–1719.
- Pinheiro, Flavia R, Peter J van Leeuwen, and Gernot Geppert (2019). "Efficient nonlinear data assimilation using synchronization in a particle filter". In: *Quarterly Journal of the Royal Meteorological Society* 145.723, pp. 2510–2523.
- Pitt, Michael K and Neil Shephard (1999). "Filtering via simulation: Auxiliary particle filters". In: *Journal of the American statistical association* 94.446, pp. 590–599.
- Planton, Yann, Jérôme Vialard, Eric Guilyardi, Matthieu Lengaigne, and Takeshi Izumo (2018). "Western Pacific oceanic heat content: A better predictor of La Niña than of El Niño". In: *Geophysical Research Letters* 45.18, pp. 9824–9833.
- Plesca, Elina, Verena Grützun, and Stefan A Buehler (2018). "How robust is the weakening of the Pacific Walker circulation in CMIP5 idealized transient climate simulations?" In: *Journal of Climate* 31.1, pp. 81–97.
- Poli, Paul, Hans Hersbach, Dick P Dee, Paul Berrisford, Adrian J Simmons, Frédéric Vitart, Patrick Laloyaux, David GH Tan, Carole Peubey, Jean-Noël Thépaut, et al. (2016). "ERA-20C: An atmospheric reanalysis of the twentieth century". In: *Journal of Climate* 29.11, pp. 4083–4097.

- Power, S, T Casey, C Folland, A Colman, and V Mehta (1999). "Inter-decadal modulation of the impact of ENSO on Australia". In: *Climate Dynamics* 15.5, pp. 319–324.
- Power, Scott and Rob Colman (2006). "Multi-year predictability in a coupled general circulation model". In: *Climate Dynamics* 26.2-3, pp. 247–272.
- Power, Scott, François Delage, Christine Chung, Greg Kociuba, and Kevin Keay (2013). "Robust twenty-first-century projections of El Niño and related precipitation variability". In: *Nature* 502.7472, pp. 541–545.
- Power, Scott B and Greg Kociuba (2011). "What caused the observed twentieth-century weakening of the Walker circulation?" In: *Journal of Climate* 24.24, pp. 6501–6514.
- Power, Scott B, Francois PD Delage, Christine TY Chung, Hua Ye, and Bradley F Murphy (2017). "Humans have already increased the risk of major disruptions to Pacific rainfall". In: *Nature communications* 8.1, pp. 1–7.
- Pulkkinen, Tuija (2007). "Space weather: terrestrial perspective". In: *Living Reviews in Solar Physics* 4.1, p. 1.
- Puy, Martin, Jérôme Vialard, Matthieu Lengaigne, Eric Guilyardi, Pedro N DiNezio, Aurore Voldoire, Magdalena Balmaseda, Gurvan Madec, Christophe Menkes, and Michael J Mcphaden (2017). "Influence of Westerly Wind Events stochasticity on El Niño amplitude: the case of 2014 vs. 2015". In: *Climate Dynamics*, pp. 1–20.
- Rahn, David A and René D Garreaud (2014). "A synoptic climatology of the near-surface wind along the west coast of South America". In: *International Journal of Climatology* 34.3, pp. 780–792.
- Rayner, NAA, De E Parker, EB Horton, Chris K Folland, Lisa V Alexander, DP Rowell, EC Kent, and A Kaplan (2003). "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century". In: *Journal of Geophysical Research: Atmospheres* 108.D14.
- Reason, CJC and M Rouault (2002). "ENSO-like decadal variability and South African rainfall". In: *Geophysical Research Letters* 29.13, pp. 16–1.
- Reynolds, Richard W, Nick A Rayner, Thomas M Smith, Diane C Stokes, and Wanqiu Wang (2002). "An improved in situ and satellite SST analysis for climate". In: *Journal of climate* 15.13, pp. 1609–1625.
- Rhein, Monika, Stephen R Rintoul, Shigeru Aoki, Edmo Campos, Don Chambers, Richard A Feely, Sergey Gulev, Gregory C Johnson, Simon A Josey, Andrey Kostianoy, et al. (2013). "Observations: ocean". In:
- Riahi, Keywan, Arnulf Grübler, and Nebojsa Nakicenovic (2007). "Scenarios of long-term socio-economic and environmental development under climate stabilization". In: *Technological Forecasting and Social Change* 74.7, pp. 887–935.
- Ribes, Aurélien, Jean-Marc Azaïs, and Serge Planton (2009). "Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate". In: *Climate Dynamics* 33.5, pp. 707–722.
- Robock, Alan (2000). "Volcanic eruptions and climate". In: *Reviews of geophysics* 38.2, pp. 191–219.

- Roemmich, D, J Gilson, R Davis, P Sutton, S Wijffels, and S Riser (2007). "Decadal spinup of the South Pacific subtropical gyre". In: *Journal of Physical Oceanography* 37.2, pp. 162–173.
- Rotstayn, Leon D and Ulrike Lohmann (2002). "Tropical rainfall trends and the indirect aerosol effect". In: *Journal of Climate* 15.15, pp. 2103–2116.
- Rousseaux, Cecile SG, Ryan Lowe, Ming Feng, Anya M Waite, and Peter A Thompson (2012). "The role of the Leeuwin Current and mixed layer depth on the autumn phytoplankton bloom off Ningaloo Reef, Western Australia". In: *Continental Shelf Research* 32, pp. 22–35.
- Ruddiman, William F et al. (2001). *Earth's climate: past and future*. Macmillan.
- Ruprich-Robert, Yohan, Thomas Delworth, Rym Msadek, Frederic Castruccio, Stephen Yeager, and Gokhan Danabasoglu (2018). "Impacts of the Atlantic multidecadal variability on North American summer climate and heat waves". In: *Journal of Climate* 31.9, pp. 3679–3700.
- Rykaczewski, Ryan R and David M Checkley (2008). "Influence of ocean winds on the pelagic ecosystem in upwelling regions". In: *Proceedings of the National Academy of Sciences* 105.6, pp. 1965–1970.
- Sansó, Bruno, Chris E Forest, Daniel Zantedeschi, et al. (2008). "Inferring climate system properties using a computer model". In: *Bayesian Analysis* 3.1, pp. 1–37.
- Sardeshmukh, Prashant D and Philip Sura (2009). "Reconciling non-Gaussian climate statistics with linear dynamics". In: *Journal of Climate* 22.5, pp. 1193–1207.
- Schlesinger, Michael E and Navin Ramankutty (1994). "An oscillation in the global climate system of period 65–70 years". In: *Nature* 367.6465, pp. 723–726.
- Schneider, Tapio and Isaac M Held (2001). "Discriminants of twentieth-century changes in Earth surface temperatures". In: *Journal of Climate* 14.3, pp. 249–254.
- Schurer, Andrew P, Simon FB Tett, and Gabriele C Hegerl (2014). "Small influence of solar variability on climate over the past millennium". In: *Nature Geoscience* 7.2, pp. 104–108.
- Seager, Richard and Isla R Simpson (2016). "Western boundary currents and climate change". In: *Journal of Geophysical Research: Oceans* 121.9, pp. 7212–7214.
- Semenov, Vladimir A, Mojib Latif, Dietmar Dommenges, Noel S Keenlyside, Alexander Strehz, Thomas Martin, and Wonsun Park (2010). "The impact of North Atlantic–Arctic multidecadal variability on Northern Hemisphere surface air temperature". In: *Journal of Climate* 23.21, pp. 5668–5677.
- Sen Gupta, Alexander and Matthew H England (2006). "Coupled ocean–atmosphere–ice response to variations in the southern annular mode". In: *Journal of Climate* 19.18, pp. 4457–4486.
- Sepulchre, Pierre, Lisa C Sloan, Mark Snyder, and Jerome Fiechter (2009). "Impacts of Andean uplift on the Humboldt Current system: A climate model sensitivity study". In: *Paleoceanography* 24.4.
- Servonnat, J, P Yiou, Myriam Khodri, D Swingedouw, and S Denvil (2010). "Influence of solar variability, CO₂ and orbital forcing between 1000 and 1850 AD in the IPSLCM4 model." In: *Climate of the Past* 6.3.

- Slutz, Ralph J, Sandra J Lubker, Jane D Hiscox, Scott D Woodruff, Roy L Jenne, Dennis H Joseph, Peter M Steurer, and Joe D Elms (1985). "Comprehensive ocean-atmosphere data set: Release 1". In: *Boulder, Colo.[NOAA and NCDC]*.
- Smerdon, Jason E (2012). "Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments". In: *Wiley Interdisciplinary Reviews: Climate Change* 3.1, pp. 63–77.
- Smerdon, Jason E, Alexey Kaplan, Diana Chang, and Michael N Evans (2010). "A pseudo-proxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium". In: *Journal of Climate* 23.18, pp. 4856–4880.
- Smith, Doug M, Ben BB Booth, Nick J Dunstone, Rosie Eade, Leon Hermanson, Gareth S Jones, Adam A Scaife, Katy L Sheen, and Vikki Thompson (2016). "Role of volcanic and anthropogenic aerosols in the recent global surface warming slowdown". In: *Nature Climate Change* 6.10, pp. 936–940.
- Snyder, Chris, Thomas Bengtsson, Peter Bickel, and Jeff Anderson (2008). "Obstacles to high-dimensional particle filtering". In: *Monthly Weather Review* 136.12, pp. 4629–4640.
- Sohn, Byung-Ju, Sang-Wook Yeh, Ahreum Lee, and William KM Lau (2019). "Regulation of atmospheric circulation controlling the tropical Pacific precipitation change in response to CO₂ increases". In: *Nature communications* 10.1, pp. 1–8.
- Solomon, Amy and Matthew Newman (2012). "Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record". In: *Nature Climate Change* 2.9, pp. 691–699.
- Son, S-W, EP Gerber, J Perlwitz, Lorenzo M Polvani, NP Gillett, K-H Seo, Veronika Eyring, TG Shepherd, D Waugh, H Akiyoshi, et al. (2010). "Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment". In: *Journal of Geophysical Research: Atmospheres* 115.D3.
- Spiller, Elaine T, Amarjit Budhiraja, Kayo Ide, and Chris KRT Jones (2008). "Modified particle filter methods for assimilating Lagrangian data into a point-vortex model". In: *Physica D: Nonlinear Phenomena* 237.10-12, pp. 1498–1506.
- Staten, Paul W, Jian Lu, Kevin M Grise, Sean M Davis, and Thomas Birner (2018). "Re-examining tropical expansion". In: *Nature Climate Change* 8.9, pp. 768–775.
- Storch, Hans von, Heike Langenberg, and Frauke Feser (2000). "A spectral nudging technique for dynamical downscaling purposes". In: *Monthly weather review* 128.10, pp. 3664–3673.
- Strub, P Ted (1998). "Coastal ocean circulation off western South America". In: *The global coastal ocean. Regional studies and syntheses*, pp. 273–315.
- Suarez, Max J and Paul S Schopf (1988). "A delayed action oscillator for ENSO". In: *Journal of the atmospheric Sciences* 45.21, pp. 3283–3287.
- Suess, Hans E (1980). "The radiocarbon record in tree rings of the last 8000 years". In: *Radiocarbon* 22.2, pp. 200–209.
- Sutton, Rowan T and Daniel LR Hodson (2005). "Atlantic Ocean forcing of North American and European summer climate". In: *science* 309.5731, pp. 115–118.
- Sverdrup, Harald Ulrik (1937). *On the process of upwelling*. Sears Foundation.

- Swingedouw, Didier, Juliette Mignot, Eric Guilyardi, Sébastien Nguyen, and Lola Ormières (2017). "Tentative reconstruction of the 1998–2012 hiatus in global temperature warming using the IPSL–CM5A–LR climate model". In: *Comptes Rendus Geoscience* 349.8, pp. 369–379.
- Takahashi, Chiharu and Masahiro Watanabe (2016). "Pacific trade winds accelerated by aerosol forcing over the past two decades". In: *Nature Climate Change* 6.8, pp. 768–772.
- Talagrand, Olivier (1997). "Assimilation of observations, an introduction (gtspecial issuel-talagrand data assimilation in meteorology and oceanography: Theory and practice)". In: *Journal of the Meteorological Society of Japan. Ser. II* 75.1B, pp. 191–209.
- Talagrand, Olivier (1999). "Evaluation of probabilistic prediction systems". In: *Workshop proceedings "Workshop on predictability", 20-22 October 1997, ECMWF, Reading, UK.*
- Taylor, Karl E, Ronald J Stouffer, and Gerald A Meehl (2012). "An overview of CMIP5 and the experiment design". In: *Bulletin of the American Meteorological Society* 93.4, pp. 485–498.
- Terray, Laurent (2012). "Evidence for multiple drivers of North Atlantic multi-decadal climate variability". In: *Geophysical Research Letters* 39.19.
- Thompson, David WJ and Susan Solomon (2002). "Interpretation of recent Southern Hemisphere climate change". In: *Science* 296.5569, pp. 895–899.
- Thompson, David WJ, John J Kennedy, John M Wallace, and Phil D Jones (2008). "A large discontinuity in the mid-twentieth century in observed global-mean surface temperature". In: *Nature* 453.7195, pp. 646–649.
- Thompson, David WJ, John M Wallace, Phil D Jones, and John J Kennedy (2009). "Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights". In: *Journal of Climate* 22.22, pp. 6120–6141.
- Tipton, John, Mevin Hooten, Neil Pederson, Martin Tingley, and Daniel Bishop (2016). "Reconstruction of late Holocene climate based on tree growth and mechanistic hierarchical models". In: *Environmetrics* 27.1, pp. 42–54.
- Tokinaga, Hiroki, Shang-Ping Xie, Clara Deser, Yu Kosaka, and Yuko M Okumura (2012). "Slowdown of the Walker circulation driven by tropical Indo-Pacific warming". In: *Nature* 491.7424, pp. 439–443.
- Torrence, Christopher and Gilbert P Compo (1998). "A practical guide to wavelet analysis". In: *Bulletin of the American Meteorological Society* 79.1, pp. 61–78.
- Tourre, Yves M, Balaji Rajagopalan, Yochanan Kushnir, Mathew Barlow, and Warren B White (2001). "Patterns of coherent decadal and interdecadal climate signals in the Pacific basin during the 20th century". In: *Geophysical Research Letters* 28.10, pp. 2069–2072.
- Trenberth, Kevin E and Julie M Caron (2001). "Estimates of meridional atmosphere and ocean heat transports". In: *Journal of Climate* 14.16, pp. 3433–3443.
- Trenberth, Kevin E and Jerry G Olson (1988). "An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts". In: *Bulletin of the American Meteorological Society* 69.9, pp. 1047–1057.

- Trepte, Charles R and Matthew H Hitchman (1992). "Tropical stratospheric circulation deduced from satellite aerosol data". In: *Nature* 355.6361, pp. 626–628.
- Trepte, Charles R, Robert E Veiga, and M Patrick McCormick (1993). "The poleward dispersal of Mount Pinatubo volcanic aerosol". In: *Journal of Geophysical Research: Atmospheres* 98.D10, pp. 18563–18573.
- Tung, Ka Kit and Charles D Camp (2008). "Solar cycle warming at the Earth's surface in NCEP and ERA-40 data: A linear discriminant analysis". In: *Journal of Geophysical Research: Atmospheres* 113.D5.
- Tung, Ka-Kit and Jiansong Zhou (2013). "Using data to attribute episodes of warming and cooling in instrumental records". In: *Proceedings of the National Academy of Sciences* 110.6, pp. 2058–2063.
- Tziperman, Eli, Laure Zanna, and Cecile Penland (2008). "Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere GCM". In: *Journal of Physical Oceanography* 38.3, pp. 588–604.
- Ueno, Genta, Tomoyuki Higuchi, Takashi Kagimoto, and Naoki Hirose (2010). "Maximum likelihood estimation of error covariances in ensemble-based filters and its application to a coupled atmosphere–ocean model". In: *Quarterly Journal of the Royal Meteorological Society* 136.650, pp. 1316–1343.
- Undorf, S, MA Bollasina, BBB Booth, and GC Hegerl (2018). "Contrasting the Effects of the 1850–1975 Increase in Sulphate Aerosols from North America and Europe on the Atlantic in the CESM". In: *Geophysical Research Letters* 45.21, pp. 11–930.
- Uppala, Sakari M, PW Kållberg, AJ Simmons, U Andrae, V Da Costa Bechtold, M Fiorino, JK Gibson, J Haseler, A Hernandez, GA Kelly, et al. (2005). "The ERA-40 re-analysis". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.612, pp. 2961–3012.
- Valcke, S (2013). "The OASIS3 coupler: a European climate modelling community software". In: *Geoscientific Model Development* 6.2, p. 373.
- Van Leeuwen, Peter Jan (2009). "Particle filtering in geophysical systems". In: *Monthly Weather Review* 137.12, pp. 4089–4114.
- Van Leeuwen, Peter Jan, Hans R Künsch, Lars Nerger, Roland Potthast, and Sebastian Reich (2019). "Particle filters for high-dimensional geoscience applications: A review". In: *Quarterly Journal of the Royal Meteorological Society* 145.723, pp. 2335–2365.
- Van Oldenborgh, Geert Jan, Sjoukje Y Philip, and Matthew Collins (2005). "El Nino in a changing climate: a multi-model study". In:
- Van Vuuren, Detlef P, Jae Edmonds, Mikiko Kainuma, Keywan Riahi, Allison Thomson, Kathy Hibbard, George C Hurtt, Tom Kram, Volker Krey, Jean-Francois Lamarque, et al. (2011). "The representative concentration pathways: an overview". In: *Climatic change* 109.1-2, p. 5.
- Vecchi, Gabriel A and Andrew T Wittenberg (2010). "El Niño and our future climate: Where do we stand?" In: *Wiley Interdisciplinary Reviews: Climate Change* 1.2, pp. 260–270.

- Vecchi, Gabriel A, Brian J Soden, Andrew T Wittenberg, Isaac M Held, Ants Leetmaa, and Matthew J Harrison (2006). "Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing". In: *Nature* 441.7089, pp. 73–76.
- Vecchi, Gabriel A, Amy Clement, and Brian J Soden (2008). "Examining the tropical Pacific's response to global warming". In: *Eos, Transactions American Geophysical Union* 89.9, pp. 81–83.
- Verdon-Kidd, DC and AS Kiem (2009). "On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall." In: *Hydrology & Earth System Sciences* 13.4.
- Vialard, Jérôme and Pascale Delecluse (1998). "An OGCM study for the TOGA decade. Part II: Barrier-layer formation and variability". In: *Journal of physical oceanography* 28.6, pp. 1089–1106.
- Villamayor, Julián, Elsa Mohino, Myriam Khodri, Juliette Mignot, and Serge Janicot (2018). "Atlantic control of the late nineteenth-century Sahel Humid period". In: *Journal of Climate* 31.20, pp. 8225–8240.
- Vimont, Daniel J, John M Wallace, and David S Battisti (2003). "The seasonal footprinting mechanism in the Pacific: Implications for ENSO". In: *Journal of Climate* 16.16, pp. 2668–2675.
- Vuille, Mathias, Eric Franquist, René Garreaud, Waldo Sven Lavado Casimiro, and Bolivar Caceres (2015). "Impact of the global warming hiatus on Andean temperature". In: *Journal of Geophysical Research: Atmospheres* 120.9, pp. 3745–3757.
- Wang, Chunzai, Clara Deser, Jin-Yi Yu, Pedro DiNezio, and Amy Clement (2012). "El Niño and southern oscillation (ENSO): a review". In: *Coral reefs of the Eastern Pacific* 8, pp. 3–19.
- Wang, Chunzai, Liping Zhang, Sang-Ki Lee, Lixin Wu, and Carlos R Mechoso (2014). "A global perspective on CMIP5 climate model biases". In: *Nature Climate Change* 4.3, pp. 201–205.
- Watanabe, Masahiro and Hiroaki Tatebe (2019). "Reconciling roles of sulphate aerosol forcing and internal variability in Atlantic multidecadal climate changes". In: *Climate Dynamics* 53.7-8, pp. 4651–4665.
- Weisberg, Robert H and Chunzai Wang (1997). "A western Pacific oscillator paradigm for the El Niño-Southern Oscillation". In: *Geophysical research letters* 24.7, pp. 779–782.
- Werner, Johannes P, Juerg Luterbacher, and Jason E Smerdon (2013). "A pseudoproxy evaluation of Bayesian hierarchical modeling and canonical correlation analysis for climate field reconstructions over Europe". In: *Journal of Climate* 26.3, pp. 851–867.
- White, Warren B, Judith Lean, Daniel R Cayan, and Michael D Dettinger (1997). "Response of global upper ocean temperature to changing solar irradiance". In: *Journal of Geophysical Research: Oceans* 102.C2, pp. 3255–3266.
- Whitfield, Paul H, RD Moore, Sean W Fleming, and Alexi Zawadzki (2010). "Pacific Decadal Oscillation and the hydroclimatology of western Canada—Review and prospects". In: *Canadian Water Resources Journal* 35.1, pp. 1–28.

- Widmann, Martin, Hugues Goosse, Gerard van der Schrier, Reiner Schnur, and Jan Barkmeijer (2010). "Using data assimilation to study extratropical Northern Hemisphere climate over the last millennium." In: *Climate of the Past* 6.5.
- Willeit, M, A Ganopolski, R Calov, and Victor Brovkin (2019). "Mid-Pleistocene transition in glacial cycles explained by declining CO₂ and regolith removal". In: *Science Advances* 5.4, eaav7337.
- Winkler, Christopher R, Matthew Newman, and Prashant D Sardeshmukh (2001). "A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill". In: *Journal of climate* 14.24, pp. 4474–4494.
- Woodruff, Scott D, Ralph J Slutz, Roy L Jenne, and Peter M Steurer (1987). "A comprehensive ocean-atmosphere data set". In: *Bulletin of the American meteorological society* 68.10, pp. 1239–1250.
- Woodruff, SD, HF Diaz, EC Kent, RW Reynolds, and SJ Worley (2008). "The evolving SST record from ICOADS". In: *Climate Variability and Extremes during the Past 100 Years*. Springer, pp. 65–83.
- Wu, Zhaohua, Norden E Huang, John M Wallace, Brian V Smoliak, and Xianyao Chen (2011). "On the time-varying trend in global-mean surface temperature". In: *Climate dynamics* 37.3-4, p. 759.
- Xie, Shang-Ping, Clara Deser, Gabriel A Vecchi, Jian Ma, Haiyan Teng, and Andrew T Wittenberg (2010). "Global warming pattern formation: Sea surface temperature and rainfall". In: *Journal of Climate* 23.4, pp. 966–986.
- Xie, Shang-Ping, Bo Lu, and Baoqiang Xiang (2013). "Similar spatial patterns of climate responses to aerosol and greenhouse gas changes". In: *Nature Geoscience* 6.10, pp. 828–832.
- Xu, Haiming, Yuqing Wang, and Shang-Ping Xie (2004). "Effects of the Andes on eastern Pacific climate: A regional atmospheric model study". In: *Journal of climate* 17.3, pp. 589–602.
- Xue, Yan, Ants Leetmaa, and Ming Ji (2000). "ENSO prediction with Markov models: The impact of sea level". In: *Journal of Climate* 13.4, pp. 849–871.
- Yang, Hu, Gerrit Lohmann, Wei Wei, Mihai Dima, Monica Ionita, and Jiping Liu (2016). "Intensification and poleward shift of subtropical western boundary currents in a warming climate". In: *Journal of Geophysical Research: Oceans* 121.7, pp. 4928–4945.
- Ying, Jun, Ping Huang, Tao Lian, and Hongjian Tan (2019). "Understanding the effect of an excessive cold tongue bias on projecting the tropical Pacific SST warming pattern in CMIP5 models". In: *Climate Dynamics* 52.3-4, pp. 1805–1818.
- Zalasiewicz, Jan, Colin N Waters, Mark Williams, Anthony D Barnosky, Alejandro Cearreta, Paul Crutzen, Erle Ellis, Michael A Ellis, Ian J Fairchild, Jacques Grinevald, et al. (2015). "When did the Anthropocene begin? A mid-twentieth century boundary level is stratigraphically optimal". In: *Quaternary International* 383, pp. 196–203.
- Zanchettin, Davide, Oliver Bothe, Hans F Graf, Nour-Eddine Omrani, Angelo Rubino, and Johann H Jungclaus (2016). "A decadal delayed response of the tropical Pacific to Atlantic multidecadal variability". In: *Geophysical Research Letters* 43.2, pp. 784–792.

- Zanna, Laure (2012). "Forecast skill and predictability of observed Atlantic sea surface temperatures". In: *Journal of Climate* 25.14, pp. 5047–5056.
- Zhang, Honghai, Amy Clement, and Pedro Di Nezio (2014). "The South Pacific meridional mode: A mechanism for ENSO-like variability". In: *Journal of Climate* 27.2, pp. 769–783.
- Zhang, Liping and Chunzai Wang (2013). "Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations". In: *Journal of Geophysical Research: Oceans* 118.10, pp. 5772–5791.
- Zhang, Rong and Thomas L Delworth (2006). "Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes". In: *Geophysical Research Letters* 33.17.
- Zhang, Rong, Thomas L Delworth, Rowan Sutton, Daniel LR Hodson, Keith W Dixon, Isaac M Held, Yochanan Kushnir, John Marshall, Yi Ming, Rym Msadek, et al. (2013). "Have aerosols caused the observed Atlantic multidecadal variability?" In: *Journal of the Atmospheric Sciences* 70.4, pp. 1135–1144.
- Zhang, Rong, Rowan Sutton, Gokhan Danabasoglu, Thomas L Delworth, Who M Kim, Jon Robson, and Stephen G Yeager (2016). "Comment on "The Atlantic Multidecadal Oscillation without a role for ocean circulation"". In: *Science* 352.6293, pp. 1527–1527.
- Zheng, Xiao-Tong, Chang Hui, and Sang-Wook Yeh (2018). "Response of ENSO amplitude to global warming in CESM large ensemble: uncertainty due to internal variability". In: *Climate Dynamics* 50.11-12, pp. 4019–4035.

Annexe A: Tableaux et Figures supplémentaires

A.1 Chapitre 3

Table A.1: Liste des observations utilisées avec les références associées, la période couverte par chacune d'elle et les variables considérées.

Observation	Référence	Période	Variable
OISSTv2	Reynolds et al., 2002	1981-présent	SST
ERSSTv4	Huang et al., 2015	1854-présent	SST
COADS	Slutz et al., 1985; Woodruff et al., 1987	Climatologie de la période 1945-1989	SST et stress du vent
HadISSTv1.1	Rayner et al., 2003	1871-présent	SST
NOAA-20thCv2c	Compo et al., 2011	1854-2014	Vents de surface
GPCP	Adler et al., 2003	1979-présent	précipitations

Table A.2: Liste des réanalyses utilisées avec les références associées, la période couverte par chacune d'elle et la méthode d'assimilation de données utilisée.

Réanalyses	Référence	Période	Méthode
ERA-Interim	Dee et al., 2011	1979-2019	4DVAR
ERA-20C	Poli et al., 2016	1900-2010	4DVAR
NOAA-CIRES 20th	Compo et al., 2011	1851-2014	Filtre de Kalman d'ensemble
NCEP	Kalnay et al., 1996	1979-2018	3DVAR
SODA3.3.1	Carton and Giese, 2008	1980-2015	OI
ORAS4	Balmaseda et al., 2013	1958-2013	3DVAR
ORA-20C	Boisséson et al., 2018	1900-2010	3DVAR

Table A.3: Liste des modèles CMIP5 utilisés dans le chapitre 3 et groupes de modélisations. Les deux dernières colonnes indiquent la période disponible pour les simulations historiques (Hist) et scénarios futurs (RCP8.5).

	Modèle	Groupe	Hist	RCP8.5
1.	ACCESS1-0	Commonwealth Scientific and Industrial Research Organisation and Bureau of Meteorology	1850-2005	2006-2100
2.	ACCESS1-3	Commonwealth Scientific and Industrial Research Organisation and Bureau of Meteorology	1850-2005	2006-2100
3.	bcc-csm1-1	Beijing Climate Center, China Meteorological Administration	1850-2012	2006-2300
4.	bcc-csm1-1-m	Beijing Climate Center, China Meteorological Administration	1850-2012	2006-2100
5.	CanESM2	Canadian Center for Atmospheric Research	1850-2005	2006-2100
6.	CNRM-CM5	Centre National de Recherche Meteorologiques/ Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	1850-2005	2006-2300
7.	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	1850-2005	2006-2300
8.	FGOALS-s2	Institute of Atmospheric Physics and Chinese Academy of Sciences	1850-2005	2006-2100
9.	GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory	1860-2005	2006-2100
10.	GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	1861-2005	2006-2100
11.	GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	1861-2005	2006-2100
12.	GISS-E2-R	Goddard Institute for Space Studies	1850-2005	2006-2300
13.	GISS-E2-R-CC	Goddard Institute for Space Studies	1850-2010	2006-2100
14.	inmcm4	Institute for Numerical Mathematics	1850-2005	2006-2100
15.	IPSL-CM5A-LR	Institut Pierre-Simon Laplace	1850-2005	2006-2300
16.	IPSL-CM5A-MR	Institut Pierre-Simon Laplace	1850-2005	2006-2100
17.	IPSL-CM5B-LR	Institut Pierre-Simon Laplace	1850-2005	2006-2100
18.	MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies and Japan Agency for Marine-Earth Science and Technology	1850-2012	2006-2100
19.	MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1850-2005	2006-2100
20.	MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1850-2005	2006-2100
21.	MPI-ESM-LR	Max Planck Institute for Meteorology	1850-2005	2006-2300
22.	MPI-ESM-MR	Max Planck Institute for Meteorology	1850-2005	2006-2100
23.	MRI-CGCM3	Meteorological Research Institute	1850-2005	2006-2100
24.	MRI-ESM1	Meteorological Research Institute	1851-2005	2006-2100
25.	NorESM1-M	Norwegian Climate Centre	1850-2005	2006-2100

A.2 Chapitre 6

Table A.4: Corrélations et CE des températures de surface moyennes prédites par le LIM et la Persistance par hémisphère à l'horizon 12 mois avec les températures de surface réelles.

	Corr		CE	
	LIM	Persistance	LIM	Persistance
Nord (au-dessus de 25° N)	0.26	0.2	0.08	-0.22
Tropique (entre 25° S-25° N)	0.44	0.26	0.21	-0.27
SH (en dessous de 25° S)	0.31	0.26	0.09	-0.15

Table A.5: Corrélations et CE du SIR-LIM pour les expériences Dense et Pages2k des températures moyennes globales (Global), dans l'Hémisphère Nord (NH), dans les tropiques (Tropique) et dans l'Hémisphère Sud (SH). Les corrélations et CE obtenus pour les indices Niño3.4, TPI non filtrés et l'AMV sont également présentés. L'ensemble de ces scores sont calculés pour des reconstructions d'une période de 150 ans prise sur VAL.

	Pages2k		Dense	
	Corr	CE	Corr	CE
Global	0.83	0.63	0.90	0.78
NH [au-dessus de 25° N]	0.72	0.52	0.75	0.57
Tropique [entre 25° S et 25° N]	0.85	0.65	0.90	0.79
SH [en dessous de 25° S]	0.57	0.17	0.82	0.60
Nino3.4	0.79	0.59	0.87	0.75
TPI	0.84	0.64	0.90	0.78
AMV	0.55	0.28	0.62	0.33

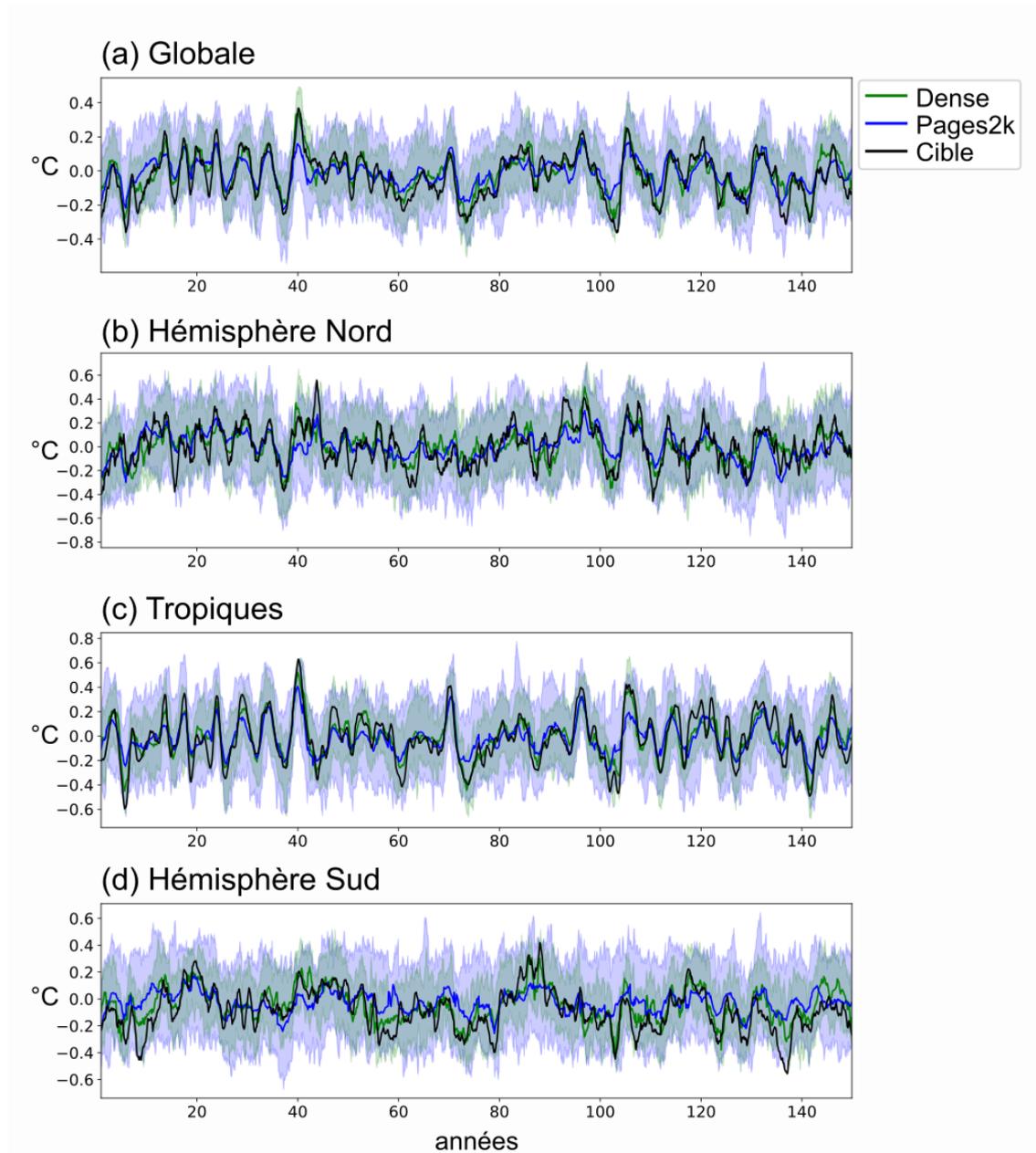


Figure A.1: Reconstruction par le SIR-LIM pour les expériences Pages2k (en bleu) et Dense (en vert) des moyennes de températures (a) globales, (b) de l'Hémisphère Sud [en dessous de 25°S], (c) de l'Hémisphère Nord [au-dessus de 25°N], (d) des Tropiques [entre 25°S et 25°N]. Les traits pleins noirs indiquent la cible à reconstruire. Les traits pleins bleus et verts indiquent les reconstructions obtenues par la moyenne de l'ensemble des particules et les enveloppes bleues et vertes décrivent la dispersion totale de l'ensemble des particules.

Table A.6: Même chose que Tableau A.5 pour SIR-0

	Pages2k		Dense	
	Corr	CE	Corr	CE
Global	0.14	-0.06	0.11	-0.19
NH [au-dessus de 25°N]	0.15	0	0.10	-0.20
Tropique [entre 25°S et 25°N]	0.14	0	0.12	-0.11
SH [en dessous de 25°S]	0.07	-0.34	0.08	-0.46
Nino3.4	0.13	0	0.12	-0.13
TPI	0.15	0	0.12	-0.13
AMV	0.11	0.03	0.06	-0.23

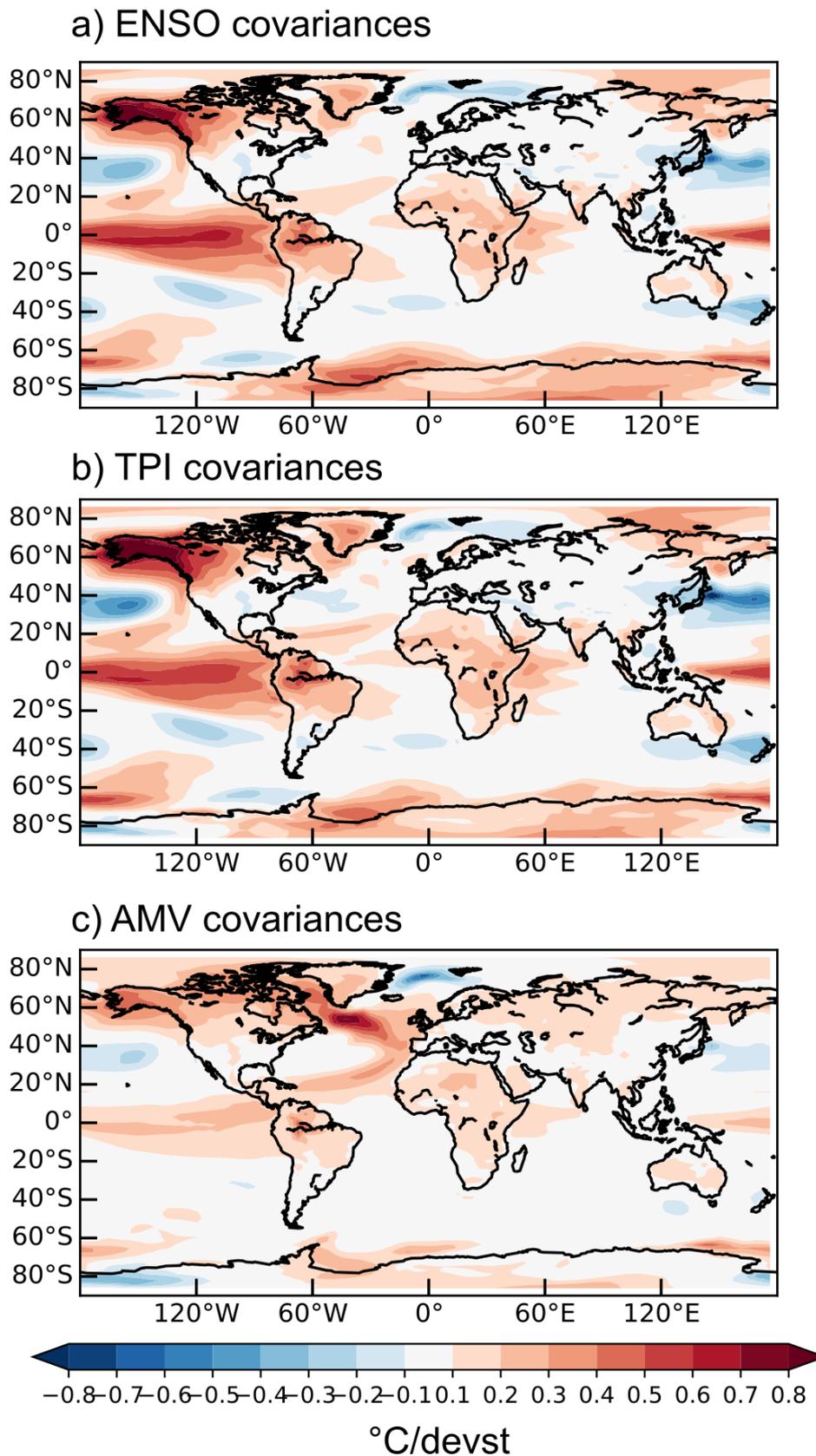


Figure A.2: (a-c) Covariances des indices (a) d'ENSO, (b) du TPI et (e) d'AMV avec les températures de surface.

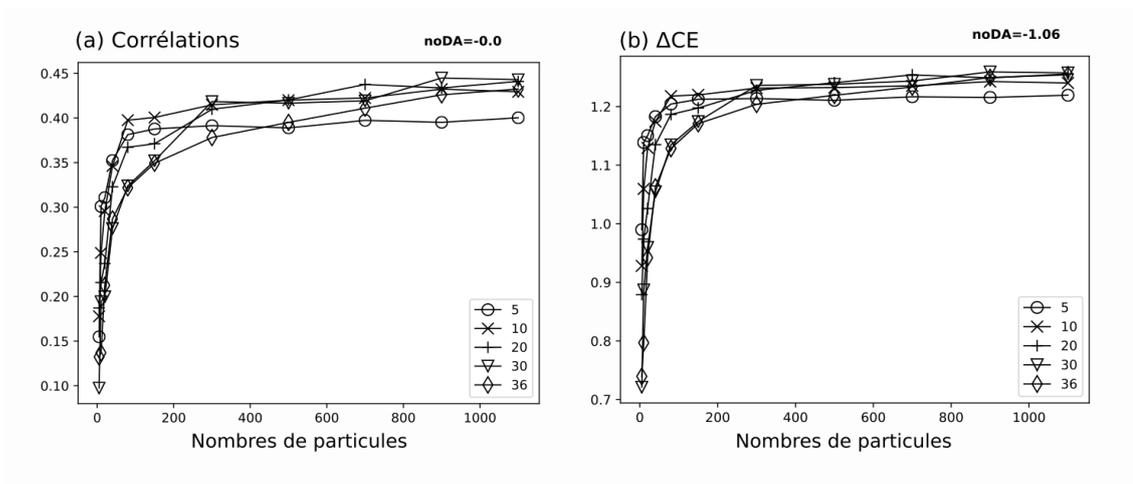


Figure A.3: Expériences de sensibilité du SIR-LIM au nombre d'EOFs et nombre de particules utilisées pour $\sigma = 1.5$. Les expériences sont réalisées avec le sous-échantillonnage spatial Pages2k pour des reconstructions d'une période de 100 ans prise dans VAL. Les scores indiqués correspondent aux moyennes globales de (a) corrélations moyennes avec la corrélation de noDA indiquée en haut à droite, et (b) différence de CE avec le CE noDA indiquée en haut à droite. Les scores de noDA indiqués en haut à droite permettent de retrouver les RMSE et CE d'origines en respectivement multipliant par la RMSE de noDA et ajoutant le CE de noDa.

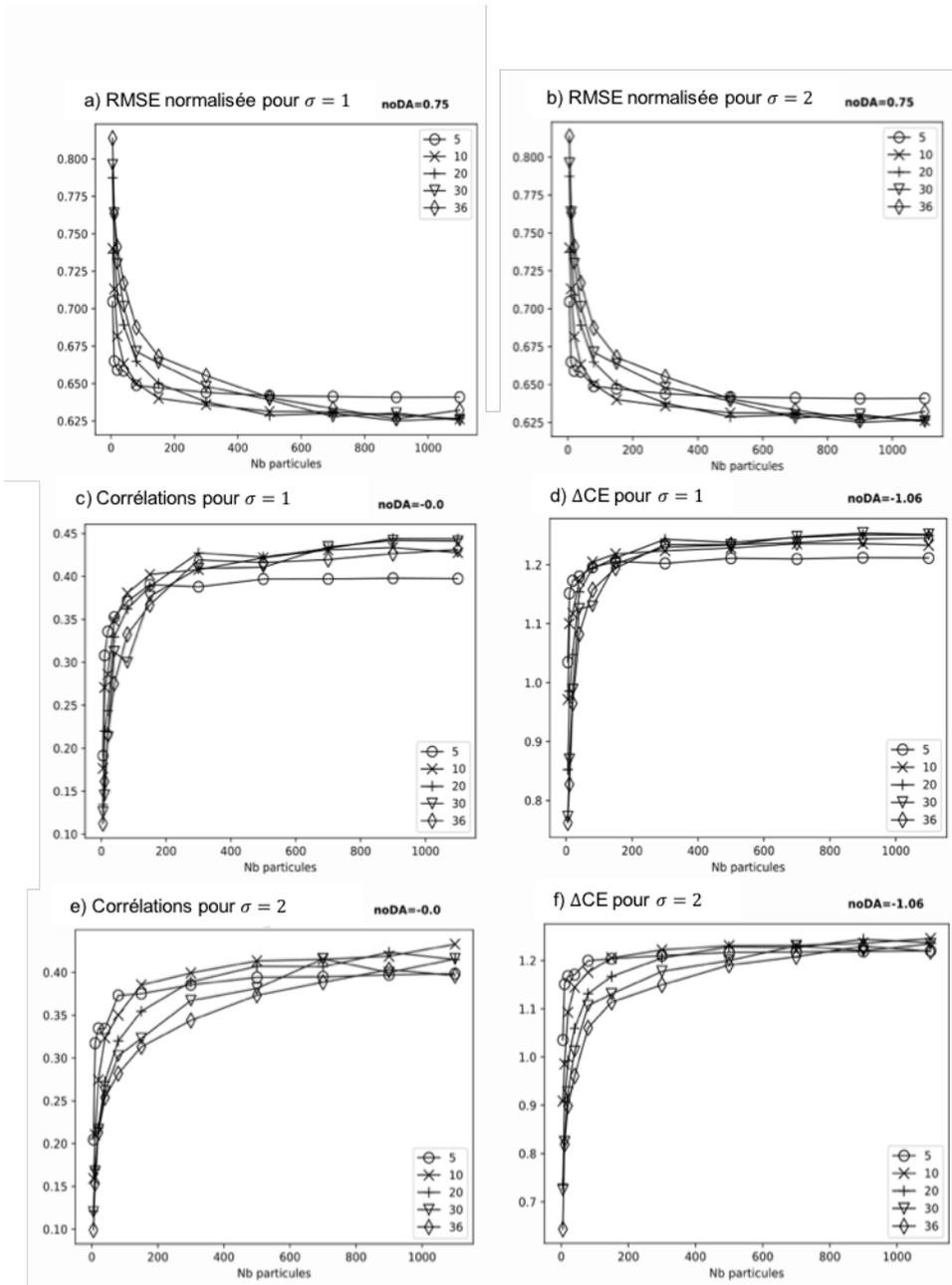


Figure A.4: (a,b) Même chose que Fig. 5.23 pour $\sigma = 1$ (première colonne) et $\sigma = 2$ (seconde colonne). (c,d) Même chose que A.3 pour $\sigma = 1$ (e,f) Même chose que A.3 pour $\sigma = 2$.

Annexe B: SomVar

SELF ORGANIZING MAP FOR FEATURE CLUSTERING - APPLICATION TO CLIMATE DATA

Jebri Beyrem¹ · Niang Ndeye² · Zebiri Aniss³ · Khodri Myriam¹ · Sylvie Thiria^{1,4}

Abstract We present a new neural algorithm (SomVar) able to cluster a large set of variables into a reduced number of classes having a well organized ordering. These clusters composed of variables highly correlated are associated with latent variables (referents) and share the topological proprieties with the classical SOM (Self Organizing Map) algorithm. SomVar is a robust algorithm, weakly sensitive to the initial conditions. It retrieves the hidden structures of the variables and can handle a large number of referents. Moreover SomVar provides an order that can be used to understand the deep intrinsic meaning of the data under study. As an illustration, SomVar was used to analyze an 80 years-long Sea Surface Temperature (SST) data set in the South East Pacific region as simulated by the IPSL-CM5A coupled Ocean-Atmosphere General Circulation Model over the 1940-2020 historical period. The SomVar analysis allows successfully clustering coherent model grid points based on SST co-variability at decadal timescale over the simulated historical period.

Keywords Clustering of variables · Principal component analysis · Climatological data · Self-organizing maps

1 Introduction

Cluster analysis whose objective is to evidence significantly related groups of individuals inside a data base, is one of the major issue of data mining. It encompasses a lot of different methods and is used in many fields as engineering, biology, environment, business, social sciences and climate or meteorology [Risien et al., 2004, Liu and Weisberg, 2005, Liu et al., 2006]. The two most popular approaches to

clustering are i) the agglomerative hierarchical clustering that proceeds iteratively by merging small clusters into larger ones, and ii) the direct partitioning methods such as the widely used K-means and the SOM [Kohonen, 1982, Kohonen, 1998] in the framework of neural networks. A survey on clustering techniques can be found in [Jain, 2010]. These methods are also briefly reviewed in [Kotsiantis et al., 2006] with a focus on recent methods of ensemble clustering.

Most of the time, clustering techniques are used to reduce a large number of individuals described by multi-dimensional variables into coherent clusters of individuals sharing some similarities, the different clusters being well separated. If the number of variables is large and there are numerous co-linearities between them, it may be important to reduce the number of variables by providing latent variables, each one being a representative of a cluster of the initial variables. This is achieved by some existing methods like Empirical Orthogonal Function (EOF) or by k-means like algorithms. Organizing the variables into significant structures associated with latent variables presents some major advantages for the classification and the interpretation. It has to be noticed that bi-clustering methods dedicated to the clustering of both individuals and variables are not considered in this paper.

[Vigneau and Qannari, 2003] proposed a method for clustering variables around latent components (ClustVarLV) according to their degree of correlation. Following the partitioning algorithm proposed by [Vigneau and Qannari, 2003], which is described in the next section, we propose an extension based on the SOM partitioning technique for feature clustering. The new proposed method (SomVar) provides more robust clustering which are less sensitive to local minima. Indeed ClustVarLV which is a K-means like algorithm, owing to the gradient descent use, is highly sensitive to the initialization phase and to the initial value of the cluster centers [Celebi et al., 2013]. It therefore may converge toward secondary local minima. Moreover the topological order provided by the SOM allows a clear representation of the clusters: close neurons on the SOM map represent close clusters of the variables. This is of great interest for the user when analyzing the proposed clustering in order to interpret the initial data base.

The paper is organized in four sections. In the background section (section 2), self organizing map and the clustering variables around latent component methods are briefly reviewed. Then, the proposed methodology is presented in section 3. In the applications section (section 4), the interest of the proposed method and its properties are illustrated with both simulated data sets and a real problem related to decadal climate variability: the Sea Surface Temperature variability over the last 80 years as simulated by the IPSL-CM5A coupled ocean-atmosphere general circulation model (OAGCM, Dufresne et al. 2013) that participated to the last Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al. 2012). Finally conclusions and directions for future work are given (section 5).

2 Background

2.1 Self-organizing maps

The Self-Organizing Maps (SOMs), [Kohonen, 1998], are powerful unsupervised learning neural networks that achieve both tasks of clustering and non-linear projection. They are used to reduce, visualize and interpret high dimensional data in a low-dimensional space. In the process, a multidimensional database is partitioned into clusters and simultaneously non linearly projected onto a lower dimensional map of neurons (denoted \mathcal{C}) characterized by a topological order. We note \mathcal{D} the matrix of data whose columns are formed with the variables $v_j \in V$ and the lines with the individuals $z_i \in \mathcal{Z}$ where \mathcal{Z} represents the set of n individuals, and V the set of p variables. \mathcal{C} is a regular grid of usually one or two dimensions considered as an indirect graph. A topological relation between neurons is described by a neighborhood kernel function $\mathcal{K}^T(\delta(c, r))$ where δ evaluates a distance between the neurons c and r corresponding to the length of the shortest path on the grid between the two neurons. The SOM enables a partitioning of \mathcal{Z} into clusters such as each subset is associated with a neuron c of the map \mathcal{C} and is represented by a so called referent vector $w_c \in W$ with W corresponding to the set of referents associated with the map \mathcal{C} . Each vector $z_i \in \mathcal{Z}$ is assigned to the neuron whose referent is the closest according to the euclidian distance, and is called the projection of the vector z_i on the map. The referent vectors W , reflect the distribution of the input vectors. A fundamental property of the SOM is the topological order provided at the end of the learning phase and the strong constraint on the distribution of the map in the data space. Indeed, the organization is such that two “relatively” close neurons (with respect to δ) on the map are associated with close vectors of \mathcal{Z} . The estimation of the referent vectors of a SOM and the topological order are made through a learning algorithm applied to \mathcal{Z} through a minimization process. The cost function is of the form:

$$J_{\text{Som}}^T(\chi, W) = \sum_{z_i \in \mathcal{Z}} \sum_{c \in \mathcal{C}} K^T(\delta(c, \chi(z_i))) \|z_i - w_c\|^2 \quad (1)$$

In (eq. 1), the topological order is introduced by using the neighbourhood function $K^T(\delta(c, \chi(z_i)))$, where K^T is a kernel function generally taken as a Gaussian. T is a neighborhood parameter that changes with the iterations and converges toward zero allowing the control of the neighborhood radius at each iteration. Each vector z_i is assigned to the closest neuron, represented by its referent $w_{\chi(z_i)} \in \mathcal{W}$, using the assignment function $\chi(z_i)$. At the end of the minimization, the value of a referent associated with a neuron is a weighted average of the vectors assigned to this neuron and their neighbourhood on the map (see eq. 2).

$$w_c = \frac{1}{\sum_{z_i \in \mathcal{Z}} K^T(\delta(c, \chi(z_i)))} * \left(\sum_{z_i \in \mathcal{Z}} K^T(\delta(c, \chi(z_i))) * z_i \right) \quad (2)$$

2.2 ClustVarLV

[Vigneau and Qannari, 2003] proposed a method for clustering variables which consists in directly searching a specific partitioning (clusters) for which the data

subsets are highly correlated to their referents. ClustVarLV works in the case for which both positive and negative correlations are present. The method searches K clusters C_1, \dots, C_k and K corresponding latent variables w_1, w_2, \dots, w_K according to their degree of correlation. The aim is to maximize the correlation between a cluster of variables and the corresponding latent variables by maximizing a cost function J_{cv} . Since we focused on algorithms dedicated to the clustering of variables, the data set should be normalized through a pre-processing phase, in which the standard deviation σ_{v_j} of each variable v_j is set equal 1. In this version of the algorithm, J_{cv} clusters the variables that share similar information by maximizing the correlation whatever its sign, the distinction between positive or negative correlation being addressed later on.

$$J_{cv} = \sum_{k=1}^K \sum_{j=1}^P \delta_{kj} Cov^2(v_j, w_k) = \sum_{k=1}^K \sum_{j=1}^P \delta_{kj} Corr^2(v_j, w_k) \quad (\text{with } \sigma_{v_j} = \sigma_{w_k} = 1) \quad (3)$$

where w_k is latent variable (referent) of cluster k , and $\delta_{kj} = 1$ if the j^{th} variable belongs to the cluster C_k and $\delta_{kj} = 0$.

The algorithm maximizes J_{cv} using an iterative procedure repeated until convergence, i.e. when a local maximum is reached. The ClustVarLV algorithm is a two step algorithm: the assignment step gathers variables that are similar while the maximization step provides for each subset a variable representative of this new cluster (the referent). After convergence, each latent variable is expressed as a linear combination of a subset of the initial variables. The partitioning of the variables into a disconnected partition makes more easy the interpretation of the underlying structure of the data.

ClustVarLV algorithm:

1. Start with K groups of variables which are obtained by a random allocation of the variables into K groups. (C_1, \dots, C_K), or are derived from a hierarchical clustering method.
2. Until $J_{CV}(\chi^{t-1}, W^{t-1}) - J_{CV}(\chi^t, W^t) = 0$ iterate:
 - (a) For each k , compute the new referent w_k as the first standardized principal component of the matrix D_k , corresponding to the standardized eigen vector associated to the largest eigenvalue of the matrix $D_k D_k'$ where D_k is the sub matrix of D composed only by the variables belonging to C_k and D_k' its transpose.
 - (b) Assign each variable to the cluster C_k , corresponding to the the highest squared correlation $corr(v_j, w_k)^2$. This determines the new clusters of variables.

In the ClustVarLV algorithm, [Vigneau and Qannari, 2003] proposed to add, if necessary, a step with a hierarchical clustering method in order to reduce the number of cluster C_k . They used the same strategies as those used for individual clustering: Ward criterion, minimum linkage, maximum linkage or average linkage.

3 SomVar approach

In the present paper, we propose a new method (SomVar) for clustering variables sharing the topological order of the Kohonen maps. The topological structure of the map \mathcal{C} introduces a discrete distance δ between referents. This topological order is introduced in a new cost function $J_{SV}^T(\chi, W)$ by the neighborhood function K^T and the parameter T . $J_{SV}^T(\chi, W)$ uses the neighborhood function in a similar manner as in $J_{SOM}^T(\chi, W)$, but maximizes the correlations between the variables and the latent variables(referents):

$$J_{SV}^T(\chi, W) = \sum_{v_j \in V} \sum_{c \in C} K^T(\delta(c, \chi(v_j))) * Corr(v_j, w_c)^2 \quad (4)$$

where $\delta(c, \chi(v_j))$ is the discrete distance between the neuron c and the neuron associated with v_j ($\chi(v_j)$).

The SomVar algorithm can be summarized as follows:

1. Start with K clusters of variables which may be obtained by a random allocation of the variables into K groups, or from a hierarchical clustering method
2. Until $K^T(\delta(k, j)) = \delta_{kj}$ (see eq. 3) and $J_{SV}^T(\chi^{t-1}, W) - J_{SV}^T(\chi^t, W) = 0$, iterate:
 - (a) In cluster C_k , w_k is the standardized eigen vector associated to the highest eigen value of DM^kD' where M^k is a diagonal matrix of weights. In other words, w_k is the first eigen vector of the D' weighted PCA using M^k as metric (see eq. 5).

$$M^k = \begin{cases} m_{ji}^k = 0, i \neq j \\ m_{jj}^k = \frac{1}{\sum_j K^T(\delta(\chi(v_j), k))} * K^T(\delta(\chi(v_j), k)) \end{cases} \quad (5)$$

- (b) The new clusters of variables are determined by assigning a variable to the cluster k with the highest weighted squared correlation with the referent w_k . The assignment function is then:

$$\chi(v_j) = \operatorname{argmax}_r \sum_{c \in C} K^T(\delta(r, c)) * corr(v_j, w_c)^2 \quad (6)$$

- (c) Decrease the value of T slowly.

Convergence of the SomVar method at T fixed

We denote χ^{t-1}, W^{t-1} (resp. χ^t, W^t), the assignment function and the set of referents at iteration t-1 (resp. time t). For a fixed T , each step of the algorithm (Maximization and Assignment) allows an increase of the criterion $J_{SV}^T(\chi, W)$ since a variable switches from one cluster to another only if its correlation with the new referent is higher than the old one.

From the iteration t-1 to the iteration t we have:

- Using the assignment function χ^{t-1} , the Maximization looks for each cluster C_k , the new w_k^t that maximizes the quantity:

$$\sum_{v_j \in D} K^T(\delta(k, \chi^{t-1}(v_j))) * \text{corr}(v_j, w_k^t)^2 \quad (7)$$

thus

$$J_{SV}^T(\chi^{t-1}, W^{t-1}) \leq J_{SV}^T(\chi^{t-1}, W^t) \quad (8)$$

The solution of the maximization step is the eigen vector associated with the largest eigenvalue of the weighted PCA, where the weights are given by the M^k matrix (see eq. 5).

- From this new optimal set of referents W^t , the Assignment step assigns a variable v_j whose the associated referent is w_k to a new referent $w_{k'}$ with $k \neq k'$ if and only if:

$$\sum_{c \in C} K^T(\delta(c, k)) * \text{corr}(v_j, w_k)^2 < \sum_{c \in C} K^T(\delta(c, k')) * \text{corr}(v_j, w_{k'})^2 \quad (9)$$

thus

$$J_{SV}^T(\chi^{t-1}, W^t) \leq J_{SV}^T(\chi^t, W^t) \quad (10)$$

For a given value of T , if we iterate the two step procedure, SomVar provides an increase of the cost function J_{SV}^T until convergence.

$$J_{SV}^T(\chi^{t-1}, W^{t-1}) \leq J_{SV}^T(\chi^{t-1}, W^t) \leq J_{SV}^T(\chi^t, W^t) \quad (11)$$

At the end of the process, SomVar is similar to ClustVarLV algorithm for which the cost function J_{cv} is maximized and allows to increase the correlation between the variables of each class and their referents. The neighborhood parameter T , which progressively decreases to zero, provides two improvements with respect to the ClustVarLV algorithm. Firstly, by contrast with SomVar, ClustVarLV which is a K-means like algorithm, is more sensitive to the initialization step [Celebi et al., 2013]. By forcing an early individual convergence of the referents, the K-means process is known to be highly sensitive to the initial values of the referents and may converge towards non significant local minima. On the con-

trary, as showed by [Bação et al., 2005], the topological maps are more robust and allow to avoid local minima by constraining the referents to consider the global data space and to adjust together during the learning phase before an individual convergence of the referents at the end of the learning phase. Moreover the second major benefit is the topology that presents a clear representation of the relationship between the clusters: close neurons on the SomVar map represent close clusters of variables. This topological map is of great interest for the user when analyzing and interpreting his data, many applications in various fields take advantage of it [Risien et al., 2004, Liu and Weisberg, 2005, Liu et al., 2006, Park et al., 2006].

4 Application

In order to evaluate the usefulness of SomVar, we first (section 4.1) simulated different classification problems by varying the number n of classes and compared the performances with those of ClustVarLV. In a second part (section 4.2), we applied the SomVar approach to a real environmental problem: The Pacific sea surface temperature variability simulated by the IPSL-CM5A OAGCM from 1940-2020. This second application highlights the benefit of using the topological order for the understanding of geophysical phenomenon and more generally of real data. We showed that the information provided by the topological order on the SomVar map is useful for understanding and interpreting the functioning of the geophysical phenomenon under study.

4.1 Simulated data

ClustVarLV has already been proved to give good performances when compared to other variable clustering methods like the well known VarClus function of the SAS computer application [Vigneau and Qannari, 2003, Sarle, 1990]. In the following we compared SomVar and ClustVarLV using the same number of clusters. In the experiments, ClustVarLV uses $k=16$ clusters and SomVar map is a 2D-grid of 4×4 neurons, the kernel function used for the topological order is the gaussian function [Lee and Verleysen, 2002, Ritter et al., 1992].

We simulated 20 random data sets of variables for each of the n -class classification problem, I1 (resp F1) stands for the 2 class problem; I2 to I10 (resp F2 to F10) represent the 3-class to 11-class problems. All the data sets I1 to I10 are made of 500 individuals and 50 variables while the data sets F1 to F10 are composed of 500 variables and 50 individuals.

For each data set p and each class c we randomly simulated a mean variable ($mean_varclass_p^c$). The $mean_varclass_p^c$ is simulated taking at random the first individual ($mean_varclass_p^c(0)$) according to the $\mathcal{N}(0, 1)$ distribution. The following individuals of the $mean_varclass_p^c$ are then simulated:

$$mean_varclass_p^c(i) = mean_varclass_p^c(i - 1) + \mathcal{N}(0, 1) \quad (12)$$

Then a variable that belongs to the same class is simulated, by adding a $\mathcal{N}(0, 1)$ noise to $mean_varclass_p^c(i)$ individual by individual. To illustrate the classification problems we displayed in Fig1 the correlations matrix for the 50 variables from I1 to I4 data sets and the 500 variables from F1 to F4 data sets. The variables being ordered by classes, we can easily see the squares of high correlation level on the second diagonals that correspond to correlation of variables belonging to the same class.

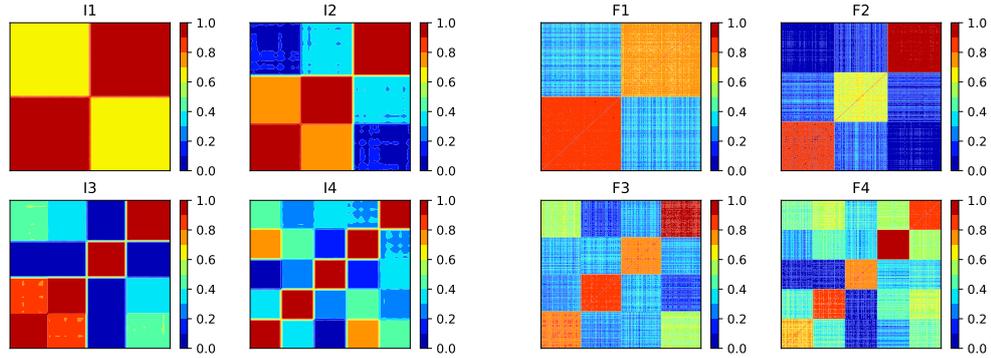


Fig. 1 Square correlations between the 50 variables for I1 to I4 and the 500 variables for F1 to F4 . The origine of the figure is the left bottom corner, the variables are ordered from the origine. The correlations of the variables are plotted from the left to the right for the x-axis and upward for the y-axis. The high correlations on the second diagonal emphasis the intra-class correlations.

As we know that the initialization of the learning phase could influence the results, we made for each data set and each algorithm (SomVar, ClustVarLV) 20 runs with a random initialization. At the end of the learning phase we evaluate the ability of both algorithms to retrieve the simulated classes. In the final classification, each cluster takes the label given by the majority of the individuals it has captured. The performances of the unsupervised classification is estimated by computing the miss classification rate for each of the independent run, Tab.1 presents the mean for the 20 runs for each method and each problem.

Table 1 Comparison of ClustVarLV and SomVar misclassification percentage: mean of 20 independante runs.

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
SomVar	0	0	0	0	0	0	0	10	10	2
ClustVar	0	0	6	3	10.6	16.2	19.2	14	17	26
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
SomVar	0	0	0	0	0	0	0	0	0	0
ClustVar	0	0	0	0	0.19	0.01	3.71	6.2	1.52	7.26

Tab.1 shows that, for both algorithms, the miss-classification rate increases when the number of classes increases. Independently of the number of variables or indi-

viduals SomVar always outperform ClustVarLV. We also note that the percentage of miss-classified variables are weaker for the F_i data sets where we have more variables than individuals.

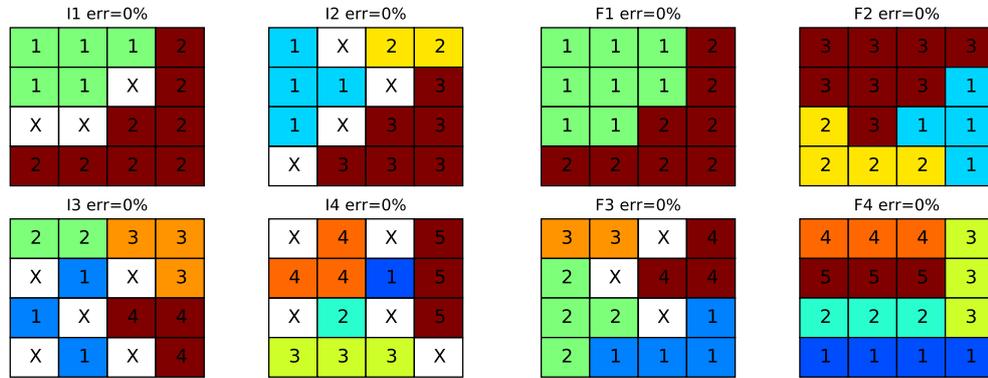


Fig. 2 SomVar Map of size 4x4 for a particular run on I1 to I4 and F1 to F4. The number and color corresponding to each neuron indicates the label taken by the neuron when applying the majority vote.

In Fig 2, we illustrated the SomVar topology order by displaying an example of the SomVar Maps obtained indicating the name and the color code of the class for each neuron. We can see that each class is represented by close neurons on the maps. We also see that the SomVar topological order give empty neurons located at the boundaries of the different classes enlightening the multi-dimensional distribution of the data. The visualizations on the Maps permit a good assessment of the modes of the distributions under study.

4.2 Sea Surface Temperature variability simulated by a CMIP5 coupled model

In this section we applied SomVar to a specific actual problem linked to climate sciences: The SST decadal variability in the Pacific as simulated by an OAGCM. OAGCMs are usually used to simulated recent climate change and provide future climate estimates and are based on physical laws that control the circulation and thermodynamics of the atmosphere, ocean, land and sea ice. We analyzed results of an ensemble of 30 historical simulations performed with the Institut Pierre Simon Laplace (IPSL) coupled Ocean-Atmosphere General Circulation Model, the IPSL-CM5A-LR OAGCM [Dufresne et al., 2013] that participated to the last Coupled Model Intercomparison Project Phase 5 (CMIP5; [Taylor et al., 2012]). The simulations cover the 1940-2020 period and use the CMIP5 historical natural (volcanoes) and anthropogenic greenhouse gases (GHG), tropospheric aerosols and ozone forcing from 1940 to 2005 and RCP8.5 from 2006 to 2020 [Riahi et al., 2011]. Volcanic stratospheric aerosol forcing is applied by prescribing the aerosol optical depth, as given in [Sato et al., 1993] (2012 updated version). Each member of the

historical ensemble was initialized from a set of 6 historical simulations with random initial conditions in ocean and atmosphere starting in 1940. In the following analyses, we focus only on the 30-members ensemble mean in order to increase the signal-to noise ratio of the forced variability due to applied external forcings with respect to the noise of internal variability. We analyzed the SST variability in the Pacific Ocean between 30N-40S and 140E-65W, for the winter (from December to February, DJF hereafter) and summer (from June to August, JJA hereafter) seasons. The winter and summer learning data sets correspond to matrix each composed of 80 regional maps including 11088 grid points (each grid points corresponding to a SST time series of 80 years). In order to better evidence the SST variations, we computed the SST anomalies by subtracted to each SST time-series their respective averaged values over the 1940-1960 period. DJF and JJA SST are matrices with 11088 columns (grids points considered as the variables) and 80 lines (the number of years). Our aim is to determine geographical regions where SST co-varies in time throughout the period. We applied the SomVar to both DJF and JJA matrix in order to get regions (set of variables) which co-vary in time and maximize the correlation with their corresponding latent variables during the last 80 years. It is found that SomVar is able to determine these regions and that the SomVar topological order can be interpreted in term of geophysics processes.

4.3 Results

Similarly to the precedent experiments we used a 4x4 SomVar Maps. Results obtained for DJF and JJA seasons are shown on (Fig 3) and (Fig 4) respectively. We observe that the SomVar clustering gives geographical regions well defined. The times series corresponding to each neuron reveal a positive SST trend which is consistent with increasing atmospheric GHG forcing due to burning fossil fuel by human activities throughout the simulated period for both seasons. We also see three abrupt SST drops lasting only a few years caused by the volcanic forcing of three volcanic eruptions, Agung (February 1963), El chichon (March 1982) and Pinatubo (June 1991). In DJF (JJA), the neuron 1 (neuron 3) reveals a geographical region with the smallest SST positive trend contrasting with the tendencies of the others neurons [Falvey and Garreaud, 2009].

Table 2 The colors, names and referents number composing the different geographical areas presented in figures 3 and 4 for the 2 seasons DJF and JJA

Color	DJF	JJA
Red	Southeastern Pacific (1)	Southeastern Pacific (3)
Blue	South 2 (13,14,15,16)	X
Magenta	Tropic (2,5,6,9)	South (4,7,8,11)
Orange	North (10, 11, 12)	North (9,10,12,13,14,15,16)
Brown	South 1 (3,4,7,8)	Tropic (1,2,5,6)

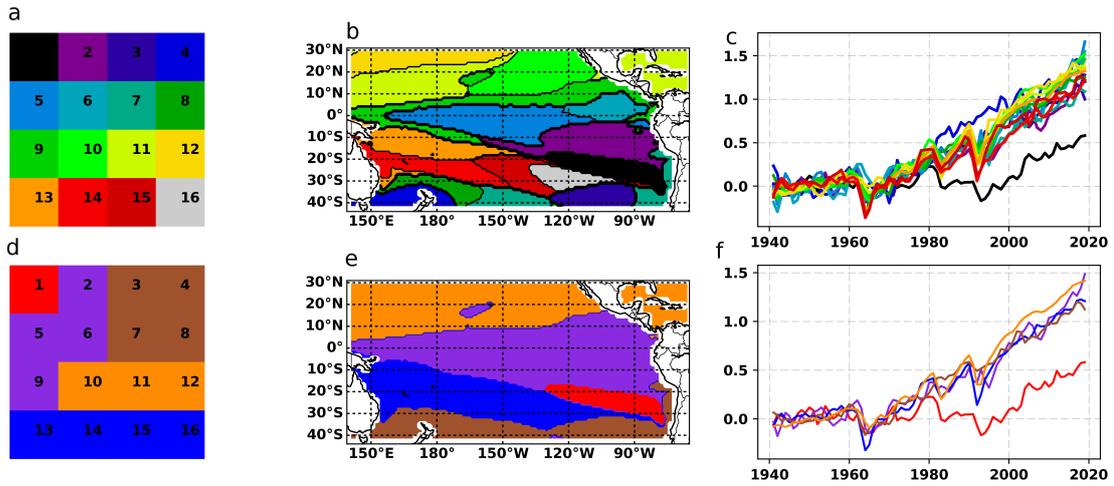


Fig. 3 DJF SomVar clustering. **a)** SomVar maps and the colors associated with each neuron. **b)** Geographic areas corresponding to each neuron. **c)** Time series corresponding to each neuron. **d-e)** Same presentation as (a-c) but after grouping the neurons.

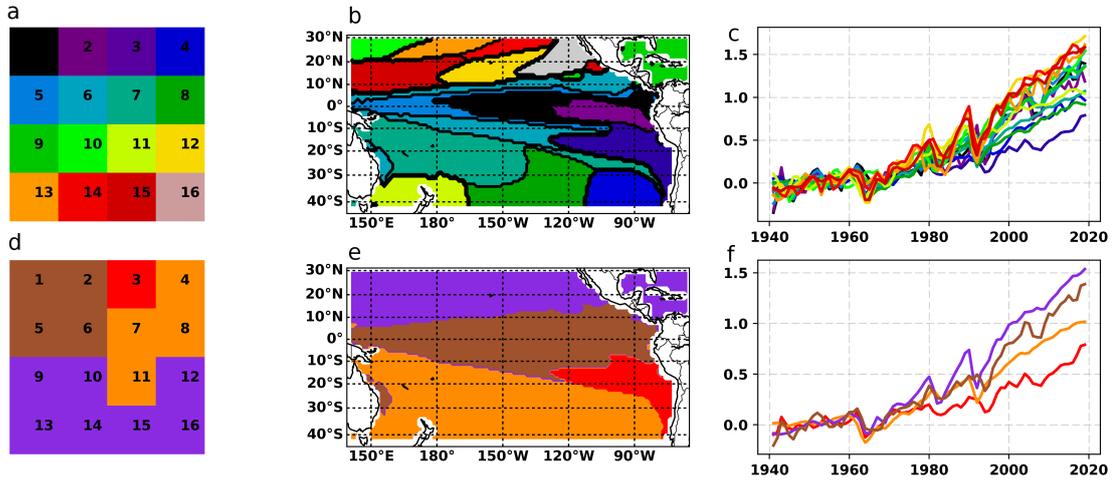


Fig. 4 Same as figure 3, but for JJA.

Inspection of the geographic map shows that closest clusters (Fig. 3d and Fig. 4d) are grouped by the SomVar clustering (Fig. 3e and Fig. 4e) based on close co-variability consistent with tropical and extra tropical Pacific climatic regions. If we consider the groups of neurons as described in Table 2, we get well defined large scale oceanic patterns (Northern extra tropics, Southern extra tropics, Tropic and the Southeastern Pacific region (SEP)) with distinct responses to the external forcing.

We note that the South Pacific region is represented by two distinct groups of neurons (or clusters) for DJF and JJA which is consistent with the different climate conditions during both seasons. As evidenced by the time series of the different ocean areas, GHG and volcanic forcing do not have the same signature in each regions that were discriminated by SomVar. We clearly see than the SST increase

is not uniform around the world; it is larger in the North Pacific than in the South [Boer et al., 2000] for DJF and JJA with a minimum SST increasing trend close to the South American coast compared to regions much offshore in the Pacific Ocean [Falvey and Garreaud, 2009]. We also clearly detect the cooling impacts of the three major volcanic eruptions which differs from one region to another. The volcanic impact is much more stronger in the South, in cluster 2 as compared to cluster 1 in DJF. This explains why in DJF the southern regions have been clustered in two groups of referents which are not close to each other in the SomVar map. The different groups (five for DJF and four for JJA), clearly indicate that close neurons on the SomVar map are linked with close geographical areas sharing similar geophysical properties: the topological order accurately partitions the continuous variability of the geophysical phenomena.

In this Section, we showed that SomVar is able to delimit coherent ocean areas with specific SST variations. Moreover the topological order given by the SOM-map is coherent with the SST co-variabilities. The SomVar SST study allowed us to evidence the Pacific sea surface temperature response to climate changes over the historical period. To better understand the physical processes responsible of the SST variability simulated by the IPSL-CM5A-LR model since 1940, it should be necessary to analyze other variables related to the Southeastern Pacific dynamics, such as surface wind stress and the ocean vertical transport. Obviously, SomVar is able to analyze real data by classifying the variables to give clusters with a significant physical meaning.

5 Conclusion

In this paper we have presented a new neural algorithm (SomVar) that uses a correlation distance to reduce a large set of variables into clusters having a well defined topological ordering. In each cluster, the variables are highly correlated. They are represented by a referent (or latent variable) that is a linear combination of the initial variables. SomVar is a robust algorithm, weakly sensitive to the initial conditions, which retrieves the hidden structures of the variables and can handle a large number of referents. Besides SomVar provides an ordering that can be interpreted by taking into account the physical significance of the problem described by the data. SomVar and its topological order permitted to analyze the regional responses of Pacific ocean SST variations to external forcing in the context of climate change. The SST time series associated with these regions revealed the different regional responses to volcanism on one hand and to increasing atmospheric GHG concentrations on the other hand. Our analyses allowed identifying the very specific ocean response in the Southeastern Pacific.

It is clearly demonstrated in this paper that SomVar is able to extract information from climate timeseries. A weakness of this method is that we only consider the eigenvectors associated with the highest eigenvalues without considering the percentage of explained variance. Moreover as climate studies should take into account a large variety of distinct variables, it will be of interest to study how to introduce this constraint in the clustering process. An improvement of this method

could be to take into consideration other Principal Components according to their explained variance.

References

- [Bação et al., 2005] Bação, F., Lobo, V., and Painho, M. (2005). Self-organizing maps as substitutes for k-means clustering. In *International Conference on Computational Science*, pages 476–483. Springer.
- [Boer et al., 2000] Boer, G., Flato, G., and Ramsden, D. (2000). A transient climate change simulation with greenhouse gas and aerosol forcing: projected climate to the twenty-first century. *Climate Dynamics*, 16(6):427–450.
- [Celebi et al., 2013] Celebi, M. E., Kingravi, H. A., and Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert systems with applications*, 40(1):200–210.
- [Dufresne et al., 2013] Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., et al. (2013). Climate change projections using the ipsl-cm5 earth system model: from cmip3 to cmip5. *Climate Dynamics*, 40(9-10):2123–2165.
- [Falvey and Garreaud, 2009] Falvey, M. and Garreaud, R. D. (2009). Regional cooling in a warming world: Recent temperature trends in the southeast pacific and along the west coast of subtropical south america (1979–2006). *Journal of Geophysical Research: Atmospheres*, 114(D4).
- [Jain, 2010] Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666.
- [Kohonen, 1982] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69.
- [Kohonen, 1998] Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, 21(1-3):1–6.
- [Kotsiantis et al., 2006] Kotsiantis, S. B., Zaharakis, I. D., and Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review*, 26(3):159–190.
- [Lee and Verleysen, 2002] Lee, J. A. and Verleysen, M. (2002). Self-organizing maps with recursive neighborhood adaptation. *Neural Networks*, 15(8-9):993–1003.
- [Liu and Weisberg, 2005] Liu, Y. and Weisberg, R. H. (2005). Patterns of ocean current variability on the west florida shelf using the self-organizing map. *Journal of Geophysical Research: Oceans*, 110(C6).
- [Liu et al., 2006] Liu, Y., Weisberg, R. H., and Mooers, C. N. (2006). Performance evaluation of the self-organizing map for feature extraction. *Journal of Geophysical Research: Oceans*, 111(C5).
- [Park et al., 2006] Park, Y.-S., Tison, J., Lek, S., Giraudel, J.-L., Coste, M., and Delmas, F. (2006). Application of a self-organizing map to select representative species in multivariate analysis: a case study determining diatom distribution patterns across france. *Ecological Informatics*, 1(3):247–257.
- [Riahi et al., 2011] Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P. (2011). Rcp 8.5—a scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1-2):33.
- [Risien et al., 2004] Risien, C. M., Reason, C., Shillington, F., and Chelton, D. B. (2004). Variability in satellite winds over the benguela upwelling system during 1999–2000. *Journal of Geophysical Research: Oceans*, 109(C3).
- [Ritter et al., 1992] Ritter, H., Martinetz, T., Schulten, K., Barsky, D., Tesch, M., and Kates, R. (1992). *Neural computation and self-organizing maps: an introduction*. Addison-Wesley Reading, MA.
- [Sarle, 1990] Sarle, W. (1990). The varclus procedure. sas/stat user’s guide. sas institute. Inc., Cary, NC, USA.
- [Sato et al., 1993] Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B. (1993). Stratospheric aerosol optical depths, 1850–1990. *Journal of Geophysical Research: Atmospheres*, 98(D12):22987–22994.
- [Taylor et al., 2012] Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

[Vigneau and Qannari, 2003] Vigneau, E. and Qannari, E. (2003). Clustering of variables around latent components. *Communications in Statistics-Simulation and Computation*, 32(4):1131–1150.

